

DOI:10.13476/j.cnki.nsbdkk.2022.0047

刘欣欣,马细霞,程旭,等.月径流随机模拟序列优选方法[J].南水北调与水利科技(中英文),2022,20(3):468-477. LIU X X, MA X X, CHENG X, et al. Optimization method for monthly runoff stochastic simulation sequence[J]. South-to-North Water Transfers and Water Science & Technology, 2022, 20(3): 468-477. (in Chinese)

月径流随机模拟序列优选方法

刘欣欣¹, 马细霞^{1,2}, 程旭¹, 王倩丽¹, 张静文¹

(1. 郑州大学水利科学与工程学院, 郑州 450001; 2. 郑州大学黄河生态保护与区域协调发展研究院, 郑州 450001)

摘要:针对月径流随机模拟序列检验指标偏少、检验结果未能全面反映径流过程复杂变化的问题,在传统检验指标的基础上,引入集中度、不均匀系数和样本熵等指标,提出截口-复杂性指标体系,基于该指标体系对径流序列截口特性、复杂变化特性进行分析,通过灰色关联分析法对月径流模拟序列进行综合评价和优选,并与传统指标优选的结果进行对比及合理性分析。以淮河流域干支流 3 个研究站点资料为例建立季节性自回归模型,对随机生成的多组月径流模拟序列进行应用研究,结果表明:与传统评价方法相比,截口-复杂性指标体系优选出的模拟序列在保证年内分配多样性的同时,能够减少年内分配不合理的情况,较好地反映径流序列的复杂特性,为识别更为准确、可靠的月径流随机模拟序列提供了一种新途径。

关键词:随机模拟;模拟检验;指标体系;灰色关联分析;序列优选

中图分类号:TV213 文献标志码:A 开放科学(资源服务)标志码(OSID):



径流序列作为水资源系统的输入项,通常用来分析系统未来的发展趋势。通过水文随机模型生成足够长的径流序列,可以弥补实测资料代表性不足的缺陷,已广泛应用于水库(群)优化调度^[1-2]、水资源系统风险分析^[3-4]等领域。

目前,水文随机模型的精度评价常采用“统计一致性”准则,即生成的模拟序列必须和实测资料具有相近的统计特征。如:Weiss 等^[5]以均值、均方差、滞后 1 和滞后 2 的自相关系数作为检验指标,对实测序列和模拟长序列的高分位数进行比较,评估季节性自回归模型[SAR(p)]保持实测序列统计特征方面的性能。Hao 等^[6]以科罗拉多河为研究对象,通过分析月径流模拟序列基本统计参数的相对误差,发现平均值、标准差和偏度的相对误差均小于 10%,滞后 1 的自相关系数相对误差均小于 20%,由此得出了模拟序列能较好地保持实测序列截口统计特征的结论。Hamid 等^[7]采用平均绝对百分

比误差(E_{MAP})等方法来分析模拟结果的均值、标准差、偏态系数和最大值,根据 E_{MAP} 的最小值表示最佳模式,优选出最适合的季节性自回归综合移动平均模型。高瑞忠等^[8]采用平均相对误差和平均绝对误差方法,评价巴拉格尔河流域年季尺度下的径流量模拟效果,据此得出了神经网络方法对该流域径流量进行模拟总体适用性较好的结论。传统评价方法在保证模拟系列均值、方差等基本统计特征较为一致的情况下,能够模拟出具有不同时空分配特性的径流过程,对水库(群)优化调度、水资源系统风险分析、水资源工程规模确定具有重要意义。

径流过程是十分复杂的随机过程,但也具有一定的统计规律,譬如年内分配过程具有汛期和非汛期的交替变化特征,同一地区汛期最大 4 个月所处月份基本固定,多年平均最大 4 个月水量占全年平均水量的比例也较为稳定。径流模拟序列具有不同

收稿日期:2021-06-01 修回日期:2021-11-17 网络出版时间:2021-11-23

网络出版地址:https://kns.cnki.net/kcms/detail/13.1430.TV.20211123.1103.002.html

基金项目:河南省科技攻关项目(192102310228);中国工程院重大咨询研究项目(2020-ZD-18-5)

作者简介:刘欣欣(1996—),女,河南济源人,主要从事水文学及水资源研究。E-mail:654566831@qq.com

通信作者:马细霞(1963—),女,河南叶县人,教授,硕士生导师,主要从事水文学及水资源教学与科研工作。E-mail:maxx@zhu.edu.cn

的时空分配特性,若仅保证模拟系列均值、方差等基本统计特征较为一致,则有可能存在某些年份的年内分配过程不合理问题,因此基于此模拟序列获得的应用成果存在一定的安全风险。为此,马细霞等^[9]引入样本熵指标对径流序列的复杂特性进行检验,结果表明样本熵指标能大体诊断出模拟序列内较大月径流量和位置的异常情况,但其合理性没有得到充分论证。

以往评价方法的不足之处在于未能全面检验模拟序列年内复杂变化的特性,因此,有必要增加一些反映径流时序变化的特征指标,如连续最大4个月径流百分率^[10]、径流年内分配不均匀系数、径流年内集中度^[10-11]等,在保证模拟序列年内分配多样性的同时,排除年内分配异常的情况。此外,同一随机模型所生成的模拟序列在时程变化上差异较大^[9],以往的评价方法在工程规划设计和风险分析方面,未能筛选出最适宜的径流模拟序列。鉴于此,本文提出截口-复杂性指标体系,以淮河流域河南段干支流上的典型水文研究站月径流随机模拟为应用实例,对季节性自回归模型模拟序列的截口特性和复杂特性进行全面检验,并基于灰色关联分析法对以往指标体系和截口-复杂性指标体系的计算结果进行优选和分析,以期为月径流随机模型模拟序列优选提供理论依据。

1 模拟序列评价指标体系及指标量化

1.1 评价指标体系构建

为全面评价模拟序列,本文从以下2个方面构建截口-复杂性指标体系。截口特征方面:均值、变差系数、偏差系数反映序列截口统计特性,相邻截口滞时1、2的自相关系数反映径流序列的时间相依特性。这些指标在随机模拟序列检验中已得到广泛应用,故本文仍将其作为月径流随机模拟效果评价的主要指标。复杂特性方面:河川径流的时空分布规律是水资源评价的重要内容,也是河流类型和水利规划的一项重要指标^[12],本文引入连续最大4个月径流百分率、集中度和不均匀系数作为反映径流变化过程的指标,以期更好地反映实测序列与模拟序列年内分配的复杂特征。样本熵表示非线性动力学系统产生新信息的概率,主要用来定量地刻画系统的规则度及复杂度^[13],当用于度量时间序列复杂性时,其值越大,序列产生新信息的概率越高,序列越复杂^[14]。由于其较少地依赖于时间序列长度,所需数据序列短、计算成果更为稳定,已在水文气象研究领域广泛应用,因此可以作为描述径流序列复杂性

特征的指标^[15]。传统评价指标与截口-复杂性指标体系对比见表1。

表1 传统指标与截口-复杂性指标体系对比

Tab.1 Comparison of traditional index and the index system		
评价方面	传统评价指标体系	截口-复杂性指标体系
截口统计特性	均值(\bar{x})	均值(\bar{x})
	变差系数(C_v)	变差系数(C_v)
	偏态系数(C_s)	偏态系数(C_s)
	一阶自相关系数(r_1)	一阶自相关系数(r_1)
	二阶自相关系数(r_2)	二阶自相关系数(r_2)
复杂变化特性		连续最大4个月径流百分率(Q_4)
		集中度(C_d)
		不均匀系数(C_i)
		样本熵(H)

1.2 评价指标量化方法

截口参数采用统计学方法计算^[16],以下主要介绍样本熵、集中度和不均匀系数的计算方法。

1.2.1 样本熵计算

设原始数据为 $X(1), X(2), \dots, X(N)$, 共 N 点,可按如下步骤计算其样本熵^[17]。

按序号连续顺序,组成一组 m 维矢量,从 $X_m(1)$ 到 $X_m(N-m+1)$,其中

$$X_m(i) = [x(i), x(i+1), \dots, x(i+m-1)], \quad i=1 \sim N-m+1 \quad (1)$$

定义 $X(i)$ 与 $X(j)$ 之间的距离 $d[X_m(i), X_m(j)]$ 为两者对应元素中差值最大的一个,即

$$d[X_m(i), X_m(j)] = \max(|x(i+k) - x(j+k)|), \quad k=0 \sim m-1; i, j-1 \sim N-m+1; i \neq j \quad (2)$$

给定阈值 r ,对每一个 $1 \leq i \leq N-m$ 值,统计 $d[X_m(i), X_m(j)] < r$ 的数目(模板匹配数)及此数目与距离总数 $N-m+1$ 的比值,记作 $B_{m,i}(r)$

$$B_{m,i}(r) = \{d[X_m(i), X_m(j)] < r\} / (N-m+1), \quad j=1 \sim N-m, j \neq i \quad (3)$$

求其对所有 i 的平均值,即

$$B_m(r) = \{1/(N-m)\} \sum_{i=1}^{N-m} B_{m,i}(r) \quad (4)$$

将维数增加1,变为 $m+1$ 维矢量,重复式(1)~(3)的步骤后,得

$$B_{m+1}(r) = \{1/(N-m)\} \sum_{i=1}^{N-m} B_{m+1,i}(r) \quad (5)$$

式中: $B_m(r)$ 和 $B_{m+1}(r)$ 分别为 m 点和 $m+1$ 点两序列相似的概率,则序列理论上的样本熵

$$H(m, r) = \lim_{N \rightarrow \infty} \{-\ln[B_{m+1}(r)/B_m(r)]\} \quad (6)$$

当 N 为有限时,得出序列样本熵估计值

$$H(m, r, N) = -\ln[B_{m+1}(r)/B_m(r)] \quad (7)$$

参数 m, r 的选择是样本熵估计的关键,但目前尚无最佳标准,通常 $m=2, r=0.2S_D$ (S_D 为原始序

列的标准差)^[17]。

1.2.2 集中度计算

径流在年内的集中程度用集中度(C_d)表示,计算公式^[18]为

$$C_d = \frac{\sqrt{R_x^2 + R_y^2}}{R} \quad (8)$$

式中: R 为多年平均月径流计算得出的年径流量; R_x 、 R_y 分别为 12 个月份($i=1,2,\dots,12$)的分量之和所构成的水平、垂直分量,其中

$$R_x = \sum_{i=1}^{12} r_i \sin \theta_i, R_y = \sum_{i=1}^{12} r_i \cos \theta_i \quad (9)$$

式中: r_i 为第 i 月平均月径流; θ_i 为第 i 月份对应的方位角,以各月月中值代表的角度数值来表示。

1.2.3 不均匀系数计算

不均匀系数(C_t)的计算公式^[19]为

$$C_t(t) = \frac{\sigma(t)}{Q(t)} = \frac{\sqrt{\sum_{\lambda=1}^m [Q(t, \lambda) - Q(t)]^2 / m}}{\sum_{\lambda=1}^m Q(t, \lambda) / m} \quad (10)$$

式中: $C_t(t)$ 为第 t 年的径流年内分配不均匀系数; $\sigma(t)$ 为第 t 年的年内时段平均流量的均方差; $Q(t)$ 为第 t 年的平均流量, m^3/s ; λ 为月份, $\lambda=1,2,\dots,m$, 此处 $m=12$ 。径流年内分配不均匀系数反映了对径流调控的难易程度。对于 n 年系列的不均匀系数的均方差,其值越大,则表示径流不均匀性的年际变化越大。

1.3 评价指标分析方法

平均绝对百分比误差(E_{MAP})属于无量纲统计量,可以在不同变量间进行比较,其值越小,说明模型模拟或预测的精确度越高,因此可表征各评价指标中模拟序列的精度。 E_{MAP} 的计算公式为

$$E_{MAP} = \frac{\sum_{i=1}^n \frac{|A_i - F_i|}{A_i}}{n} \times 100\% \quad (11)$$

式中: A_i 为实测序列第 i 个指标计算值; F_i 为模拟序列第 i 个指标计算值。

2 随机模拟序列的灰色系统评价方法

随机模拟生成径流序列最基本的假定是未来事件应具有和实测序列相同的随机性质,这就要求生成的模拟序列必须和实测资料具有相近的统计特征,这也是优选随机模型和评价生成序列的主要依据。灰色关联分析法是一种衡量因素的发展态势和特征变化相似程度的方法,通过灰色关联度表征数据序列之间联系的紧密程度。本文采用灰色关联分

析法,计算模拟序列与实测资料在截口、复杂性指标的综合关联度,优选较为理想的随机模拟序列。灰色系统关联分析的具体步骤如下^[20]。

确定参考数列与比较数列。本文采用截口-复杂性指标体系中的 9 个指标作为月径流序列的特征指标,由实测序列统计所得的 9 个特征指标值构成参考数列 X_0 ,第 i 个随机模拟序列统计所得的 9 个特征指标值构成比较数列 X_i ($i=1,2,\dots,n$)。

$$X_0 = \{x_0(1), x_0(2), \dots, x_0(9)\} \quad (12)$$

$$X_i = \{x_i(1), x_i(2), \dots, x_i(9)\} \quad (13)$$

式中: $x_0(1), x_0(2), \dots, x_0(9)$ 分别为参考数列 X_0 的 9 个特征指标值; $x_i(1), x_i(2), \dots, x_i(9)$ 分别为比较数列 X_i 的 9 个特征指标值。

指标值的无量纲化处理。计算公式为

$$y_i(k) = \frac{x_i(k)}{x_0(k)}, x_i(k) \leq x_0(k) \quad (14)$$

$$y_i(k) = \frac{x_0(k)}{x_i(k)}, x_i(k) > x_0(k) \quad (15)$$

式中: $x_0(k)$ 为参考数列 X_0 的第 k 个特征指标值; $x_i(k)$ 为比较数列 X_i 的第 k 个特征指标值; $y_i(k)$ 为比较数列 X_i 第 k 个特征指标的效果测度; $k=1,2,\dots,9$ 。

求参考数列与比较数列的灰色关联系数。计算公式为

$$\xi_i(k) = \frac{\min_i \min_k |y_i(k) - 1| + \rho \max_i \max_k |y_i(k) - 1|}{|y_i(k) - 1| + \rho \max_i \max_k |y_i(k) - 1|} \quad (16)$$

式中: $\xi_i(k)$ 为参考数列 X_0 与比较数列 X_i 第 k 个特征指标的关联系数; ρ 为分辨系数,表示平均分辨水平,其值为 $0 \sim 1$,一般取 0.5 ^[20]。

求参考数列与比较数列的加权关联度。为便于从整体上比较模拟序列,有必要将各指标的关联系数集中为一个值,即加权关联度为 r_i ,其表达式为

$$r_i = \sum_{k=1}^9 \omega_i(k) \xi_i(k) \quad (17)$$

式中: $\omega_i(k)$ 为比较数列 X_i 第 k 个特征指标的权重,考虑到截口特性、复杂变化特性方面的各指标重要性相当,故本文采用等权处理,即: $\omega_i(k) = 1/9$ 。加权关联度越大,说明相应的模拟序列在各指标的综合评价中表现越好。

3 实例分析与讨论

3.1 研究区概况及数据来源

淮河是河南省内流域面积最大的水系,境内控制流域面积 8.83 万 km^2 ,占淮河流域总面积的

32%,占河南省总面积的53%。为验证本文方法的合理性和有效性,在淮河干流、沙河和史河支流上,分别选取流域面积相差较大的息县(10 190 km²)、漯河(12 150 km²)和蒋家集水文研究站(5 930 km²)月径流随机模拟为应用实例,以期为淮河流域中上游径流模拟序列选取提供理论支撑。以上3站实测年月径流资料(1951—2018,共68年)均来源于《河南省水情手册》,这些资料均通过了“三性”审查,可以用于模拟计算。

3.2 季节性自回归模型建立

根据RIC定阶方法^[21],计算漯河站、息县站、蒋家集站SAR(p)模型不同月份的RIC值。其中:漯河站和息县站12个月的RIC最小值对应阶数均为1阶;蒋家集站除11月份外,其余11个月RIC最小值对应的阶数均为1阶。因此,根据RIC准则,3个水文站均采用SAR(1)模型进行随机模拟。

3.3 随机模拟序列精度分析

通过所建SAR(1)模型,随机生成10组样本容量为680年的模拟序列,样本的选取具有随机性。为便于分析,将漯河站10组模拟序列记为1-1、1-2……1-10,息县站记为2-1、2-2……2-10,蒋家集站记为3-1、3-2……3-10。采用相对误差和平均绝对百分比误差(E_{MAP}),从截口、复杂变化方面分析模拟序列的拟合效果,并结合模拟序列多年平均及各水平年年内分配情况,验证截口、复杂变化特性指标评价的合理性。

3.3.1 截口统计特性精度分析

模拟序列截口指标的相对误差和 E_{MAP} 结果见表2。由表2可知,各水文站10组模拟序列在截口特性方面的精度存在差异,3个水文研究站的10组模拟序列的均值和变差系数相对误差均在15%范围内,偏态系数的误差相对较大,一阶自相关系数的误差平均值在8%以内,二阶自相关系数的误差平均值在15%范围内。 E_{MAP} 计算结果表明:漯河站 E_{MAP} 值最小的是1-5序列,其次是1-8、1-10、1-9序列;息县站 E_{MAP} 值最小的是2-1、2-10序列,其次是2-9序列;蒋家集站 E_{MAP} 值最小的是3-4、3-7序列,其次是3-1序列。与同一水文研究站的其他模拟序列相比,这些模拟序列在截口统计特性评价中表现较好。

3.3.2 复杂变化特性精度分析

模拟序列复杂变化指标的相对误差和 E_{MAP} 结果见表3。可以看出,同一水文研究站各模拟序列在复杂变化特性方面的精度存在差异。其中:3个

水文站连续最大4个月径流百分率、不均匀系数的相对误差在13%范围内;集中度、样本熵指标变化幅度较大,说明模拟序列在集中程度和复杂性方面差异较大。 E_{MAP} 计算结果表明:漯河站 E_{MAP} 值最小的是1-5序列,其次是1-2序列;息县站 E_{MAP} 值最小的是2-1序列,2-2、2-3序列次之;蒋家集站 E_{MAP} 值最小的是3-7序列,其次是3-9序列。与同一水文研究站的其他模拟序列相比,这些模拟序列较好保持了实测序列在形态变化方面的特性。

表2 模拟序列截口指标相对误差及 E_{MAP} 结果统计

Tab.2 Statistics of RE and E_{MAP} results of simulated sequence section index %

研究站点	模拟序列	截口指标					
		\bar{x}	C_v	C_s	r_1	r_2	E_{MAP}
漯河站	1-1	4	-14	-37	-14	-21	18
	1-2	5	-10	-27	-8	-19	14
	1-3	-1	-11	-34	-3	-3	10
	1-4	1	-10	-26	-9	-10	11
	1-5	5	-1	-6	-4	-16	6
	1-6	-6	-8	-15	-13	-20	12
	1-7	-2	-8	-23	-16	-20	14
	1-8	8	-5	-4	-4	-13	7
	1-9	-8	-11	-16	-6	-2	9
	1-10	2	-12	-20	-5	-1	8
息县站	2-1	-2	2	-21	-7	-3	7
	2-2	3	-12	-34	24	0	15
	2-3	5	-1	-14	20	35	15
	2-4	5	-11	-27	-12	-37	18
	2-5	2	2	-20	15	24	13
	2-6	2	-4	-20	-13	15	11
	2-7	-9	-2	-27	2	11	10
	2-8	-7	1	-20	-19	-13	12
	2-9	3	-5	-22	6	-4	8
	2-10	2	-3	-18	9	2	7
蒋家集站	3-1	6	-4	-14	-9	-7	8
	3-2	-1	-1	-19	22	17	12
	3-3	10	4	-20	7	10	10
	3-4	2	-5	-16	-3	-10	7
	3-5	5	-14	-31	5	23	16
	3-6	11	-6	-19	-4	39	16
	3-7	2	1	-4	-6	-22	7
	3-8	-2	-8	-26	20	27	17
	3-9	10	6	-7	20	27	14
	3-10	1	-8	-22	-6	-20	11

表 3 模拟序列复杂变化指标相对误差及 E_{MAP} 统计

Tab. 3 Statistics of RE and E_{MAP} results of complex change index of simulation sequence %

研究站点	复杂变化指标	模拟序列				
		Q_d	C_d	C_r	H	E_{MAP}
漯河站	1-1	-2	-12	3	30	12
	1-2	-2	3	8	-10	6
	1-3	-3	7	13	-20	11
	1-4	-3	-6	7	10	7
	1-5	0	2	4	4	3
	1-6	-7	-6	6	21	10
	1-7	0	6	7	14	7
	1-8	-4	12	9	2	7
	1-9	0	-17	-5	11	8
	1-10	-1	16	6	-21	11
息县站	2-1	0	-5	4	-1	3
	2-2	-1	2	5	7	4
	2-3	-2	-7	4	2	4
	2-4	-2	0	3	18	6
	2-5	-2	-6	5	-16	7
	2-6	-4	13	13	10	10
	2-7	5	-1	12	8	7
	2-8	4	-3	5	12	6
	2-9	2	11	8	1	6
	2-10	2	6	12	-12	8
蒋家集站	3-1	-3	6	12	-12	8
	3-2	-3	7	-4	-9	6
	3-3	-4	8	-1	-7	5
	3-4	-2	11	8	1	6
	3-5	-4	3	-6	16	7
	3-6	-12	-1	0	-1	4
	3-7	-2	-4	2	-1	2
	3-8	-2	-3	-4	11	5
	3-9	-7	3	-1	-2	3
	3-10	-2	3	-7	8	5

3.3.3 模拟序列精度综合分析

为验证截口-复杂性指标体系的合理性,采用距平百分率法对径流序列进行丰、平、枯划分,统计各站模拟序列多年平均及丰、平、枯水年年内分配的相对误差,结果见图 1。可以发现:漯河站模拟序列丰、平水年以及多年平均月径流的拟合效果相对较好,相对误差均在 20% 范围内,枯水年各序列间相对误差差异较大,年内分配总体拟合较好的是 1-5、1-7 序列;息县站模拟序列各水平年及多年平均的相对误差在 25% 范围内,丰水年的相对误差均在 12% 范围内,年内分配总体拟合较好的是 2-1 序列;蒋家集站模拟序列多年平均及各水平年的相对误差整体在 25% 范围内,各序列相对误差差异明显,年内分配总体拟合较好的是 3-4、3-5、3-7 序列。

对比模拟序列截口、复杂性精度分析结果与水平年精度分析结果(见表 4),可以看出:同一水文研究站的模拟序列在截口、复杂变化特性方面拟合效果相差各异。例如:漯河站模拟序列 1-8、1-9、1-10 在截口统计特性评价中 E_{MAP} 值较小,说明其较好的保持了实测序列截口统计特性,但在复杂变化特性和年内分配过程方面 E_{MAP} 值均较大;蒋家集站 3-7、3-9 序列复杂变化指标的 E_{MAP} 值相差较小,分别为 2% 和 3%,截口指标的 E_{MAP} 值分别为 7% 和 14%,上述序列在复杂变化特性方面拟合效果相差较小,但保持实测序列截口统计特性的效果相差较大。单独使用截口或复杂性指标进行评价,不能保证模拟序列年内分配方面的拟合效果。例如息县站的 2-2、2-10 序列,分别在复杂性、截口指标中 E_{MAP} 值较小,但其平、枯水年及多年平均情况下的相对误差均较大。综合上述分析,需要全面考虑模拟序列截口、复杂变化特性方面的拟合效果。

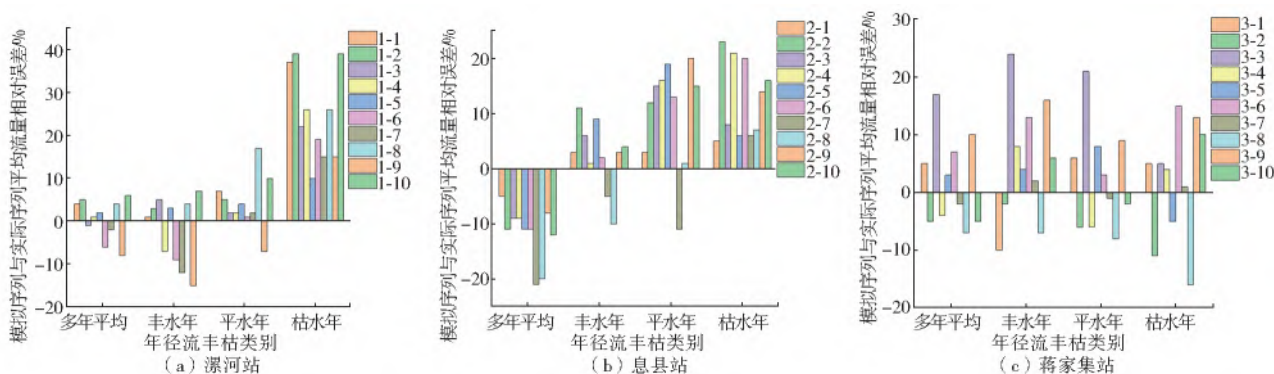


图 1 各站模拟序列多年平均以及丰/平/枯水年相对误差

Fig. 1 The relative errors of the simulated series of each station in annual average, wet, normal and dry years

表 4 模拟序列不同角度精度分析结果对比

Tab. 4 Comparison of accuracy analysis results of different angles of simulation sequence

研究站点	分析角度	优选序列
漯河站	截口指标	1-5、1-8、1-9、1-10
	复杂性指标	1-2、1-5
	水平年分配情况	1-5、1-7
息县站	截口指标	2-1、2-9、2-10
	复杂性指标	2-1、2-2、2-3
	水平年分配情况	2-1
蒋家集站	截口指标	3-1、3-4、3-7
	复杂性指标	3-7、3-9
	水平年分配情况	3-4、3-5、3-7

3.4 模拟序列优选

基于截口-复杂性指标体系,采用灰关联分析法对模拟序列进行优选,结果见表 5。由表 5 可以看出漯河站、息县站和蒋家集站最优模拟序列分别为 1-5 序列、2-1 和 3-7 序列。

表 5 模拟序列的关联度

Tab. 5 Correlation degrees of simulated sequences

研究站点	模拟序列	关联度	优选序列
漯河站	1-1	0.646 5	1-5
	1-2	0.676 4	
	1-3	0.754 1	
	1-4	0.746 0	
	1-5	0.774 2	
	1-6	0.685 0	
	1-7	0.696 3	
	1-8	0.754 4	
	1-9	0.681 9	
	1-10	0.770 5	
息县站	2-1	0.794 7	2-1
	2-2	0.751 1	
	2-3	0.763 9	
	2-4	0.682 4	
	2-5	0.764 1	
	2-6	0.707 2	
	2-7	0.776 1	
	2-8	0.728 6	
	2-9	0.678 7	
	2-10	0.763 8	
蒋家集站	3-1	0.677 7	3-7
	3-2	0.716 5	
	3-3	0.764 6	
	3-4	0.723 8	
	3-5	0.625 6	
	3-6	0.665 4	
	3-7	0.800 6	
	3-8	0.640 4	
	3-9	0.700 4	
	3-10	0.673 1	

注:加粗表示关联度最大。

为进一步说明截口-复杂性指标体系用于模拟序列评价的合理性,对传统指标、样本熵最小指标^[9]以及截口-复杂性指标体系优选出的模拟序列进行对比,截口、复杂变化特性方面的相对误差和 E_{MAP} 见表 6 和表 7,多年平均以及丰、平、枯水年的拟合效果见图 2。

表 6 3 种评价指标优选序列截口指标的相对误差及 E_{MAP} 结果统计Tab. 6 Statistics of RE and E_{MAP} results of optimized sequence section indexes of three evaluation indexes %

研究站点	最优模拟序列			截口指标		E_{MAP}		
	模拟序列	评价指标	\bar{x}	C_v	C_s		r_1	r_2
漯河站	1-10	传统	2	-12	-20	-5	-1	8
	1-8	样本熵	8	-5	-4	-4	-13	7
	1-5	截口-复杂性	5	-1	-6	-4	-16	6
息县站	2-10	传统	2	-3	-18	9	2	7
	2-9	样本熵	3	-5	-22	6	-4	8
	2-1	截口-复杂性	-2	2	-21	-7	-3	7
蒋家集站	3-1	传统	6	-4	-14	-9	-7	8
	3-4	样本熵	2	-5	-16	-3	-10	7
	3-7	截口-复杂性	2	1	-4	-6	-22	7

表 7 3 种评价指标优选序列复杂性指标的相对误差及 E_{MAP} 结果统计Tab. 7 Statistics of RE and E_{MAP} results of complex change index of three evaluation index optimization series %

研究站点	最优模拟序列		复杂变化指标				E_{MAP}
	模拟序列	评价指标	Q_1	C_d	C_t	H	
漯河站	1-10	传统	-1	16	6	-21	11
	1-8	样本熵	-4	12	9	2	7
	1-5	截口-复杂性	0	2	4	4	3
息县站	2-10	传统	2	6	12	-12	8
	2-9	样本熵	2	11	8	1	6
	2-1	截口-复杂性	0	-5	4	-1	3
蒋家集站	3-1	传统	-3	6	12	-12	8
	3-4	样本熵	-2	11	8	1	6
	3-7	截口-复杂性	-2	-4	2	-1	2

由表 6 和表 7 可以看出:3 种指标优选的模拟序列截口指标的 E_{MAP} 值差异较小,最大为 2%;复杂变化指标的 E_{MAP} 值差异较大,截口-复杂性指标体系和传统指标优选序列的 E_{MAP} 值最大相差 8%,与样本熵指标优选序列的 E_{MAP} 值最大相差 4%。由图 2 可以看出,截口-复杂性指标体系优选序列的各水平年年内分配相对误差总体小于传统指标和样本熵指标优选序列的结果,且平、枯水年相对误差差异最为明显。

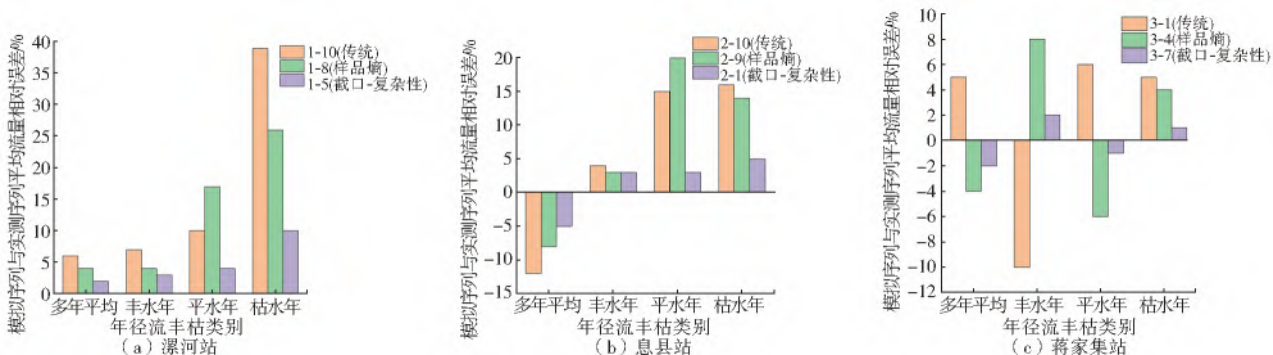


图 2 3 种评价指标优选序列多年平均及丰/平/枯水年相对误差

Fig. 2 The relative errors of three evaluation index optimization series in annual average, wet, normal and dry years

综合上述分析,3 种指标评价方法优选序列的精度为: 截口-复杂性指标 > 样本熵指标 > 传统指标。这说明本文提出的截口-复杂性指标体系是切实可行的, 有利于全面评价径流模拟序列, 实现不同随机模拟序列之间的比较与优选。

3.5 讨论

截口、复杂性及水平年内分配方面的对比结果表明, 3 种指标优选的序列均较好地保持了实测序列的截口特性, 但传统指标、样本熵最小指标体系不能全面评价模拟序列, 有可能遗漏径流序列重要的细节信息。以下从年内分配异常情况和时程分配(过程)的差异性两方面进行探讨。

根据 3 个水文站历年逐月实测径流资料统计结果, 漯河站、息县站、蒋家集站多年平均汛期最大 4 个月(6—9 月)径流量占全年径流量分别为 67%、64%、61%。现将连续最大 4 个月径流出现在非汛期视为异常情况, 统计传统截口指标优选序列、截口-复杂性指标体系优选序列年内连续最大 4 个月径流在非汛期的百分率, 并与实测资料进行对比(表 8), 结果发现: 传统指标优选序列的百分率在 3 种序列类型中最大, 且与截口-复杂性指标优选序列的情况相差较大。这说明只保证模拟系列均值、方差等基本统计特征较为一致情况下, 模拟序列会出现较多的年内

分配异常情况, 而截口-复杂性指标体系有助于排除年内分配异常的情况。

表 8 3 种径流序列连续最大 4 个月径流在非汛期的百分率

Tab. 8 Percentage of the maximum 4-month continuous runoff of three runoff series in non-flood season

研究站点	序列类型	非汛期连续最大 4 个月径流百分率/%
漯河站	实测序列	0.75
	传统指标优选	3.32
	截口-复杂性指标优选	0.94
息县站	实测序列	0.58
	传统指标优选	1.70
	截口-复杂性指标优选	0.57
蒋家集站	实测序列	0.65
	传统指标优选	1.96
	截口-复杂性指标优选	0.66

表 9 统计了漯河站 3 种序列的年内分配指标, 可以看出: 截口-复杂性指标体系优选序列的标准差数值整体大于传统指标优选序列, 说明本文指标体系增加的复杂变化指标并没有影响模拟序列各子序列年内分配特征的多样性; 两种指标优选序列与实测序列的标准差差异较小, 说明传统指标体系和截口-复杂性指标体系优选的长序列模拟序列中, 各子序列的多样性受多年平均计算条件下各指标约束的影响较小。

表 9 3 种径流序列年内分配指标分析

Tab. 9 Analysis of annual distribution index of three runoff series

研究站点	序列类型	不均匀系数		集中度		连续最大 4 个月径流百分率/%	
		均值	标准差	均值	标准差	均值	标准差
漯河站	实测序列	1.18	0.45	49.38	18.94	67	15
	传统指标优选	1.34	0.40	54.33	18.84	72	13
	截口-复杂性指标优选	1.29	0.39	53.14	18.91	71	13
息县站	实测序列	1.08	0.35	46.46	14.10	65	11
	传统指标优选	1.20	0.32	53.58	14.20	72	10
	截口-复杂性指标优选	1.23	0.43	50.09	18.80	68	23
蒋家集站	实测序列	1.20	0.51	47.15	18.51	66	13
	传统指标优选	1.16	0.39	47.71	15.99	66	15
	截口-复杂性指标优选	1.20	0.53	46.39	18.70	67	21

模拟序列截口、复杂变化特性的评价以及年内分配异常情况的检验表明,引入表征复杂变化特性的指标(Q_4 、 C_d 、 C_t 、 H),可以在保证模拟序列多样性的同时,减少年内分配的异常情况,能够描述传统评价指标所忽略的细节特征,较好地体现原始序列的形态复杂变化特征。因此,截口-复杂性指标体系更准确、真实地刻画了月径流模拟序列的年内复杂变化特征,优选出的模拟序列更能保持实测序列的时间变化特征。

4 结 论

本文从径流序列非线性及复杂特性出发,提出了月径流随机模拟序列评价指标体系,并将其用于3个水文站季节性自回归模型的模拟序列效果评价中,得到以下结论。

只考虑截口统计参数的评价方法,忽视了径流序列的复杂变化特性,可能遗漏重要的细节信息,从而影响模拟序列检验的精度。截口-复杂性指标体系中表征复杂变化特性的指标(Q_4 、 C_d 、 C_t 、 H)能简单有效地区分模拟序列保持实测资料复杂变化特性的能力,优选序列在年内分配方面的 E_{MAP} 值最小,总体拟合效果最好。

对传统指标、样本熵最小指标以及截口-复杂性指标优选的序列进行了分析,结果表明截口-复杂性指标体系优选的序列能较好地反映实测序列的变化过程和复杂特性,综合截口和复杂变化特征方面的分析,月径流随机模拟序列优选方法的优劣顺序为截口-复杂性指标>样本熵指标>传统指标。

经过截口和复杂变化特征方面的模拟序列评价、年内分配异常情况分析以及时程分配(过程)差异性分析的多重验证,截口-复杂性指标体系在保证模拟序列年内分配多样性的同时可以减少年内分配异常的情况,有助于水文工作者掌握不同模拟序列之间的差异,更加全面地评价径流模拟序列,从而提供更加准确的水文模拟序列。

为进一步提高径流序列模拟效果评价的准确性,未来将考虑采用主成分分析法找出相互独立的主要影响指标,以此构成更具代表性的指标体系,并由此对随机模拟序列进行检验,通过与其他指标体系检验结果的对比分析其代表性。

参考文献(References):

[1] 刘强,钟平安,徐斌,等.三峡及金沙江下游梯级水库群蓄水期联合调度策略[J].南水北调与水利科技,2016,14(5):62-70. (LIU Q,ZHONG P A,XU B,et al. Joint

operation strategy of water impoundment period of Three Gorges-Jinsha River lower reach cascade reservoirs[J]. South-to-North Water Transfers and Water Science & Technology, 2016, 14(5): 62-70. (in Chinese)) DOI: 10. 13476/j. cnki. nbsdqk. 2016. 05. 010.

- [2] 马川惠.基于径流随机模拟的龙滩水库优化调度研究[D].西安:西安理工大学,2019. (MA C H. The research of Longtan Reservoir's optimal scheduling based on runoff stochastic simulation[D]. Xi'an: Xi'an University of Technology, 2019. (in Chinese))
- [3] LIU P, LI L, GUO S, et al. Optimal design of seasonal flood limited water levels and its application for the Three Gorges Reservoir [J]. Journal of Hydrology, 2015, 527: 1045-1053. DOI: 0. 1016/j. jhydrol. 2015. 05. 055.
- [4] ZHANG H, SINGH V P, WANG B, et al. CEREF: A hybrid data-driven model for forecasting annual streamflow from a socio-hydrological system [J]. Journal of Hydrology, 2016, 540: 246-256. DOI: 10. 1016/j. jhydrol. 2016. 06. 029.
- [5] WEISS J, BERNARDARA P, ANDREEWSKY M, et al. Seasonal autoregressive modeling of a skew storm surge series [J]. Ocean Modelling, 2012, 47: 41-54. DOI: 10. 1016/j. ocemod. 2012. 01. 005.
- [6] HAO Z, SINGH V P. Entropy-copula method for single-site monthly streamflow simulation [J]. Water Resources Research, 2012, 48(6): 6604. DOI: 10. 1029/2011WR011419.
- [7] HAMID H A, YAHAYA A S, RAMLI N A, et al. Short Term prediction of PM₁₀ concentrations using seasonal time series analysis [J]. MATEC Web of Conferences, 2016, 47. DOI: 10. 1051/mateconf/20164705001.
- [8] 高瑞忠,白勇,刘廷玺,等.内蒙古高原典型草原内河流域径流的时序演变特征[J].南水北调与水利科技,2018,16(3):10-17. (GAO R Z, BAI Y, LIU Y X, et al. Evolution characteristic of runoff in the typical grassland inland river basin of Inner Mongolia Plateau [J]. South-to-North Water Transfers and Water Science & Technology, 2018, 16(3): 10-17. (in Chinese)) DOI: 10. 13476/j. cnki. nbsdqk. 2018. 0062.
- [9] 马细霞,朱慧青.基于信息熵和统计参数的径流随机模拟序列检验方法[J].水电能源科学,2019,37(4):5-8. (MA X X, ZHU H Q. Runoff random simulation sequence test method based on information entropy and statistical parameters [J]. Water Resources Power, 2019, 37(4): 5-8. (in Chinese)) DOI: CNKI: SUN: SD-NY. 0. 2019-04-002.
- [10] 甘容,陈长征.沙颍河流域径流过程模拟与径流组分变化特征[J].南水北调与水利科技(中英文),2021,

- 19(1): 83-91. (GAN R, CHEN C Z. Simulation of runoff process and variation characteristic in Shaying River basin[J]. South-to-North Water Transfers and Water Science & Technology, 2021, 19(1): 83-91. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2021. 0008.
- [11] 莫崇勋, 刘朋, 朱新荣, 等. 澄碧河流域径流年内分配变化规律分析[J]. 南水北调与水利科技, 2019, 17(4): 63-70. (MO C X, LIU P, ZHU X R, et al. Interannual variation of runoff in the Chengbi River basin [J]. South-to-North Water Transfers and Water Science & Technology, 2019, 17(4): 63-70. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2019. 0085.
- [12] 梅超, 尹明万, 洪林, 等. 两种月径流随机模拟方法应用比较[J]. 中国农村水利水电, 2016(12): 93-96, 100. (MEI C, YIN M W, HONG L. A comparison of two stochastic simulation methods for monthly runoff [J]. China Rural Water Hydropower, 2016(12): 93-96, 100. (in Chinese)) DOI: CNKI: SUN: ZNSD. 0. 2016-12-020.
- [13] 孙周亮, 刘艳丽, 刘冀, 等. 1960—2012 年湄公河干流径流时空演变[J]. 南水北调与水利科技(中英文), 2020, 18(1): 51-65. (SUN Z L, LIU Y L, LIU J, et al. Spatiotemporal evolution of runoff in the mainstream of the Mekong River from 1960 to 2012[J]. South-to-North Water Transfers and Water Science & Technology, 2020, 18(1): 51-65. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2020. 0007.
- [14] 彭涛, 陈晓宏, 庄承彬. 基于样本熵的东江月径流序列复杂性分析[J]. 生态环境学报, 2009, 18(4): 1379-1382. (PENG T, CHEN X H, ZHUANG C B. Analysis on complexity of monthly runoff series based on sample entropy in the Dongjiang River[J]. Ecological Environment, 2009, 18(4): 1379-1382. (in Chinese)) DOI: 10. 3969/j. issn. 1674-5906. 2009. 04. 031.
- [15] 李勋贵, 魏宁, 魏霞. 系统复杂性测度参数确定的新方法及其应用[J]. 系统工程理论与实践, 2018, 38(1): 252-262. (LI X G, WEI N, WEI X. A new method for determining parameters of system complexity measure and its application [J]. Systems Engineering-Theory & Practice, 2018, 38(1): 252-262. (in Chinese)) DOI: 10. 12011/1000-6788(2018)01-0252-11.
- [16] 刘梦洋, 尹义星, 韩翠, 等. 基于地区线性矩法的江西省极值降水频率分析[J]. 水电能源科学, 2018, 36(9): 1-5. (LIU M Y, YIN Y X, HAN C, et al. Frequency analysis of extreme precipitation in Jiangxi Province based on regional L-moments method [J]. Water Resources and Power, 2018, 36(9): 1-5. (in Chinese)) DOI: CNKI: SUN: SDNY. 0. 2018-09-001.
- [17] MA X X, ZHU H Q, XIAO Y, et al. Inspection method for random simulation of runoff sequence based on sample entropy and statistical parameters [J]. Journal of Hydrology, 2020, 581. DOI: 10. 1016/j. jhydrol. 2019. 124392.
- [18] 裴宏伟, 杨佳, 张红娟, 等. 变化环境下清水河流域径流演变特征及驱动力[J]. 南水北调与水利科技(中英文), 2020, 18(2): 1-13. (PEI H W, YANG J, ZHANG H J, et al. Characteristics and driving forces of runoff evolution under changing climate in the Qingshui River basin [J]. South-to-North Water Transfers and Water Science & Technology, 2020, 18(2): 1-13. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2020. 0022.
- [19] 张文浩, 瞿思敏, 徐瑶, 等. 泼水水库对黄河径流过程及水文情势的影响[J]. 水资源保护, 2021, 37(3): 61-65. (ZHANG W H, ZHAI S M, XU Y, et al. Influence of Pohe reservoir on runoff and hydrological regime of Huanghe River [J]. Water Resources Protection, 2021, 37(3): 61-65. (in Chinese)) <http://kns. cnki. net/kcms/detail/32. 1356. TV. 20210112. 1912. 010. html>.
- [20] 张黎鸣, 赵岩, 王红瑞, 等. 基于信息熵与灰关联的西安市城市经济与用水结构的耦合度研究[J]. 南水北调与水利科技, 2017, 15(4): 187-192, 202. (ZHANG L M, ZHAO Y, WANG H R, et al. Study of the relationship between urban economic development and the utilization of water resources in Xi'an City based on information entropy and grey correlation grade [J]. South-to-North Water Transfers and Water Science & Technology, 2017, 15(4): 187-192, 202. (in Chinese)) DOI: 10. 13476/j. cnki. nsbdqk. 2017. 04. 030.
- [21] 李雅晴, 谢平, 桑燕芳, 等. 水文序列相依变异识别的 RIC 定阶准则: 以自回归模型为例[J]. 水利学报, 2019, 50(6): 721-731. (LI Y Q, XIE P, SANG Y F, et al. RIC criterion for identifying dependent variation of hydrological time series: With a case study of autoregressive model [J]. Journal of Hydraulic Engineering, 2019, 50(6): 721-731. (in Chinese)) DOI: 10. 13243/j. cnki. slxb. 20181069.

Optimization method for monthly runoff stochastic simulation sequence

LIU Xinxin¹, MA Xixia^{1,2}, CHENG Xu¹, WANG Qianli¹, ZHANG Jingwen¹

(1. School of Water Conservancy Engineering, Zhengzhou University, Zhengzhou 450001, China; 2. Yellow River Institute for Ecological Protection & Regionally Coordinated Development, Zhengzhou University, Zhengzhou 450001, China)

Abstract: Hydrologic stochastic model can provide a scientific and decision-making basis for alleviating water resources problems

by simulating a large number of random sequences for hydrologic system. However, the simulation series generated by the same stochastic model was quite different in time and process changes and the traditional statistical parameter test method ignored the variation characteristics of runoff as a time series, which can not fully reflect the complex variation characteristics of flow in time series distribution, and failed to select the most suitable runoff simulation series for engineering planning and design and risk analysis. Therefore, a set of evaluation index system of monthly runoff random generation series is proposed, which incorporates statistical characteristics of runoff simulation series at different time intersections and the complex variation characteristics of runoff series. Three hydrological stations in the main and tributary streams of the Huaihe River basin are taken as the research object to establish a seasonal autoregressive model, three evaluation methods are selected based on constructed index system and multiple sets of randomly generated monthly runoff simulation sequences. The sequence comparison aimed to provide a more accurate and reliable new way to identify the random simulation sequence of monthly runoff.

Starting from the nonlinear and complex characteristics of runoff series, indicators such as concentration, uneven coefficient and sample entropy based on traditional test indicators, and a set of index systems is proposed that characterize the statistical parameter characteristics and complex changes of the runoff series. The accuracy of the simulation sequence was analyzed, and the simulation sequence was optimized by the gray correlation analysis method. The fitting results of the three evaluation methods (traditional index, minimum entropy index of sample and index) were compared and analyzed to illustrate the rationality of the index. Moreover, the advantages of the proposed evaluation index system were further verified from the two aspects of the abnormal situation of the year allocation and the difference of annual time history distribution (process) of runoff.

Comparison results of the optimized sequences of the three evaluation methods showed that: (1) In the evaluation of statistical characteristics of runoff simulation series at different time intersections, the difference of the average absolute percentage error of the simulation sequence optimized by the three indexes was small, and the maximum error was 2%. (2) In the evaluation of complex change characteristics, the difference of average absolute percentage error was large. The maximum error difference between the optimization sequence based on the index and the traditional index was 8%, and the maximum error difference between the optimization sequence based on the index and the sample entropy index was 4%. The selected sequence performed well in the evaluation of statistical parameter characteristics and complex change characteristics, and the most significant difference was in complex change characteristics compared with the other two evaluation methods. (3) The results showed that the performance of the series selected was good in the annual distribution characteristics of each level, and the relative error was generally smaller than that of the series selected by traditional index and sample entropy index, and the difference between the relative error in normal and dry years was the most obvious.

In conclusion, the three evaluation methods were effective in maintaining the characteristics of the runoff series, but the previous index system can not comprehensively evaluate the simulation series, which may omit the important details of the runoff series, thus affecting the accuracy of the simulation series. In this index system, the indexes which represented the complex change characteristics can simply and effectively distinguish the ability of simulation series in maintaining the complex change characteristics of measured series. The selected series can better reflect the change process and complex characteristics of the measured series, and the overall fitting effect of annual distribution was also the best. The order of merit and demerit of the three methods was the proposed index of this paper > the index of sample entropy > the traditional index. The results showed that the proposed index system comprehensively considered the statistical parameters of the section and the characteristics of complex changes was feasible. The abnormal situation of the year allocation and the difference of annual time history distribution (process) of runoff were analyzed, the results showed that the advantage of proposed index, which is beneficial to a comprehensive evaluation of runoff simulation series and realizes the random simulation comparison and optimization among different series, thereby provided more accurate hydrological simulation sequences.

Key words: random simulation; simulation test; index system; grey correlation analysis; sequence optimization