

淮河流域洪水极值非平稳性特征*

孙 鹏^{1,2}, 孙玉燕^{1,3,4}, 张 强^{4,5**}, 姚 蕊^{1,3}, 温庆志^{1,3}, 王友贞², 蒋尚明²

(1: 安徽师范大学地理与旅游学院, 芜湖 241002)

(2: 安徽省水利部淮河水利委员会水利科学研究院, 水利水资源安徽省重点实验室, 蚌埠 233000)

(3: 自然灾害过程与防控研究安徽省省级重点实验室, 芜湖 241002)

(4: 北京师范大学环境演变与自然灾害教育部重点实验室, 北京 100875)

(5: 北京师范大学地表过程与资源生态国家重点实验室, 北京 100875)

摘 要: 基于淮河流域 9 个水文站的月径流量数据, 采用 Pettitt 非参数检验法、GAMLSS 模型与洪水频率分析模型等方法, 揭示了淮河中上游洪水频率的演变规律, 分析基于平稳性和非平稳性条件下的洪水发生强度及洪涝灾害所带来的影响。研究发现, 潢川、横排头 and 蚌埠站点未发生明显变异, 其余 6 个站点发生均值或方差变异, 变异时间主要集中在 2000 年左右。淮河流域的最优拟合分布函数是 Weibull; 班台、蒋家集和横排头站适宜于非平稳性模型, 其余站点选择平稳性模型。各站点非平稳性条件下 10 年和 20 年一遇设计流量值与平稳性条件下皮尔逊Ⅲ型分布设计流量值相差不大, 但 30 年一遇、50 年一遇和 100 年一遇的设计流量相差逐渐变大。横排头站和蚌埠站洪水放大因子随着时间增加呈上升趋势且大于 1, 百年一遇重现期不足 80 年。各站点年最大洪峰流量与淮河流域、安徽省水灾面积通过了 95% 或 99% 的显著性检验。

关键词: 非平稳性; GAMLSS 模型; 洪水放大因子; 重现期; 淮河流域

Evaluation on non-stationarity assumption of annual maximum peak flows during 1956–2016 in the Huaihe River Basin

SUN Peng^{1,2}, SUN Yuyan^{1,3,4}, ZHANG Qiang^{4,5**}, YAO Rui^{1,3}, WEN Qingzhi^{1,3}, Wang Youzhen² & JIANG Shangming²

(1: School of Geography and Tourism, Anhui Normal University, Wuhu 241002, P.R.China)

(2: Key Laboratory of Water Conservancy and Water Resources of Anhui Province, Water Resources Research Institute of Anhui Province and Huaihe River China, Bengbu 233000, P.R.China)

(3: Anhui Key Laboratory of Natural Disaster Process and Prevention, Wuhu 241002, P.R.China)

(4: Key Laboratory of Environmental Change and Natural Disaster, Ministry of Education, Beijing Normal University, Beijing 100875, P.R.China)

(5: State Key Laboratory of Surface Process and Resource Ecology, Beijing Normal University, Beijing 100875, P.R.China)

Abstract: Based on the monthly runoff data of 9 hydrological stations in the Huaihe River Basin from 1956 to 2016, Pettitt non-parametric test, GAMLSS model and flood frequency analysis model were used to reveal the variation of flood frequency in the middle and upper reaches of the Huaihe River. Sexual and non-stationary conditions under the conditions of flood intensity and floods caused by the impact. The results showed that: Huangchuan, Hengpaitou and Bengbu Stations did not change significantly, the other six stations of the mean or variance variation, variation time is mainly concentrated in 2000 or so. The optimal fitting distribution function of Huaihe River Basin is Weibull, followed by Lognormal distribution. The non-stationary model of Bantai, Jiangjiaji and Hengpaitou Stations were selected, and the other six stations choose the stationary model. While the optimal fitting function

* 国家自然科学基金项目(41601023, 41771536)、国家杰出青年科学基金项目(51425903)、地表过程与资源生态国家重点实验室开放基金资助项目(2017-KF-04)、中国水利水电科学研究院流域水循环模拟与调控国家重点实验室开放基金项目(IWHR-SKL-201720)和安徽省自然科学基金项目(1808085QD117)联合资助。2017-12-14 收稿; 2018-01-11 收修改稿。孙鹏(1986 ~), 男, 博士, 副教授; E-mail: sun68peng@163.com。

** 通信作者; E-mail: zhangq68@bnu.edu.cn。

based on GAMLSS model has good fitting effect. The divergence curves of the Hengpaitou Station and the Jiangjiayi Station in the nonstationary station show a fluctuating trend with the time series. The site distribution of the station is 25% and 75%, and the actual distribution is not very reasonable. The rest of the site stability of the overall distribution of the site is more reasonable. The design value of the Pearson type III distribution design under the condition of 10 years and 20 years of design flow rate and stationary conditions is not very different, but in 30 years, 50 years and 100 years the difference between the design flow of a gradual increase. County, Wangjiaba, Lutaizi and Bengbu Stations design flow value of more than 10000 m³/s. flooding station and Bengbu Station flood amplification factor increases with time and is greater than 1, a hundred years in the event of less than 80 years. Huangchuan Station, Wangjiaba Station and Fuyang Station flood amplification factor are less than 1, and the 100-year re-emergence period will be more than 300 years. The maximum annual peak flow of each station is highly correlated with the Huaihe River Basin and the flood area in Anhui Province, and basically has passed the 95% or 99% significance test.

Keywords: Nonstationarity; GAMLSS model; flooding factor; reproduction period; Huaihe River Basin

近年来,在气候变化和人类活动的综合影响下,极端气象水文事件(如暴雨和洪水等)频发^[1],致使河流水文情势发生显著变化。据联合国公布资料显示,洪涝灾害所造成的损失已占全球自然灾害损失的 20%,使洪水风险管理受到国际科研和管理机构高度关注^[2-5]。变化环境下所导致的水文变异使得流域径流序列形成的环境背景“平稳性”已发生显著变异。Milly 等^[6]认为,变化环境下的“平稳性”假设已不再适合作为水资源风险评估的默认假设,采用现有水文过程分析方法将会面临由变化环境带来的设计频率失真风险。因此,非平稳性水文序列频率的研究近年来受到格外关注。Rigby 等^[7]提出广义可加模型(简称 GAMLSS),该模型可以构建分布函数参数和多个解释变量的线性、非线性、参数和非参数关系,并且扩展了分布函数的选择范围,具有很大的灵活性;Cunderlik 等^[8]提出非平稳条件下汇流的洪水频率模型,该模型将分位数函数分为本地时间相关分量(包括位置和尺度分布参数)以及在二阶非平稳假设下与时间变化无关的区域分量,主要应用均质区域。Chen 等^[9]结合非参数离散小波变换(DWT)和综合经验模态分解(EEMD),采用参数加权最小二乘法(WLS),通过评估台湾主要气象站年平均最大日降水量的非平稳性来识别这 3 种方法的适用性。另外非平稳性研究方法也引起国内学者的广泛关注^[10-13]。

气候变化和人类活动的共同影响下,淮河流域降水和径流的时空特征发生了显著变化。杜鸿等^[14]通过对淮河流域极端径流的时空变化规律分析发现,淮河流域的极端径流事件以 1960s 和 1970s 居多,主要来自淮河干流、淮南山区及伏牛山区;魏凤英等^[15]采用小波变换和广义极值分布等方法分析了 1922—2007 年淮河流域夏季降水的年际和年代际震荡特征。1990s 末以来,淮河流域夏季降水准 2 年振荡特征突出,极端强降水事件的概率显著增加;刘思敏等^[16]基于淮河流域内 229 个气象站点 1950—2012 年的实测逐小时降水数据研究发现,淮河流域暴雨事件高发区域呈现出从流域部分地区向全流域扩张的趋势。极端降水和气候事件的增加,导致淮河流域洪涝灾害频发,近 60 年先后出现了 17 个洪涝年份(发生频率为 4 年一次),年均受涝面积 269.8 万 hm²,成灾面积 140.8 万 hm²,分别占全流域耕地面积的 21% 和 11%。严重的洪涝灾害不仅影响工农业生产、人民生活 and 生态环境,给国民经济也造成重大损失。

因地处南北气候、中低纬度和海陆相三种过渡带的重叠地区,淮河流域天气系统复杂多变,形成了该区上游区域“无降水旱、有降水涝、强降水洪”的典型区域旱涝特征^[14]。淮河流域水利工程众多,极大地改变了河川的平稳性,河川径流的平稳假设受到挑战。然而目前对淮河流域径流变化的研究主要是基于平稳性条件下开展的,且多集中于降水、气候变化和农业干旱风险评价等方面,对于非平稳性条件下的径流变化特征的研究几乎未涉及。随着极端气候事件的增加,致使极端水文事件增加将近 1 倍。在变化环境下将径流和其他水文过程做稳定性过程处理可能不再合理,尤其是依据水文过程稳定性假设构建的传统频率分析模型计算的设计标准修建的流域开发利用工程、防洪和抗旱工程等,将面临由变化环境引起的水文变异导致的风险。淮河径流是否具有平稳性,然而对于淮河流域的非平稳性洪水的研究并不多见。因此,本文基于 GAMLSS 模型与洪水频率分析模型开展淮河流域的非平稳性洪水特征研究。该研究可为提高淮河流域的防洪减灾能力,促进流域经济社会的可持续发展提供有力的支撑。

1 研究区域与数据

淮河流域(30°55′~36°36′N, 111°55′~121°25′E)(图 1),地跨鲁、皖、苏、豫和鄂五省;介于长江和黄河

两流域之间,总面积 27 万 km²,人口约 1.8 亿,耕地面积 1333 万 hm²,流域以占全国 10% 的耕地面积生产全国近 20% 的粮食,平均每年向国家提供商品粮约占全国商品粮的 1/4,在我国农业生产中的地位举足轻重.因此,本文选取淮河流域位于干流和主要支流上的 9 个水文控制站点的年最大日流量数据进行统计分析,数据起止年限为:王家坝站,1964—2016 年;潢川站和横排头站,1980—2016 年;其余各站点均为 1956—2016 年.

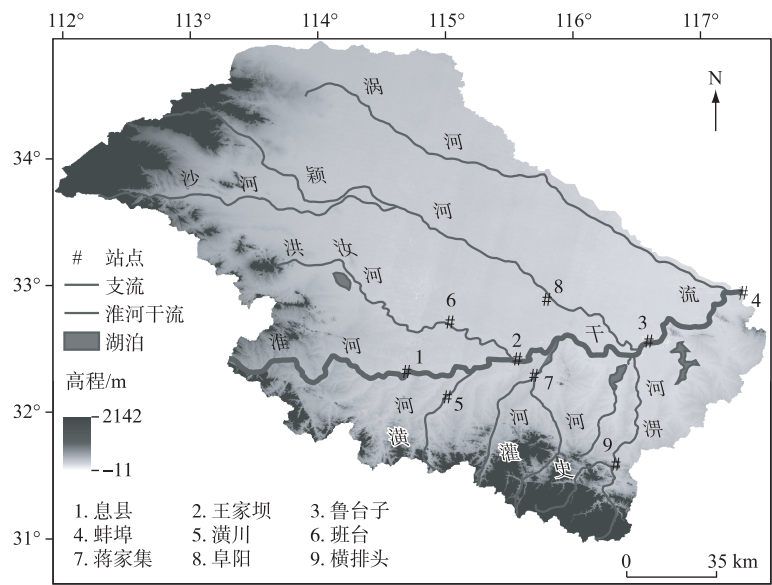


图 1 研究区及水文站点位置
Fig.1 Locations of the study region and hydrological stations

2 研究方法

2.1 变异点分析

Pettitt 非参数检验法是基于 Mann-Whitney 的统计函数^[17]. 均值和方差变异是水文变异的两种基本类型,是用于描述时间序列模型的参数变化的单因素变量^[18-19]. 该方法通过检验时间序列要素均值和方差变化进行变异点分析,认为两个样本 x_1, \dots, x_t 和 x_{t+1}, \dots, x_n 均来自同一体:

$$U_{t,n} = U_{t-1,n} + \sum_{j=1}^n \text{sgn}(x_t - x_j) \tag{1}$$

式中, x_t 为水文序列中第 t 个点的值, x_j 为水文序列中的第 j 个点的值, $t = 2, 3, \dots, n$. 式(1)用于均值突变检测,用 Pettitt 法检测序列方差变异,需要对水文序列进行处理^[17]:

$$Y_i = (x_i - L_i)^2 \tag{2}$$

用式(2)进行突变检测,若存在突变点则为方差突变. x_1, x_2, \dots, x_n 表示实测洪水极值序列; L 代表参考函数 Loess 函数; Y 代表残差平方和序列.

2.2 GAMLSS 模型

将时间序列均值和方差突变及时间趋势纳入广义可加模型 (generalized additive models for location scale and shape, GAMLSS 模型) 的框架中进行突变和时间趋势检测. GAMLSS 模型可用于检测时间序列的平稳性和非平稳性,是一种半参数回归模型^[20].

在 GAMLSS 模型中,假设同一时间序列 y_1, y_2, \dots, y_n 相互独立并且服从分布函数 $F_Y(y_i | \theta_i)$, $\theta_i = (\theta_1, \theta_2, \dots, \theta_p)$ 表示 p 个参数 (位置、尺度和形状参数) 形成的向量. 记 $g_k(\cdot)$ 表示 θ_k 与解释变量 X_k 和随机效应项之间的单调函数关系:

$$g_k(\theta_k) = \eta_k = X_k \cdot \beta_k + \sum_{j=1}^{J_k} Z_{jk} \cdot \gamma_{jk} \quad (3)$$

式中, η_k 和 θ_k 是长度为 n 的向量, $\beta_k^T = \{\beta_{1k}, \beta_{2k}, \dots, \beta_{J_k k}\}$ 是长度为 J_k 的参数向量, X_k 是长度为 $n \times J_k$ 的解释变量矩阵, Z_{jk} 是已知的 $n \times q_{jk}$ 固定设计矩阵, γ_{jk} 是正态分布随机变量. 如果不考虑随机效应对分布参数的影响时, 即令 $J_k = 0$, 式(3)就变成一个全参数模型:

$$g_k(\theta_k) = \eta_k = X_k \cdot \beta_k \quad (4)$$

当解释变量为时间 t 时, 解释变量矩阵 X_k 可以表示为:

$$X_k = \begin{bmatrix} 1 & t & \cdots & t^{l_k-1} \\ 1 & t & \cdots & t^{l_k-1} \\ 1 & t & \cdots & t^{l_k-1} \\ 1 & t & \cdots & t^{l_k-1} \end{bmatrix}_{n \times l_k} \quad (5)$$

将式(5)代入式(4)可以得到分布参数与解释变量时间 t 的函数关系:

$$\begin{cases} g_1(\theta_1(t)) = \beta_{11} + \beta_{21}t + \cdots + \beta_{l_{11}} \cdot t^{l_{11}-1} \\ g_2(\theta_2(t)) = \beta_{12} + \beta_{21}t + \cdots + \beta_{l_{12}} \cdot t^{l_{12}-1} \end{cases} \quad (6)$$

M

本文主要探讨均值和方差(分别对应位置参数和尺度参数)的平稳性, 选择 4 种最常用的两参数极值分布进行分析: Gumbel (GU)、Gamma (GA)、Lognormal (LOGNO) 和 Weibull (WEI). 以时间 t 作为唯一的解释变量, 构造参数 θ_1 (均值) 和 θ_2 (方差) 与时间 t 的线性函数, 由公式(6)可以得出:

$$g_1(\theta_1^i) = t_i \cdot \beta_1 \quad (7)$$

$$g_2(\theta_2^i) = t_i \cdot \beta_2 \quad (8)$$

本文通过 GAMLSS 模型, 建立时间序列分布矩与时间的函数关系, 将趋势和突变统一纳入到非平稳性框架中进行分析. 主要用两参数模型分析未发生突变点的站点年最大洪峰流量序列, 分为以下 4 种模型: ① 平稳性模型, θ_1 和 θ_2 均为常数; ② θ_1 非平稳, θ_1 是时间 t 的线性函数; ③ θ_2 非平稳, θ_2 是时间 t 的线性函数; ④ θ_1 、 θ_2 均非平稳, θ_1 和 θ_2 均为时间 t 的线性函数. 用赤池信息量准则 AIC (Akaike Information Criterion) 值选择最优拟合分布模型和函数, 用残差诊断图 (worm 图) 分析模型拟合质量. 通过这种方式, 可以比较不同概率分布、趋势和突变点 (均值/方差) 的序列在非平稳性框架中的效果.

2.3 非平稳性洪水频率分析模型

Vogel 提出一个结合两参数对数正态分布函数所构造的非平稳性洪水频率分析模型, 本文仅给出模型的基本表达式、洪水放大因子和重现期表达式, 模型的具体推导和构建过程见文献[21].

模型基本表达式:

$$x_p(t) = \exp \left[\bar{y} + \hat{\beta} \left(t - \frac{t_1 + t_n}{2} \right) + z_p \cdot s_y \right] \quad (9)$$

式中, \bar{y} 为年最大日流量系列对数值的均值, $\hat{\beta}$ 为模型参数 β 的估计值, t 为年最大日流量序列的时间; t_1 和 t_n 分别为年最大日流量系列的起始和终止时间; z_p 为标准正态分布逆函数值; s_y 为年最大日流量系列对数值的标准差. $x_p(t)$ 为第 t 年设计标准为 p 的设计流量值.

洪水放大因子: 现在设计洪水必须乘以洪水放大系数获得未来跟现在洪水同量级的设计洪水值^[21]. 洪水放大因子大于 1, 表明未来设计洪水值要高于现在设计值, 意味着现有的防洪过程设计标准可能无法满足未来防洪需求; 洪水放大因子小于 1, 则相反.

$$M = \frac{x_p(t + \Delta t)}{x_p(t)} = \exp(\hat{\beta} \cdot \Delta t) \quad (10)$$

式中, M 为洪水放大因子; Δt 为时间间隔; 其他变量意义同上.

重现期: 现在发生的洪水在间隔 t 年后, 其重现期大小.

$$T_f = \frac{1}{1 - \Phi\left[z_{p0} - \frac{\hat{\beta} \cdot \Delta t}{s_y}\right]} \quad (11)$$

式中, T_f 为未来重现期; $\Phi(\cdot)$ 为标准正态分布累积概率分布函数, 其他变量意义同上。

3 结果

3.1 变异点分析

淮河流域的王家坝、班台和阜阳站年最大洪峰流量序列发生了均值变异和方差变异, 息县、蒋家集和鲁台子年最大洪峰流量仅发生了方差变异, 而蚌埠、潢川和横排头年最大洪峰流量的均值和方差均没有发生变异(图 2a、b)。淮河流域继 2000 年发生干旱后, 2001 年发生了春夏秋冬连续干旱的大旱年, 7 月份息县站降水较常年同期减少 99%, 大旱年也导致了淮河干流和史灌河的水文站的方差发生变异。蚌埠、潢川和横排头水文站的年最大洪峰流量未发生变异, 与其他站点相比, 淮河南岸的潢川和横排头站所在的潢河和淠河的流域面积是最小的, 分别为 2400 和 6000 km²。流域面积越小, 其径流的变化越大, 但是潢河(石山口水库)和淠河(佛子岭水库、响洪甸水库、磨子潭水库和白莲崖水库)上游有大型水库, 其调节拦蓄了其流域的大量洪水。虽然淠河属于山区型河道, 坡陡水浅水流急, 易受洪水侵袭^[22], 但是其水库防洪库容远较大, 因此气候变化和水利工程的调节使得潢河和淠河未发生变异。而蚌埠站未发生变异, 主要是因为蚌埠站控制流域面积大, 且蚌埠到王家坝之间分布了大量的城东湖、瓦埠湖、高塘湖等天然湖泊和荆山湖等行蓄滞洪区, 因此蚌埠年最大洪峰流量在极端降水和暴雨事件增多的前提下, 因为水利工程的调节未发生变异。

淮河流域 9 个水文站点中未出现仅存在均值变异的站点, 除了王家坝的均值变异(1992 年)、阜阳的均值/方差变异(2000 年)和班台的均值/方差变异(2009 年/1969 年)以外, 其他站点的方差变异时间均发生在 2001 年, 这主要是因为近 60 年来, 全流域年平均气温具有升高趋势, 尤其是在 1990s 之后增暖趋势更加明显, 并于 1990s 中后期发生暖化突变^[23]; 淮河流域降水量年际波动较为强烈, 1990 之前降水量基本呈下降趋势, 而 2000 年后明显上升^[24]。2000 年是旱涝急转发生面积最大(48.08%)的年份, 其中王家坝南岸和北岸旱涝急转发生频率高达 40% 左右^[25]。因而王家坝站在 1992 年发生均值变异主要是受到旱涝急转的影响, 其是在 1991 年爆发全流域性大洪水而在 1992 年发生干旱所导致。

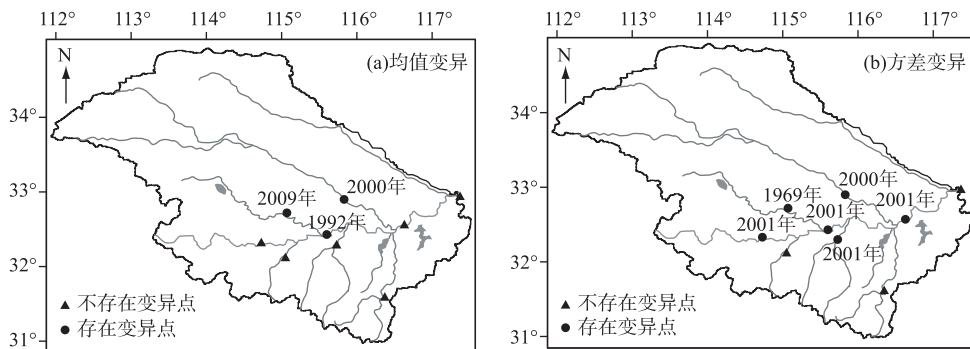


图 2 淮河流域年最大洪峰流量序列变异点分布

Fig.2 The distribution of annual maximum peak flow change points in the Huaihe River Basin

3.2 时间趋势分析

表 1 中仅为方差变异的水文站年最大洪峰流量在变异后呈减小趋势, 但是减小趋势并不显著; 方差和均值变异的阜阳水文站变异前的年最大洪峰流量通过了信度 95% 显著性检验, 减小趋势显著, 班台站年最大洪峰流量趋势恰好与阜阳站相反, 呈增加趋势, 且变异前的增加趋势显著; 王家坝站与其他 5 个变异站点不同, 变异前年最大洪峰流量呈减小趋势, 变异点后呈增加趋势。整体上, 淮河流域年最大洪峰流量除了阜阳和潢川下降趋势显著外, 其他站点呈不显著性下降趋势, 而横排头却呈不显著性上升趋势。通过分析变异

点对趋势变化的影响发现(图3),息县、蒋家集、阜阳和鲁台子站的变异点对趋势变化的影响大体一致,除阜阳站在变异前存在显著时间趋势性外,其余各站点在变异前、后序列均无显著趋势性.班台站整体序列呈微弱上升趋势,受变异点影响,变异前上升趋势较为明显(图3b).王家坝站(图3c)整体序列呈下降趋势,受变异点影响,变异前呈下降趋势,变异后呈上升趋势.由此可见,趋势分析时进行变异点识别是至关重要的,如果不考虑变异点的影响,趋势分析结果将会误导对序列统计特征的判断^[26].

表1 存在突变点序列变异点前、后子序列趋势检验结果
Tab.1 Trend test results for subseries posterior and prior to the change points

| | 站点 | 突变点 | 变异前 | | 变异后 | | 整体趋势 |
|------|-----|--------|----------|----|---------|----|----------|
| | | | MK | 方向 | MK | 方向 | |
| 均值变异 | 王家坝 | 1992 年 | -0.89 | - | 0.35 | + | -1.19 |
| | 班台 | 2009 年 | 1.39 * | + | 0.12 | + | -0.01 |
| | 阜阳 | 2000 年 | -1.96 ** | - | -1.52 * | - | -1.86 ** |
| 方差变异 | 息县 | 2001 年 | -1.19 | - | -1.04 | - | -0.82 |
| | 蒋家集 | 2001 年 | -1.10 | - | -0.36 | - | -0.64 |
| | 鲁台子 | 2001 年 | -1.21 | - | -0.59 | - | -0.36 |
| 未变异 | 蚌埠 | / | / | / | / | / | -0.10 |
| | 潢川 | / | / | / | / | / | -1.40 * |
| | 横排头 | / | / | / | / | / | 0.13 |

* 表示 MK 检验统计值 Z 的绝对值在 1.28 通过了信度 90% 的显著性检验, ** 表示 MK 检验统计值 Z 的绝对值在 ≥ 1.64 通过了信度 95% 的显著性检验.“+”表示呈增加趋势,“-”表示呈减小趋势,“/”表示趋势不变.

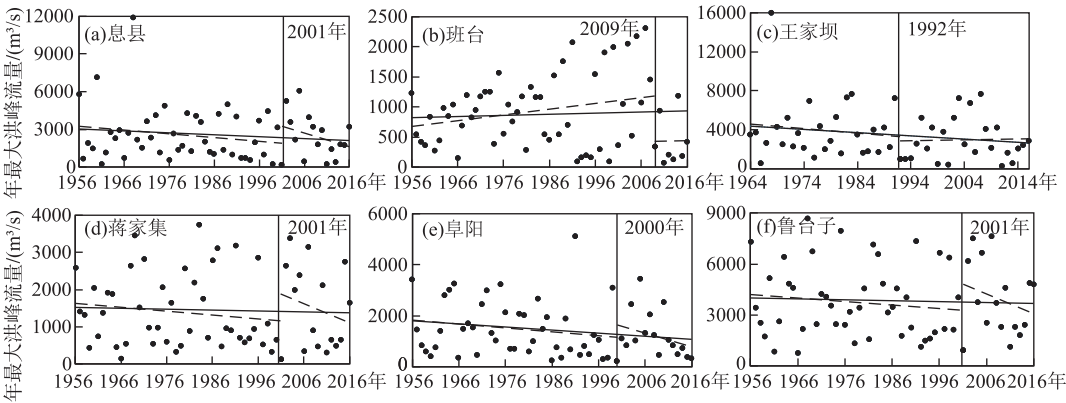


图3 变异点对趋势分析的影响
Fig.3 The impact of change points on trend test results

3.3 GAMLSS 模型分析

由上述分析可知,突变点和时间趋势可能导致年最大洪峰极值序列具有非平稳性,本文用 GAMLSS 模型检测洪峰极值序列的变化. 主要用两参数模型分析流域的年最大洪峰流量序列,由表 2 可知,在没有发生突变的 3 个站点中,潢川站和蚌埠站年最大洪峰流量的最佳拟合分布函数为两参数的 Weibull 分布,而 Log-normal 分布是横排头站的最佳拟合函数. 从序列平稳性来分析,潢川站和蚌埠站最优模型为平稳性模型,横排头站选择 θ_1 、 θ_2 均非平稳(表 2). 潢川站非平稳性模型与平稳性模型 AIC 值相差范围在 1.73~5.22 之间;蚌埠站和横排头站差值分别在 1.97~5.54 和 0.61~2.50 之间. 而横排头站非平稳模型仅比平稳性模型小 0.61. 对于发生突变的 6 个站点(表 2),Weibull 分布函数为最佳选择,符合息县、班台和王家坝站年最大洪峰流量的拟合模拟;Gamma 分布函数次之,其是蒋家集和班台站的最优拟合函数;Gumbel 分布函数则没有

适合的站点. 从序列平稳性来看,有 4 个站点最优模型为平稳性模型,而班台站和蒋家集站分别选择 θ_1 非平稳与 θ_2 非平稳模型,无站点选择 θ_1 、 θ_2 均非平稳. 从 AIC 值来看(表 3),发生突变的 6 个站点中阜阳站平稳性与非平稳性的 AIC 差值的范围最小,仅在 0.90~2.802 之间,班台站的 θ_1 非平稳性模型与平稳性模型的 AIC 值差值最大,达到 8.15.

从整体上来看,淮河中上游的 9 个站点中,有 5 个站点的最优拟合分布函数是 Weibull 函数,其次是 Lognormal 分布函数;班台、蒋家集和横排头站的选择的最优分布模型为非平稳性模型,其余 6 个站点选择的最优模型是平稳性模型. 未发生突变的 3 个站点差值范围相对而言较小,其平稳性模型和非平稳性模型区别并不明显,而发生突变的 6 个站点的平稳性模型与非平稳性模型 AIC 差值范围普遍较大. 因而,GAMLSS 模型拟合的结果与突变点、趋势分析的结果相吻合,未发生突变的站点其趋势变化不明显.

表 2 GAMLSS 模型分析结果*
Tab.2 Analysis results of GAMLSS model

| 站点 | 函数模型 | 平稳性模型 | θ_1 非平稳 | θ_2 非平稳 | θ_1 、 θ_2 均非平稳 |
|-----|-------|-------|----------------|----------------|------------------------------|
| 息县 | WEI | Y | — | — | — |
| 潢川 | WEI | Y | — | — | — |
| 班台 | WEI | — | Y | — | — |
| 王家坝 | WEI | Y | — | — | — |
| 蒋家集 | GA | — | — | Y | — |
| 阜阳 | LOGNO | Y | — | — | — |
| 横排头 | LOGNO | — | — | — | Y |
| 鲁台子 | GA | Y | — | — | — |
| 蚌埠 | WEI | Y | — | — | — |

* Y 表示选择的表中 4 种模型的一种,—表示没有选择的模型; AIC 值作为优选模型和最优分布选择的依据.

表 3 与表 2 相对应的各站点最优概率分布的 AIC 值
Tab.3 The AIC values of the best probability distribution for the corresponding stations in Tab.2

| 站点 | AIC | | | |
|-----|--------|----------------|----------------|------------------------------|
| | 平稳性模型 | θ_1 非平稳 | θ_2 非平稳 | θ_1 、 θ_2 均非平稳 |
| 息县 | 1076.8 | 1078.7 | 1081.2 | 1083.2 |
| 潢川 | 564.3 | 566.0 | 567.7 | 569.5 |
| 班台 | 942.1 | 934.0 | 941.2 | 935.4 |
| 王家坝 | 965.4 | 967.3 | 967.8 | 969.4 |
| 蒋家集 | 1004.3 | 1005.6 | 1002.8 | 1010.2 |
| 阜阳 | 997.1 | 999.0 | 998.0 | 999.9 |
| 横排头 | 597.3 | 599.2 | 598.6 | 596.7 |
| 鲁台子 | 1102.9 | 1104.8 | 1107.4 | 1109.2 |
| 蚌埠 | 1093.4 | 1095.4 | 1097.2 | 1099.0 |

3.4 基于 GAMLSS 模型残差分析

残差图的分布趋势可以帮助判断所拟合的线性模型是否满足有关假设,残差序列的分布状况是评估模型拟合效果的重要依据^[27]. 为检验表 3 最终确定的回归模型是否合理,图 4 是将各站点年最大洪峰流量序列服从表 3 最优分布模型条件下的 GAMLSS 模型进行残差分析. 各站点的样本点沿着红色曲线并且位于两条黑色曲线(95%置信曲线)中间,显示 GAMLSS 模型所对这 9 个站点拟合较好(图 4). 为了更直观地反映残差的分布情况,绘出残差正态 QQ 图(图 5). GAMLSS 模型拟合的各个站点正态 QQ 图的残差点沿着理论直线分布,基本保持一致状态,而且概率点据相关系数 R 也非常接近 1,这说明实际残差序列与理论残差序列有非常好的相关关系. 因此从残差分布的角度,各站点基于 GAMLSS 模型选择的最优拟合函数对淮河中

上游 9 个站点具有很好的拟合效果.

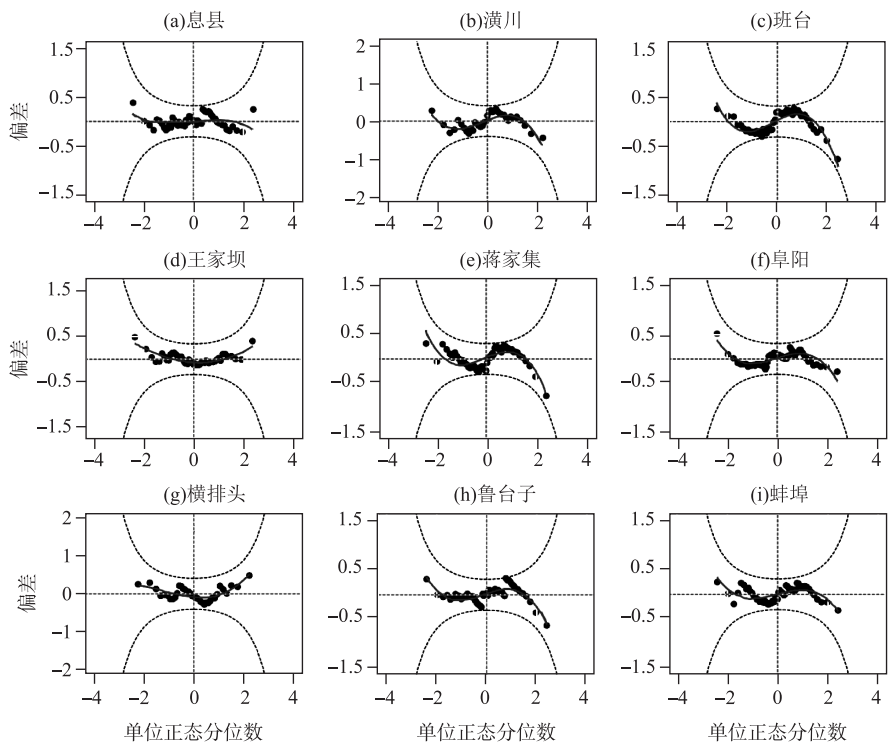


图 4 各站点 GAMLSS 拟合的残差检测图

Fig.4 Residual detection of GAMLSS fit at each site

基于 GAMLSS 模型得到年最大洪峰流量序列的最优拟合结果(图 4、图 5),根据参数的计算结果做出分位图(图 6),然后统计落在每条分位曲线下方实际的频率值(表 4).各站点年最大洪峰流量序列的分位图中,非平稳性站点中横排头站和蒋家集站的分位曲线随时间序列变化呈现波动趋势,班台站的点虽然波动趋势不明显,但其在 25% 和 75% 分位线上产生较大偏差,均为 6.15%,实际频率分布不是很合理.其余平稳性站点中除阜阳站在 25% 和潢川站在 75% 偏差较大外,总体频率分布较为合理,而且落在各条分位曲线下方

表 4 实际频率与理论分位数曲线概率对比

Tab.4 Comparison of actual frequency and theoretical quantile curve probability

| 水文序列 | 分布类型 | 分位曲线/% | | | | |
|------|-------|--------|-------|-------|-------|-------|
| | | 5 | 25 | 50 | 75 | 95 |
| 息县 | WEI | 6.56 | 27.87 | 49.18 | 68.85 | 96.72 |
| 潢川 | WEI | 8.10 | 29.73 | 48.65 | 67.57 | 97.30 |
| 班台 | WEI | 3.28 | 31.15 | 47.54 | 68.85 | 100 |
| 王家坝 | WEI | 3.77 | 20.75 | 54.72 | 77.36 | 94.34 |
| 蒋家集 | GA | 3.28 | 29.51 | 52.46 | 67.21 | 100 |
| 阜阳 | LOGNO | 6.56 | 32.79 | 44.26 | 68.85 | 96.72 |
| 横排头 | LOGNO | 2.70 | 24.32 | 54.05 | 78.38 | 91.89 |
| 鲁台子 | GA | 4.92 | 26.23 | 49.18 | 73.77 | 98.36 |
| 蚌埠 | WEI | 4.92 | 29.51 | 50.82 | 72.13 | 96.72 |

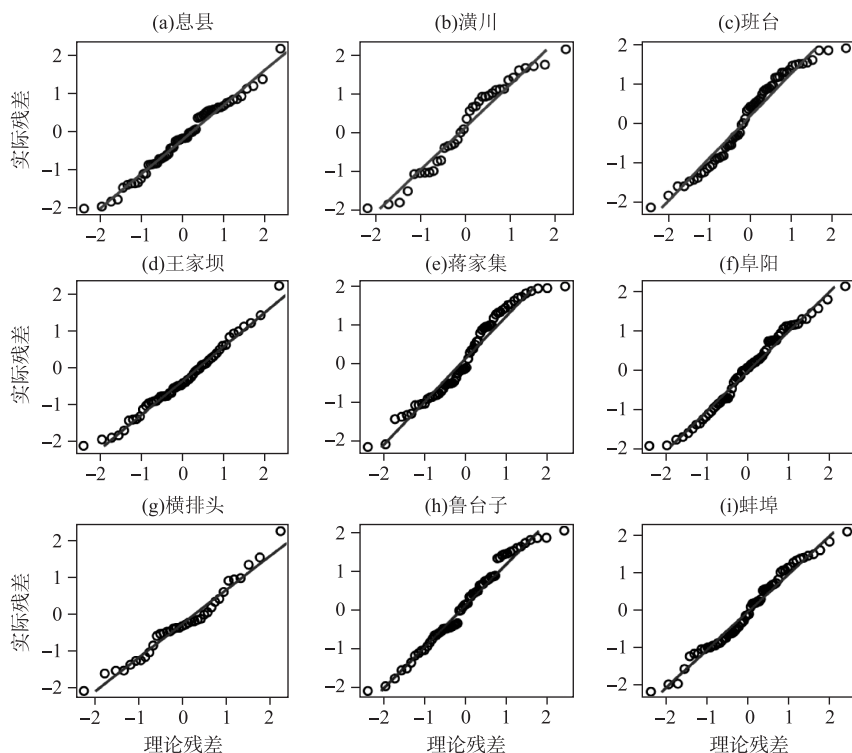


图 5 基于 GAMLSS 各站点线性回归正态 QQ 图

Fig.5 Based on GAMLSS site linear regression normal QQ map

实际的频率与理论概率小,偏差值仅为 0.08%。由于水文序列趋势变化可能为整体趋势,也可能为局部趋势,所以对于水文长序列变化分析仅用平稳性条件下的线性趋势反映整个序列的变化并不合适。本文选用的水文序列较长,如果用平稳性的线性描述较长序列的趋势就可能造成偏差,淮河中上游非平稳性站点序列的实际和理论分位图在局部时段出现了明显的波动偏差(图 6)。故而基于 GAMLSS 模型的两参极值分布可能会使站点频率曲线分布产生偏差,则需要通过 GAMLSS 模型检验站点的平稳性,然后根据实际与理论曲线的偏差大小,选择是否要引入更多的参数来描述序列非线性趋势变化。

3.5 非平稳性洪水频率分析

图 7 为非平稳性洪水频率分析模型对淮河流域洪水频率分析,结果发现同一站点的年最大洪峰流量在不同的年份其设计流量值是不同的。横排头和蚌埠水文站年不同重现期的设计流量随时间增加而增加,其他站点不同重现期的设计流量则恰好相反。图 8 为各站点年最大洪峰流量累积频率曲线变化图,通过皮尔逊Ⅲ型分布模型计算所得,其是经验频率曲线拟合最优的一种模型^[28]。因经验点与 $Cs=2Cv$ 曲线拟合较好,选择此曲线作为皮尔逊Ⅲ型分布累积频率曲线。将皮尔逊Ⅲ型分布模型与非平稳性模型计算所得流量设计值相比较分析显示(图 7、图 8 和表 5),淮河流域非平稳设计流量随着时间跨度呈减小趋势。而各站点在 10 年一遇与 20 年一遇的非平稳设计流量值与皮尔逊Ⅲ型分布设计流量值相差不大,但随着时间跨度的增加,30 年一遇、50 年一遇和 100 年一遇的设计流量相差却越来越大。尽管王家坝、鲁台子、蚌埠和横排头站使用非平稳模型计算的不同重现期的设计流量随着时间变化呈降低趋势,但是其非平稳性模型计算的不同重现期的设计流量远高于平稳性模型,说明目前现有的防洪工程设计标准较低,对于未来洪水抵御能力弱;息县、潢川、班台和蒋家集在 2010 年之后,非平稳性的重现期 10 年和 20 年下的设计流量低于平稳性的。淮河流域应当提高超过 30 年一遇的的防洪标准,提高淮河流域应对大的洪水事件能力。

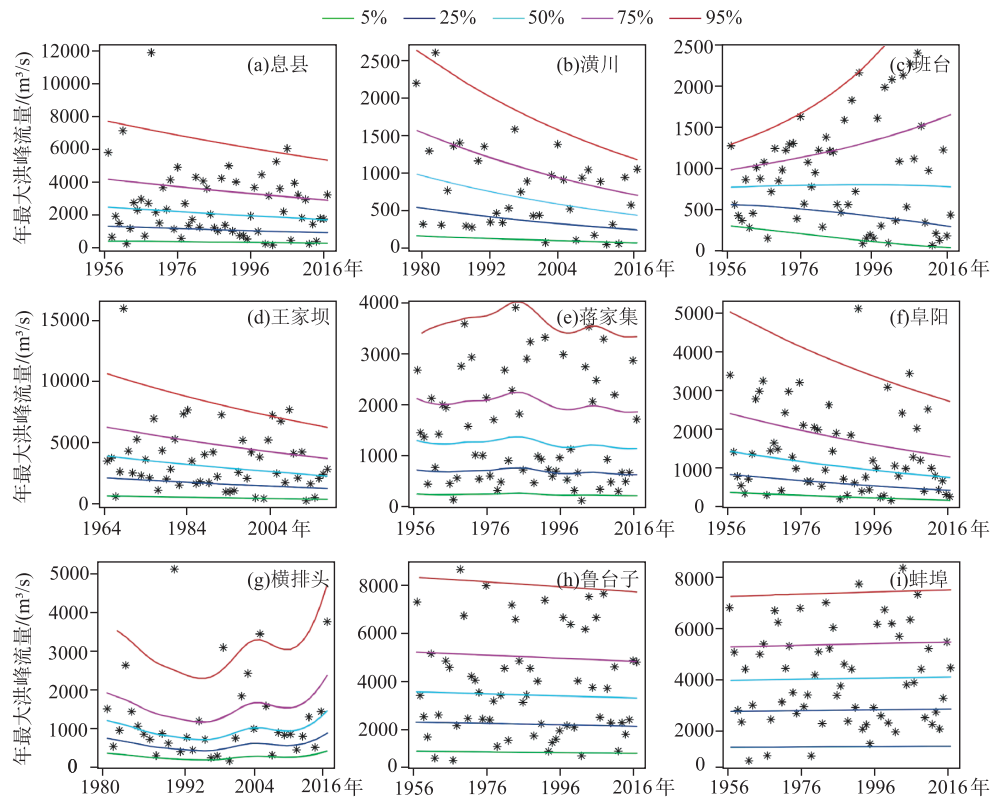


图 6 各站点年最大洪峰流量序列分位图

Fig.6 The annual maximum peak flow sequence of each station

表 5 基于皮尔逊Ⅲ不同重现期的年最大洪峰流量(m³/s)

Tab.5 Annual maximum peak flow in different return periods based on Pearson type Ⅲ

| 重现期 | 息县 | 王家坝 | 鲁台子 | 蚌埠 | 潢川 | 班台 | 蒋家集 | 阜阳 | 横排头 |
|---------|------|-------|-------|------|------|------|------|------|------|
| 10 年一遇 | 5160 | 6950 | 6650 | 6650 | 1540 | 1650 | 2810 | 2870 | 2660 |
| 20 年一遇 | 6640 | 8910 | 8060 | 7830 | 1970 | 2080 | 3470 | 3560 | 3460 |
| 30 年一遇 | 7800 | 10250 | 8790 | 8410 | 2340 | 2340 | 3920 | 4020 | 4090 |
| 50 年一遇 | 8240 | 11010 | 9330 | 8840 | 2400 | 2480 | 4190 | 4340 | 4350 |
| 100 年一遇 | 9500 | 12590 | 10400 | 9670 | 2750 | 2830 | 4760 | 5000 | 5030 |

图 9 给出了洪水放大因子和年最大洪峰流量百年一遇重现期随时间的变化情况(30 年间隔). 横排头站和蚌埠站(图 9)经过 30 年的变化,其洪水放大因子随着时间增加呈上升趋势且大于 1,百年一遇重现期不足 80 年,意味着现有的防洪工程设计标准可能无法满足未来防洪需求. 潢川站、王家坝站和阜阳站洪水放大因子均小于 1,这些站点经过 30 年的变化百年一遇重现期将均超过 300 年,淮河流域的防洪取得了显著的成就. 同时洪涝灾害频发对农业生产和社会经济发展起阻碍作用,洪涝灾害的受灾面积及成灾面积综合反应了灾害对粮食产量波动的危害程度. 通过对淮河流域 9 个水文站点与洪涝灾害成灾及受灾面积做相关性分析可知,各站点年最大洪峰流量与淮河流域、安徽省灾害面积相关性基本上都通过了 95% 的显著性检验,王家坝、蒋家集、鲁台子和蚌埠站与淮河流域和安徽省洪涝灾害的相关性均突破了 99% 的显著性检验(图 10, 图 11). 尤其是蚌埠站的相关系数最大,其位于淮河中游与洪泽湖之间,控制的集水面积高达 121330 km²,淮河蚌埠河段河床平缓,易受到洪涝灾害的影响. 从区域性差异来看,淮河流域各站点年最大洪

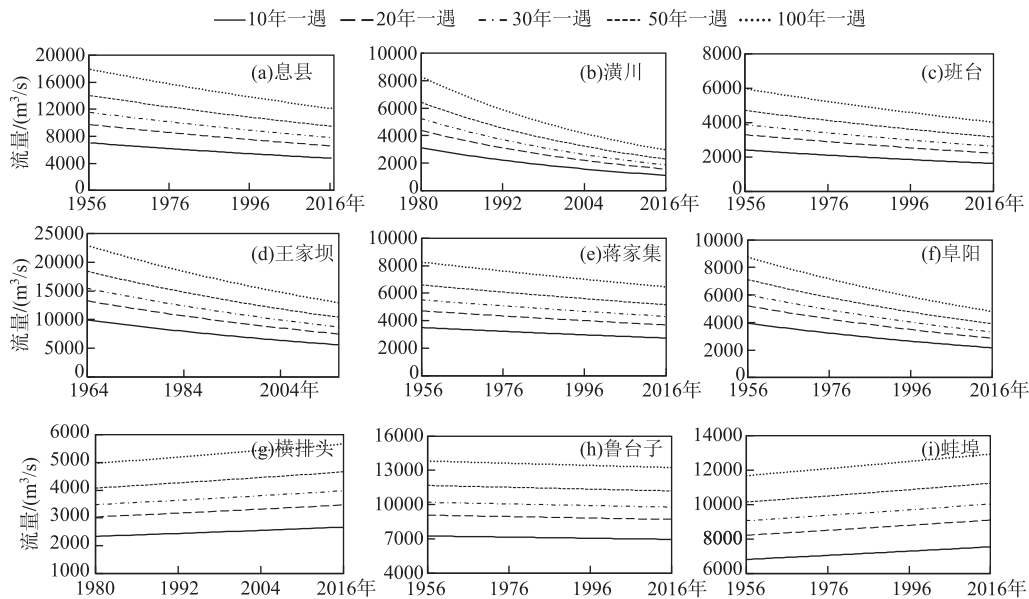


图 7 10 年、20 年、30 年、50 年一遇和百年一遇设计流量值随时间变化

Fig.7 Design flow value of annual maximum peak flow in different return periods with time

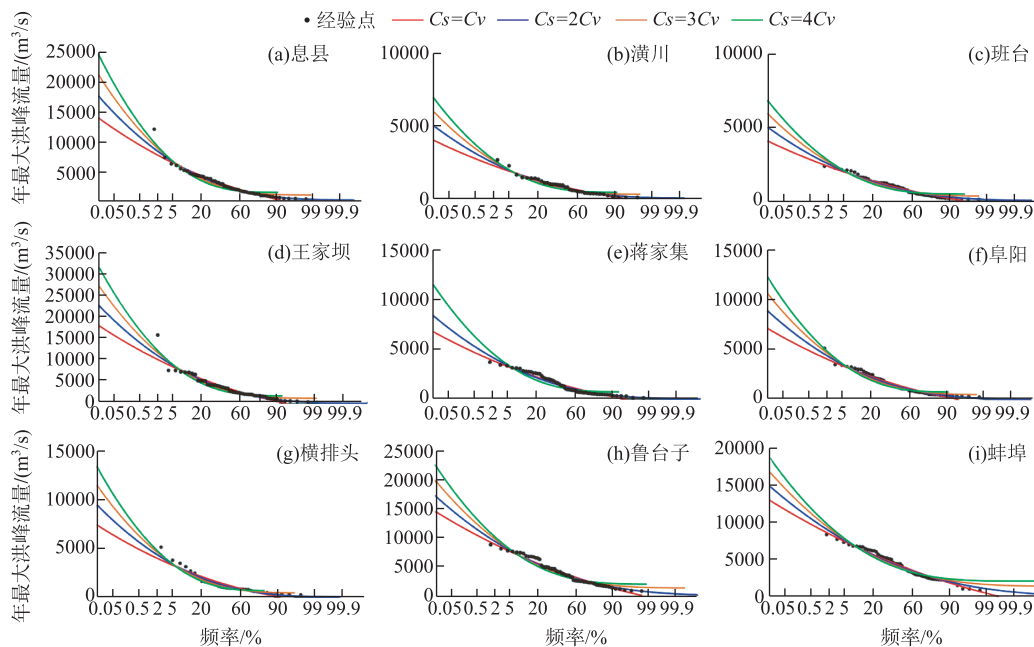


图 8 基于皮尔逊Ⅲ各站点年最大洪峰流量累积频率曲线

Fig.8 Cumulative frequency curves of annual maximum peak flow at each station based on Pearson type Ⅲ

峰流量变化对安徽省水灾受灾面积影响较大,这主要是因为淮河流域暴雨洪涝灾害主要集中在流域中上游安徽阜南县蒙洼蓄洪区及周边地势低洼地,使得安徽省成为淮河流域洪涝灾害受灾最严重的省区. 安徽皖北地区的耕地面积占全省总耕地面积 56.2%,由于耕地抵御洪水灾害能力弱,易受淹农田比例非常高^[29].

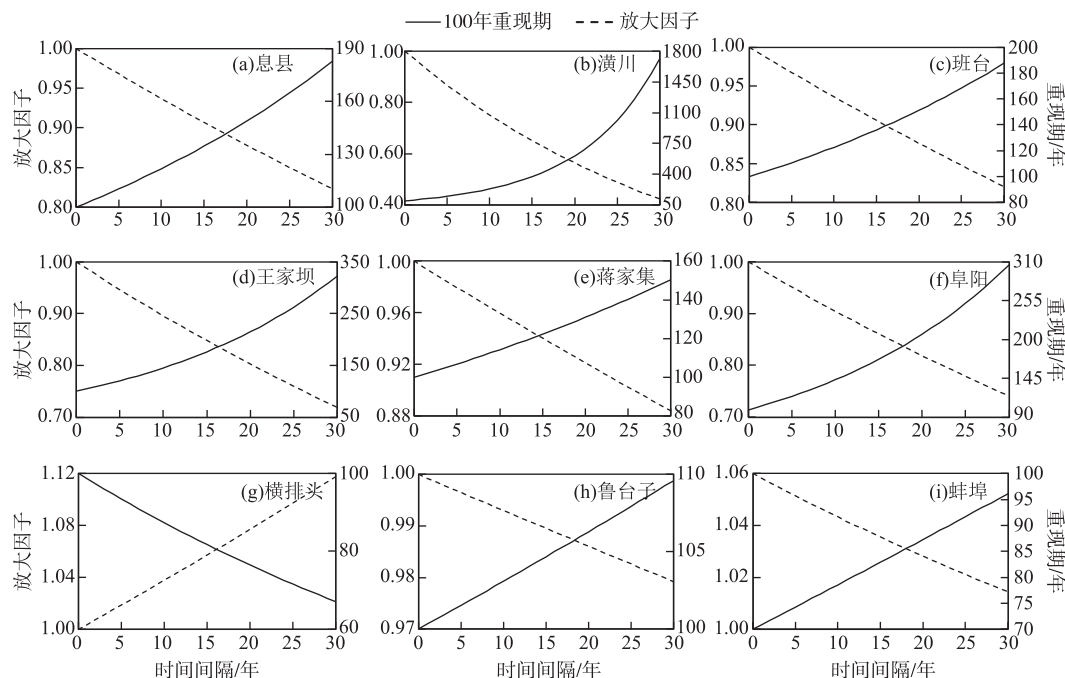


图9 洪水放大因子和百年一遇重现期随时间的变化

Fig.9 The flood magnification factors and 100-year return period changing with time

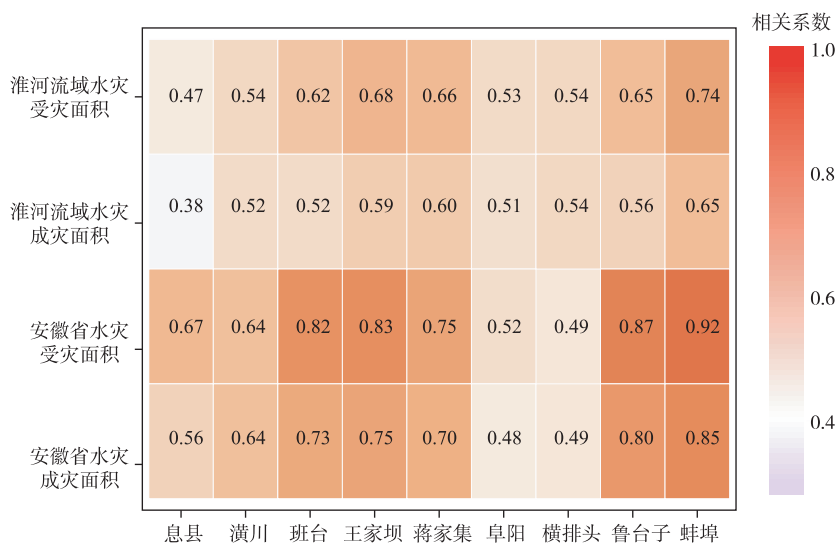


图10 各站点年最大洪峰流量与淮河流域、安徽省灾害面积相关性

Fig.10 The correlation between maximum annual peak flow of each station and the disaster area of Huaihe River Basin, Anhui Province

总的来说,淮河流域不仅是中国各大流域人口密度最高的地区,也是重要粮食生产基地和能源基地,1983—2014年淮河流域粮食播种面积从 $16.4 \times 10^4 \text{ km}^2$ 增加到了 $19.2 \times 10^4 \text{ km}^2$. 大量种植的水稻、小麦等粮

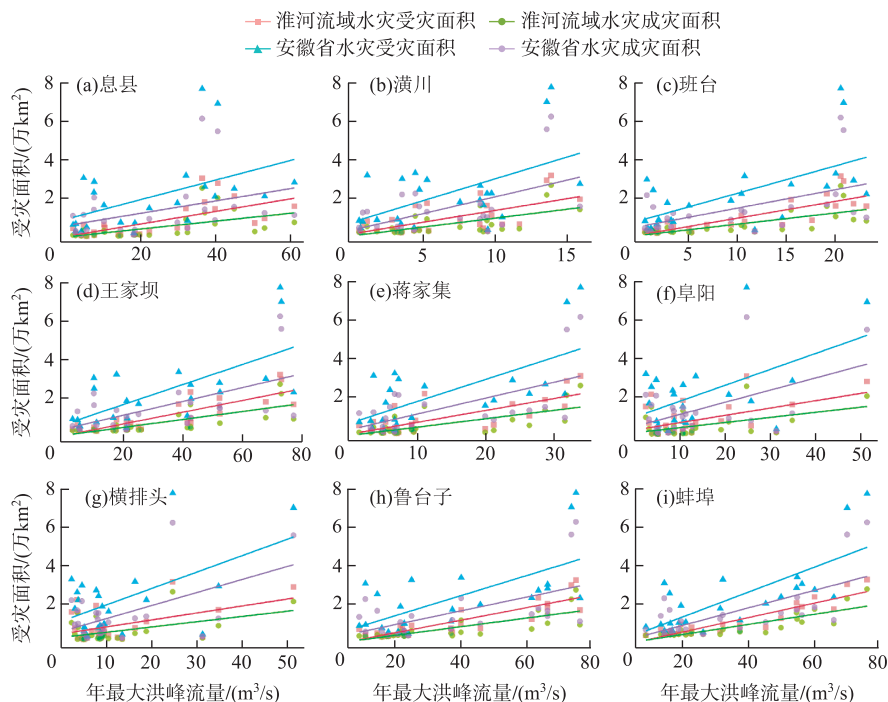


图 11 各站点年最大洪峰流量与淮河流域、安徽省灾害面积散点图

Fig.11 The scatter plot between maximum annual peak flow of each station and the disaster area of Huaihe River Basin and Anhui Province

农作物是洪涝灾害最重要的承灾体. 淮河流域自 2000 年以来 GDP 和人口也在不断增加,截至 2014 年 GDP 增加了 55846.03 亿元,人口也增加了 1719.99 万人. 这在一定程度上改变了淮河流域水文情势,以致淮河流域年最大洪峰流量的设计流量在极端事件增加的情况下随着时间跨度呈减小趋势. 但值得注意的是,人们在推进社会经济发展的同时,盲目围垦造田、乱砍乱伐等引起河道变更、水土流失也会造成淮河自身抗涝能力减弱.

4 结论

通过 Pettitt 与 M-K 非参数检验法、GAMLSS 模型和非平稳性洪水频率分析模型等方法,对淮河流域 9 个水文站点年最大洪峰极值序列的非平稳性假设进行了全面分析,得出以下结论:

1) 通过 Pettitt 法检验淮河中上游 9 个水文站点时间序列要素的均值和方差变化进行变异点分析发现:潢川、横排头和蚌埠站点的年最大洪峰流量未发生明显变异;共有 6 个站点的年最大洪峰流量发生均值或方差变异,变异时间主要集中在 2000 年左右. 其中班台、王家坝和阜阳站的年最大洪峰流量发生均值变异,变异时间分别为 2009、1992 和 2000 年;班台站、王家坝站、阜阳站的年最大洪峰流量发生方差和均值变异.

2) 淮河中上游 9 个水文站点中有 5 个站点的最优拟合分布函数是两参数的 Weibull 函数,其次是 Log-normal 分布函数;班台、蒋家集和横排头站的选择的最优分布模型为非平稳性模型,其余 6 个站点选择的最优模型是平稳性模型. 各站点基于 GAMLSS 模型选择的最优拟合函数对淮河流域站点具有很好的拟合效果. 非平稳性站点中横排头站和蒋家集站的分位曲线随时间序列变化呈现波动趋势,班台站在 25% 和 75% 分位线上产生较大偏差,实际分布不是很合理. 其余平稳性站点中除阜阳站在 25% 和潢川站在 75% 偏差较大外,总体频率分布较为合理.

3) 通过比较各站点在皮尔逊Ⅲ型分布模型和非平稳性洪水频率分析模型的设计流量值发现,尽管淮河

流域非平稳设计流量随着时间跨度呈减小趋势,各站点在 10 年一遇与 20 年一遇的非平稳设计流量值与皮尔逊Ⅲ型分布设计流量值相差不大,但 30 年一遇、50 年一遇和 100 年一遇的设计流量相差却越来越大,淮河干流的王家坝、鲁台子和蚌埠在非平稳性的设计流量远高于在皮尔逊Ⅲ型的设计流量。

4)横排头站和蚌埠站经过 30 年变化,其洪水放大因子随着时间增加呈上升趋势且大于 1,百年一遇重现期不足 80 年,意味着现有的防洪过程设计标准可能无法满足未来防洪需求。潢川站、王家坝站和阜阳站经过 30 年变化,百年一遇重现期将均超过 300 年。各站点年最大洪峰流量与淮河流域、安徽省水灾面积相关性基本上都通过了 95% 的显著性检验,王家坝、蒋家集、鲁台子和蚌埠站与淮河流域和安徽省洪涝灾害的相关性均突破了 99% 的显著性检验。

5 参考文献

- [1] Schiermeier Q. Increased flood risk linked to global warming. *Nature*, 2011, **470**(7334): 316.
- [2] Reinsurance M. Topics geo: natural catastrophes 2014. Munich Re, 2015: 26-27.
- [3] European Commission. Directive 2007 /60 /EC of the European parliament and of the council of 23 October 2007 on the assessment and management of flood risk. *Official Journal of the European Union*, 2007, **288**: 27-34.
- [4] Schelfaut K, Panneman B, Craats IVD *et al.* Bringing flood resilience into practice: The FREEMAN Project. *Environmental Science & Policy*, 2011, **14**(7): 825-833.
- [5] Priest SJ, Suykens C, Rijswick HFMWV *et al.* The European Union approach to flood risk management and improving societal resilience: Lessons from the implementation of the floods directive in six European countries. *Ecology and Society*, 2016, **21**(4): 4-8.
- [6] Milly PCD, Betancourt J, Falkenmark M *et al.* On critiques of "Stationarity is Dead: Whither Water Management?". *Water Resources Research*, 2015, **51**: 7785-7789.
- [7] Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society*, 2010, **54**(3): 507-554.
- [8] Cunderlik JM, Burn DH. Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, 2003, **276**(14): 210-223.
- [9] Chen PC, Wang YH, You JY *et al.* Comparison of methods for non-stationary hydrologic frequency analysis: Case study using annual maximum daily precipitation in Taiwan. *Journal of Hydrology*, 2016, **545**: 197-211.
- [10] Ren LL, Shen HR, Yuan F *et al.* Hydrological drought characteristics in the Weihe catchment in a changing environment. *Advances in Water Science*, 2016, **27**(4): 492-500. [任立良, 沈鸿仁, 袁飞等. 变化环境下渭河流域水文干旱演变特征剖析. 水科学进展, 2016, **27**(4): 492-500.]
- [11] Lu F, Xiao WH, Yan DH *et al.* Progresses on statistical modeling of non-stationary extreme sequences and its application in climate and hydrological change. *Journal of Hydraulic Engineering*, 2017, **48**(4): 379-389. [鲁帆, 肖伟华, 严登华等. 非平稳时间序列极值统计模型及其在气候-水文变化研究中的应用综述. 水利学报, 2017, **48**(4): 379-389.]
- [12] Sang YF, Xie P, GU HT *et al.* Discussion on several major issues in the studies of hydrological nonstationarity. *Chinese Science Bulletin*, 2017, (4): 254-261. [桑燕芳, 谢平, 顾海挺等. 水文过程非平稳性研究若干问题探讨. 科学通报, 2017, (4): 254-261.]
- [13] Liang ZM, Hu YM, Wang J *et al.* Estimation of design flood using equivalent reliability method under changing environment. *Advances in Water Science*, 2017, **28**(3): 398-405. [梁忠民, 胡义明, 王军等. 基于等可靠度法的变化环境下工程水文设计值估计方法. 水科学进展, 2017, **28**(3): 398-405.]
- [14] Du H, Xia J, Zeng SD *et al.* Temporal and spatial variations and statistical models of extreme runoff in Huaihe River Basin. *Acta Geographica Sinica*, 2012, **67**(3): 398-409. [杜鸿, 夏军, 曾思栋等. 淮河流域极端径流的时空变化规律及统计模拟. 地理学报, 2012, **67**(3): 398-409.]
- [15] Wei FY, Zhang T. Study on oscillation characteristics of summer precipitation in Huaihe River Basin and its relationship with climate background. *Science China Earth Sciences*, 2009, **10**: 1360-1374. [魏凤英, 张婷. 淮河流域夏季降水的振荡特征及其与气候背景的联系. 中国科学: 地球科学, 2009, **10**: 1360-1374.]
- [16] Liu SM, Wang H, Yan DH *et al.* Spatial-temporal evolution of isolated rainstorm events in Huai River basin under the background of climate change. *Glacier Permafrost*, 2016, **38**(5): 1264-1272. [刘思敏, 王浩, 严登华等. 气候变化背景下淮河流域场次暴雨事件时空演变分析. 冰川冻土, 2016, **38**(5): 1264-1272.]

- [17] Pettitt AN. Anon-parametric approach to the change-point problem. *Journal of the Royal Statistical Society*, 1979, **28**(2): 126-135.
- [18] Zhang HB, Yu YH, Nan ZN *et al.* TFPW-BS-Pettitt method for detection of multiple change-points in the mean of hydrological series. *Journal of Hydroelectric Engineering*, 2017, **36**(7): 14-22. [张洪波, 余茆皓, 南政年等. 基于 TFPW-BS-Pettitt 法的水文序列多点均值跳跃变异识别. 水利发电学报, 2017, **36**(7): 14-22.]
- [19] Beaulieu C, Chen J, Sarmiento JL. Change-point analysis as a tool to detect abrupt climate variations. *Philosophical Transactions*, 2012, **370**: 1228-1249.
- [20] Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society*, 2010, **54**(3): 507-554.
- [21] Vogel RM, Yaindl C, Walter M. Nonstationarity: Flood magnification and recurrence reduction factors in the United States. *Jawra Journal of the American Water Resources Association*, 2011, **47**(3): 464-474.
- [22] Anhui Local Chronicles Compilation Committee ed. Anhui volunteers: Natural environment. Beijing: Fangzhi Press, 1999. [安徽省地方志编委会. 安徽省志: 自然环境志. 北京: 方志出版社, 1999.]
- [23] Ye JY, Huang Y, Zhang CL *et al.* Spatial-temporal variations of climate change of the Huai River Basin during recent 50 years. *Journal of Eco-environment*, 2016, **25**(1): 84-91. [叶金印, 黄勇, 张春莉等. 近 50 年淮河流域气候变化时空特征分析. 生态环境学报, 2016, **25**(1): 84-91.]
- [24] Lu YY, Wu BW, Tian H *et al.* Spatial and temporal variability characteristic of precipitation in Huai River Basin during 1961-2005. *Resources and Environment in the Yangtze Basin*, 2011, **20**(5): 567-573. [卢燕宇, 吴必文, 田红等. 基于 Kriging 插值的 1961-2005 年淮河流域降水时空演变特征分析. 长江流域资源与环境, 2011, **20**(5): 567-573.]
- [25] Huang R. Research on evolution and countermeasures of droughts-floods abrupt alternation events in Huai River basin [Dissertation]. Beijing: China Water Resources and Hydropower Research Institute, 2015. [黄茹. 淮河流域旱涝急转事件演变及应对研究[学位论文]. 北京: 中国水利水电科学研究院, 2015.]
- [26] Gu XH, Zhang Q, Wang ZZ. Evaluation on stationarity assumption of annual maximum peak flows during 1951-2010 in the Pearl River Basin. *Journal of Natural Resources*, 2015, **30**(5): 824-835. [顾西辉, 张强, 王宗志. 1951-2010 年珠江流域洪水极值序列平稳性特征研究. 自然资源学报, 2015, **30**(5): 824-835.]
- [27] Jiang C, Xiong LH. Trend analysis for the annual discharge series of the Yangtze River at the Yichang Hydrological Station based on GAMLSS. *Acta Geographica Sinica*, 2012, **67**(11): 1505-1514. [江聪, 熊立华. 基于 GAMLSS 模型的宜昌站年径流序列趋势分析. 地理学报, 2012, **67**(11): 1505-1514.]
- [28] Ministry of Water Resources of the People's Republic of China ed. Water and Hydropower Engineering Design Flood Calculation Specification SL44-2006. Beijing: China WaterPower Press, 2006. [中华人民共和国水利部. 水利水电工程设计洪水计算规范 SL44-2006. 北京: 中国水利水电出版社, 2006.]
- [29] Xiang SK. Land resources overview and analyses of structure and condition of land use in Anhui Province. *Guangdong Agricultural Sciences*, 2010, **37**(8): 349-353. [项思可. 安徽土地资源概况及土地利用结构与状况分析. 广东农业科学, 2010, **37**(8): 349-353.]