

Analysis of Rainfall Extreme in Han River Basin Based on PDS/GP Model

Dan Li, Shenglian Guo, Xingjun Hong

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan Hubei
Email: 54dli@163.com

Received: Mar. 28th, 2015; accepted: Apr. 21st, 2015; published: Apr. 27th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The monthly rainfall of 15 meteorological stations in Han River basin from 1961 to 2010 was taken as case study. Two methods of mean residual life and change of parameters were combined to determine the proper thresholds for each data series, and the fit residuals method was adopted to reduce the uncertainty during the process. Partial duration series (PDS) and annual maximum (AM) series were obtained and fitted by General Pareto (GP) and Pearson Type III (PIII) distributions, respectively. The PDS/GP and AM/PIII models were constructed and used to analyze rainfall extreme variations and predict rainfall quantiles with 20a, 50a, 100a and 200a return periods for each station. The results and comparison indicated that the estimates supplied by AM/PIII model were mostly larger than that of PDS/GP model, while the differences were not so much significant because of the temporal and spatial variations of the monthly rainfall extremes.

Keywords

Rainfall Extreme, Han River Basin, Return Period, PDS/GP Model, AM/PIII Model

基于PDS/GP模型分析汉江流域降雨极值特征

李丹, 郭生练, 洪兴骏

武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉
Email: 54dli@163.com

收稿日期: 2015年3月28日; 录用日期: 2015年4月21日; 发布日期: 2015年4月27日

作者简介: 李丹(1980-), 女, 高级工程师, 博士后, 主要从事水文学及水资源配置、水资源管理等方面的研究工作。

摘要

以汉江流域15个气象站1961~2010年的实测逐月降雨系列为例,将平均超出量函数和参数估计量变化两种阈值选取方法相结合,确定了各系列的合理阈值,并采用拟合残差法来减小阈值选取的不确定性。采用超定量(PDS)和年最大(AM)取样方法,选择广义Pareto (GP)和皮三型(PIII)分布,分别构建PDS/GP和AM/PIII模型,分析汉江流域降雨极值变化特征,预测汉江流域各站20年、50年、100年和200年一遇的月降雨极值。对比分析两种模型的计算结果,AM/PIII模型的估计值普遍大于PDS/GP模型,但受月降雨极值年际变化和空间差异的影响,差别不显著。

关键词

降雨极值, 汉江流域, 重现期, PDS/GP模型, AM/PIII模型

1. 引言

联合国气候变化政府间专门委员会 IPCC (International Panel on Climate Change)第五次发布的特别报告指出,大量观测资料表明,最近 60 年无论是在全球尺度还是区域尺度上,降水、气温还有极端事件发生频率等方面都发生了显著变化[1]。张利平等把极值定义为在给定的时期内,或该时期的一定月份或季节内观测到的水文要素的最高值或最低值[2]。但在现代极值理论研究中,不同的建模方法对可以作为极值的观测数据有不同要求[3]。目前,极端水文事件是国内外研究的一个热点问题,而作为重要影响因子之一的降雨量的极值变化特征研究也在相关领域得到关注和重视。翟盘茂等利用中国 296 个分布均匀的测站的逐日降水资料,研究了我国过去 45 年中降水量、降水频率、降水强度等方面的极值变化趋势[4];黄琰等从极值概率角度研究了我国极端气温和降水的空间分布随时间变化的情况[5];姜彤等运用广义极值分布、广义帕累托分布、广义逻辑分布和韦克比分布等 4 种分布函数定量拟合了长江流域降水极值的概率分布[6],都促进了极值理论在降水极值变化趋势分析方面的应用,但这些研究多采用最大值系列作为样本进行分析,而忽略了时段内发生的其他独立极值样本。本文拟基于超定量系列建立降雨极值分析模型,并以汉江流域为例,通过探讨汉江流域降雨极值的时空分布规律和统计概率特征,为进一步深入研究气候变化对汉江流域极端事件的影响提供一些思路和参考。

2. 研究区概况

汉江是长江中游最大的支流,流域面积约 15.9 万 km^2 ,干流全长 1577 km,流经陕西、湖北两省,于武汉市注入长江。汉江干流根据其河道特点,分为上、中、下三个河段,丹江口以上为上游,河段长 925 km,占汉江全长的 59%,控制流域面积 9.52 万 km^2 ;丹江口至钟祥碾盘山为中游,全长 270 km,占汉江全长的 17%,区间流域面积 4.68 万 km^2 ;碾盘山以下为下游,长 382 km,占汉江总长的 24%,集水面积 1.7 万 km^2 。

汉江流域处于东亚副热带季风区,其降水主要来源于东南和西南两股暖湿气流。汉江流域多年平均降水量为 904 mm,折合降水总量为 1399.5 亿 m^3 ,占长江流域降水量的比重约为 7%。由于纬度和地形条件的差异,年降水量呈现南岸大于北岸,上、下游大,中游小的地区分布规律,流域内多年平均降水量为 700 mm~1800 mm。汉江上游年内降水有三个集中时段:4月下旬至5月下旬为春汛,6月下旬至7月下旬为夏汛,8月下旬至10月为秋汛,其中夏汛时段雨量最大,秋汛次之,但遇降雨天气有异时,秋汛雨量也会超过夏汛。相比而言,下游地区春汛、秋汛主峰不如上游明显。流域降水年内分配不均匀,5~10月降水占全年的 70%~80%,7月、8月、9月这三个月的降水量占年降水量的 40%~60%,年降水量系列 C_v 值上游地区为 0.20,中、下游地区为 0.25 [7]。

本次采用汉江流域 15 个气象站 1961~2010 年共 50 年的逐月降水系列进行降雨极值特征分析,站点分布见图 1。

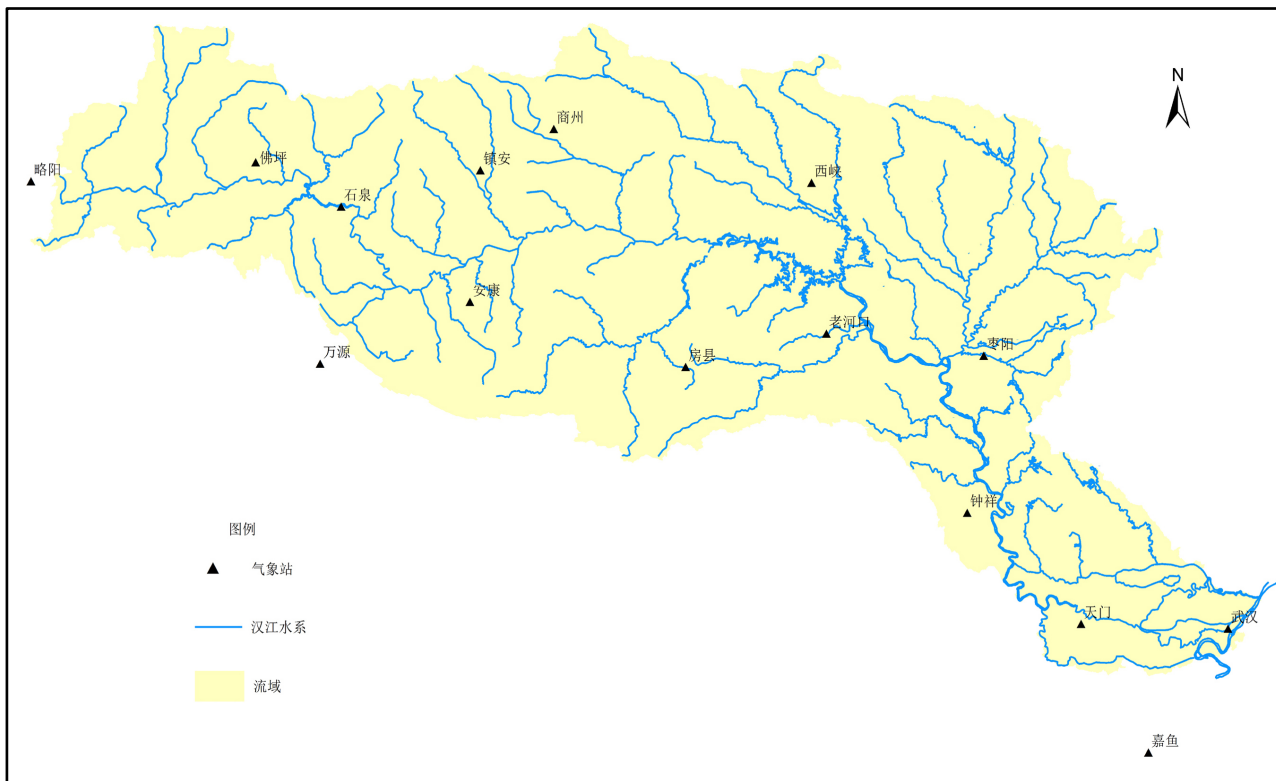


Figure 1. Spatial distribution of the meteorological stations in Han River basin

图 1. 汉江流域气象站点分布示意图

3. 模型与方法

3.1. PDS/GP 模型

PDS (Partial Duration Series)在水文上也称为超定量 POT (Peak Over Threshold)系列[8], 超定量系列样本数为超定量发生次数[9]。经典的 PDS 模型假定: ① 超定量发生次数服从 Poisson 分布; ② 超定量样本独立; ③ 超定量样本服从指数分布(EXP)。但是, 近十几年的研究表明, PDS 超定量样本符合 GP (General Pareto)分布, 指数分布只是其特例。因此, 本文采用 PDS/GP 模型[10]分析汉江流域降雨极值特征。由于本次所用资料系列的时间尺度为月, 根据 Willem's 等的研究[11]-[13], 可认为样本满足独立性条件。

按照基本假定 1, 单位时间 t 内发生 n 次超定量的概率服从 Poisson 分布, 即有:

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} \exp(-\lambda t) \quad n = 0, 1, 2, \dots \quad (1)$$

式中, λ 为单位时间内超定量平均发生次数。

符合 GP 分布的 PDS 模型表达为[14]:

$$F(x) = \begin{cases} 1 - \left[1 - \frac{\xi}{\sigma} \times (x - u) \right]^{1/\xi} & \xi \neq 0 \\ 1 - \exp\left(-\frac{x - u}{\sigma}\right) & \xi = 0 \end{cases} \quad (2)$$

式中, σ 为尺度参数, ξ 为形状参数, u 为选定的阈值。

3.2. 阈值选取

由于阈值选取的重要性, 已经有很多文献对此进行了讨论[15]-[17]。但最常用的方法仍然是平均超出量函数(或称平均剩余寿命函数)法, 即确定一个关于 u 的线性函数 $e(u)$:

$$e(u) = E(X - u | X > u) = \frac{\sigma + \xi u}{1 - \xi}, \quad \xi < 1 \quad (3)$$

对于给定的样本 X_1, \dots, X_n , 定义样本的超出量函数为:

$$e_n(u) = \frac{1}{N_u} \sum_{i \in \Delta_n(u)} (X_i - u), \quad u > 0 \quad (4)$$

式中, N_u 为超出量的个数, $\Delta_n(u) = \{i: X_i > u\}$ 表示超过阈值 u 的观测值的下标集合。如果 u_0 是适当的阈值, 则对大于 u_0 的 u , $e_n(u)$ 的斜率将保持不变。

另外一种选取 u 的常用方法是判断 u 值改变引起的参数估计量的变化, 即在一个 u 值取值范围内, 利用超出量估计 σ 和 ξ 。由于 σ 的估计值是 u 的函数, 有必要对其进行修正, 令 σ^* 为修正的尺度参数, 并有:

$$\sigma^* = \sigma - \xi u \quad (5)$$

当 $u > u_0$ 时, $\sigma = \sigma_{u_0} + \xi(u - u_0)$, 带入式(5)可以得到:

$$\sigma^* = \sigma_{u_0} - \xi u_0 \quad (6)$$

此时可以发现, 修正的尺度参数 σ^* 已经与 u 无关。因此, 如果 u_0 是适当的阈值, 相应的超出量服从 GP 分布, 则对大于 u_0 的 u , 相应的 σ^* 和 ξ 的估计量将保持不变。

本次将上述两种方法结合使用来选定合适的阈值, 并通过拟合残差平方和最小的方法来减小主观定性判断选取阈值的不确定性。此外, 为了充分体现 PDS/GP 模型预测极值的优越性, 在阈值选取时确保每年至少有一个数据。

3.3. 参数估计

GP 分布常用的参数估计方法有矩法(MOM)、极大似然法(ML)、概率权重矩法(PWM)、线性矩法(L-M)等。综合比较各种方法的适用性和优劣后, 本次选用 ML 方法进行参数估计。

对于满足 GP 分布的超定量样本分布函数, 其对数似然函数为:

$$L(\sigma, \xi) = N_u \log \xi - N_u \log \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i \in \Delta_n(u)} \log \left(1 + \frac{\xi}{\sigma} x_i\right) \quad (7)$$

式中符号意义同上。使函数 L 达到最大的 σ 和 ξ 即为参数的极大似然估计 $\hat{\sigma}^*$ 和 $\hat{\xi}$ 。

4. 结果和讨论

4.1. 统计特征初步分析

分析近 50 年的资料, 汉江流域 15 个站点的最大月降雨量除了嘉鱼站出现在 1964 年, 其它均出现在 80 年代以后, 直方图显示样本序列均为非正态分布。从图 2 可以看出, 有 11 个站点的峰度系数大于 3, 表明所采用的资料系列具有明显的尖峰厚尾特征; 在所有站点中, 武汉站具有最大的极差值, 万源站次之。(受篇幅限制, 本次仅在上、中、下游各取一个代表站显示结果)。

4.2. 降雨极值分析

根据式(4)计算各站降雨系列在不同阈值下的平均超出量, 并绘制平均剩余寿命图(图 3, 红线标注可能合理

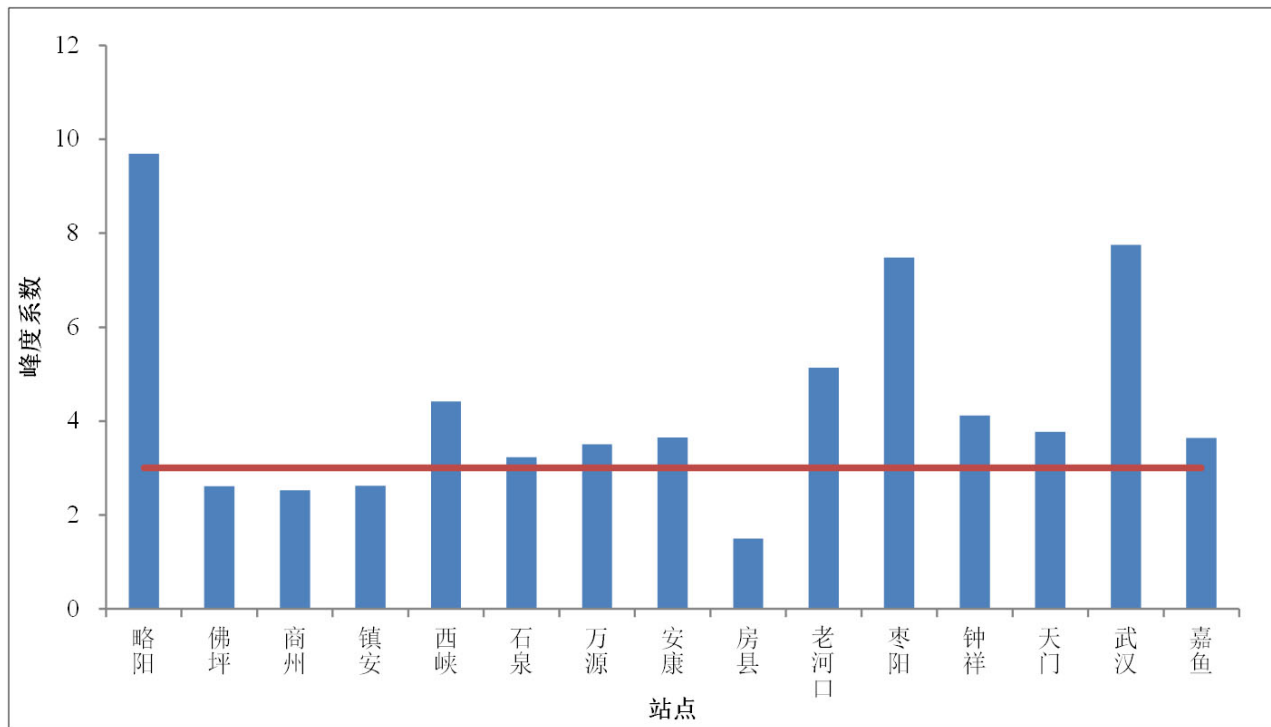


Figure 2. Kurtosis of monthly rainfall series at each station in Han River basin

图 2. 汉江流域各站点降雨系列峰度系数比较图

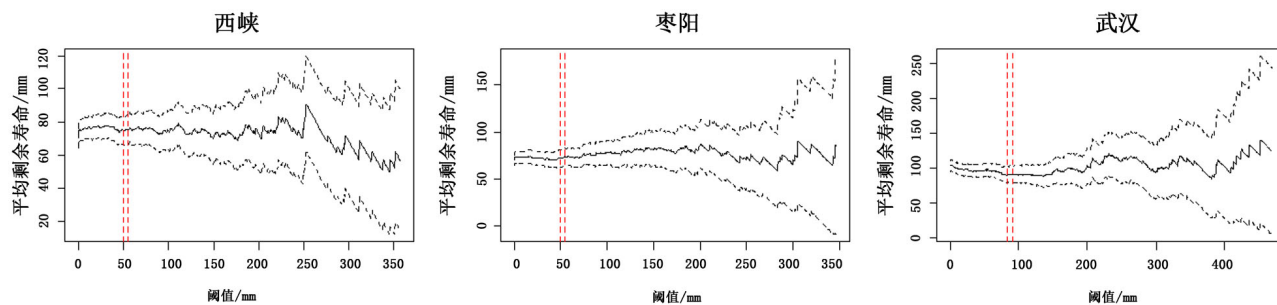


Figure 3. Average residual life plot of monthly rainfall at representation stations in Han River basin

图 3. 汉江流域代表站月降雨平均剩余寿命图

阈值区间，黑色虚线为 95%置信区间)。根据 15 个站点的平均剩余寿命图，进一步确定平均超出量与阈值大致呈线性关系的区间。由于近似线性为定性判断，存在一定误差，本次通过取残差平方和最小的方法进一步缩小阈值的取值范围，提高阈值选取的稳定性。以武汉站为例，当阈值在[83, 112]区间时，平均超出量与阈值大致呈线性关系，进一步将该区间内的数据升序排列后划分为 3 个子区间，分别为[83, 92]、[93, 102]和[103, 112]，并求出每一组数据中的第一点和最后一个点连成的直线的斜率，然后求出每一组数据对该直线的拟合残差平方和，最后选取拟合残差值最小的[83, 92]区间为可能合理阈值区间。用同样的方法可确定其它站点的可能合理阈值区间，结果详见表 1。

在平均剩余寿命图确定的可能合理阈值区间基础上，根据模型的尺度参数和形状参数估计值随阈值的变化情况，找出参数估计值稳定不变的区间。最后利用每个可能的合理阈值计算得到相应的参数估计值，并通过模拟检验找出模拟效果最好的阈值及相应阈值下的 $\hat{\sigma}^*$ 和 $\hat{\xi}^*$ 。

以武汉站为例，在第一种方法确定的[83, 92]区间内， $\hat{\sigma}^*$ 和 $\hat{\xi}^*$ 值在[85.5, 92]区间内基本保持不变(图 4，蓝线

Table 1. Reasonable threshold and corresponding parameters

表 1. 最终确定的合理阈值及相应参数估计值

站名	可能合理阈值区间	u_0	样本数 n	σ^*	ξ
略阳	[30, 35]	31.7	337	80.484	-0.036
佛坪	[57, 66]	61.1	261	95.810	-0.140
商州	[41, 46]	42.9	288	64.076	-0.124
镇安	[44, 50]	48.8	292	71.896	-0.096
西峡	[50, 55]	52.8	281	76.343	-0.016
石泉	[71, 77]	72.0	231	82.762	-0.071
万源	[38, 50]	42.1	357	135.550	-0.113
安康	[78, 80.5]	79.7	213	59.653	0.011
房县	[66, 70]	66.3	257	67.749	-0.147
老河口	[76, 82]	76.1	213	59.522	0.000
枣阳	[49, 54]	50.9	287	67.090	0.069
钟祥	[38, 41]	39.8	370	81.722	-0.043
天门	[41, 45]	41.4	405	89.143	-0.056
武汉	[83, 92]	85.5	278	85.741	0.058
嘉鱼	[83, 89]	83.5	331	96.989	-0.047

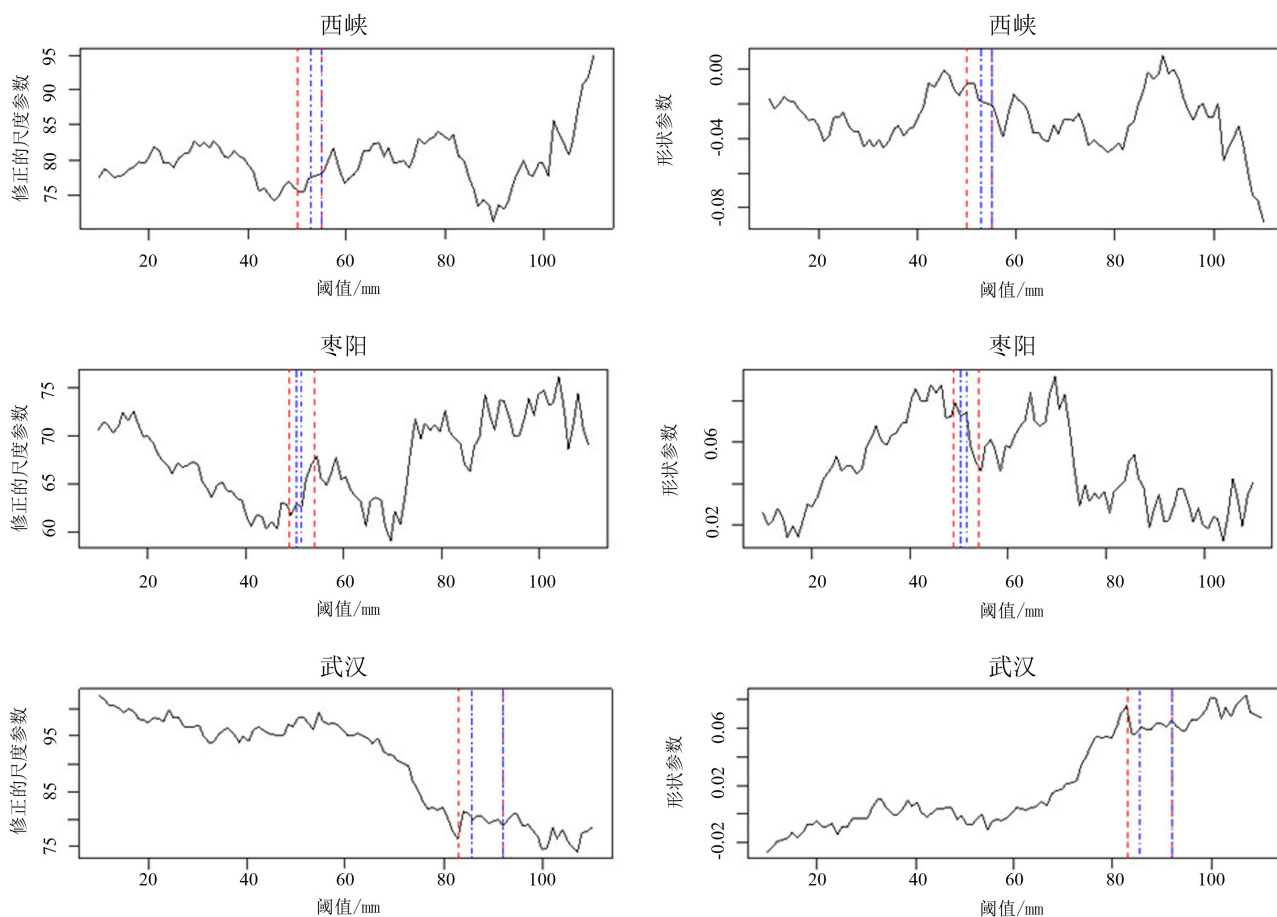


Figure 4. Change of modified scale and shape parameters with the threshold at representative stations

图 4. 汉江流域代表站模型参数随阈值变化情况图

标注可能合理阈值区间), 此时可确定 19 个可能成为合理阈值的月降雨量。通过比较发现, 当阈值为 85.5 时, 模型模拟效果最好(图 5), 相应的 $\hat{\sigma}^*$ 和 $\hat{\xi}$ 分别为 85.741 和 0.058, 标准误分别为 7.378 和 0.062。尺度参数估计值的标准误较大表明超过阈值的月降雨量波动较大。由实测资料分析, 武汉站超阈值样本的最大值与最小值之比为 8.83, 说明参数估计值符合实际情况。用同样的方法可确定其它站点的合理阈值以及相应的参数估计值, 结果详见表 1。

4.3. 重现水平估计

随机变量 x 的 T 年重现水平为:

$$\hat{x}_T = u_0 + \frac{\hat{\sigma}^*}{\hat{\xi}} \left((\lambda T)^{\hat{\xi}} - 1 \right) \tag{8}$$

式中, u_0 为选定的阈值, T 为重现期, 其他符号意义同上。

根据式(8)计算了各站点不同重现期的月降雨极大值, 并利用 Kriging 方法进行空间插值(图 6), 从图中可以看出, 汉江流域月降雨极值分布的空间差异性较大, 总体看来呈下游大、上游其次、中游最小的分布趋势。受极大值影响, 相同重现期下万源、武汉、嘉鱼站的月降雨极值预测结果远远大于其他站点。由于本次所采用的资料系列较短, 随着重现期的增大, 模型预测的不确定性逐渐增加。通过将 15 个站的计算结果取平均求得汉江流域 20 年、50 年、100 年和 200 年一遇的月降雨极值分别为 403.3 mm、462.4 mm、506.0 mm 和 548.8 mm。

4.4. 与 AM/PIII 模型的对比分析

为了分析 PDS/GP 模型的可靠性, 本文将模型估计结果与采用年最大月降雨量系列的 PIII 型分布(AM/PIII)的计算结果进行了对比(见图 7)。分析发现, 相同重现期下 AM/PIII 模型的计算结果普遍大于 PDS/GP 模型预测值。以武汉站为例, PDS/GP 模型预测的 20 年、50 年、100 年和 200 年一遇月降雨量分别为 550.3 mm、656.5 mm、740.7 mm 和 828.4 mm; AM/PIII 模型预测的 20 年、50 年、100 年和 200 年一遇月降雨量分别为 562.3 mm、671.3 mm、752.7 mm 和 833.5 mm, 比前者增大 12.0 mm、14.8 mm、12.0 mm 和 5.1 mm。但总体看来, 差别不是特别明显, 这一方面是因为研究流域内各站点的月降雨极值年际变化和空间差异较大, 另一方面则与本次研究所采用的资料序列过短有关。

5. 结论

(1) 汉江流域 15 个站点的 50 年逐月降雨量系列统计参数表明, 资料系列具有明显的尖峰厚尾特征, 选用合

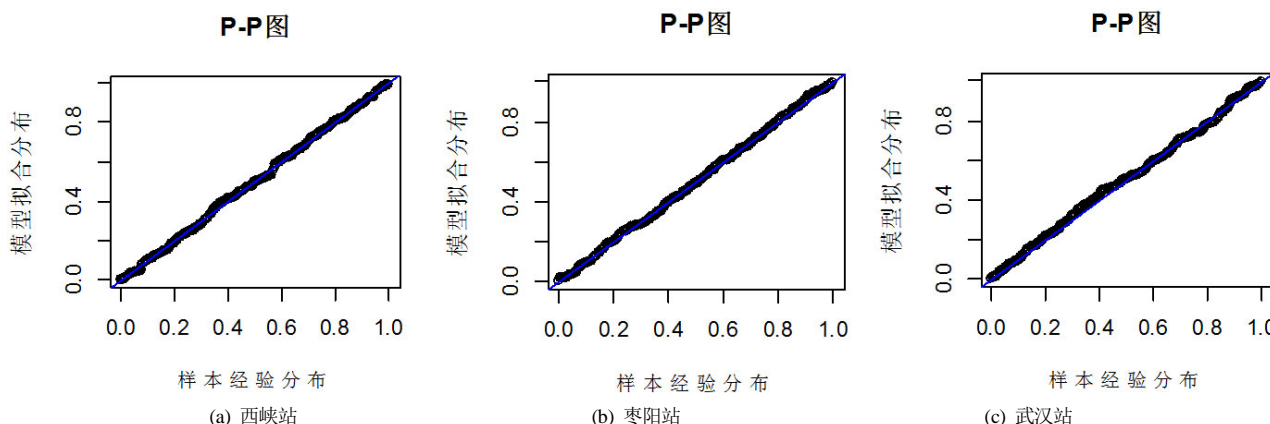


Figure 5. Change of modified scale and shape parameters with the threshold at representative stations

图 5. 汉江流域代表站模型参数随阈值变化情况图

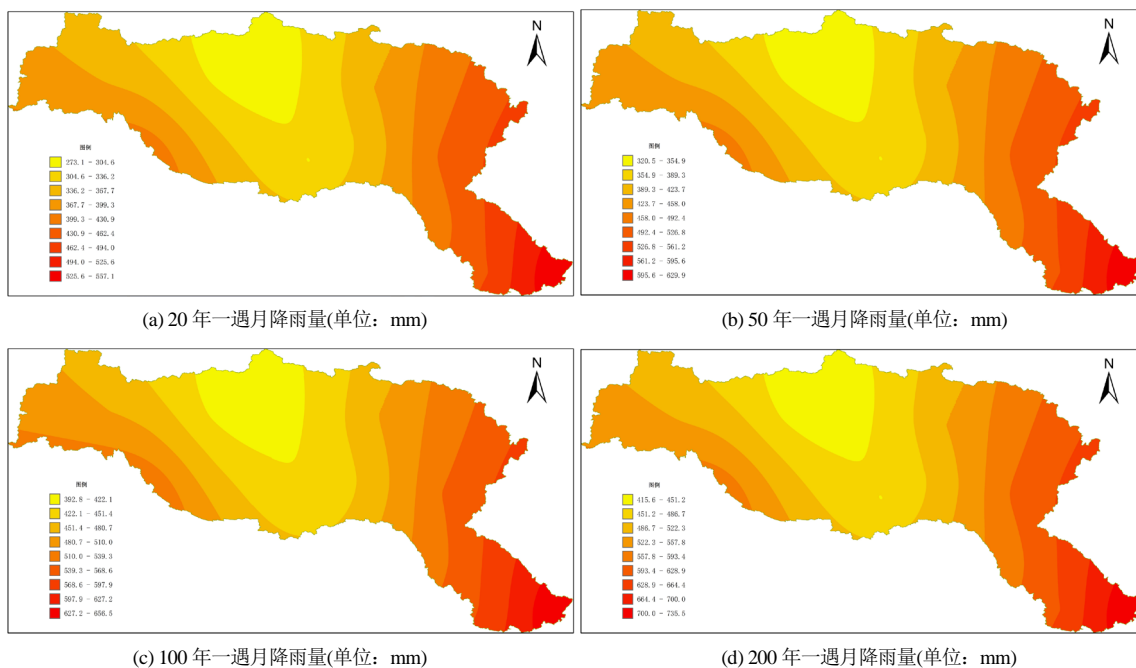


Figure 6. Spatial distribution of monthly rainfall extremes for different return periods in Han River basin
图 6. 汉江流域不同重现期月降雨量空间分布图

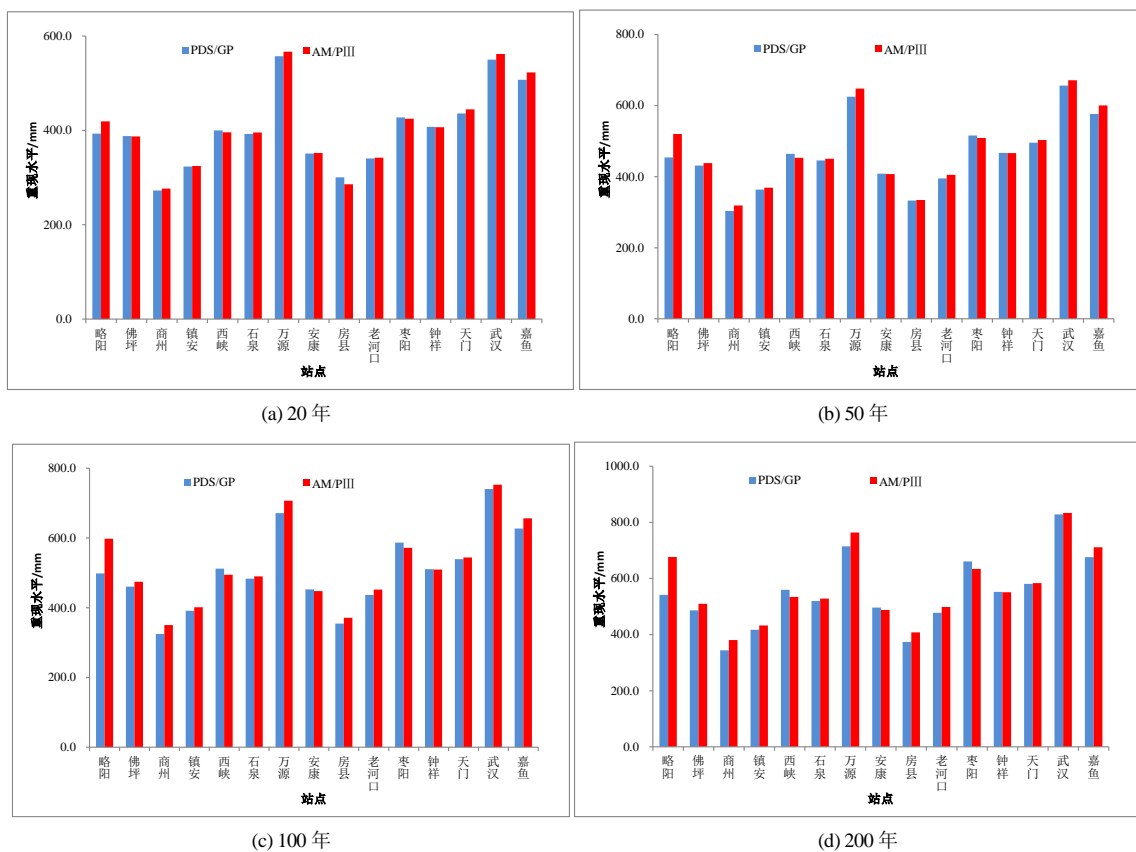


Figure 7. Results and comparison between PDS/GP model and AM/PIII model in Han River basin
图 7. 汉江流域各站 PDS/GP 模型和 AM/PIII 模型计算结果对比图

适的极值模型估计降雨极值重现水平优于传统的频率分析模型。

(2) PDS/GP 模型采用了超定量样本,较常规的年最大值系列样本数多,能更充分地利用数据中的极值信息。将 PDS/GP 模型计算的不同重现期月降雨量与年最大值系列拟合 PIII 型曲线所得的结果对比分析,前者的估计值受样本最大值影响相对较小,预测结果普遍小于后者。

(3) 要建立合理的 PDS/GP 模型,很重要的一步是阈值确定,本文将平均超出量函数和参数估计量变化两种阈值确定方法结合使用,大大缩小了可能合理阈值范围,能在一定程度上减少计算工作量;同时采用拟合残差法减小阈值选取的不确定性。

(4) 汉江流域 20 年、50 年、100 年和 200 年一遇的月降雨极值分别为 403.3 mm、462.4 mm、506.0 mm 和 548.8 mm。15 个站点中,除了略阳和房县,其它站点的月降雨量极大值 50 年一遇的 95% 置信区间都包含了 50 年实测资料系列的最大值,PDS/GP 模型拟合度较高。

参考文献 (References)

- [1] IPCC, 2012: Summary for policymakers. In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 1-19.
- [2] 张利平, 杜鸿, 夏军, 等. 气候变化下极端水文事件的研究进展[J]. 地理科学进展, 2011, 30(11): 1370-1379.
ZHANG Liping, DU Hong, XIA Jun, et al. Progress in the study of extreme hydrologic events under climate change. *Progress in Geography*, 2011, 30(11): 1370-1379. (in Chinese)
- [3] 史道济. 实用极值统计方法[M]. 天津: 天津科技出版社, 2006.
SHI Daoji. *Practical methods of extreme value statistics*. Tianjin: Tianjin Science and Technology Press, 2006. (in Chinese)
- [4] 翟盘茂, 任福民, 张强. 中国降水极值变化趋势检测[J]. 气象学报, 1999, 57(2): 208-216.
ZHAI Panmao, REN Fumin and ZHANG Qiang. Detection of trends in China's precipitation extremes. *Acta Meteorologica Sinica*, 1999, 57(2): 208-216. (in Chinese)
- [5] 黄琰, 封国林, 董文杰. 近 50 年中国气温、降水极值分区的时空变化特征[J]. 气象学报, 2011, 69(1): 125-136.
HUANG Yan, FENG Guolin and DONG Wenjie. Temporal changes in the patterns of extreme air temperature and precipitation in the various regions of China I recent 50 years. *Acta Meteorologica Sinica*, 2011, 69(1): 125-136. (in Chinese)
- [6] 姜彤, 苏布达, Marco Gemmer. 长江流域降水极值的变化趋势[J]. 水科学进展, 2008, 19(5): 650-655.
JINAG Tong, SU Buda and MARCO Gemmer. Trends in precipitation extremes over the Yangtze River basin. *Advances in Water Science*, 2008, 19(5): 650-655. (in Chinese)
- [7] 王俊, 郭生练. 南水北调中线工程水源区汉江水文水资源分析关键技术研究与应用[M]. 北京: 中国水利水电出版社, 2010.
WANG Jun, GUO Shenglian. *Research and application of key technique for the Hydrological Features and Water Resources of Han River in the Middle Route of South-to-North Water Transfer Project*. Beijing: China Water Power Press, 2010. (in Chinese)
- [8] 郭生练. 设计洪水研究进展与评价[M]. 北京: 中国水利水电出版社, 2005.
GUO Shenglian. *Advance and assessment of design flood methods*. Beijing: China Water Power Press, 2005. (in Chinese)
- [9] 王善序. 洪水超定量系列频率分析[J]. 人民长江, 1999, 30(8): 24-26.
WANG Shanxu. Flood frequency analysis based on super-quantitative series. *Yangtze River*, 1999, 30(8): 24-26. (in Chinese)
- [10] 戴昌军, 梁忠民, 栾承梅, 等. 洪水频率分析中 PDS 模型研究进展[J]. 水科学进展, 2006, 17(1): 136-140.
DAI Changjun, LIANG Zhongmin, LUAN Chengmei, et al. Advance in flood frequency analysis for partial duration series. *Advances in Water Science*, 2006, 17(1): 136-140. (in Chinese)
- [11] WILLEMS, P. Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. *Journal of Hydrology*, 2000, 233(1-4): 189-205.
- [12] NTEGEKA, V., WILLEMS, P. Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. *Water Resources Research*, 2008, 44(7): W07402.
- [13] NYEKO-OGIRAMOI, P., WILLEMS, P. and NGIRANE-KATASHAYA, G. Trend and variability in observed hydrometeorological extremes in the Lake Victoria basin. *Journal of Hydrology*, 2013, 489: 56-73.
- [14] VAN MONTFORT, M. A. J., WITTER, J. V. Testing exponentiality against generalised Pareto distribution. *Journal of Hydrology*, 1985, 78(3-4): 305-315.
- [15] DUPUIS, D. J. Exceedances over high thresholds: A guide to threshold selection. *Extremes*, 1999, 1(3): 251-261.

- [16] BEHRENS, C. N., LOPES, H. F. and GAMERMAN, D. Bayesian analysis of extreme events with threshold estimation. *Statistical Modelling*, 2004, 4(3): 227-244.
- [17] SCARROTT, C., MACDONALD, A. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT—Statistical Journal*, 2012, 10(1): 33-60.