

洞庭湖区极端降水统计分布模型分析与特征重现

周念清¹, 李章平¹, 沈新平², 刘晓群²

(1. 同济大学 水利工程系, 上海 200092; 2. 湖南省洞庭湖水利工程管理局, 长沙 410007)

摘 要: 以洞庭湖区汉寿、临湘、湘阴 3 个城市 1990 年至 2012 年逐日降水资料为基础, 选取 3 参数广义极值 (GEV) 分布和 2 参数耿贝尔 (Gumbel) 分布两种统计分布, 采用极大似然法估计模型参数, 阐述了两统计分布对极端日降水拟合过程, 使用 χ^2 拟合优度检验法和均方根误差 (RMSE) 评判拟合效果, 确定洞庭湖区极端日降水统计分布模型, 并分析极端日降水重现特征。研究结果表明: 洞庭湖最大日降水统计分布为 GEV 分布, 洞庭湖区最大日降水重现水平差异很明显, 为当地排水工程设计提供了科学依据。

关键词: 极端日降水; 统计分布; 重现特征; 洞庭湖

中图分类号: P426.6

文献标识码: A

文章编号: 1672-643X(2014)04-0011-05

Analysis of statistical distribution model of extreme precipitation and characteristics return in Dongting lake

ZHOU Nianqing¹, LI Zhangping¹, SHEN Xinping², LIU Xiaoqun²

(1. Department of Hydraulic Engineering, Tongji University, Shanghai 200092, China

2. Dongting Lake Water Resources Administration Bureau of Hunan Province, Changsha 410007, China)

Abstract: Based on the daily precipitation data of Hanshou, Linxiang and Xiangyin in Dongting Lake from 1990 to 2012, the paper choose generalized extreme value distribution with three parameters and Gumbel distribution with two parameters in detail and took the maximum likelihood method to estimate model parameters, and elaborated the process of extreme daily precipitation of two statistics distribution, and used χ^2 chi-square goodness-of-fit test and RMSE to evaluate imitative effect and decide statistical distribution model for extreme precipitation in Dongting Lake, and analyzed the return characteristics of extreme daily rainfall. The results show that the statistical distribution model of the maximum daily precipitation for Dongting Lake is GEV distribution. The difference of return level about maximum daily rainfall in Dongting Lake is obvious. The result can provide scientific basis for the design of local drainage engineering.

Key words: extreme daily precipitation; statistical distribution; return characteristics; Dongting Lake

随着全球气候变暖, 我国降水量有逐渐增加的趋势^[1], 有些地区在降水总量增加的同时, 极端降水量也显著增加^[2], 极端降水事件对人类生活和社会经济有着重要的影响^[3]。在许多工程设计中, 往往要考虑极端气候事件的重现期, 如在排水工程设计中一般以最大日降水重现期作为设计标准, 以满足排水要求。受实际水文气象资料长度限制, 用现有资料经验分布推算重现期比较困难。目前国内外多采用物理因子放大法和统计分布拟合法来推算重现期, 因物理因子放大法只能推算较长重现期, 统计

分布拟合法应用相对比较广泛。由于地形地貌复杂, 经纬度位置不同, 区域间水文气象要素存在差异, 极端降水的概率分布也不尽相同。苏布达等对长江流域降水进行统计分析, 得到概率分布为韦克比 (Wakeby) 分布^[4], 广东不存在普遍适用的极端降水概率分布模式^[5], 印度西北部极端降水服从广义极值 (GEV) 分布^[6]。部分学者直接采用某种统计分布模型推算极端降水重现期, 如荣艳淑等^[7]直接用 GEV 分布估算淮海流域最大日降水重现期和重现水平, 韩军彩等^[8]直接用耿贝尔 (Gumbel) 分布计

收稿日期: 2014-02-28; 修回日期: 2014-04-06

基金项目: 国家自然科学基金项目 (41272249); 高等学校博士学科点专项科研基金 (20110072110020)

作者简介: 周念清 (1964-), 男, 湖南石门人, 博士, 教授, 博士生导师, 主要研究方向为水文学及水资源、地下水数值模拟与水文环境评价。

算重现水平,其合理性有待研究。应该对具体区域极端降水进行统计分布拟合,确定最优拟合分布模型后,才能用上述分布模型计算重现期及重现水平。

洞庭湖是我国第二大淡水湖,有关洞庭湖的研究成果较多,但降水方面的研究多集中在洞庭湖流域降水时空分布特征^[9]和持续性暴雨形成机制^[10],而针对洞庭湖区极端降水统计分布及重现特征方面的研究尚不多见。GEV 分布和 Gumbel 分布被广泛应用于水文气象极端事件研究^[11]。本文选取 3 参数 GEV 分布和 2 参数 Gumbel 分布拟合洞庭湖汉寿、临湘、湘阴 3 个城市最大日降水量,确定洞庭湖区最大日降水概率分布模型,为分析洞庭湖降水统计特征提出了一种概率方法,计算重现期,并分析重现特征,为当地排水工程设计提供参考依据。

1 资料来源及研究方法

1.1 资料来源

选取分布于洞庭湖区的汉寿、临湘、湘阴 3 个城市的日降水量资料作为研究对象,资料来源于湖南省防汛系统,时间为 1990-01-01–2012-12-31。地理位置如图 1 所示。

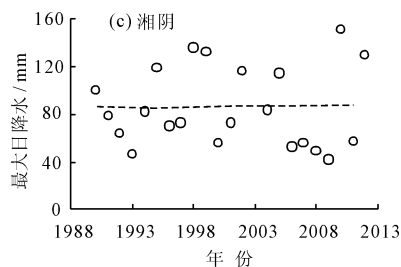
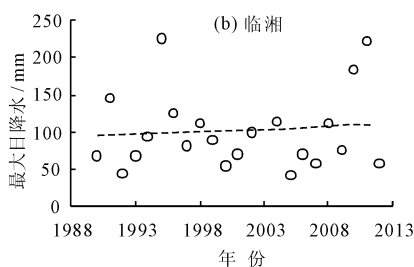
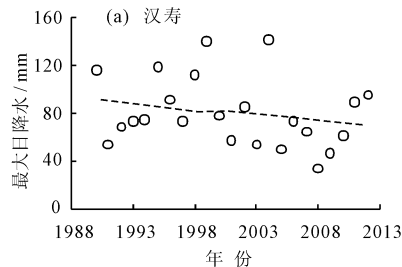


图 2 各城市最大日降水序列

1.2 统计分布模型

1.2.1 三参数 GEV 分布 用 $X_1, X_2, \dots, X_n$ 表示日降水量,根据传统极值的定义,极大值 $M_n = \max(X_1, X_2, \dots, X_n)$ 。当 $n = 365$ 时,极值 M_n 为年最大日降水,采用 GEV 分布来进行拟合。

GEV 分布密度函数表达式为:

$$f(x) = \frac{1}{\beta} (1 + Z)^{-1 - \frac{1}{k}} e^{-\frac{1}{k} (1 + Z)^{-1/k}} \quad (1)$$

$$Z = k (X - \zeta) / \beta, 1 + Z > 0 \quad (2)$$

式中: ζ 为位置参数; β 为尺度参数; k 为形状参数。概率函数表达式为:

$$F(x) = e^{-\frac{1}{k} (1 + Z)^{-1/k}} \quad (3)$$

反函数为:

$$X = F^{-1}(P) = \zeta + \frac{\beta}{k} \{ [-\ln(P)]^{-k} - 1 \} \quad (4)$$

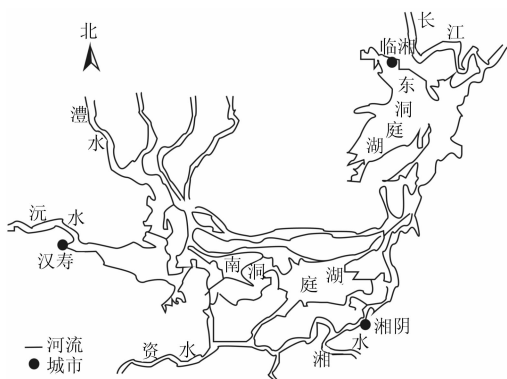


图 1 洞庭湖区地理位置图

不同研究者对极端降水的定义各不相同。百分位法和传统极值法是定义极端降水最为普遍的两种方法^[12],本文采用传统极值法将年最大日降水定义为极端降水,图 2 给出了以上 3 个城市最大日降水序列。从图中可看出:洞庭湖东北部(临湘)最大日降水量最大,平均值达 102 mm,西部(汉寿)和南部(湘阴)最大日降水量较小,平均值约 80 mm;临湘和湘阴 2 个城市最大日降水量在近 30 a 有增加的趋势,汉寿最大日降水量表现出减少的趋势,但增加和减少的趋势均不显著。

式中: $P = F(x)$ 。 $k < 0$ 为极值 II 型,即 Frechet 分布; $k > 0$ 为极值 III 型,即 Weibull 分布。

1.2.2 二参数 Gumbel 分布 GEV 模型中 $k = 0$ 为极值 I 型,即 Gumbel 分布,此时分布函数中含 2 个参数。概率密度函数表达式为:

$$f(x) = \frac{1}{\beta} e^{(-e^{-Z} - Z)} \quad (5)$$

$$Z = (X - \zeta) / \beta \quad (6)$$

概率函数表达式为:

$$F(x) = e^{-e^{-Z}} \quad (7)$$

1.3 分布模型拟合优度评判

本文分别采用 χ^2 拟合优度检验法和均方根误差(RMSE) 对各统计分布模型进行拟合优度检验。

(1) χ^2 拟合优度检验法。

检验假设 H^0 : 样本总体服从理论分布。把实轴划

分成 m 个互不相交的区间 $D_1, D_2, \dots, D_m$, 容量为 n 的样本 $(X_1, X_2, \dots, X_n)$, 落在区间 $D_j (j = 1, 2, m)$ 的实际值和理论值分别为 N_j 和 np_j , 计算检验统计量:

$$\chi^2 = \sum_{j=1}^k ((N_j - np_j)^2 / np_j) \tag{8}$$

取显著性水平 $\alpha = 0.05$, 如果 $\chi^2 < \chi^2_{1-\alpha}$, 则接受 H^0 ;
(2) 均方根误差 (RMSE)

$$RMSE = \sqrt{\sum_{i=1}^n (P_i - P_{ai})^2 / n} \tag{9}$$

式中: P_i 为理论累积概率; P_{ai} 为实测累积概率; n 为样本容量。 $RMSE$ 小于 0.1 表示拟合效果较好, 小于 0.05 表示拟合效果非常好。

2 极端降水统计分布模型分析

2.1 统计分布参数估计

参数估计的常用方法有极大似然 (ML)、矩法 (MOM)、概率权值矩 (PWM)、线性矩 (LM)、最小二乘法 (LS) 以及间隔最大积 (MPS) 等估计方法。极大似然参数估计法适用性较强, 参数估计精度较高^[13]。本文选取极大似然参数估计法估计统计模型参数, 结果见表 1 所示。由表 1 可知, 汉寿最大日降水 GEV 分布形状参数 k 为负值, 此表明为极值 II 型分布, 即 Frechet 分布。临湘和湘阴最大日降水 GEV 分布形状参数 k 均为正值, 此表明为极值 III 型分布, 即 Weibull 分布。

表 1 两种统计分布参数估计结果

分布函数	城市名	形状参数 k	尺度参数 β	位置参数 ζ
GEV	汉寿	-0.0195	22.9390	67.4580
	临湘	0.2520	30.0670	75.6080
	湘阴	0.0990	24.2380	69.3910
Gumbel	汉寿	0.0000	30.6070	95.5010
	临湘	0.0000	60.1650	129.3890
	湘阴	0.0000	32.5200	102.6270

2.2 统计分布拟合

GEV 分布和 Gumbel 分布广泛应用于水文气象中, 最大日降水作为极端降水的定义与极值分布理论的定义完全一致, 因而用这两种统计分布来拟合最大日降水理论上是可行的, 首先通过拟合分布密度函数和分布函数直观分析洞庭湖区最大日降水统计分布模型。图 3 为最大日降水实测分布与 GEV 分布以及 Gumbel 分布拟合的密度函数。从图 3 可看出, 3 个城市 GEV 分布密度曲线都与实测分布密度直方图吻合良好, 而 Gumbel 分布密度曲线与实测分布密度直方图不相吻合。最大日降水实测分布与 GEV 分布以及 Gumbel 分布拟合的分布函数如图 4 所示。

由图 4 可知, GEV 分布曲线基本上与实测分布曲线重合, 拟合效果较好, 两分布曲线存在微小差距可能是抽样误差所致, Gumbel 分布拟合效果都很差。

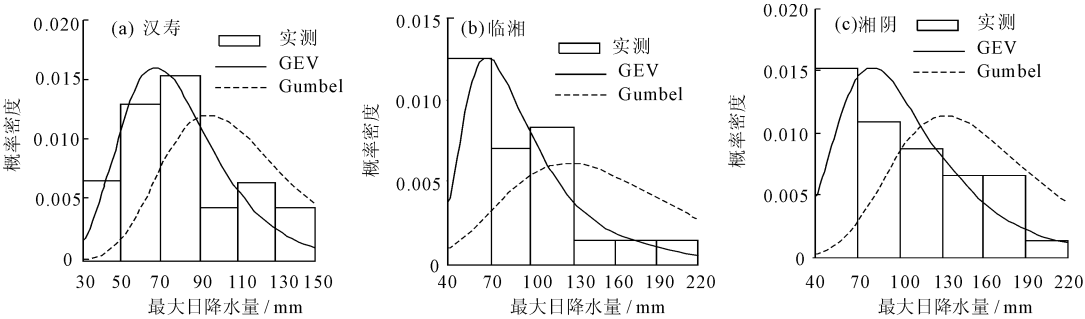


图 3 最大日降水统计分布密度函数拟合

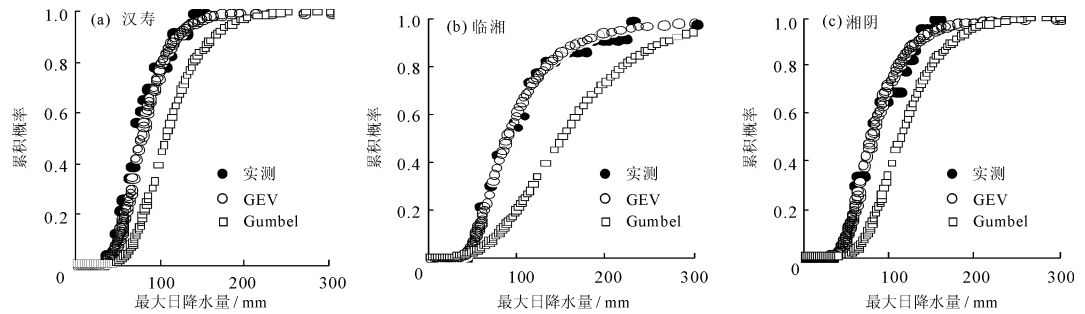


图 4 最大日降水统计分布函数拟合

2.3 拟合优度检验

为了客观评判两种统计分布模型的拟合效果,按照本文中给出的检验法和评判指标计算公式进行计算,得到相应结果(表2)。由表2可知, χ^2 拟合优度检验计算结果表明,洞庭湖区3个城市最大日降水完全服从3参数GEV分布,不服从2参数Gumbel分布。RMSE计算结果表明,GEV分布对汉寿和临湘最大日降水拟合效果非常好,对湘阴最大日降水拟合效果较好,Gumbel分布对洞庭湖区3个城市最大日降水拟合效果不理想。通过以上研究,结果表明:洞庭湖区最大日降水存在普遍统计分布模型,这个统计分布模型符合3参数GEV分布。确定统计分布模型后,可推算重现期并分析重现特征。

表2 拟合优度检验结果

分布函数	城市名	χ^2	$\chi^2_{1-\alpha}$	RMSE
GEV	汉寿	3.2480	5.9910	0.0350
	临湘	3.3310	5.9910	0.0310
	湘阴	3.5070	5.9910	0.0530
Gumbel	汉寿	42.4510	7.8150	0.2630
	临湘	47.2360	7.8150	0.3030
	湘阴	82.1550	7.8150	0.2720

3 极端降水特征重现

3.1 最大日降水重现期估算

重现期降水量是水利水电工程设计洪水推算中的一个重要参数,也是排水工程设计的设计标准。在排水工程设计中常选用重现期较短的最大日降水量作为设计标准,根据确定好的三参数GEV分布模型,本文以3、5、7、10、12、15、18和20 a 8个时间段为重现期尺度,计算各个城市最大日降水量,计算结果见表3。

表3 不同重现期最大日降水量

重现期	降水量(汉寿)	降水量(临湘)	降水量(湘阴)
3	88	106	92
5	101	130	109
7	110	147	119
10	118	167	130
12	122	177	136
15	127	190	144
18	131	201	149
20	134	209	153

从表3可知,在不同的重现期尺度下,最大日降水量的分布均类似,最大日降水量的极大值均集中

在洞庭湖东北部(临湘)。重现期与重现事件发生的概率是对应的,重现期越长,事件发生的概率越小,降水量越大。

3.2 最大日降水重现特征

洞庭湖区最大日降水的重现水平差异明显。因此,对于最大日降水的重现特征要给予足够的重视。最大日降水量与重现期的关系曲线如图5所示。

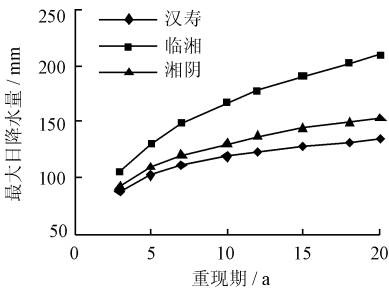


图5 最大日降水量与重现期的关系曲线

从图5可以看出,无论重现期为何值,3个城市对应最大日降水量从小到大为临湘、湘阴、汉寿,即洞庭湖区东北部(临湘)最大日降水量最大,南部(湘阴)次之,西部(汉寿)最小。当重现期增大时,对应最大日降水量增长趋势不同:湘阴和汉寿2个城市有缓慢增大趋势,临湘有快速增长趋势;在重现期增大过程中,当重现期达到10 a以后,对应的最大日降水量有较慢增长趋势。或者说,洞庭湖东北部当重现期增大到10 a过程中,重现期越大,对应的最大日降水量明显增加;而洞庭湖西部和南部当重现期超过10 a时,重现期越大,对应最大日降水量增加不明显。

4 结 语

(1)通过研究发现,洞庭湖区最大日降水存在普遍适用的3参数GEV分布模型,对西部(汉寿)和东北部(临湘)拟合效果非常好,对南部(湘阴)拟合效果较好。

(2)由于受实际资料限制,存在抽样误差,拟合分布模型理论值与实测值存在一定误差。如果增加实际监测资料长度,可以减小这种误差的产生。

(3)3参数GEV分布对洞庭湖区最大日降水拟合效果较好,而2参数Gumbel分布不能拟合。如果直接采用2参数Gumbel分布估算重现期是不合理、不科学的,需针对具体地区先进行拟合,再根据相应分布估算重现期。

(4)洞庭湖区最大日降水重现水平差异很明显,东北部(临湘)降水偏强,南部(湘阴)和西部(汉

寿)降水偏弱。在重现期增大过程中,存在一个 10 a 的临界值,前后对应最大日降水量增加速率不同:临界点前最大日降水量增加明显,临界点后最大日降水量增加不明显。

参考文献:

- [1] 闵 岫,钱永甫. 我国近 40 年各类降水事件的变化趋势[J]. 中山大学学报(自然科学版),2008,47(3):105 - 111.
- [2] Meehl G A, Karl T, Easterling D R, et al. An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections [J]. Bulletin of the American Meteorological Society, 2000, 81(3): 413 - 416.
- [3] 黄 琰,董文杰,支 蓉,等. 强降水持续过程对上海市内交通经济损失评估模型初探[J]. 物理学报,2011,60(4):803 - 812.
- [4] 苏布达,姜 彤,董文杰. 长江流域极端强降水分布特征的统计拟合[J]. 气象科学,2008,28(6):625 - 629.
- [5] 陈子燊,路剑飞,刘曾美. 广东极端降水的三参数概率分布模式对比[J]. 中山大学学报(自然科学版),2012,51(1):102 - 106.
- [6] Iqbal M J, Ali M. A probabilistic approach for estimating return period of extreme annual rainfall in different cities of

Punjab [J]. Arabian Journal of Geosciences, 2013, 6(7): 2599 - 2606.

- [7] 荣艳淑,王 文,王 鹏,等. 淮河流域极端降水特征及不同重现期降水量估计[J]. 河海大学学报(自然科学版),2012,40(1):1 - 8.
- [8] 韩军彩,王传辉,陈 静,等. 石家庄市暴雨变化特征及降水极值重现期研究[J]. 干旱区研究,2013,30(5):796 - 801.
- [9] 李景刚,黄诗峰,李纪人,等. 1960 - 2008 年间洞庭湖流域降水变化时空特征分析[J]. 中国水利水电科学研究院学报,2011,8(4):275 - 280.
- [10] 张小玲,陶诗言,张顺利,等. 1996 年 7 月洞庭湖流域持续性暴雨过程分析[J]. 应用气象学报,2004,15(1):21 - 31.
- [11] Hennessy K J, Gregory J M, Mitchell J F B. Change in daily precipitation under enhanced greenhouse conditions [J]. Climate Dynamics, 1997, 13(9):667 - 680.
- [12] Fowler H J, Kilsby C G. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000 [J]. International Journal of Climatology, 2003, 23(11):1313 - 1334.
- [13] 卢安平,赵 林,郭增伟,等. 基于 Monte Carlo 法的极值分布类型及其参数估计方法比较[J]. 哈尔滨工业大学学报,2013,45(2):88 - 95.

(上接第 10 页)

- [5] 严登华,王 浩,王 芳,等. 我国生态需水研究体系及关键研究命题初探[J]. 水利学报,2007,38(3):267 - 273.
- [6] 杨爱民,唐克旺,王 浩,等. 生态用水的基本理论与计算方法[J]. 水利学报,2004,35(12):39 - 45.
- [7] 张 丽,李丽娟,梁丽乔,等. 流域生态需水的理论及计算研究进展[J]. 农业工程学报,2008,24(7):307 - 312.
- [8] 徐志侠,陈敏建,董增川. 河流生态需水计算方法评述[J]. 河海大学学报(自然科学版),2004,32(1):5 - 9.
- [9] 陈敏建. 水循环生态效应与区域生态需水类型[J]. 水利学报,2007,38(3):282 - 288.
- [10] 陈敏建,王 浩. 中国分区域生态需水研究[J]. 中国水利,2007(9):31 - 37.

- [11] 苏 飞,陈敏建,董增川,等. 辽河河道最小生态流量研究[J]. 河海大学学报(自然科学版),2006,34(2):136 - 139.
- [12] 王高旭,陈敏建,丰华丽,等. 黄河中下游河道生态需水研究[J]. 中山大学学报(自然科学版),2009,48(5):125 - 130.
- [13] 陈敏建,丰华丽,王立群,等. 适宜生态流量计算方法研究[J]. 水科学进展,2007,18(5):745 - 750.
- [14] 张长春,邵景力,李慈君,等. 地下水位生态环境效应及生态环境指标[J]. 水文地质工程地质,2003,30(3):6 - 11.