

# 广东极端降水的三参数概率分布模式对比\*

陈子燊<sup>1</sup>, 路剑飞<sup>1</sup>, 刘曾美<sup>2</sup>

(1. 中山大学水资源与环境系, 广东 广州 510275;  
2. 华南理工大学水利水电工程系, 广东 广州 510640)

**摘要:** 采用包括广义极值分布、威布尔分布、广义逻辑斯特分布和皮尔逊3型分布在内的4种三参数概率分布模式推算广东省85个气象站历年日最大降水不同重现水平的分位数。主要结果如下: 1) 参数估计结果表明不存在普遍适用的极端降水概率分布模式; 2) 广东极端降水以珠江三角洲和沿海较为突出。其中, 清远、汕尾、阳江为广东三大极端降水中心区; 3) 采用4种概率分布模式推算的极端降水重现水平差别很大。

**关键词:** 三参数概率分布模式; 极端降水; 拟合优度检验; 重现水平; 广东

**中图分类号:** TV12212 **文献标志码:** A **文章编号:** 0529-6579(2012)01-0102-05

## Comparative Analysis on Three-parameter Probability Distribution Models for Extreme Precipitation in Guangdong Province

CHEN Zishen<sup>1</sup>, LU Jianfei<sup>1</sup>, LIU Zengmei<sup>2</sup>

(1. Department of Water Resource and Environment, Sun Yat-sen University, Guangzhou 510275, China;  
2. Department of Water Conservancy and Hydropower Engineering, South China University of Technology, Guangzhou 510640, China)

**Abstract:** Four probability distribution models, generalized extreme value distribution, Weibull distribution, log-logistic distribution and Pearson type-III distribution, were used to calculate the different return levels of maximum daily precipitation for 85 stations in Guangdong province. The main results are as follows: 1) The results of parameter estimation indicate that there is no universally applicable extreme precipitation probability distribution model; 2) The extreme precipitation occurs more frequently in the Pearl River Delta and the coastal areas of Guangdong among which Qingyuan, Shanwei and Yangjiang are three extreme precipitation centers; 3) The return levels of extreme precipitation differ greatly between the calculations of four probability distribution models.

**Key words:** three-parameter probability distribution models; extreme precipitation; goodness-of-fit test; return levels; Guangdong

全球气候变化异常, 极端干旱和强降水引发的旱涝灾害等水文气象事件的频率和强度趋于增大。有关分析指出, 尽管广东数十年来年降水总量没有明显变化, 但近些年雨日减少、雨强加大, 极端降水频发, 强降水过程导致流域洪水和城市内涝问题加重。为此, 应用能更精确地描述极端事件的概率

分布模型推断稀遇概率重现水平, 已成为应对气候变化与防灾减灾研究的重要研究内容, 并可为工程设施风险管理与防灾预警方案设计提供重要的科学依据。

通常三参数概率分布模式比二参数分布模式能更好地拟合有限样本的极端水文气象数据, 从而得

\* 收稿日期: 2011-07-05

基金项目: 广东水利创新资助项目(2011370004209292)

作者简介: 陈子燊(1952年生), 男, 教授, 博士生导师; E-mail: eesczs@mail.sysu.edu.cn

以广泛应用。其中，皮尔逊 3 型分布 (P-III) 是列入我国工程规范与应用最为普遍的极端水文气象要素频率分析的线型<sup>[1-3]</sup>。但不同区域实际水文气象条件差异大，采用单一的分布函数描述极端水文气象事件显然并不适宜。为此，广义极值分布 (generalized extreme-value distribution, GEV)、威布尔分布和对数逻辑斯特分布 (log-logistic distribution, LLD) 等三参数概率分布模式也被大量应用于风速、海浪、洪水、降水等极端水文气象事件研究<sup>[4-15]</sup>。

本文拟分别应用广义极值分布、威布尔分布、对数逻辑斯特分布和皮尔逊 3 型分布 4 种概率分布模式推算广东日最大降水的重现水平，并通过多个拟合优度检验方法对 4 个概率分布统计推断结果加以对比，进而分析广东日降水极值的空间分布特征。

## 1 概率分布模式

### 1.1 广义极值分布

根据极值分布理论，可证明当极值的渐近分布存在且为非退化时可以将 3 种类型的经典极值分布发展为一种统一的具有三参数的极值分布函数——广义极值分布 (简记为 GEV)<sup>[16]</sup>。作为经典极值渐近分布的广义形式，其避免了单独采用某一分布的不足。

设  $X_1, \dots, X_m$  是服从 GEV 分布的独立随机变量，广义极值分布的超值累积分布函数  $F_X$  为：

$$F_X(x) = \begin{cases} \exp\left\{-\left[1 - \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{1/\xi}\right\} & \xi \neq 0 \\ \exp\left[-\exp\left(\frac{x-\mu}{\sigma}\right)\right] & \xi = 0 \end{cases}$$

当  $\xi \rightarrow 0$  为极值 I 型，即 Gumbel 分布； $\xi < 0$  为极值 II 型，即 Fréchet 分布； $\xi > 0$ ，为极值 III 型，即 Weibull 分布。广义极值分布的反函数  $x_F$  为：

$$x_F = \begin{cases} \mu + \frac{\sigma}{\xi}\{1 - [-\ln(F)]^\xi\} & \xi \neq 0 \\ \mu - \sigma \ln[-\ln(F)] & \xi = 0 \end{cases}$$

式中， $\xi$ 、 $\beta$ 、 $\mu$  分别为形状参数、尺度参数和位置参数，下同。

### 1.2 威布尔分布

威布尔分布早期被应用于质量管理上的元件和材料寿命可靠性研究，是可靠性分析和寿命检验的理论基础。近年来被广泛应用于风速和波高的极值推算<sup>[9-12]</sup>。

随机变量  $X$  的威布尔不及累积分布函数为：

$$F_X(x) = P(X < x) = \begin{cases} 1 - \exp\left[-\left(\frac{x-\mu}{\beta}\right)^\xi\right] & x \geq \mu \\ 0, & x < \mu \end{cases}$$

当  $\xi \neq 0$  时，反函数  $x_F$  为： $x_F = \mu + \beta[-\ln(1-F)]^{1/\xi}$ 。

### 1.3 对数逻辑斯特分布

国外较多利用三参数的对数逻辑斯特分布函数拟合洪水频率分布，并对此函数性质做了较为深入的分析<sup>[13-15]</sup>。随机变量  $X$  的对数逻辑斯特不及累积分布函数为：

$$F_X(x) = P(X < x) = \left[\frac{x-\mu}{\beta}\right]^\xi / \left\{1 + \left[\frac{x-\mu}{\beta}\right]^\xi\right\}$$

$$x > \mu; \xi \geq 1, \beta > 0$$

反函数  $x_F$  为： $x_F = \mu + \beta[F/(1-F)]^{1/\xi}$

### 1.4 P-III 型分布

P-III 型累积分布函数为：

$$F_X(x) = P(X < x) = \frac{\beta^\xi}{\Gamma(\xi)} \int_{-\infty}^x (x-a_0)^{\xi-1} \exp[-\beta(x-a_0)] dx; (\xi, \beta > 0, x > a_0)$$

式中， $a_0$  为 P-III 型分布的位置参数。

上述概率分布函数的参数估计采用 Greenwood 等 (1979)<sup>[17]</sup> 定义的概率权重矩和 Hosking (1986)<sup>[18]</sup> 的线性矩计算方法。研究表明，由概率权重矩 (PWM) 或其线性组合 (LM) 估计的分布参数具有良好的无偏性和较高的精度<sup>[19]</sup>。

## 2 拟合优度检验

分别采用均方根误差 (RMSE)：RMSE =

$$\sqrt{\sum_{i=1}^n ((\hat{x}_i - x_i)/x_i)^2/n}; \text{经验频率和理论频率拟}$$

合误差平方和 (Q)： $Q = \sum_{i=1}^n (P_i - P_{ei})^2$ ；概率点据

相关系数 (PPCC)： $PPCC = \sum_{i=1}^n (x_i - x_m)(\hat{x}_i - \hat{x}_m) / \left[ \sum_{i=1}^n (x_i - x_m)^2 \sum_{i=1}^n (\hat{x}_i - \hat{x}_m)^2 \right]^{1/2}$  对各概率分布

模式做拟合优度检验。式中， $P_{ei}$  为经验累积频率； $P_i$  为理论累积频率； $n$  为样本容量。在样本容量相同的情况下，均方根误差和  $Q$  值越小，相关系数 PPCC 越大，样本与理论线形拟合越好。

## 3 广东降水极值分位数及其空间分布

广东年均雨量大，但 80% 降雨高度集中在汛期以洪涝肆虐的形式出现，一直成为困扰经济社会

发展的三大自然灾害之一。利用广东省 85 个国家气象基准站或基本站 1951 - 2009 年的逐日降水观测值采用年最大值抽样方法得到长度 43 ~ 59 a 不等的各测站日最大降水量样本。

四种概率分布的参数估计结果存在以下特点: 绝大部分测站 (共 62 个站) 日最大降水广义极值分布的形态参数  $\xi < 0$ , 此表明日最大降水序列以极值 II 型分布为主, 即 Frechet 分布。极值 II 型分布属于正偏右长尾, 表明出现极值分位数的可能性大于极值 I 型和极值 III 型。P-III 型分布参数估计结果有部分测站 (8 个站) 偏态系数  $C_s$  值小于 2

倍的变差系数  $C_v$  值, 由此可推得分布函数的位置参数为负值, 从而有违实际物理含义, 此外另有 7 个站偏态系数  $C_s$  值大于 2。此说明这部分测站日极端降水不服从 P-III 型分布。使用对数逻辑斯特分布也有 27 个测站的位置参数为负值, 同样表明不符合实际物理含义。威布尔分布的参数估计仅东莞测站拟合的位置参数为负值。以上说明, 为了能恰当地描述不同区域不同极端水文气象事件长期分布特征, 确实需要采用多种概率分布模式加以对比选择。部分测站 4 种概率分布参数和拟合优度指标见表 1。

表 1 广东部分测站日最大降水四种概率分布参数和拟合优度指标  
Table 1 Parameters of four probability distribution and indicators of goodness of fit for maximum daily precipitation of some selective stations in Guangdong

分布函数	站名(样本数)	形态参数	尺度参数	位置参数	RMSE	Q 统计值	PPCC
GEV	韶关(59)	-0.072	24.510	82.194	3.49	0.02	0.997
	梅县(57)	-0.138	23.284	77.105	5.74	0.04	0.992
	高要(56)	-0.106	27.411	91.573	5.37	0.04	0.992
	清远(51)	-0.326	38.091	121.094	31.75	0.02	0.955
	广州(59)	-0.155	32.803	98.961	7.52	0.03	0.992
	汕头(59)	0.028	44.865	121.667	4.17	0.02	0.998
	深圳(58)	0.017	52.869	147.111	6.94	0.04	0.995
	汕尾(58)	-0.042	57.627	150.683	17.32	0.05	0.978
	湛江(59)	0.086	50.686	128.178	8.33	0.05	0.991
	阳江(56)	-0.028	80.242	167.276	15.55	0.03	0.989
WBL	韶关(59)	1.432	52.344	50.670	4.27	0.03	0.993
	梅县(57)	1.265	47.613	49.971	6.73	0.04	0.985
	高要(56)	1.320	56.246	58.780	4.74	0.05	0.994
	清远(51)	0.893	69.626	87.377	33.79	0.07	0.935
	广州(59)	1.218	65.950	61.977	7.52	0.05	0.991
	汕头(59)	1.731	103.107	54.458	4.06	0.01	0.998
	深圳(58)	1.683	119.700	69.881	6.38	0.04	0.996
	汕尾(58)	1.508	125.096	73.576	18.87	0.05	0.972
	湛江(59)	2.042	128.040	40.019	9.19	0.04	0.989
	阳江(56)	1.512	172.227	60.537	12.69	0.03	0.993
LLD	韶关(59)	4.584	78.037	13.718	6.37	0.03	0.986
	梅县(57)	3.842	64.509	21.893	6.66	0.04	0.985
	高要(56)	4.178	81.111	21.281	11.31	0.05	0.964
	清远(51)	2.571	78.905	58.371	15.00	0.02	0.986
	广州(59)	3.677	87.600	24.536	13.75	0.03	0.969
	汕头(59)	6.341	185.896	-47.378	10.67	0.03	0.984
	深圳(58)	6.081	211.521	-44.471	15.78	0.06	0.974
	汕尾(58)	4.987	195.096	-22.190	14.96	0.08	0.983
	湛江(59)	8.054	258.165	-111.376	6.91	0.06	0.994
	阳江(56)	5.187	279.084	-81.046	32.09	0.05	0.957
P - III	韶关(59)	2.400	0.045	45.456	3.84	0.02	0.995
	梅县(57)	1.653	0.036	48.394	6.29	0.04	0.987
	高要(56)	1.973	0.035	54.492	4.72	0.04	0.994
	清远(51)	0.700	0.010	91.264	33.45	0.18	0.935
	广州(59)	1.514	0.024	60.330	7.18	0.04	0.992
	汕头(59)	4.877	0.040	24.594	4.04	0.02	0.998
	深圳(58)	4.450	0.032	37.985	6.53	0.04	0.995
	汕尾(58)	2.909	0.022	55.223	18.17	0.05	0.974
	湛江(59)	8.306	0.049	-15.863	8.49	0.05	0.991
	阳江(56)	3.199	0.017	27.633	13.64	0.04	0.992

图 1 显示，广东极端降水总体分布以珠江三角洲和沿海高，粤北山区低为主要特征。其中，清远、汕尾、阳江为广东 3 大极端降水中心所在地。清远地区是广东北江流域锋面雨主要影响区域，海滨城市汕头、汕尾、阳江和湛江极端降水主要是影响广东沿海的台风暴雨造成。对 3 个极端降水中心极值分位数计算结果的拟合优度检验对比表明，采用的 4 种概率分布模式总体上都具有优良的统计特性，但以 RMSE、 $Q$  和 PPCC 检验值为择优标准衡量，汕尾日极端降水更符合广义极值分布；阳江日极端降水更符合威布尔分布；清远日极端降水更符

合对数逻辑斯特分布（见表 1 黑体字）。

拟合优度差异导致推算的极端降水重现水平值影响显著。三个极端降水中心按不同的概率分布模式推算的重现水平值差别很大，其 100 a、50 a 一遇的极值降水分位数见表 2 黑体字。以清远站百年一遇日降水量为例，按 LLD 推算为 530 mm，和 GEV 分布推算结果相近，而 P-III 分布为 477 mm，二者相差达 53 mm。阳江站和汕尾站存在类似问题。此进一步表明，不存在普遍适用的极端降水概率分布模式，应择优使用概率分布模式推算极端降水重现水平。

表 2 广东部分测站不同概率分布日最大降水的重现水平 mm

Table 2 Return level values of four probability distribution for maximum daily precipitation of some selective stations in Guangdong

站名	GEV			P-III			WBL			LLG		
	100a	50a	20a	100a	50a	20a	100a	50a	20a	100a	50a	20a
韶关	216	193	163	207	189	164	203	186	163	226	196	162
梅县	227	197	163	214	193	164	209	190	163	235	200	161
高要	254	224	187	242	219	188	238	217	188	265	227	185
清远	528	421	312	477	413	328	473	408	325	530	417	306
广州	319	275	223	299	268	225	293	264	224	330	277	220
汕头	315	287	250	310	284	249	304	281	249	336	296	248
深圳	381	347	300	373	342	300	367	339	300	406	357	299
汕尾	443	395	333	427	388	333	418	383	333	468	404	330
湛江	321	296	261	319	295	260	311	290	259	345	307	261
阳江	561	498	416	542	489	416	533	485	416	596	510	411

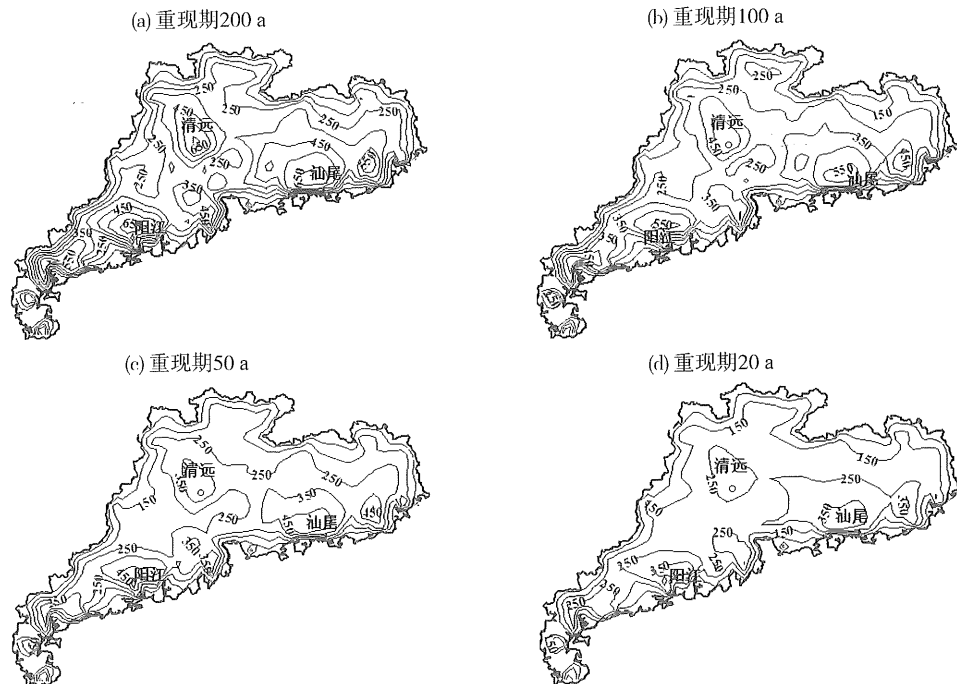


图 1 广东日最大降水 GEV 分布的重现水平（图中数字单位：mm）

Fig. 1 The return levels of GEV distribution for maximum daily precipitation of Guangdong

## 4 结 论

本文利用四种概率分布模式拟合广东日最大降水值, 获得以下主要结论:

1) 参数估计结果表明不存在普遍适用的降水极值分布函数, 经拟合优度检验择优的概率分布函数还应符合实际物理含义;

2) 广东极端降水以珠江三角洲和沿海较为突出。其中, 清远、汕尾、阳江为广东三大极端降水中心;

3) 以四种概率分布模式推算的重现水平值差别很大。按相对最优概率分布模式推算的清远、汕尾、阳江百年一遇日降水量分别为 530、443 和 533 mm, 50 a 一遇日降水量分别为 417、395 和 485 mm。

### 参考文献:

- [1] 中华人民共和国行业标准. JTJ213-98 海港水文规范[S]. 北京: 人民交通出版社, 1998.
- [2] 黄振平, 林小丽, 侯云青, 等. P-III 型分布参数的矩估计与设计洪水的计算频率[J]. 河海大学学报: 自然科学版, 2005, 33(1): 49-51.
- [3] 周芬, 郭生练, 肖义, 等. P-III 型分布参数估计方法的比较研究[J]. 水电能源科学, 2003, 21(3): 10-13.
- [4] 刘聪, 秦伟良, 江志红. 基于广义极值分布的设计基本风速及其置信限计算[J]. 东南大学学报: 自然科学版, 2006, 36(2): 331-334.
- [5] 陈元芳, 李兴凯, 陈民, 等. 可考虑历史洪水信息的广义极值分布线性矩法的研究[J]. 水文, 2008, 28(3): 8-13.
- [6] 陈子燊, 刘曾美, 路剑飞. 广义极值分布参数估计方法的对比分析[J]. 中山大学学报: 自然科学版, 2010, 49(6): 105-109.
- [7] 夏华永, 李树华. 广西沿海年极值波高分析[J]. 热带海洋学报, 2001, 20(2): 1-7.
- [8] 段忠东, 周道成. 极值概率分布参数估计方法的比较研究[J]. 哈尔滨工业大学学报, 2004, 36(12): 1605-1609.
- [9] 张秀芝. Weibull 分布参数估计方法及其应用[J]. 气象学报, 1996, 54(1): 108-116.
- [10] 邓建, 古德生, 李夕兵. 确定可靠性分析 Weibull 分布参数的概率加权矩法[J]. 计算力学学报, 2004, 21(5): 609-613.
- [11] NEELAMANI S, SALEM K AL, RAKHA K. Extreme waves for Kuwaiti territorial waters[J]. Ocean Engineering, 2007, 34(10): 1496-1504.
- [12] SEGURO J V, LAMBERT T W. Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2000, 1(85): 75-84.
- [13] HAKTANIR T, HANS B H. Evaluation of various distributions for flood frequency analysis[J]. Hydrological Sciences Journal, 1993, 38: 1: 15-32.
- [14] ROWINSKI P M, STRUPCZEWSKI W G, SINGH V P. A note on the applicability of log-Gumbel and log-logistic probability distributions in hydrological analyses: I. Known pdf[J]. Hydrological Scientific Journal, 2002, 47(1): 107-122.
- [15] FAHIM A, SMAIL M. Fitting the log-logistic distribution by generalized moments[J]. Journal of Hydrology, 2006, 328: 694-703.
- [16] COLES S. An introduction to statistical modeling of extreme values[M]. New York: Springer Verlag, 2001: 36-78.
- [17] GREENWORD J A, LANDWEHR J M, MATALAS N C, et al. Probability weighted moments: Definition and relation to parameters of distribution expressible in inverse form[J]. Water Resources Research, 1979, 15(5): 1049-1054.
- [18] HOSKING J R M. L-moments: Analysis and estimation of distributions using linear combinations of order statistics[J]. J R Stat Soc, Ser B, 1990, 52: 105-124.
- [19] 金光炎. 矩、概率权重矩与线性矩的关系分析[J]. 水文, 2005, 25(5): 1-6.