

不同概率分布函数降雨极值的适用性分析

张玉虎¹, 王琛茜^{1,2}, 刘凯利³, 陈秋华³

(1. 首都师范大学资源环境与旅游学院, 北京 100048; 2. 北京建筑大学测绘与城市空间
信息学院, 北京 100044; 3. 首都师范大学数学科学学院, 北京 100048)

摘要:极端降雨极值发生的重现期是流域与城市防洪设施规划设计标准需要参考的最重要参数之一。利用常用的5种水文统计学分布函数,选取中国十大流域内10个站点不同时段的最大降雨极值序列进行拟合,并检验筛选不同站点的适用性分布函数。结果表明:10个站点拟合优度检验拟合效果较好,曲线差异度较小的分布依次为广义极值分布、对数正态分布、皮尔逊III分布;不同站点适宜性曲线的差异程度不同。研究结果可为区域降雨极值序列的拟合提供参考,即不同的区域、不同的季节、不同时间的降雨极值序列都应寻找其较适宜的分布函数并采用多种检验方法来拟合,以降低不确定性。

关键词:极端降雨;概率分布;重现期;K-S检验;A-D检验

中图分类号:P426.615 **文献标识码:**A **文章编号:**1000-0690(2015)11-1460-08

IPCC(2012)第5次评估报告中阐述“在未来气候变暖背景下,极端暖事件将进一步增多,中纬度大部分陆地区域和湿润的热带区域降水强度、降水发生频率可能增多^[1]。中国地处中纬度地区,不同地区不同季节极端降水发生的概率及强度蕴含不确定性^[2]。极端强降水频率和强度的变化,直接导致洪涝灾害的发生^[3]。水文统计方法能够客观反映水文现象的不确定性^[4-6],降雨量重现期的准确计算是洪水风险评估及防洪设施工程设计中最重要的参数。

水文气象统计学常用分析方法有:广义极值分布(GEV, Generalized Extreme Value)^[7-9]、对数正态分布^[10]、韦伯分布(Weibull)^[11]、皮尔逊III(Pearson-III)分布^[12,13]、耿贝尔分布^[14,15]、指数分布(Exponential)、韦克比分布(Wakeby, 简称WAK)^[16,17]等。自1980年以来,中国在水利水电工程水文领域及洪水设计等方面通常采用皮尔逊III型分布函数进行降雨极值的拟合计算。计算过程中经验点数据与频率曲线拟合不好时,经过论证有时也采用其他线型^[18]。未来气候变化下,探讨有效、准确评估全国洪水风险及可能灾害,需要加强降雨量重现

期拟合的准确计算。同时降雨极值方法学探讨也有助于水电工程及防洪设施标准的制定。尽管全国尺度的洪水风险已经有少数学者开展了评估,诱发洪水发生的降雨极值通常都用皮尔逊III型拟合。为探讨不同区域降雨极值序列不同概率分布函数适宜性,特从全国十大流域选取10个站点(10个站点不能代表10个流域)进行拟合。

本文选取中国范围内10个站点的日降雨量数据,建立年最大降雨量序列(AMS, annual maxima series)、夏季最大降雨量序列(SMS, summer maxima series)和冬季最大降雨量序列(WMS, winter maxima series),尝试运用不同分布函数对最大1、2、3 d降雨量进行拟合,通过K-S(Kolmogorov-Smirnov)与A-D(Anderson-Darling)方法检验产生不同站点的最优拟合函数,对降水极值序列分布函数的重现期曲线做出比较,探索不同线型对站点的适用性。

1 数据与方法

研究采用的气象观测数据由中国气象局国家气象信息中心提供,考虑到数据的质量、数据的连续性与统计样本的最短年限,选取降雨统计数据

收稿日期:2014-09-17;修订日期:2014-11-20

基金项目:国家十二五科技支撑计划课题(No.2013BAC10B01、2012BAC19B03-05)资助。

作者简介:张玉虎(1975-),男,江苏省徐州市人,博士,讲师,主要从事水资源、水环境分析评价与灾害风险应对研究。E-mail: zhang_yuhu@163.com

的年限为 1960 年 1 月 1 日至 2010 年 12 月 31 日。在十大一级流域选取 10 个站点(表 1)。

从季节分布上划分,每个站点的日降雨量数据分为全年,夏季(6~8 月)以及冬季(12 月至次年 2 月);从降水时长上划分,从每个站点的逐日降水数据中统计每年(夏季、冬季)最大 1 d 降雨量、最大 2 d 降雨量以及最大 3 d 降雨量。对各个站点全年、夏季、冬季的最大降雨量序列进行 5 种概率分布拟合,包括广义极值分布(GEV, Generalized Extreme Value)、对数正态分布(LN, Lognormal)、韦伯分布(Weibull)、皮尔逊 III 分布(Pearson-III)、耿贝尔分布(Gumbell)。

K-S(Kolmogorov-Smirnov)检验^[19]是一种基于经验累积分布函数(ECDF)的算法,它的特性是具有分布无关性,所以适用于任何连续分布,很适合小样本。A-D(Anderson-Darling)检验^[20,21]是对 K-S 检验的一种修正,相比 K-S 检验它加重了对尾部数据的考量,K-S 检验具有分布无关性,它的临界值并不依赖被测的特定分布,而 A-D 检验使用特定分布去计算临界值,这使得 A-D 检验具有更灵敏的优势。根据 K-S 与 A-D 拟合优度检验对几种分布函数的拟合效果进行检验。数据统计、分布拟合以及优度检验通过 MATLAB 软件实现。

1.1 Kolmogorov-Smirnov 检验

K-S 检验的检验方法是以样本数据的累计频数分布与特定理论分布比较,若两者间的差距很小,则推论该样本取自某特定分布族。设 $y_1, y_2, \dots, y_i, \dots, y_n$ 是样本量为 N 的样本且已按升序进行了排序,则 ECDF 的表达式为:

$$E_N = n(i)/N \quad (1)$$

式中, $n(i)$ 是小于 y_i 的点的数量。K-S 检验的构造

如下: H_0 为样本数据,来自于某个特定分布的总体; H_1 为样本数据,不是来自于某个特定分布的总体。K-S 检验的统计量定义为:

$$D = \max_{1 \leq i \leq N} \left[F(y_i) - \frac{1-i}{N}, \frac{i}{N} - F(y_i) \right] \quad (2)$$

其中, F 是被检验的分布的理论 CDF,它必须被完全定义,即其位置、比例和形状参数不能由样本来估计。我们通过将 D 与若干已被算出的临界值相比较来得出检验的结论。

1.2 Anderson-Darling 检验

A-D 检验所用的统计量一般用 A^2 来表示,其表达式是:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \{ \ln F(y_i) + \ln [1 - F(y_{n-i+1})] \} \quad (3)$$

其中, n 是样本量, F 是特定分布的概率分布函数(CDF)。另外注意 y_i 是按大小顺序排列的。A-D 检验的构造如下: H_0 为样本数据,来自于某个特定分布的总体; H_1 为样本数据,不是来自于某个特定分布的总体。

Stephens 给出的正态分布的临界值如下: $\alpha=0.10$ 时, $A=0.656$; $\alpha=0.05$ 时, $A=0.787$; $\alpha=0.025$ 时, $A=0.918$; $\alpha=0.01$ 时, $A=1.092$ 。

对于被检验的分布,在确定检验的置信度 α 后,计算统计量 A 并将其和相应置信度下的临界值进行比较,得出拒绝或接收原假设的结论。

2 结果与讨论

2.1 降水极值统计分析

站点降雨极值统计信息如表 2。10 个站点中全年 1 d 最大降雨量最高值位于珠江流域,广西省梧州站点,1966 年 6 月 12 日记录降雨量达到 334.5

表 1 逐日降水站点信息

Table 1 Summary of daily rainfall at meteorological site

站台编号	站名	省名	经度	纬度	海拔高度(m)	起始年月	终止年月	所属一级流域
50745	齐齐哈尔	黑龙江	123°55'	47°23'	147.1	1960.1	2010.12	松花江
54342	沈阳	辽宁	123°27'	41°44'	44.7	1960.1	2010.12	辽河
57051	三门峡	河南	111°12'	34°48'	409.9	1960.1	2010.12	黄河
58203	阜阳	安徽	115°44'	32°52'	32.7	1960.1	2010.12	淮河
53798	邢台	河北	114°30'	37°04'	77.3	1960.1	2010.12	海河
57461	宜昌	湖北	111°18'	30°42'	133.1	1960.1	2010.12	长江
59265	梧州	广西	111°18'	23°29'	114.8	1960.1	2010.12	珠江
58737	建瓯	福建	118°19'	27°03'	154.9	1960.1	2010.12	东南诸河
56227	波密	西藏	95°46'	29°52'	2736.0	1960.1	2010.12	西南诸河
52679	武威	甘肃	102°40'	37°55'	1531.5	1960.1	2010.12	西北诸河

表2 站点降雨极值统计信息

Table 2 Maximum precipitation for 1-, 2- and 3-day duration at meteorological sites

降雨极 值序列	连续降雨 日数(d)	站台 编号	最小值 (mm)	最大值 (mm)	均值 (mm)	标准差 (mm)	站台 编号	最小值 (mm)	最大值 (mm)	均值 (mm)	标准差 (mm)
全年	1	50745 齐齐哈尔	25.60	135.50	53.75	22.12	57051 三门峡	30.80	115.80	59.29	22.47
	2		34.00	159.70	66.62	26.57		32.70	153.10	70.27	24.42
	3		37.40	165.10	71.72	27.59		40.00	213.70	78.40	28.89
夏季	1		20.50	135.50	53.00	22.68		18.40	112.80	54.16	23.88
	2		22.00	159.70	65.30	27.77		18.80	153.10	62.93	26.36
	3		26.80	165.10	70.32	28.82		18.80	213.70	69.20	31.59
冬季	1		0.20	7.90	2.31	1.56		0.60	14.40	5.93	3.28
	2		0.20	8.20	2.72	1.79		0.60	17.00	7.02	4.06
	3		0.20	8.30	2.82	1.87		0.60	21.40	7.59	4.63
全年	1	52679 武威	8.30	62.70	22.12	11.67	57461 宜昌	32.10	229.10	96.37	39.53
	2		11.50	79.10	26.78	13.55		39.60	229.10	113.50	39.80
	3		11.70	87.80	29.98	15.25		52.40	229.10	126.57	40.25
夏季	1		4.50	62.70	20.21	12.70		32.10	229.10	89.72	38.20
	2		4.50	79.10	24.30	15.18		37.10	229.10	107.27	39.19
	3		4.50	87.80	26.74	16.74		51.50	229.10	119.91	40.56
冬季	1		0.10	5.30	2.11	1.33		3.90	36.90	15.95	7.97
	2		0.10	6.90	2.36	1.52		4.00	53.10	21.23	10.90
	3		0.10	8.50	2.65	1.92		4.50	73.40	24.35	14.12
全年	1	53798 邢台	25.40	304.30	80.46	47.50	58203 阜阳	38.20	226.10	95.16	43.56
	2		29.20	431.50	104.58	66.90		56.30	249.80	120.75	49.50
	3		30.00	570.20	116.66	86.14		57.10	334.30	133.22	52.95
夏季	1		25.40	304.30	77.16	48.08		36.00	226.10	89.45	46.28
	2		29.20	431.50	100.16	67.65		38.40	249.80	115.67	52.74
	3		30.00	570.20	111.52	87.47		44.20	334.30	127.47	57.05
冬季	1		0.10	16.20	5.54	3.85		4.80	53.90	17.91	9.17
	2		0.10	18.90	6.59	4.66		4.80	54.20	23.77	11.90
	3		0.10	21.30	7.17	5.36		4.80	65.60	27.15	14.24
全年	1	54342 沈阳	32.50	215.50	75.28	33.78	58737 建瓯	52.40	183.40	96.03	29.19
	2		44.10	216.40	91.68	36.88		75.60	331.70	132.67	48.62
	3		44.40	216.40	100.15	37.93		84.30	398.50	160.41	69.67
夏季	1		28.10	215.50	74.89	34.25		33.00	183.40	83.74	33.77
	2		34.00	216.40	91.05	37.71		41.90	331.70	116.92	55.22
	3		41.70	216.40	99.64	38.55		43.60	398.50	139.40	75.59
冬季	1		2.00	26.80	9.05	6.00		14.30	97.80	32.81	14.00
	2		2.00	27.50	10.49	6.82		16.50	125.30	46.13	19.55
	3		2.00	27.50	11.04	6.90		16.90	135.00	53.98	22.34
全年	1	56227 波密	22.10	111.70	43.41	17.40	59265 梧州	48.60	334.50	108.95	50.34
	2		28.30	146.50	63.49	23.10		71.80	353.60	133.90	53.40
	3		33.90	174.20	76.11	28.53		81.90	368.60	149.30	55.54
夏季	1		10.50	75.20	34.04	15.56		33.80	334.50	92.42	55.74
	2		19.50	107.50	51.12	23.30		49.40	353.60	113.89	61.02
	3		20.50	137.40	61.77	26.90		50.50	368.60	127.45	63.06
冬季	1		1.10	29.90	9.57	6.31		5.60	100.10	31.62	20.32
	2		1.40	46.20	13.76	9.89		5.90	139.40	42.28	25.43
	3		1.40	60.10	15.94	11.92		8.60	182.10	48.95	31.68

mm,同样为夏季1 d降雨量最高值;全年2 d最大降雨量最高值为431.5 mm,位于海河流域,河北省邢台站点,同样为夏季2 d最大降雨量最高值;全年3 d最大降雨量与夏季3 d最大降雨量的最高值也发生在此点,降雨量值达到570.2 mm。冬季1、2、3 d最大降雨量最高值均位于珠江流域,广西省梧州站点,1969年1月30日记录降雨量为100.1 mm。

从最大降雨量统计值最大值看,站点的年最大降雨量(1、2、3 d)大多发生在夏季,除波密站点(位于西北诸河流域,西藏)年最大1 d降雨量于1998年10月20日达到最大值111.7 mm,夏季最大1 d降雨量于1982年6月10日达到最大值75.2 mm,相应年最大2、3 d降雨量与夏季最大值时间点也不重合;三门峡站点(黄河流域,河南)年最大1 d降雨量为1972年9月1日,达到115.8 mm,夏季最大1 d降雨量记录时间点为1960年7月15日,降雨量为112.8 mm。

从最大降雨量统计值标准差比较,一般情况为夏季略高于全年,远高于冬季,这与中国地理位置与季风气候密切相关。但波密站点(位于西北诸河流域,西藏)与宜昌站点(长江流域,湖北)全年标准差略高于夏季,说明其最大降雨量统计值离散程度全年略高于夏季。

从降雨量最大值差异度来看,全年与夏季最大值差异较小(0~32.7%),全年(夏季)与冬季最大值差异较大(46.7%~96.3%);全年与冬季最大1 d降雨量差异度在邢台站点(海河流域,河北)为最大值点,达到94.7%,但全年与冬季最大1 d降雨量在梧州站点(珠江流域,广西)达最大值,分别为334.5与100.1 mm。最大2 d降雨量和最大3 d降雨量其差异度最大值与雨量最大值重合于同一站点。

2.2 拟合优度检验

采用5种分布函数与2种检验方法对10个站点最大降雨量序列进行拟合并做优度检验。

1) 从10个站点拟合检验结果可以看出,拟合效果较好的为广义极值分布(GEV, Generalized Extreme Value)、对数正态分布(LN, Lognormal)和皮尔逊III分布(Pearson-III)。其中,广义极值分布(GEV, Generalized Extreme Value)拟合效果最佳,在通过一种以上检验的降水序列中,全年最大降雨量序列为23个,占60%,夏季最大降雨量序列为14个,占35%,冬季最大降雨量序列为18个,占46%。其次为对数正态分布(LN, Lognormal),全

年最大降雨量序列为9个(24%),夏季为15个(38%),冬季为6个(15%)。再次为皮尔逊III分布(Pearson-III)全年最大降雨量序列为5个(13%),夏季为9个(23%),冬季为7个(18%)。

2) 全年与夏季的降水序列中,通过两种拟合优度检验分布函数仍是广义极值分布(GEV)、对数正态分布(LN)与皮尔逊III分布(Pearson-III);冬季最大降水序列中拟合较好的依次为广义极值分布(GEV, Generalized Extreme Value)(52%)、皮尔逊III分布(Pearson-III)(19%)和韦伯分布(Weibull)(14%)。

3) 同时适用于最大1、2和3 d降雨量序列且通过两种检验的概率分布函数有广义极值分布(GEV)(3个最大降雨量序列)和皮尔逊III分布(Pearson-III)(1个最大降雨量序列)。其中,武威站(西北诸河流域,甘肃,52679)年降雨量序列、建瓯站(东南诸河流域,福建,58737)冬季降雨量序列与梧州站(珠江流域,广西,59265)年降雨量序列,均以GEV拟合全年最大降雨量序列效果最优,邢台站(海河流域,河北,53798)以皮尔逊III分布拟合冬季最大降雨量序列效果最优。

4) 从10个站点分析结果可以看出,同一个站点其全年、夏季、冬季适宜性曲线不尽相同。由降雨量极大值统计分析已知,所选站点降雨量均集中于夏季,降雨量序列的离散程度与偏度导致这3个时期所适宜的分布函数有所差异。仅有一个站点建瓯站(东南诸河流域,福建,58737)的最大2、3 d降雨量在全年,夏季与冬季均以GEV拟合最佳,且满足两种检验。

2.3 重现期比较分析

以每个站点通过检验的概率函数拟合不同重现期下降雨极值的线型,图1~3展示重现期为1~100 a时部分站点的降雨量曲线。通过K-S与A-D两种优度检验的最大值降雨量序列用最优线型拟合,仅通过一种检验的序列尝试两种线型的拟合并对比差异性。结果表明:

1) 在所有的季节序列(AMS、SMS、WMS)中,曲线之间的差异度在重现期25 a以内均小于7.8%,且大多数拟合曲线在重现期25 a以内几乎重合(差异度小于4%)。仅站点53798(海河流域,的河北,邢台站)以广义极值分布(GEV)、对数正态分布(LN)两种曲线拟合夏季最大3 d降雨量差异度为7.8%;站点56227(西南诸河流域,西藏,波

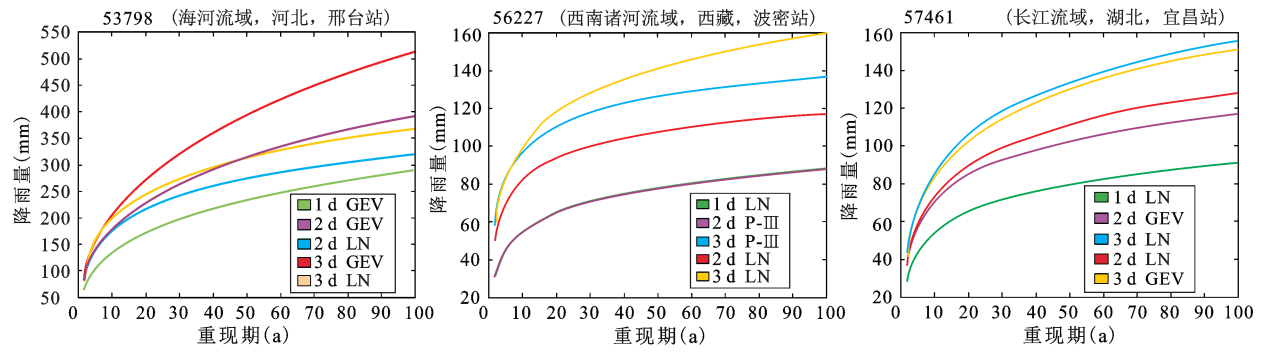


图1 拟合差异度大于6%的站点

Fig.1 The best-fit distributions of site with differences more than 6%

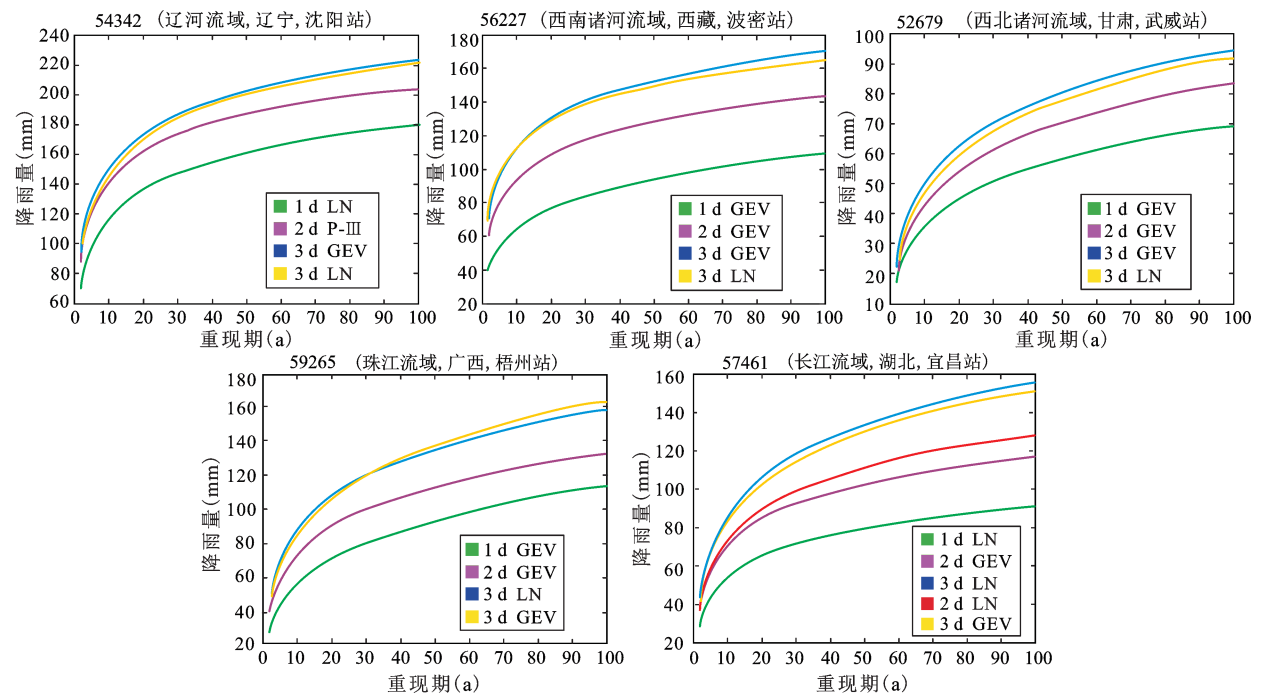


图2 GEV与LN拟合贴合度高的站点

Fig.2 High similarity curves estimated by GEV and LN

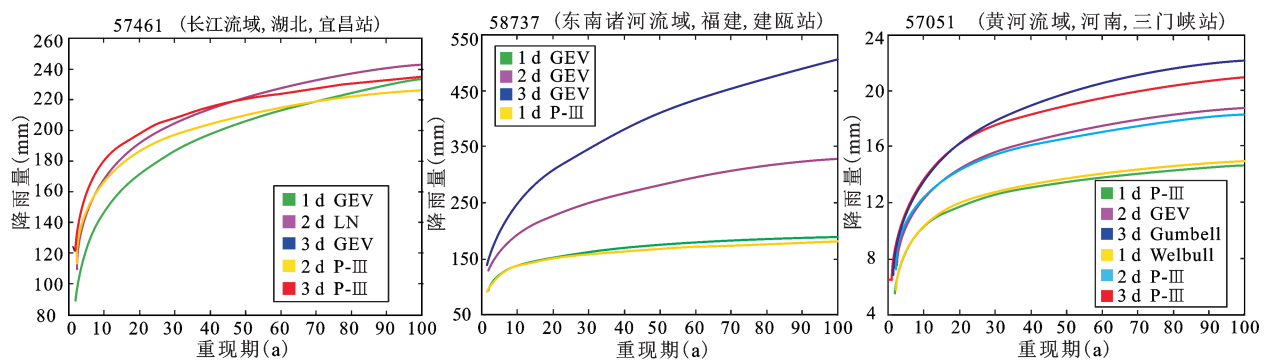


图3 GEV与Pearson-III拟合贴合度高的站点

Fig.3 High similarity curves estimated by GEV and Pearson-III

密站点)夏季最大3 d降雨量以广义极值分布(GEV)、对数正态分布(LN)拟合差异度为6%,冬季最大3 d降雨量以对数正态分布(LN)、皮尔逊III分布(Pearson-III)拟合差异度为6%;另外,站点57461(长江流域,湖北,宜昌站)冬季最大2 d降雨量拟合曲线以广义极值分布(GEV)、对数正态分布(LN)拟合在重现期为25 a时差异度为6.4%(见图1)。

2) 在一个降雨量极值序列的两种线型拟合中,一些站点表现出较好的贴合效果,即使重现期延伸到100年一遇,其曲线差异量仅为几毫米,这种线型相似度使得小概率降雨事件的计算结果更为可靠。在站点54342、56227年最大3 d降雨量序列,52679夏季最大3 d降雨量序列与57461、59265冬季最大3 d降水序列中广义极值分布(GEV)与对数正态分布(LN)曲线贴合效果较好(图2);在站点57461年最大3 d降雨量序列、58737年最大1 d降雨量序列与57051冬季最大2 d降雨量序列中,广义极值分布(GEV)与皮尔逊III分布(Pearson-III)线型分布几乎一致(图3);站点57461夏季最大1 d降雨量序列用对数正态分布(LN)与皮尔逊III分布(Pearson-III)分别拟合,线型重合度较高,重现期23 a时出现略微分歧,但重现期为100 a时拟合降雨量差异为3 mm;站点57051冬季最大1 d降雨量序列用皮尔逊III分布(Pearson-III)与韦伯分布(Weibull)拟合相差最大0.3 mm;站点58203冬季最大2 d降雨量序列用广义极值分布(GEV)与耿贝尔分布(Gumbell)拟合,两种线型基本重合。在11组贴合度较高的曲线中不难看出:广义极值分布(GEV),对数正态分布(LN),皮尔逊III分布(Pearson-III)三者之间组合的降雨量序列占82%,其中广义极值分布(GEV)参与组合的比例占73%,对数正态分布(LN)参与组合的比例占55%,皮尔逊III分布(Pearson-III)参与组合的比例占27%,这与前文10个站点拟合检验结果相吻合。

3) 个别站点降雨量序列以不同函数进行拟合呈现较大分歧。全年与夏季降水极值序列拟合曲线差异度最大值均在站点53798(海河流域,河北,邢台站),全年最大2 d降雨量以GEV、LN最大差异度为10.7%,夏季最大3 d降雨量以GEV、LN拟合,最大差异度为28.8%;另外在站点57461(长江流域,湖北,宜昌站),最大1、3 d拟合曲线贴合度较好,最大2 d拟合曲线以不同分布函数频率计

算结果大多存在偏差,全年最大2 d降雨量LN、P3拟合差异度最大为7%,夏季最大2 d降雨量LN、P3拟合差异度最大为8.8%,冬季最大2 d降雨量GEV、LN拟合差异度最大为9.4%。

4) 拟合曲线分歧呈现季节差异,夏季差异较高。对于拟合降雨量值的最大差异度,当拟合频率为0.01(重现期为100年一遇)时,夏季最大1、2、3 d降雨量差异度最大值(28.8%)大于年最大1、2、3 d降雨量差异度最大值(10.7%),冬季差异度最小(9.4%),与前文不同季节最大降雨量序列方差大小一致,即所选站点所代表的降雨量在夏季离散程度较大,夏季不同拟合曲线易呈现不同的结果。

5) 不同日数最大降雨量拟合效果不同。在1、2、3 d最大降雨量序列的拟合比较中,1 d拟合效果相对较好,除站点50745(夏季曲线最大差异度为8.7%)之外,数据序列中仅6个未一致通过两类检验,但两种线型都比较贴近。最大3 d降雨量序列的拟合效果次之,仅站点56227与站点53798的100年一遇夏季最大3 d降雨量相差13.8%与28.8%。最大2 d降雨量有10个数据序列未一致通过两种检验,站点53798与站点57461的全年、夏季与冬季拟合曲线中,100年一遇降雨量值拟合差异范围为7%~17.9%。

3 结 论

应用5种概率分布函数对全国范围内10个站点进行降雨极值拟合,通过K-S、A-D检验做出适宜函数的筛选,分析比较了降雨极值的分布规律,对站点的适宜曲线做拟合优度检验并比较不同重现期的曲线间差异,得出以下初步结论:

1) 通过10个站点的降雨量最大值统计,得出年最大降雨量多发生于夏季。各站点降雨量极值的标准差(离散程度)说明降雨量存在季节性差异(分析了冬、夏季差异),尤其是夏季降雨量的波动比较。

2) 两种方法(K-S、A-D)检验后,拟合优度检验结果说明10个站点广义极值分布(GEV, Generalized Extreme Value)拟合效果最佳,其次为对数正态分布(LN, Lognormal),再次为皮尔逊III分布(Pearson-III)。

3) 此次计算分析得出皮尔逊III分布不是最优拟合曲线,分析结果进一步说明降雨极值拟合时应采用多种概率分布函数及检验方法,提高降

雨极值计算结果的可靠性。

4) 以两种分布拟合且差异度较小的组合中, 广义极值分布(GEV), 对数正态分布(LN), 皮尔逊 III 分布(Pearson-III)三者之间的组合占 82%, 这种线型相似度使得小概率降雨事件的计算结果更为可靠。

研究进一步证明, 在进行水文气象极值统计分析时, 应采取多种概率分布函数及检验方法进行计算, 降低单一计算方法的不确定性, 本研究随机 10 个站点的试验分析就说明皮尔逊 III 分布并非适宜所有站点。不同季节不同站点的适宜性曲线存在差异。由此建议防洪规划设计部门在防洪基础设施设计时应寻找最优极值分析特征方法计算极端极值。

参考文献:

- [1] IPCC.Climate change 2013:the physical science basis[M/OL]. Cambridge:Cambridge University Press,in press.2013-09-30. <http://www.ipcc.ch/report/ar5/wg1/#.UqTd7kbrrl>.
- [2] 王 庆,马倩倩,夏艳玲,等.最近 50 年来山东地区夏季降水的时空变化及其影响因素研究[J].地理科学,2014,34(2):220~228.
- [3] 张 婷,魏凤英.华南地区汛期极端降水的概率分布特征[J].气象学报,2009,67(3):442~451.
- [4] 张 强,孙 鹏,白云岗,等.塔河流域枯水流量概率特征及成因与影响研究[J].地理科学,2013,33(4):465~472.
- [5] 王国庆,王兴泽,张建云,等.中国东北地区典型流域水文变化特性及其对气候变化的响应[J].地理科学,2011,31(6):641~646.
- [6] 周玉良,袁潇晨,金菊良,等.基于 Copula 的区域水文干旱频率分析[J].地理科学,2011,31(11):1383~1388.
- [7] 李 建,宇如聪,孙 颢.从小时尺度考察中国中东部极端降水的持续性和季节特征[J].气象学报,2013,71(4):652~659.
- [8] Sevruck B,Geigerh.Selection of distribution types for extremes of precipitation[M].World Meteorological Organization,Operational Hydrology Report,1982:15.
- [9] Viatcheslav V K,Francis W Z.Estimating extremes in transient climate change simulations[J].Journal of Climate,2005,18:1156-1172.
- [10] 屠其璞,丁裕国.气象应用概率统计学[M].北京:气象出版社,1984:32.
- [11] Wiks D S.Raifall intensity,the Weibull distribution and estimation of daily surface runoff[J].Journal of ApplMeteor,1989,28(1):52-58.
- [12] 马开玉.气候统计原理与方法[M].北京:气象出版社,1993:69~76.
- [13] 张耀存,张录军.东北气候和生态过渡区近 50 年来降水和温度概率分布特征变化[J].地理科学,2005,25(5):561~566.
- [14] 冯 平,王仲珏.基于二维 Gumbel 分布的降雨径流频率分析模型及其应用[J].干旱区资源与环境,2007,21(10):68~72.
- [15] 蔡 敏,丁裕国,江志红.我国东部极端降水时空分布及其概率特征[J].高原气象,2007,26(2):309~318.
- [16] Park J S,Jung H S,Kim R S et al.Modelling summer extreme rainfall over the Korean Peninsula using Wakebydistribution[J].International Journal of Climatology, 2002,21:1371-1384.
- [17] 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.
- [18] 黄振平,陈元芳.水文统计学[M].北京:中国水利水电出版社,2011:159.
- [19] Justel A,Peña D,Zamar R.A multivariate Kolmogorov-Smirnov test of goodness of fit[J].Statistics & Probability Letters,1997,35(3):251-259.
- [20] Anderson K,Darling D A.Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes[J].Ann. Math. Stat.,1952,23:193-212.
- [21] Anderson T W,Darling D A. A test of goodness-of-fit.J[J].Am. Stat.Assoc.,1954,49:765-769.

Applicability of Different Probability Distributions to Estimated Extreme Rainfall

ZHANG Yu-hu¹, WANG Chen-xi^{1,2}, LIU Kai-li³, CHEN Qiu-hua³

(1. *College of Resources Environment and Tourism, Capital Normal University, Beijing, 100048, China;*

2. *Institute of Surveying and City Spatial Information, Beijing University of Civil Engineering and Architecture, Beijing 100044, China;* 3. *Institute of Math Science, Beijing University of Civil Engineering and Architecture, Beijing 100044, China*)

Abstract: Heavy precipitation is a crucial nature factor of flood. The return period of the extreme value of precipitation depth is the most significant reference of the design standard of flood prevention facilities in an urban or a basin. In this article, the series of annual, summer and winter maxima of precipitation depths for 1-day, 2-day and 3-day durations measured at ten selected stations in China are analyzed, using five commonly used hydrological statistical distribution functions. The distribution functions applicable for these stations were measured using the Kolmogorov Smirnov (K-S) and the Anderson Darling (A-D) tests. The results show that: 1) The summer maxima series shows higher standard deviation and larger differences between distributions than other maxima series and the annual maxima occur mostly in summer; 2) The Generalized Extreme Value (GEV) distribution, the lognormal (LN) distribution and the Pearson III distribution perform were better in the imitative effect test of goodness of fit, and the degree of curve difference is smaller; 3) Differences between estimates of rainfall with return periods were shorter than 25 years are smaller; 4) Estimates of precipitation can change significantly depending on the probability distribution being used, particularly for the summer series; 5) Suitability curves present seasonal difference. By statistical analysis of precipitation maxima, the precipitation is concentrated in summer; due to the disperse and skewness of precipitation series, and the appropriate distribution functions are quite different in different season periods; 6) In some extreme rainfall sequences, two curves of linear fitting are almost the same. Even if the return period extending to 100 years, the difference quantity of two curves is only a few millimeters. In this situation, the results of small probability rainfall events are more reliable; 7) There are differences among 1-day, 2-day and 3-day durations of precipitation depths, the probability distribution of 1-day maximum precipitation fits better. When carrying out statistical analysis of hydro-meteorological extremes, various probability distribution function and test methods should be taken for calculating, to reduce the uncertainty of single calculation. In this study, experimental analysis of 10 sites demonstrated that the Pearson III is not suitable for all sites. It is suggested here that the estimation of extreme precipitation should take into consideration the range of extreme values estimated by the best-fit distributions identified by more than one test as an approach to assess uncertainties related to extreme rainfall analysis.

Key words: extreme precipitation; probability distribution; return periods; Kolmogorov-Smirnov test; Anderson-Darling test