基于极值理论的北京市极端降水概率分布研究

宋晓猛^{1,2,3*}, 张建云^{2,4}, 孔凡哲¹

1. 中国矿业大学资源与地球科学学院, 徐州 221116;
2. 南京水利科学研究院, 水文水资源与水利工程科学国家重点实验室, 南京 210029;
3. 武汉大学水资源与水电工程科学国家重点实验室, 武汉 430072;
4. 水利部应对气候变化研究中心, 南京 210029

* E-mail: xmsong@cumt.edu.cn

收稿日期: 2017-05-02; 接受日期: 2017-08-28; 网络版发表日期: 2018-02-09

国家自然科学基金(批准号: 51609242)、中国矿业大学中央高校基本科研业务费(编号: 2015XKMS034)、水资源与水电工程科学国家重点实验室开放基金(编号: 2015SWG02)和水文水资源与水利工程科学国家重点实验室开放基金(编号: 2015490411)资助项目

摘要 以北京市作为研究区域, 基于极值理论分析北京市降水极值的概率统计特征, 探索不同极值分布在北京市降水极值研究中的适用性. 选择北京市45个雨量站的1960~2012年汛期逐日降水数据, 基于年最大值法和超门限阈值法(百分位法)筛选降水极值样本序列(AM和POT序列). 利用广义极值分布(GEV)和广义Pareto分布(GPD)拟合两组样本序列, 采用L矩法估计两种分布函数的拟合参数. 基于Kolmogorov-Smirnov检验结果确定最优拟合分布函数, 并根据最优拟合分布函数推求不同重现期水平下的降水量. 结果表明: (1) 除半壁店和沿河城外, AM序列的均值大于POT序列, 所有站点AM序列的最小值小于POT序列; (2) GEV分布在AM序列的最优拟合分布函数中占优(40个站点, 占全部站点的88.9%), 而GEV和GPD分布均可用于POT序列降水极值的拟合, 其中GPD分布的拟合结果较好(38个站点, 占比为84.4%); (3) 两种序列得到的不同重现期水平的降水量空间分布基本相似, 总体上利用AM序列推求不同重现期水平下的降水量略大于POT序列, 即基于AM序列推算一定重现期水平下的降水极值更有利于规避风险.

关键词 极端降水, 年最大值, 超门限阈值, 概率分布, 北京

1 引言

根据联合国人类住区规划署发布的《2011世界人口状况报告》指出, 到2011年底全球约50%的人口居住在城市, 预计未来城市化进程将继续加快, 城市人口持续增长, 特别是发展中国家和地区城市人口增长最为显著^[1,2]. 与此同时, 全球变化背景下城市化发展

一定程度上增加了城市暴雨洪涝风险^[3]. 近年来, 受全球气候变化和高强度人类活动影响, 极端降水事件发生频次增加, 给区域防洪安全乃至区域经济发展带来了严峻挑战^[4,5], 特别是高度城市化地区. 一般来说高强度(或极端)降水事件往往控制着城市洪涝灾害事件的发生与演变, 分析极端降水变化特征成为城市防洪减灾工作的重要内容之一^[6,7].

引用格式: 宋晓猛, 张建云, 孔凡哲. 基于极值理论的北京市极端降水概率分布研究. 中国科学: 技术科学, 2018, 48: 639–650
Song X M, Zhang J Y, Kong F Z. Probability distribution of extreme precipitation in Beijing based on extreme value theory (in Chinese). Sci Sin Tech, 2018, 48: 639–650, doi: 10.1360/N092017-00129

伴随着城市化进程的不断加快, 城市区域极端降水事件频发, 使得城市化地区极端降水变化研究也逐渐受到广泛关注^[6]. 如Mishra等人^[8]基于全球217个城市气象站点分析全球主要城市的极端气候事件变化特征, 与极端气温事件相比, 极端降水变化趋势表现出明显的空间分布差异, 印度和南美洲部分站点存在极端降水频率显著性增加, 而其他区域则并未表现出显著变化特征, 总体上约有17%的站点日极端降水频率表现出显著增加趋势; Ali等人^[9]分析了印度57个主要城市极端降水变化规律; Han等人^[10]调查了长江三角洲城市群地区的极端降水变化特征; Wahl等人^[11]指出极端降水事件使得美国主要城市的洪涝风险日益凸显; Min等人^[12]调查了城市化发展过程中人类对极端降水事件的贡献程度, Miao等人^[13]以北京2006年8月1日的暴雨事件分析了北京城市化进程对夏季降水的影响; Wang等人^[14]探讨了珠江三角洲地区城市化对于极端降水的增强影响机制. 北京作为全球典型的超大城市, 近年来环境变化条件下极端降水事件频发, 由此引发的城市洪涝灾害严重威胁北京地区社会经济安全^[15,16]. 目前国内外针对北京降水极值问题的研究主要围绕极端降水长期变化趋势^[6]和单次极端事件及其城市化的影响^[13,17]等方面, 但针对城市化地区降水极值分布和重现期特征的研究相对较少. 本文以北京作为研究区域, 利用极值理论分析极端降水序列的最优拟合函数, 分析不同重现期条件下的极端降水特征分布, 以期在北京地区应对极端降水和城市洪涝灾害提供科学依据.

2 资料来源与方法

2.1 研究区域与数据

北京市位于华北平原的东北边缘(39°28′~41°05′N, 115°25′~117°30′E), 地处海河流域中部, 毗邻渤海湾, 上靠辽东半岛, 下临山东半岛. 西部为太行山山脉的西山, 北部是燕山山脉的军都山, 山区面积约占总面积的61%. 平原地区的海拔高度在20~60 m, 山地一般海拔1000~1500 m(图1). 北京属于典型的北温带半湿润大陆性季风气候, 夏季高温多雨, 冬季寒冷干燥, 降水季节分配不均, 全年降水主要集中在夏季(6~8月)和汛期(6~9月). 据1950~2012年北京平均降水量统计, 夏季降水约占全年的72.3%(39.5%~87.4%), 汛期占81.5%

图1 (网络版彩图)研究区雨量站点空间分布

(60.1%~91.5%)^[18].

选择北京范围内数据较为完整的45个雨量站汛期(6~9月)逐日降水数据(数据来源: 北京水文总站和《海河流域水文资料》), 站点布设如图1所示. 45个雨量站点分布较为均匀, 可较好地控制北京地区的降水空间分布特征. 根据各雨量站观测资料长度, 本文采用的研究序列为1960~2012年, 所有数据均经过严格质量控制, 满足国家水文数据质量控制要求.

2.2 研究方法

2.2.1 样本筛选

目前国内外主要采用年最大值法(annual maximum, AM)和超门限阈值法(peak-over-threshold, POT)选择极端降水样本^[5]. AM序列由每年的最大日降水量组成, 但考虑到气候条件的年际差异, 多水年可能出现多次极端降水事件, 少水年可能存在无强降水事件, 由此导致AM序列可能存在漏选或掺入假信息的问题. 因此, 本文在考虑AM序列的同时也采用POT序列筛选极端降水事件, 构建极端降水序列. 对于POT序列, 本文采用百分位阈值法确定门限值^[6], 从而确定极端降水序列样本, 具体步骤是: 将多年逐日降水量从小到大排序, 取某一特定百分位下的日降水量作为极端降水阈值, 若某日降水量超过该阈值, 则设定为极端降水日. 本文以95%阈值进行样本选择分析. 针对AM序列和POT序列的一致性检验, 本文采用Mann-Kendall

检验方法进行序列趋势分析, 根据检验结果(限于篇幅, 具体图略)可知对于AM序列而言有13个站点存在显著性变化趋势, 而对于POT序列则有5个站点存在显著性变化趋势, 但考虑到上述站点占比相对较少, 因此本研究中未考虑样本序列的非平稳性特征, 将在今后研究中开展非平稳性条件下的极端降水事件频率分析及风险评估.

2.2.2 极值分布

目前用于研究极端气候事件的统计分布模型主要包括广义极值分布(GEV)、广义Pareto分布(GPD)、Gamma分布和广义Logistic分布(GLO)以及广义正态分布(GNO)等, 其中, GEV模型综合了Weibull, Gumbel, Frechet极值分布, 可以得到一定置信水平下重现期的置信区间, 能满足实际工程设计需要, 在拟合极端降水方面有较好的应用; GPD模型是一种专门描述超过门限峰值的极值分布, 在极端降水与洪水序列统计分析等方面也存在较多的应用. 因此, 本文采用GEV分布和GPD分布进行北京地区极端降水拟合研究, 两种分布的理论基础如下.

GEV分布的分布函数为

$$F(x | \mu, \sigma, \xi) = \exp \left[- \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}} \right], \quad (1)$$

$$1 + \xi \frac{x - \mu}{\sigma} > 0,$$

式中, x 为极端降水序列, μ , σ 和 ξ 分别为位置参数、尺度参数和形状参数. 在计算得到函数的参数之后, 即可求出给定重现期 T 对应的分位数 X_T (即 T 年一遇的极端降水量)为

$$X_T = \mu + \frac{\sigma}{\xi} (1 - (-\ln(1 - 1/T))^{\xi}). \quad (2)$$

GPD分布的分布函数为

$$F(x | \mu, \sigma, \xi) = 1 - \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\xi}}, \quad (3)$$

$$x \geq \mu, 1 + \xi \frac{x - \mu}{\sigma} > 0.$$

同样在计算得到函数的参数之后, 即可求出给定重现期 T 对应的分位数 X_T 为

$$X_T = \mu + \frac{\sigma}{\xi} (1 - (1/T)^{\xi}). \quad (4)$$

根据分位数公式求得一定降水量的理论重现期可与经验重现期做对比, 经验重现期的计算参考Cunnane给出的公式^[19]:

$$T = \frac{1}{\frac{n}{N} \left(1 - \frac{i-0.4}{n+0.2} \right)}, \quad (5)$$

式中, n 为极端降水序列样本数, N 为极端降水序列总年数, $i = 1, 2, \dots, n$.

2.2.3 参数估计

分布函数参数估计方法主要有矩法、极大似然法、最小二乘法、概率权重矩法和L矩法等. 考虑到L矩法对序列的极大值和极小值没有常规矩那么敏感, 其求得的参数估计值比较稳健^[20], 因此本文中采用L矩法估计极值分布函数的参数, 具体计算方法如下^[20-22].

对于给定极端降水样本序列 $x = \{x_1, x_2, \dots, x_n\}$, 将样本按照从小到大的顺序排列, 即 $x_{1,n} \leq x_{2,n} \leq \dots \leq x_{n,n}$, 则样本的前4阶矩为

$$\begin{cases} L_1 = b_0, \\ L_2 = 2b_1 - b_0, \\ L_3 = 6b_2 - 6b_1 + b_0, \\ L_4 = 20b_3 - 30b_2 + 12b_1 - b_0, \end{cases} \quad (6)$$

式中,

$$b_0 = \frac{1}{n} \sum_{i=1}^n x_{i,n}, \quad b_1 = \frac{1}{n} \sum_{i=2}^n \frac{i-1}{n-1} x_{i,n},$$

$$b_2 = \frac{1}{n} \sum_{i=3}^n \frac{(i-1)(i-2)}{(n-1)(n-2)} x_{i,n},$$

$$b_3 = \frac{1}{n} \sum_{i=4}^n \frac{(i-1)(i-2)(i-3)}{(n-1)(n-2)(n-3)} x_{i,n}.$$

由此可计算得到变差系数 τ_2 、偏态系数 τ_3 和峰度系数 τ_4 分别为

$$\tau_2 = L_2 / L_1, \quad \tau_3 = L_3 / L_2, \quad \tau_4 = L_4 / L_2. \quad (7)$$

进而可求得GEV分布的L矩估计参数结果:

$$\begin{cases} \hat{\mu} = L_1 + \hat{\sigma} [\Gamma(1 + \hat{\xi}) - 1] / \hat{\xi}, \\ \hat{\sigma} = L_2 \hat{\xi} / \left[(1 - 2^{-\hat{\xi}}) \Gamma(1 + \hat{\xi}) \right], \\ \hat{\xi} \approx 7.8590z + 2.9554z^2, \\ z = 2 / (3 + \tau_3) - \ln 2 / \ln 3, \end{cases} \quad (8)$$

和GPD分布的L矩估计参数结果:

$$\begin{cases} \hat{\mu} = L_1 - (2 + \hat{\xi}) / L_2, \\ \hat{\sigma} = L_2(1 + \hat{\xi})(2 + \hat{\xi}), \\ \hat{\xi} = (1 - 3\tau_3) / (1 + \tau_3), \end{cases} \quad (9)$$

式中, $\Gamma(x)$ 为Gamma函数.

2.2.4 分布检验

对比GEV和GPD模型的理论分布与极端降水序列的经验分布的偏离程度, 采用Kolmogorov-Smirnov (KS)检验. KS检验将理论分布和经验分布的累积频数进行比较, 确定两者的最大差异值 D , D 值越小则说明越符合假定分布. 当经验分布函数与理论分布函数的最大差值 D 小于KS检验临界 D_α 时, 则认为原假设的理论分布函数与样本序列的经验分布拟合较好. 本文选择显著性水平 $\alpha=0.05$, 则KS检验的临界值 $D_\alpha=0.16$. 若存在多个拟合分布都满足显著性检验结果, 则选择 D 值最小作为最佳拟合分布.

3 结果分析

3.1 极端降水序列

利用北京1960~2012年45个雨量站的汛期逐日降水序列, 分别采用AM方法和POT方法进行样本筛选得

到两种不同极端降水序列, 即AM序列和POT序列, 得到各站点两种序列的箱线图如图2所示. 以平谷、卢沟桥、通州和右安门4个站为例, 说明两种降水极值序列的选取结果, 如图3所示. 由图2和3可知, AM序列结果部分年份年最大日降水量在95%阈值之下, 而POT序列则存在一年有多个极端降水日数. 以通州为例, AM序列样本数为53, 降水极值为22.0~303.8 mm, 而POT序列样本数92, 降水极值的阈值已达到52.0 mm, 由此可知AM抽样法遗漏了部分有效的降水极值.

针对两种极端降水样本序列, 分别统计AM序列和POT序列的均值、最大值、最小值、变差系数和偏态系数, 如表1所示. 根据统计结果可知, 两个序列的最大值是一致的, 而均值和最小值存在较大差异. 除半壁店和沿河城外, AM序列的均值均大于POT序列的均值. 而AM序列的最小值均小于相应站点POT序列的最小值, 即AM序列中部分枯水年份的年最大降水并未达到POT序列的极端降水阈值门槛. 从变差系数角度来看, AM序列的变差系数偏高, 说明AM序列的变异性较POT序列更明显. 从偏态系数来看, POT序列的右偏分布较AM序列明显.

3.2 分布函数的参数估计与拟合最优分布的确定

针对前述得到的两种极端降水序列(AM序列和

图2 AM序列和POT序列样本的箱型图. 测站序号顺序对应表1中的雨量站顺序, 实心方框代表序列均值, 方框中3条横线分别表示25%、50%和75%分位数, ×表示序列最大值和最小值

图 3 (网络彩图版)典型站点两种序列样本对比

POT序列), 采用GEV分布和GPD分布对上述两种极值序列进行拟合, 利用L矩法分别估计两种极值序列GEV分布和GPD分布的位置、尺度和形状参数, 如表2所示. 当采用GEV分布进行拟合时, AM极端降水序列中有4个站点(半壁店、凤河营、密云、密云白)形状参数为正, 属于Weibull分布, 其余41个站点形状参数为负值, 属于厚尾极值II型分布, 而POT极端降水序列全部站点的形状参数均为负, 属于厚尾极值II型分布. 当采用GPD分布进行拟合时, AM极端降水序列的形状参数有31个正数, 属于薄尾分布, 14个为负值, 属于厚尾分布, 而对于POT序列, 除了怀柔 and 密云外, 其余站点形状参数均小于0, 属于厚尾分布.

采用KS拟合检验方法计算GEV和GPD分布的KS统计值, 如表3所示. 从表3中可知, 对于AM序列, 所有站点采用GEV分布进行拟合都能通过KS检验, 而采用GPD分布进行拟合时, 则有20个站点未能通过 $\alpha=0.05$ 的KS检验. 而对POT序列而言, 全部站点都能通过GEV分布和GPD分布的KS拟合检验. 根据统计结果表

明, AM序列的最优分布有5个站(高碑店、温泉、霞云岭、右安门、斋堂)是GPD分布, 其余40个站是GEV分布, 而POT序列中有38个站的最优分布是GPD分布, 仅有7个站为GEV分布.

3.3 重现期推算

根据L矩法估计得到的AM和POT两种降水序列对应的分布函数参数, 结合分位数公式估算各站点不同重现期下的降水量, 并与样本的经验重现期作比较. 图4给出了卢沟桥站两种极值分布函数计算的重现期与样本经验重现期的对比. 对于卢沟桥站的AM序列而言, GEV分布推算的一定重现期的降水极值与经验值的拟合效果略优于GPD分布结果, 而对于POT序列, GEV和GPD分布的拟合效果基本接近, 难以评判优劣, 这与前文的KS检验结果相一致.

根据KS检验结果, 利用反距离权重法进行空间插值, 选择北京地区各站点AM和POT两种序列的最优分布推算不同重现期水平下(10年、20年、50年和100

表 1 不同抽样序列样本的统计特征

站点 (序号)	AM序列					POT序列				
	均值	最大值	最小值	变差系数	偏态系数	均值	最大值	最小值	变差系数	偏态系数
半壁店(1)	74.4	155.0	28.5	0.42	0.86	74.5	155.0	51.8	0.32	1.71
大村(2)	59.1	153.3	25.2	0.50	1.45	58.2	153.3	37.5	0.41	1.88
大孙各庄(3)	77.4	389.8	29.5	0.71	4.17	75.4	389.8	50.0	0.57	5.56
番字牌(4)	90.6	362.5	24.0	0.60	2.55	79.0	362.5	47.0	0.54	3.49
房山(5)	87.4	331.5	37.0	0.57	3.00	84.1	331.5	53.5	0.44	3.73
凤河营(6)	73.7	156.1	23.0	0.43	0.80	73.5	156.1	48.9	0.33	1.50
高碑店(7)	88.1	257.9	28.6	0.57	1.42	78.1	257.9	45.7	0.51	2.26
怀柔(8)	97.8	237.5	34.2	0.46	1.20	83.5	237.5	49.4	0.43	1.83
黄村(9)	80.9	236.5	28.2	0.52	1.75	75.0	236.5	46.9	0.45	2.43
黄花城(10)	77.7	196.4	29.3	0.49	1.34	74.2	196.4	49.9	0.38	2.24
黄松峪(11)	87.2	216.4	33.4	0.49	1.40	78.3	216.4	51.0	0.41	2.27
喇叭沟门(12)	55.5	168.0	22.5	0.51	2.59	50.8	168.0	34.3	0.42	3.67
乐家花园(13)	82.3	187.7	30.7	0.47	0.95	74.8	187.7	48.2	0.39	1.87
卢沟桥(14)	84.2	259.7	21.9	0.50	1.77	82.9	259.7	52.0	0.43	1.99
马驹桥(15)	80.8	204.9	25.2	0.51	1.16	78.5	204.9	47.0	0.43	1.90
密云(16)	85.6	191.0	28.4	0.39	0.67	76.8	191.0	48.0	0.34	1.47
密云白(17)	92.3	195.9	28.7	0.39	0.78	78.5	195.9	51.8	0.38	1.75
密云潮(18)	81.2	289.3	33.8	0.53	2.67	74.1	289.3	49.3	0.43	3.20
南各庄(19)	71.3	165.7	25.0	0.47	1.26	69.8	165.7	47.4	0.37	2.01
平谷(20)	84.2	230.1	35.0	0.45	2.07	78.8	230.1	51.1	0.38	2.74
千家店(21)	54.5	171.0	19.2	0.57	1.70	49.3	171.0	31.6	0.48	2.67
三家店(22)	102.2	268.2	26.2	0.52	1.35	87.9	268.2	49.7	0.47	2.01
沙河(23)	86.5	371.7	26.7	0.64	2.81	76.5	371.7	43.2	0.55	3.96
十三陵(24)	78.4	161.5	27.7	0.46	0.82	69.2	161.5	43.0	0.40	1.70
顺义(25)	88.6	253.9	29.9	0.51	1.73	83.8	253.9	54.0	0.39	2.53
松林闸(26)	90.8	325.3	30.0	0.57	2.31	81.1	325.3	47.2	0.48	3.33
苏庄(27)	90.8	325.7	27.7	0.60	2.18	81.6	325.7	48.1	0.49	2.99
汤河口(28)	61.7	180.4	25.3	0.51	1.83	52.4	180.4	34.1	0.49	2.65
唐指山(29)	88.4	215.5	35.4	0.46	1.22	82.6	215.5	52.6	0.40	1.91
桃峪口(30)	87.8	294.6	28.2	0.57	1.82	74.7	294.6	45.5	0.50	2.83
通州(31)	85.7	303.8	22.0	0.59	2.18	82.0	303.8	52.0	0.47	2.70
王家园(32)	78.6	250.0	20.2	0.66	1.65	67.7	250.0	38.1	0.58	2.72
温泉(33)	93.5	257.5	29.0	0.60	1.29	86.4	257.5	50.1	0.51	2.13
霞云岭(34)	96.3	298.0	30.1	0.58	1.42	83.7	298.0	50.5	0.51	2.36
下会(35)	74.5	207.4	23.8	0.48	1.88	67.9	207.4	46.2	0.42	2.65
延庆(36)	54.1	162.5	23.6	0.53	2.18	48.0	162.5	31.0	0.45	3.05
沿河城(37)	55.8	137.9	17.8	0.47	0.97	56.5	137.9	35.3	0.38	1.67
雁翅(38)	64.3	256.4	25.4	0.69	2.53	56.2	256.4	35.0	0.58	3.66
右安门(39)	86.5	288.8	31.0	0.57	1.83	76.1	288.8	46.5	0.48	2.98
榆林庄(40)	83.4	308.0	26.6	0.62	2.22	75.1	308.0	46.5	0.53	3.18
枣树林(41)	109.2	479.2	47.0	0.65	3.17	93.6	479.2	56.0	0.57	4.24
斋堂(42)	53.7	122.2	21.8	0.47	1.06	49.8	122.2	32.8	0.39	1.86
张坊(43)	94.8	387.8	24.3	0.66	2.97	82.6	387.8	49.6	0.56	4.19
张家坟(44)	94.8	218.9	23.3	0.54	0.83	81.2	218.9	46.3	0.51	1.62
镇罗营(45)	96.8	291.6	22.0	0.53	1.57	92.3	291.6	59.1	0.42	2.46

表 2 AM序列和POT序列的两种分布模型参数估计结果

站名	AM序列						POT序列					
	GEV分布			GPD分布			GEV分布			GPD分布		
	位置	尺度	形状	位置	尺度	形状	位置	尺度	形状	位置	尺度	形状
半壁店	60.16	25.97	0.031	31.00	64.11	0.478	62.88	13.63	-0.244	50.04	24.83	-0.006
大村	43.92	17.40	-0.232	27.41	32.10	0.014	45.52	10.08	-0.382	36.77	16.03	-0.222
大孙各庄	54.76	21.02	-0.339	36.06	34.82	-0.157	59.54	11.66	-0.447	49.78	17.44	-0.319
番字牌	66.16	34.41	-0.118	31.13	71.63	0.205	60.21	17.69	-0.359	44.64	28.76	-0.188
房山	64.98	25.83	-0.242	40.63	47.16	-0.003	65.68	13.18	-0.402	54.36	20.57	-0.252
凤河营	58.86	25.87	0.005	30.33	61.90	0.428	61.12	14.05	-0.254	47.96	25.35	-0.022
高碑店	63.49	33.20	-0.143	30.22	67.27	0.162	58.27	17.18	-0.368	43.23	27.68	-0.202
怀柔	76.46	32.73	-0.070	42.08	71.92	0.292	67.05	20.45	-0.237	47.70	37.53	0.005
黄村	60.15	24.90	-0.207	36.17	47.13	0.054	57.78	13.13	-0.390	46.43	20.72	-0.235
黄花城	58.88	24.62	-0.160	34.46	49.00	0.134	60.89	12.92	-0.352	49.47	21.15	-0.176
黄松峪	65.65	28.51	-0.154	37.27	57.09	0.144	62.66	14.60	-0.340	49.66	24.19	-0.157
喇叭沟门	42.65	13.92	-0.264	29.69	24.86	-0.038	41.36	8.06	-0.375	34.33	12.91	-0.212
乐家花园	63.03	29.07	-0.080	32.70	63.13	0.273	59.44	14.09	-0.335	46.87	23.44	-0.150
卢沟桥	64.15	26.42	-0.156	37.89	52.77	0.140	63.49	14.20	-0.374	51.10	22.75	-0.210
马驹桥	60.33	29.82	-0.099	29.61	63.39	0.239	61.38	16.00	-0.329	47.05	26.78	-0.141
密云	71.54	30.56	0.132	34.73	85.30	0.677	65.36	17.89	-0.146	47.47	36.12	0.157
密云白	76.05	29.78	0.034	42.54	73.78	0.484	64.92	16.24	-0.258	49.75	29.17	-0.029
密云潮	60.91	20.25	-0.304	42.51	34.74	-0.102	59.31	12.58	-0.411	48.56	19.48	-0.265
南各庄	54.29	22.25	-0.161	32.23	44.24	0.132	56.59	10.88	-0.375	47.10	17.41	-0.212
平谷	66.83	22.74	-0.159	44.26	45.32	0.136	64.71	14.08	-0.280	51.73	24.75	-0.064
千家店	39.02	17.32	-0.246	22.73	31.47	-0.010	37.84	8.39	-0.435	30.76	12.70	-0.301
三家店	76.75	37.74	-0.090	37.63	81.06	0.255	66.53	21.80	-0.300	46.67	37.54	-0.096
沙河	60.68	30.50	-0.214	31.43	57.30	0.042	56.85	17.00	-0.365	41.93	27.49	-0.196
十三陵	60.27	27.17	-0.076	31.85	59.26	0.280	55.00	13.87	-0.319	42.49	23.46	-0.125
顺义	66.56	28.22	-0.171	38.77	55.49	0.115	67.39	15.06	-0.367	54.20	24.28	-0.200
松林闸	66.46	29.89	-0.196	37.47	57.26	0.074	62.93	17.19	-0.330	47.53	28.76	-0.142
苏庄	64.51	26.83	-0.293	39.97	46.51	-0.085	62.44	14.64	-0.428	50.05	22.30	-0.290
汤河口	46.20	17.42	-0.244	29.80	31.75	-0.006	40.66	8.96	-0.438	33.12	13.52	-0.305
唐指山	68.04	27.86	-0.135	39.99	56.92	0.176	66.11	16.48	-0.301	51.11	28.35	-0.098
桃峪口	63.07	30.04	-0.201	34.03	57.21	0.064	57.45	15.71	-0.370	43.71	25.27	-0.204
通州	61.06	24.82	-0.300	38.45	42.73	-0.096	62.55	11.84	-0.499	52.93	16.91	-0.393
王家园	52.16	27.54	-0.282	26.80	48.30	-0.067	49.71	12.27	-0.467	39.55	18.04	-0.348
温泉	64.72	34.79	-0.205	31.16	66.02	0.058	64.95	17.79	-0.405	49.70	27.68	-0.257
霞云岭	62.93	32.86	-0.219	31.52	61.41	0.034	60.20	16.38	-0.454	46.54	24.36	-0.329
下会	57.18	20.14	-0.227	38.01	37.36	0.022	54.52	11.11	-0.400	44.97	17.37	-0.250
延庆	40.31	13.48	-0.314	28.13	22.90	-0.117	38.08	8.16	-0.401	31.07	12.75	-0.250
雁翅	43.74	17.10	-0.393	28.99	26.90	-0.239	42.11	8.83	-0.500	34.94	12.60	-0.395
沿河城	43.21	20.46	-0.038	21.27	46.59	0.348	44.35	10.72	-0.290	34.53	18.65	-0.079
右安门	62.53	30.22	-0.182	32.96	58.74	0.096	58.67	14.43	-0.387	46.18	22.83	-0.230
榆林庄	58.36	24.96	-0.306	35.70	42.75	-0.104	57.18	12.24	-0.471	47.08	17.92	-0.354
枣树林	76.59	30.87	-0.330	48.96	51.62	-0.143	71.13	17.83	-0.431	56.07	27.08	-0.295
斋堂	40.95	17.48	-0.135	23.34	35.73	0.176	39.91	8.42	-0.367	32.53	13.58	-0.200
张坊	67.33	27.36	-0.305	42.48	46.88	-0.104	63.71	15.37	-0.416	50.61	23.68	-0.273
张家坟	69.12	37.45	-0.100	30.56	79.52	0.237	60.59	19.01	-0.358	43.85	30.93	-0.186
镇罗营	71.62	32.68	-0.166	39.34	64.60	0.123	74.32	17.98	-0.345	58.37	29.61	-0.166

表 3 不同极端降水序列的KS检验结果^{a)}

站点	GEV	GPD	最优分布	站点	GEV	GPD	最优分布	站点	GEV	GPD	最优分布
半壁店	0.063	0.095	GEV	密云	0.062	0.12	GEV	通州	0.048	0.499*	GEV
大村	0.051	0.09	GEV	密云白	0.048	0.216*	GEV	王家园	0.057	0.166*	GEV
大孙各庄	0.043	0.231*	GEV	密云潮	0.043	0.308*	GEV	温泉	0.077	0.076	GPD
番字牌	0.063	0.123	GEV	南各庄	0.047	0.202*	GEV	霞云岭	0.058	0.05	GPD
房山	0.053	0.121	GEV	平谷	0.078	0.256*	GEV	下会	0.059	0.482*	GEV
凤河营	0.047	0.146	GEV	千家店	0.045	0.142	GEV	延庆	0.043	0.244*	GEV
高碑店	0.081	0.062	GPD	三家店	0.062	0.175*	GEV	雁翅	0.079	0.165*	GEV
AM 怀柔	0.053	0.132	GEV	沙河	0.044	0.107	GEV	沿河城	0.045	0.095	GEV
黄村	0.048	0.221*	GEV	十三陵	0.081	0.1	GEV	右安门	0.056	0.054	GPD
黄花城	0.05	0.129	GEV	顺义	0.047	0.192*	GEV	榆林庄	0.04	0.259*	GEV
黄松峪	0.067	0.098	GEV	松林闸	0.04	0.159	GEV	枣树林	0.045	0.068	GEV
喇叭沟门	0.057	0.360*	GEV	苏庄	0.076	0.324*	GEV	斋堂	0.074	0.062	GPD
乐家花园	0.042	0.059	GEV	汤河口	0.046	0.173*	GEV	张坊	0.075	0.505*	GEV
卢沟桥	0.051	0.364*	GEV	唐指山	0.036	0.118	GEV	张家坟	0.104	0.113	GEV
马驹桥	0.045	0.09	GEV	桃峪口	0.059	0.126	GEV	镇罗营	0.055	0.321*	GEV
半壁店	0.086	0.045	GPD	密云	0.073	0.046	GPD	通州	0.061	0.046	GPD
大村	0.1	0.062	GPD	密云白	0.082	0.043	GPD	王家园	0.051	0.104	GEV
大孙各庄	0.075	0.055	GPD	密云潮	0.059	0.021	GPD	温泉	0.098	0.081	GPD
番字牌	0.073	0.055	GPD	南各庄	0.044	0.044	GPD	霞云岭	0.115	0.078	GPD
房山	0.044	0.042	GPD	平谷	0.06	0.086	GEV	下会	0.077	0.041	GPD
凤河营	0.075	0.043	GPD	千家店	0.073	0.037	GPD	延庆	0.063	0.043	GPD
POT 高碑店	0.09	0.052	GPD	三家店	0.078	0.039	GPD	雁翅	0.048	0.068	GEV
怀柔	0.078	0.038	GPD	沙河	0.084	0.046	GPD	沿河城	0.067	0.04	GPD
黄村	0.053	0.052	GPD	十三陵	0.053	0.05	GPD	右安门	0.076	0.056	GPD
黄花城	0.051	0.05	GPD	顺义	0.034	0.058	GEV	榆林庄	0.046	0.055	GEV
黄松峪	0.081	0.046	GPD	松林闸	0.056	0.047	GPD	枣树林	0.067	0.036	GPD
喇叭沟门	0.057	0.052	GPD	苏庄	0.037	0.103	GEV	斋堂	0.066	0.041	GPD
乐家花园	0.058	0.044	GPD	汤河口	0.087	0.052	GPD	张坊	0.047	0.081	GEV
卢沟桥	0.083	0.052	GPD	唐指山	0.07	0.032	GPD	张家坟	0.104	0.081	GPD
马驹桥	0.074	0.053	GPD	桃峪口	0.073	0.039	GPD	镇罗营	0.053	0.048	GPD

a) “*”代表不能通过 $\alpha=0.05$ 的显著性检验

图 4 (网络版彩图)卢沟桥站的理论重新期与经验重新期. (a) AM序列; (b) POT序列

年)的降水极值, 如图5和图6所示. 整体而言, 不同重现期水平的降水量空间分布形态基本保持一致, 与年降水量和汛期降水量的空间分布相类似, 整体表现为从东南向西北方向的递减变化, 并存在多个极值中心, 如怀柔水库和密云水库附近的山前迎风坡、北京市主城区及其下风向方向和西南部山区. 对于AM序列而言(图5), 10年重现期水平下除西部个别站点降水量处于暴雨等级($50 \text{ mm} \leq \text{日降水量} < 100 \text{ mm}$), 绝大部分站点达到大暴雨级别($100 \text{ mm} \leq \text{日降水量} < 250 \text{ mm}$), 20年重现期水平下全部站点降水量属于大暴雨等级, 50

年和100年重现期水平下部分站点达到特大暴雨等级(日降水量 $\geq 250 \text{ mm}$). 对POT序列而言, 10年和20年重现期水平下分别有11个和6个站点的降水量属于暴雨级别, 其余站点为大暴雨级别, 而50年重现期和100年重现期水平下分别有1个和12个站点达到了特大暴雨等级, 其余站点为大暴雨等级.

对比图5和图6可知, AM序列和POT序列的不同重现期水平下的降水量存在明显差异, 在相同重现期水平下AM序列估计的降水量比POT序列估计值偏大. 这种差异随着重现期增大而增大, 但增加幅度相对较

图5 (网络版彩图)AM序列不同重现期水平降水极值分布(单位: mm). (a) 重现期 $T=10$ 年; (b) 重现期 $T=20$ 年; (c) 重现期 $T=50$ 年; (d) 重现期 $T=100$ 年

图 6 (网络版彩图)POT序列不同重现期水平降水极值分布(单位: mm). (a) 重现期 $T=10$ 年; (b) 重现期 $T=20$ 年; (c) 重现期 $T=50$ 年; (d) 重现期 $T=100$ 年

小. 对于重现期为10年和20年时, 平均差值为23.0和26.8 mm, 最大值为40.0和51.5 mm, 而当重现期为50年和100年时, 两个序列平均差值均为29.5 mm, 最大差值分别为66.6和88.4 mm. 同时, 两种序列差值也与站点汛期降水量密切相关, 降水量多的地区差异一般较大, 而降水量小的地区差异相对较小. 两者的差异也主要受样本筛选方法的影响, 对于不同阈值情况下得到的POT序列与AM序列的差异在今后工作中将进一步探讨.

4 结论

采用北京市45个雨量站1960~2012年的汛期逐日降水资料, 采用年最大值法和超门限阈值方法构建降水极值的两种序列, 借助L矩参数估计方法, 利用极值理论方法(GEV分布和GPD分布)分析北京市降水极值的概率分布特征, 得出如下结论.

(1) 采用年最大值法和超门限阈值方法(百分位阈值, 95%)得到的样本序列(AM序列和POT序列)存在较大差异, 根据统计特征分析可知多数站点AM序列

的样本均值大于POT序列的样本均值, AM序列的最小值则小于POT序列, 且AM序列的样本变异性较POT序列更为显著.

(2) 根据KS检验结果, 对AM序列而言, GEV分布的拟合效果较佳, GPD分布的拟合结果较差, 其中存在20个站点未能通过显著性检验, 而对于POT序列, 两种极值分布均可用于降水极值的拟合, 但整体上GPD分布的拟合结果略好. 上述结论也进一步验证了GEV模

型较适用于区块(组)最大值样本的极值拟合, 而GPD模型则可较好地拟合超门限阈值样本序列.

(3) 两种序列得到的不同重现期水平的降水量空间分布基本相似, 主要受地形、城市化发展等方面因素的影响, 总体上利用AM序列推求不同重现期水平下的降水量略大于POT序列推求结果, 从风险角度分析, 基于AM序列推算一定重现期水平下的降水极值更可靠, 更利于保证工程设计安全.

参考文献

- Grimm N B, Faeth S H, Golubiewski N E, et al. Global change and the ecology of cities. *Science*, 2008, 319: 756–760
- 张建云, 宋晓猛, 王国庆, 等. 变化环境下城市水文学的发展与挑战: I.城市水文效应. *水科学进展*, 2014, 25: 594–605
- Hallegatte S, Green C, Nicholls R J, et al. Future flood losses in major coastal cities. *Nat Clim Change*, 2013, 3: 802–806
- 宋晓猛, 张建云, 占车生, 等. 气候变化和人类活动对水文循环影响研究进展. *水利学报*, 2013, 44: 779–790
- 吴孝情, 陈晓宏, 唐亦汉, 等. 珠江流域非平稳性降雨极值时空变化特征及其成因. *水利学报*, 2015, 46: 1055–1063
- 宋晓猛, 张建云, 孔凡哲, 等. 北京地区降水极值时空演变特征. *水科学进展*, 2017, 31: 161–173
- 宋晓猛, 张建云, 王国庆, 等. 变化环境下城市水文学的发展与挑战: II.城市雨洪模拟与管理. *水科学进展*, 2014, 25: 752–765
- Mishra V, Ganguly A R, Nijssen B, et al. Changes in observed climate extremes in global urban areas. *Environ Res Lett*, 2015, 10: 024005
- Ali H, Mishra V, Pai D S. Observed and projected urban extreme rainfall events in India. *J Geophys Res Atmos*, 2014, 119: 12621–12641
- Han L, Xu Y, Pan G, et al. Changing properties of precipitation extremes in the urban areas, Yangtze River Delta, China, during 1957–2013. *Nat Hazards*, 2015, 79: 437–454
- Wahl T, Jain S, Bender J, et al. Increasing risk of compound flooding from storm surge and rainfall for major US cities. *Nat Clim Change*, 2015, 5: 1093–1097
- Min S K, Zhang X, Zwiers F W, et al. Human contribution to more-intense precipitation extremes. *Nature*, 2011, 470: 378–381
- Miao S, Chen F, Li Q, et al. Impacts of urban processes and urbanization on summer precipitation: A case study of heavy rainfall in Beijing on 1 August 2006. *J Appl Meteor Climatol*, 2011, 50: 806–825
- Wang D, Jiang P, Wang G, et al. Urban extent enhances extreme precipitation over the Pearl River Delta, China. *Atmos Sci Lett*, 2015, 16: 310–317
- 张建云, 王银堂, 贺瑞敏, 等. 中国城市洪涝问题及成因分析. *水科学进展*, 2016, 27: 485–491
- Song X, Zhang J, AghaKouchak A, et al. Rapid urbanization and changes in spatiotemporal characteristics of precipitation in Beijing metropolitan area. *J Geophys Res Atmos*, 2014, 119: 11250–11271
- Jin M, Li Y, Su D. Urban-induced mechanisms for an extreme rainfall event in Beijing China: A satellite perspective. *Climate*, 2015, 3: 193–209
- 宋晓猛, 张建云, 刘九夫, 等. 北京地区降水结构时空演变特征. *水利学报*, 2015, 46: 525–535
- Cunnane C. Unbiased plotting positions—A review. *J Hydrol*, 1978, 37: 205–222
- 蔡敏. L-矩估计方法在极端降水研究中的应用. *气象科学*, 2007, 27: 597–603
- 余敦先, 夏军, 张永勇, 等. 近50年来淮河流域极端降水的时空变化及统计特征. *地理学报*, 2011, 66: 1200–1210
- 何干皓, 李龙国, 刘铁刚, 等. 岷江流域降水极值概率分布研究. *工程科学与技术*, 2017, 49: 78–85

Probability distribution of extreme precipitation in Beijing based on extreme value theory

SONG XiaoMeng^{1,2,3}, ZHANG JianYun^{2,4} & KONG FanZhe¹

¹ School of Resources and Geosciences, China University of Mining & Technology, Xuzhou 221116, China;

² State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China;

³ State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, China;

⁴ Research Center for Climate Change, Ministry of Water Resources, Nanjing 210029, China

In this study, the statistical characteristics of the probability distribution in precipitation extremes in Beijing were explored and discussed based on the extreme value theory. The extreme precipitation series were filtered from the daily precipitation data (45 rain gauge stations) in the warm season during 1960–2012 based on the annual maximum (AM) and peak-over-threshold (POT) method. The generalized extreme distribution (GEV) and generalized Pareto distribution (GPD) were used to fit the two precipitation series and to investigate their statistic characteristics. The parameters of GEV and GPD distribution were estimated by the L-moment method. The best distribution was determined by the Kolmogorov-Smirnov test and then the best distribution was used to estimate the precipitation extremes in different return periods. The results showed that the mean values of AM series for all the rain gauges apart from Banbidian and Yanhecheng were larger than that of POT series, whereas the minimum values of AM series were lower than that of POT series. Furthermore, the optimal distribution of 40 sites for the AM series were GEV distribution, and the best distribution 38 sites for the POT series were GPD distribution, based on the results of the Kolmogorov-Smirnov test. Finally, the distributions of precipitation extremes in different return periods for this two series were similar, with the high value in the south and east and the low value in the north and west. Overall, the estimated precipitation amounts by the AM series were larger than that of POT series in the different return periods, which means the precipitation extremes based on the AM series will be relatively safer when it was used in the engineering design.

extreme precipitation, annual maximum method, peak-over-threshold method, probability distribution, Beijing

doi: [10.1360/N092017-00129](https://doi.org/10.1360/N092017-00129)