

基于线性矩的粤港澳大湾区极端降水区域频率分析*

杨芳^{1,2}, 马志鹏¹, 陈子燊³

1. 珠江水利科学研究院 / 水利部粤港澳大湾区水安全保障重点实验室, 广东 广州 510611
2. 清华大学水沙科学与水利水电工程国家重点实验室, 北京 100084
3. 中山大学地理科学与规划学院, 广东 广州 510275

摘要: 以1961—2023年粤港澳大湾区30个气象观测站年最大日降水数据为基础数据, 应用区域线性矩法分析了大湾区各子区域极端降水频率的最优分布函数与重现设计水平。研究表明: 1) 使用KNN分类法和均质性判别准则将大湾区气象站点分成5个子区域, 分别为位于西江中下游的I区、位于中北部三角洲网河区的II区、位于南部伶仃洋河口区两岸的III区、位于西南部的IV区、位于东部的V区; 2) I~V子区域的最优分布线型依次为P-III、GEV、GEV、P-III、GEV。依据区域暴雨增长曲线的方法, 分别推算了大湾区各站点的暴雨频率设计值。3) 由优选的5个分区的区域最优分布函数计算的暴雨重现水平值可作为子区域内无站点处工程设计的参考依据。

关键词: L-矩; 区域极端降水; 频率分析; KNN分类; 区域最优分布函数

中图分类号: TV122 **文献标志码:** A **文章编号:** 2097-0137(2025)02-0012-10

Linear moments-based regional frequency analysis of extreme precipitation in Guangdong-Hong Kong-Macao Greater Bay Area

YANG Fang^{1,2}, MA Zhipeng¹, CHEN Zishen³

1. Pearl River Hydraulic Research Institute / Key Laboratory of Water Security, Guangdong-Hong Kong-Macao Greater Bay Area, Ministry of Water Resources, Guangzhou 510611, China
2. State Key Laboratory of Water and Sediment Science and Water Conservancy and Hydropower Engineering, Tsinghua University, Beijing 100084, China
3. School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China

Abstract: Based on the annual maximum daily precipitation data collected from 30 meteorological observation stations in the Guangdong-Hong Kong-Macao Greater Bay Area between 1961 and 2023, the optimal distribution function and recurrence design level of extreme precipitation frequency in each sub-region of the Greater Bay Area were analyzed by using the regional linear moment method. The findings are summarized as follows: 1) The meteorological stations in the Greater Bay Area are divided into five sub-regions by KNN classification and homogeneity criterion. These sub-regions are: Region I, located in the middle and lower reaches of the Xijiang River; Region II, located in the Delta network River Area in the central and northern part of the Greater Bay Area; Region III, located on both sides of the Lingdingyang Estuary Area in the southern part of the Greater Bay Area; Region IV,

* 收稿日期: 2024-08-19 录用日期: 2024-10-12 网络首发日期: 2025-01-22

基金项目: 国家重点研发计划(2021YFC3001000); 水利青年人才发展资助项目;
国家自然科学基金(41371498)

作者简介: 杨芳(1978年生), 女; 研究方向: 河口治理与洪涝灾害防御; E-mail: zkykjc@163.com
通信作者: 马志鹏(1979年生), 男; 研究方向: 水资源调度与配置; E-mail: zhipengma@163.com

全文阅读



ZR20240260

located in the southwestern part of the Greater Bay Area; and Region V, located in the eastern part of the Greater Bay Area. 2) The optimal distribution functions of the 1st to 5th sub-regions are P-III, GEV, GEV, P-III, and GEV, respectively. Consequently, the rainstorm frequency design values of each station in the Greater Bay Area are calculated using the regional rainstorm growth curve method. 3) The rainstorm recurrence level determined by the optimal distribution functions for the five sub-regions can serve as a reference for engineering design in sub-regions that lack specific site data.

Key words: *L*-moment; regional extreme precipitation; frequency analysis; KNN classification; regional optimal distribution functions

粤港澳大湾区是我国政府推动形成全面开放新格局的重要战略,在建设具有全球影响力的国际科技创新中心和世界级城市群,包括香港特别行政区、澳门特别行政区和广东省的广州、深圳、珠海等九个地市级城市,总面积约5.6万 km^2 ,常住人口超过7000万,是中国经济活力最强、开放程度最高的区域之一。然而,粤港澳大湾区的降雨受到季风气候、台风活动、城市化进程以及海平面上升(Sweet,2014)等多重因素的影响,表现出复杂多变的气候特征。尤其是近年来频繁发生的洪涝灾害已造成了大量的经济损失和人员伤亡,如何预警及防范灾害的发生,并减少灾害造成的损失越来越成为人们所关注的焦点。而洪涝灾害的发生根源于极端降水,高强度的降水不但可以对建筑及工程设施造成很大程度的破坏,其引发的大洪水将使农业、城市、生态遭受持续的破坏和影响,因此要使灾害造成的损失减少到最低,须先对极端降水事件的发生进行风险防范与管理。Haylock(2000)认为,虽然极端降水事件在数值上没有稳定性可言,但其在时空域上出现的概率却具有某种相对稳定性。郭军等(2010)利用相应的概率分布函数来模拟极端降水事件极值的重现期与设计值,评估未来可能遭遇到风险的几率并进行防范决策。

选取恰当的极值分布函数来估算分位值需要长时间序列的观测数据,而实际上由于观测数据的短缺,采用单站点进行频率分析的准确性会受到很大影响。Kite(1988)和Guttman(1993)使用区域频率分析方法把区域作为一个整体来分析,避免了单站点观测数据的限制,从而提高了分析的可靠性与准确性。Zrinji et al.(1994)把区域频率分析结果应用于区域中无资料的地区。在总结Dalrymple(1960)、Cunnane(1988)和Stediger(1993)研究的基础上,Hosking(1993,1997)提出了基于线性矩的区域频率分析方法。Pariada(1998),

Adamowski(2000)等许多学者相继将此方法应用于洪水、枯水及降水的频率分析中,这些研究成果充分验证了这一方法的优越性。国内也陆续开展区域频率分析研究(陈元芳,2002;陈永勤等,2005; Xu et al., 2010; Yang et al., 2010a, b; Shi et al., 2010)。这些研究大多针对洪水频率,对洪涝灾害产生的源头极端降水的研究还不多。

值得提出的是,美国降雨强度的频率估计值已被土木工程师广泛应用于民用工程和大型水利工程的设计以及区域防洪规划中,用来确定都市防洪排水系统和大坝及其溢洪道的尺寸。目前,降雨强度的频率估计值的应用范围已逐渐扩大到环境生态的保护和管理工作中。美国国家海洋大气管理总署(NOAA)下属的水文局(OHD)于2006年提出了基于区域线性矩法进行暴雨频率分析的一套完整的应用系统(Lin,2006),并陆续分区分片对美国1960年代初期编制的降雨频率图集进行更新,其估算成果已经被作为国家防洪标准,应用于联邦、州、县各级政府的防洪工程设计和地区防洪规划中(Bonnin,2003,2004)。

本研究拟基于线性矩的区域频率分析方法对大湾区极端降水频率加以分析,希冀能为大湾区的水资源管理和地区防洪排涝工程设计提供参考。

1 线性矩与区域频率分析方法

1.1 线性矩

线性矩(简称*L*-矩,linear moment)源于Greenwood et al.(1979)提出的概率权重矩(PWM),Hosking(1986)将由小到大排序的序列值进行一定的线性组合定义了线性矩。与常规矩法相比,*L*-矩最大的特点是对极值事件系列中的极大值和极小值远没有常规矩那么敏感,其求得的频率曲线参数的估计值比常规矩所求得的要稳健得多,目前已经成为英、美等国洪水频率分析计算中的主要应用方法。

Hosking(1986)定义的线性矩实质是次序统计量 X 的线性组合期望值, 其 r 阶线性矩定义为

$$\lambda_r = r^{-1} \sum_{k=0}^n (-1)^k \binom{k}{r-1} E(X)_{r-k},$$

其中 $E(X)_{r-k}$ 为样本量为 r 的 $(r-k)$ 阶次序统计量的期望值。样本量为 n 的 r 阶次序统计量的期望值 $E X_{r:n}$ 的计算式为

$$E(X)_{r:n} = \frac{n!}{(r-1)!(n-r)!} \times \int_0^1 X [F(X)]^{r-1} [1-F(X)]^{n-r} dF(X).$$

作为总体线性矩的估计, r 阶样本线性矩

$$\hat{\lambda}_r = \binom{r}{n}^{-1} \sum_{1 \leq i_1 < \dots < i_r \leq n} r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} x_{i_{r-k}},$$

式中 $x_{1:n} \leq x_{2:n} \leq \dots \leq x_{n:n}$ 为样本次序统计量。与常规矩相类似, 线性矩统计特征系数定义: L -均值 (L -mean): λ_1 ; L -变差系数 (L - C_v): $\tau_2 = \lambda_2/\lambda_1$; L -偏态系数 (L - C_s): $\tau_3 = \lambda_3/\lambda_2$; L -峰度系数 (L - C_k): $\tau_4 = \lambda_4/\lambda_2$ 。

含有 N 个测站的区域 L -矩的特征参数通过对单站 L -矩特征参数的加权平均获得:

$$\hat{\tau}_r^R = \frac{1}{n_R} \sum_{i=1}^N n_i \hat{\tau}_r^{(i)}, \quad r = 2, 3, \dots$$

1.2 区域频率分析方法

一致的观测数据、均质相似的区域、恰当的分函数以及分位值估算的准确性是区域频率分析结果可靠的前提, 因此区域频率分析至少应包括 5 个步骤: ①数据一致性检验; ②区域均质性检验; ③选择区域分函数; ④利用区域频率曲线估算分位值; ⑤结果可靠性分析。

区域分函数采用广义正态分布(GNO)、广义极值分布(GEV)、广义逻辑斯特分布(GLO)、皮尔逊三型分布(PE3)和广义帕雷托分布(GPA)共 5 种 3 参数分函数作为备选分布, 对各个分区域分别计算 5 种备选分函数的 $|Z|$ 统计值加以优选。

1.2.1 区域不一致性测度 数据检测的目的是发现数据是否和该区域其他测站特征不一致, 即是否存在离群点(outlier)。这种不一致性可以根据各个站点的特征因子来求得。这些特征因子可以是对极端降水形成具有直接影响的一些区域物理特性, 或是反映极端降水规律的样本系列统计参数。

采用多变量离群点验证法计算区域不一致性测度, 具体方法为:

设区域内有 N 个测站, 各站 P 个统计参量 S_i , 以向量表示为

$$\mathbf{u}_i = [S_1^{(i)}, S_2^{(i)}, \dots, S_p^{(i)}]^T, \quad i = 1, 2, \dots, N,$$

其构成了三维空间的 N 个点。设此 N 个点独立来自同一个多变量的正态分布, 则使用 Wilk 统计分布检验是否存在离群点时, 有零假设 H_0 : \mathbf{u}_i 服从于多变量正态分布, 如果

$$\max D_i > \frac{(N-1)Z}{N-p-1+pZ},$$

则拒绝零假设。式中 Z 为 F 分布的 μ/N 的上侧分位数; $N-p-1$ 为自由度。各测站的不一致性测度 D_i 为

$$D_i = \frac{1}{3} N (\mathbf{u}_i - \bar{\mathbf{u}})^T A^{-1} (\mathbf{u}_i - \bar{\mathbf{u}}),$$

式中 $A = \sum_{i=1}^N (\bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}})^T$ 。如以 3 个线性矩 ($L-C_v, L-C_s, L-C_k$) 作为特征因子, 每个站点的样本线性矩 $L-C_v, L-C_s, L-C_k$ 可以构成一个三维向量, 记为

$$\mathbf{u}_i = [L-C_v^{(i)}, L-C_s^{(i)}, L-C_k^{(i)}]^T$$

在三维空间把 \mathbf{u}_i 当作一个点。 N 个站点则有 N 个三维向量在三维空间中构成一个点簇, 假如有个别点距离点簇的中心比较远, 可认为这个点和其他点不一致, 是离群点。记点簇的中心为 $\bar{\mathbf{u}}$,

$$\bar{\mathbf{u}} = \frac{1}{N} \sum_{i=1}^N \mathbf{u}_i.$$

Hosking et al.(1997)给出了不同 N 值情况下的临界值, 该值是站点个数 N 的函数, 随 N 增大而增大。如果 D_i 大于某一临界值, 就可判别第 i 个站点和其他点不一致, 属于离群点应予以剔除。

1.2.2 区域均质性检验 假定待检验区域内的观测站点数为 N , 第 i 个站点的样本量为 n_i , 则以样本长度为权重的区域加权平均样本线性矩系数的计算公式分别为

$$\tau_2^R = \sum_{i=1}^N n_i \tau_2^i / \sum_{i=1}^N n_i,$$

$$\tau_3^R = \sum_{i=1}^N n_i \tau_3^i / \sum_{i=1}^N n_i,$$

$$\tau_4^R = \sum_{i=1}^N n_i \tau_4^i / \sum_{i=1}^N n_i.$$

对 N 个站点进行蒙特卡洛机模拟, 每个站点均由 4 参数 Kappa 分布模拟产生与其观测样本量相等数量的模拟样本, 模拟重复次数 $N_{sim}=500$ 。根据各个站点模拟样本计算的区域平均值为 u_v , 标准差为 σ_v , 则区域均质性检验的统计量

$$H = (V - u_v) / \sigma_v.$$

根据不同的线性矩系数, 可以计算出 3 种不同

的 V 统计量, 从而有 3 种区域均质性检验统计量 H_1 、 H_2 、 H_3 。

$$V = \left\{ \frac{\sum_{i=1}^N [n_i (\tau_2^i - \tau_2^R)^2]}{\sum_{i=1}^N n_i} \right\}^{0.5},$$

$$V_2 = \frac{\sum_{i=1}^N \left\{ n_i \left[(\tau_2^i - \tau_2^R)^2 + (\tau_3^i - \tau_3^R)^2 \right] \right\}^{0.5}}{\sum_{i=1}^N n_i},$$

$$V_3 = \frac{\sum_{i=1}^N \left\{ n_i \left[(\tau_3^i - \tau_3^R)^2 + (\tau_4^i - \tau_4^R)^2 \right] \right\}^{0.5}}{\sum_{i=1}^N n_i}.$$

若统计量 $H < 1$, 则可认为该区域为均质区域; 若 $1 \leq H < 2$, 该区域可能为非均质区域; $H \geq 2$, 为非均质区域。

研究发现, 基于 V_2 和 V_3 算出来的非均质性测度 H 的区别能力没有基于 V 的非均质性测度 H 强。虽然基于 V 的 H 对样本线性矩变差系数均匀而样本线性矩偏态系数不均匀的水文分区不敏感, 不过这种不敏感对设计暴雨精度的影响并不大。Hosking(1990)推荐使用基于 V 的 H 。

1.2.3 区域分布函数选择 Hosking et al.(1993) 提出以 L -峰度系数定义的 Z 统计量来选择区域分布函数, 计算公式为

$$Z = (\tau_4^{\text{Dist}} - \tau_4^R + \beta_4) / \sigma_4,$$

$$\beta_4 = \frac{1}{N_{\text{sim}}} \sum_{m=1}^{N_{\text{sim}}} (\tau_4^{R(m)} - \tau_4^R),$$

$$\sigma_4 = \left\{ \frac{1}{N_{\text{sim}} - 1} \left[\sum_{m=1}^{N_{\text{sim}}} (\tau_4^{R(m)} - \tau_4^R)^2 - N_{\text{sim}} \beta_4^2 \right] \right\}^{0.5},$$

式中 τ_4^{Dist} 为备选分布函数的 L -峰度系数, β_4 和 σ_4 分别为蒙特卡洛随机模拟法模拟的样本与实测样本计算的区域平均 L -峰度系数的偏差和标准差。90% 置信度的 Z 临界值为 ± 1.64 , 即若备选分布函数的 $|Z| \leq 1.64$, 则该分布函数适合用作该区域的分布函数。

1.2.4 区域频率增长曲线 为了克服单站点估算分位值时由于样本系列资料短缺造成的结果不准确性, Dalrymple(1960)提出了标度洪水方法, 后人做了大量研究, 发现利用标度洪水方法估算各站点分位值相比单站点估算有更高的准确性, 并可以推广到降水和枯水的频率分析中。

区域频率分析的前提假设是区域内的频率分布与线型是完全一样的, 当这些不同站点的极端降水系列具有“水文相似性”后, 便可认为这些“水文相似”站点的极端降水系列都服从同样的总体分布, 可以采用区域频率分析方法来研究每个站点极端降水频率曲线。最常用的区域频率分析方法是各站点的样本长度作权重求得各阶 L -矩

的区域均值, 其计算步骤如下:

首先对每个站点的极端降水系列无因次化, 生成系列

$$q_{i,j} = Q_{i,j} / Q_i, \quad j = 1, 2, \dots, n; \quad i = 1, 2, \dots, N.$$

对每个站点, 分别估计出该站 q 系列的前三阶样本 L -矩, 即 L_1 、 L_2 和 L_3 。这些站点的前三阶样本 L -矩加权平均, 得到区域 L -矩

$$L_k^R = \frac{\sum_{i=1}^N n_i L_k^{(i)}}{\sum_{i=1}^N n_i}, \quad k = 1, 2, 3 \dots$$

根据区域 L -矩, 则可以估计出待使用的概率分布函数参数的区域估计值: L -均值、 L -变差系数和 L -偏态系数。定义极端降水频率设计值

$$Q_{\tau,i,j} = q_{\tau,i,j} \cdot \bar{x}_{i,j},$$

式中重现期 $T=2, 5, 10, 20, 50, 100, 200, 500$ a; 地区 $i=1, 2, \dots, N$; 站点 $j=1, 2, 3, \dots, K$ 。 $q_{\tau,i,j}$ 为一致性区内反映该区域共有的降雨特性的区域分量的频率因子, 或称区域无量纲频率分布曲线, 或称区域增长曲线, 简称区域频率因子。 $\bar{x}_{i,j}$ 为第 i 区内第 j 站点的多年降雨量平均值, 即本地或当地分量因子, 可选用当地(站)多年的年最大日降雨量的平均值作为当地因子, 多年均值可以认为是反映了某地长期气候与地形、地貌共同作用独有的特性。可得

$$q_{\tau,i} = Q_{\tau,i,j} / \bar{x}_{i,j}.$$

站点的历史雨量系列经过去均值化后可认为是反映区域共性的降雨分量。应用水文气象的区域分析法可大大降低频率设计值的不确定性, 使推求各站点的频率设计值更加可靠。

2 结果分析

2.1 日最大降雨量趋势性检验

对 1961—2023 年大湾区 30 个气象站逐日降水数据抽取的各站历年日最大降水值构成的时间序列数据, 使用 M-K 趋势检测方法, 检测结果反映所有测站都不存在显著的趋势变化, 可用于后续进一步分析。

2.2 区域站点一致性检验

根据 Hosking et al.(1997) 提出的站点不一致性检验临界值, 对大湾区 30 个站点做一致性检验, 结果如表 1 所示, 30 个站点中斗门站 D 统计量超过 3, 应作为离群点排除, 其余 29 个站点的样本可用于大湾区区域极端降水暴雨频率分析。

2.3 大湾区区域均质性检验

采用 Kappa 分布模拟检验排除斗门站后大湾区

29个站点数据的均质性,结果显示,3个指标 H_1 、 H_2 、 H_3 统计量分别为5.226、3.376、2.458,都大于2,大湾区整体上属于非均质区域,需要分成若干个子区域进行区域极端降水频率分析。

2.4 子区域划分

基本数据选用29个站点的经度、纬度和暴雨系列的变差系数 $L-1$ 、偏态系数 $L-2$ 、峰度系数 $L-3$ 作为雨量站点分类的影响因子。采用KNN法加以分类,其基本原理为:KNN(K-nearest neighbors)算法是传统机器学习中的有监督学习方法之一,用于解决分类和回归问题。分类的基本原理是通过计算待分类样本与已知类别样本集中的每个样本的距离,选取距离最近的前 k 个样本,根据这 k 个“邻居”的类别通过多数表决或加权平均等策略来决定待分类样本的类别。KNN因其简单直观且无须训练过程而在模式识别和机器学习中被广

泛应用。相关研究认为KNN算法用于气象资料分析时具有2个优势:①KNN算法是一种非参数学习方法,不需要事先对数据进行严格的假设和模型选择,能够对多种数据类型进行处理,特别适用于处理特征数和样本数相对较小、存在少量噪声和异常值(不占主要比例),以及样本之间存在明显距离或相似性关系的数据集;②KNN算法的可解释性强,分类结果可直接解释为最近邻的样本所代表的特征,便于数据分析。

大湾区气象站点极端降水的KNN分类结果与分区见表2、图1和图2。将大湾区划分5个子区域:Ⅰ西北部区,Ⅱ北部区,Ⅲ西南部区,Ⅳ南部区,Ⅴ东部区。依KNN分类划分结果客观且符合站点所在地理区位,分区Ⅰ、Ⅳ和Ⅴ归属行政区清楚,利于未来对子区域内暴雨引发的洪涝风险管控与应急响应。

表1 粤港澳大湾区30个站点不一致性检验结果

Table 1 Discordancy test results of 30 sites in the Greater Bay Area

站名	封开	德庆	怀集	广宁	四会	高要	三水	花都	从化	广州
D 统计量	0.827	0.478	1.309	0.716	0.367	0.517	1.589	0.009	1.216	0.493
站名	南海	东莞	龙门	增城	博罗	惠阳	鹤山	开平	新会	恩平
D 统计量	1.022	0.379	1.175	1.249	0.457	1.499	0.433	0.959	2.189	0.208
站名	台山	顺德	番禺	中山	斗门	珠海	惠东	深圳	香港	澳门
D 统计量	0.575	0.264	1.176	0.422	4.105	0.908	2.200	0.134	1.589	0.533

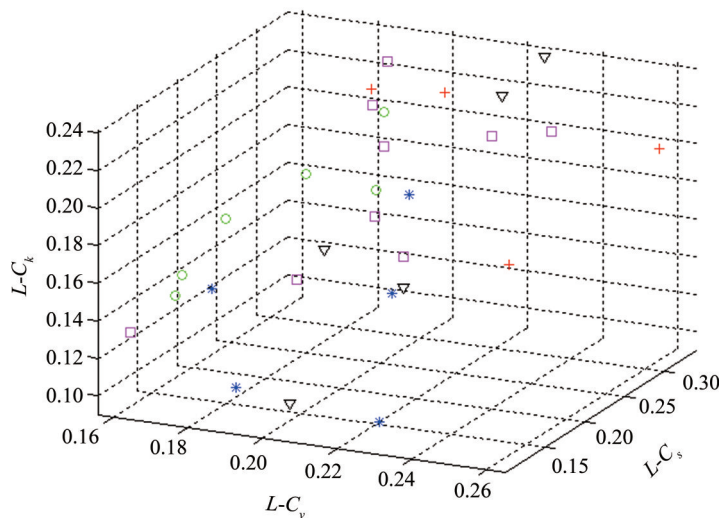


图1 KNN分类图

Fig. 1 KNN classification diagram

图1中的相同符号表示属于同一类,其站点位置和名称根据经纬度坐标确定,见图2。分区Ⅰ位于大湾区西江中下游区,由肇庆市辖区的6个站点组

成;分区Ⅱ位于中北部,由三角洲网河区9个站点组成;分区Ⅲ位于南部,由伶仃河口湾两岸的5个站点组成;Ⅳ区位于西南部,由属于江门市的5个

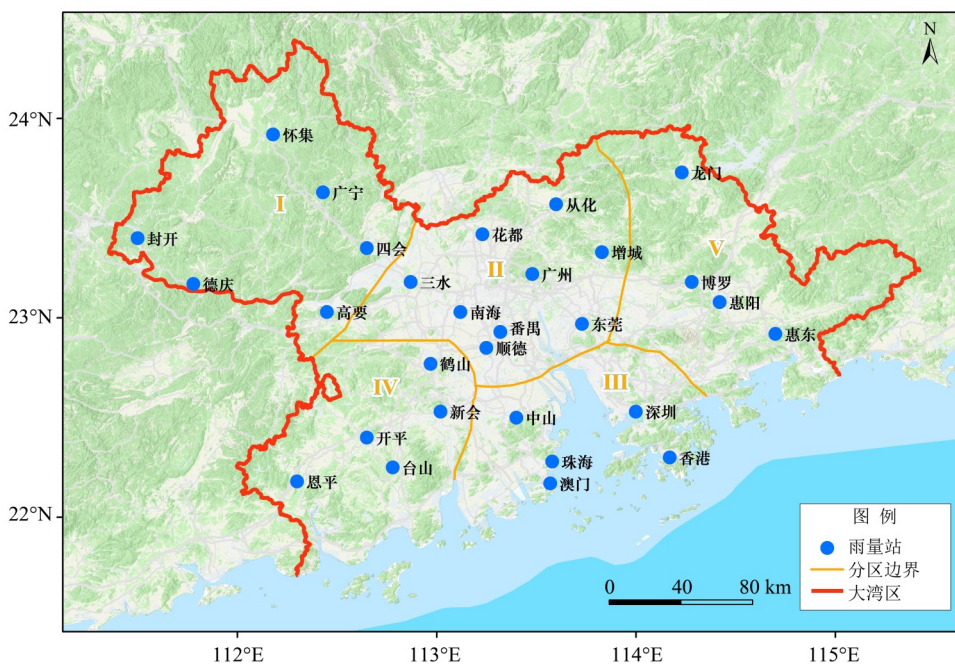


图 2 粤港澳大湾区 5 个子区域分布图

Fig. 2 Distribution diagram of the five sub-areas of the Greater Bay Area

表 2 站点 KNN 分类结果

Table 2 KNN classification results of sites

封开	德庆	怀集	广宁	四会	高要	三水	花都	从化	广州	南海	东莞	龙门	增城	博罗
1	1	1	1	1	1	2	2	2	2	2	2	5	2	5
惠阳	鹤山	开平	新会	恩平	台山	顺德	番禺	中山	珠海	惠东	深圳	香港	澳门	
5	4	4	4	4	4	2	2	3	3	5	3	3	3	

站点组成; V区位于东部, 由惠州市辖区的 4 个站点组成。

2.5 分区最优区域分布函数与暴雨重现水平

使用广义正态分布(GNO)、广义极值分布(GEV)、广义逻辑斯特分布(GLO)、皮尔逊三型分布(PE3)和广义帕雷托分布(GPA)5种3参数分布函数作为备选分布, 分别计算了5个分区的区域平均L-偏度系数和L-峰度系数及其平均值, 由此可得最优分布区域平均L-偏度系数和L-峰度系数分布图(图略)。

表 3 显示分区 I 和 IV 的均值与 PE3 分布关系密切; 分区 II、III、V 的均值与 GEV 分布密切。5 个备选分布的 5 个子区域中所有站点的 L-偏度系数和 L-峰度系数平均值也与 GEV 分布密切(图略)。表 3 区域均质性检验结果显示, 5 个子区域如统一按 H_1 统计量临界值标准, 大湾区西北部区、大湾区中北部区、大湾区西南部区和大湾区东部区为均质区域; 南部区 H_2 、 H_3 小于均质区域临界值 1, 分区包括香港和澳门在内, 也可视为均质区域用于后续的计算分析。

表 3 粤港澳大湾区 5 个子区域的均质性检验

Table 3 Homogeneity test of the five sub-areas

分区	I	II	III	IV	V
H_1	0.198	0.606	-0.814	0.663	0.103
H_2	-0.408	1.206	0.951	-0.869	-1.372
H_3	-0.676	1.695	0.883	-1.779	-1.374

采用 GNO、GEV、GLO、PE3 和 GPA 分布函数作为备选分布, 分别对各个子区域计算 5 种备选分布函数的 $|Z|$ 统计值, 结果如表 4 所示。分区 I: PE3 分布模拟得到的 Z 值最小, 选用为最优区域分布函数; 分区 II 和 III: 5 个备选分布函数的 $|Z|$ 统计

值都以 GEV 分布最小, 选作为分区的最优区域分布函数; 分区 IV: PE3 分布的 Z 值排列第二, 考虑到国家规范推荐及国内使用情况, 选用 PE3 分布为此子区域分布函数; 分区 V: GEV 分布的 $|Z|$ 最小, 可作为此区的最优区域分布函数。

表 4 各分区的 5 种备选分布函数的 $|Z|$ 统计值¹⁾Table 4 $|Z|$ statistics for the five alternative distribution functions for each partition

区域/ $ Z $ 值备选分布	I	II	III	IV	V
GNO	1.474 3	2.072 9	0.853 0	0.233 9	0.647 9
GEV	0.780 8	0.150 9*	0.326 7*	0.606 0	0.158 6*
GLO	2.725 0	1.184 7	0.890 6	2.398 4	0.852 7
PE3	0.761 7*	3.595 0	1.807 7	0.577 7*	1.521 8
GPA	4.573 3	4.042 3	3.303 2	3.451 7	2.672 0

1)*为 $|Z|$ 值在 90% 置信度临界值之内选用的最优分布函数。

由最优区域分布函数推算的降水增长曲线和各站点的重现水平设计值见表 5。由表 5 可知, 区域增长曲线推算的重现设计值有以下特点:

1) I 区偏于大湾区西北部, 受陆域地形掩护与制约, 海洋热带低压系统作用趋弱, 由区域增长曲线推算的重现水平设计值低于其余分区。

2) II 区的区域增长曲线推算的重现设计值与广州站的设计值基本一致。

3) III 区的区域增长曲线推算的重现设计值与澳门站观测数据推算的设计值基本相同, 由此由广州站和澳门站推算的暴雨设计值可代表 II 和 III 区的区域暴雨设计值。

4) 作为 IV 区的最优分布函数的 PE3, 区域增长曲线推算的小频率部分的设计值暴雨设计值会明显小于 GEV 和 GLO 分布函数推算的设计值, 此或会影响水利水务工程设计安全与投入时的抉择。

5) V 区的区域增长曲线推算的重现设计值小于惠东站, 大于其余 3 个站的设计值, 由此反映惠东受海洋低压天气系统的影响频率高。

6) 大湾区整体空间上显示各分区区域设计值由南向北趋小。

为了直观显示由各个子区域增长曲线计算的重现水平值的空间分布, 绘制了重现期 500、200、100、50、20 和 10 a 极端降水设计值, 结果见图 3。由图 3 可见, 大湾区暴雨强度分布呈现东南向西北空间变化特征。其中, 位于大湾区南部区的伶仃河口湾两侧为暴雨高强度区域, 或与此区域更易遭受海域台风天气系统影响有关。

根据蒙特卡洛随机模拟得到的样本与实测样本计算的分位值之间的相关均方根误差 (RMSE) 是一个有效评价结果可靠性的指标, 一般若 RMSE 超过 0.1 阈值则可认为结果的可靠性较低。区域频率曲线的相关均方根误差的计算式为:

$$RMSE^R(T) = \left\{ \frac{1}{N_{sim}} \sum_{m=1}^{N_{sim}} \left[\frac{\hat{p}^{R(m)}(T) - \hat{p}^R(T)}{\hat{p}^R(T)} \right]^2 \right\}^{0.5}$$

优选的 5 个分区区域分布函数计算的 RMSE 分别为 0.020 6、0.023 8、0.022 9、0.021 3、0.037 4, 都小于 0.1, 满足区域频率分析的精度要求。

3 结 论

以 1961—2023 年粤港澳大湾区 30 个气象观测站年最大日降水数据为实证对象, 应用区域线性矩法全面分析了该区各子区域极端降水频率的最优分布函数与重现设计水平, 得到如下主要结论:

1) 采用数据趋势检验和一致性检验后的大湾区 29 个气象站点, 由 KNN 分类法和均质性判别准则可将大湾区分为 5 个子区域: 位于西江中下游的 I 区、中北部三角洲网河区的 II 区、南部伶仃洋河口区两岸的 III 区、西南部的 IV 区和东部的 V 区。

2) 5 个子区域的最优分布线型依次为: PE3、GEV、GEV、PE3、GEV。由优选的区域最优分布函数模拟得到的样本与实测样本计算的分位值之间的均方根误差都小于 0.1, 满足精度要求。

3) 各子区域增长曲线计算的暴雨重现水平值可作为子区域内无站点处工程设计的参考依据。

表5 各子区域增长曲线计算的站点的暴雨重现水平值

Table 5 The rainstorm recurrence level values of the stations calculated by the growth curves of each subregion

子区域	T/a	500	200	100	50	20	10	5	2
I	区域增长曲线	2.628	2.364	2.163	1.958	1.682	1.467	1.242	0.917
	封开	253.9	228.5	209.0	189.2	162.6	141.8	120.1	88.6
	德庆	273.9	246.5	225.5	204.2	175.4	152.9	129.5	95.6
	怀集	273.7	246.3	225.3	204.0	175.2	152.8	129.4	95.5
	广宁	279.3	251.3	229.9	208.1	178.8	155.9	132.1	97.4
	四会	313.5	282.1	258.0	233.6	200.7	175.0	148.2	109.4
	高要	295.6	266.0	243.3	220.3	189.3	165.0	139.8	103.1
	区域	281.7	253.4	231.8	209.9	180.3	157.2	133.2	98.3
II	区域增长曲线	3.656	3.003	2.577	2.201	1.769	1.483	1.222	0.890
	三水	452.8	372.0	319.3	272.7	219.2	183.7	151.4	110.3
	花都	480.4	394.7	338.7	289.3	232.5	194.8	160.6	117.0
	从化	464.4	381.5	327.4	279.6	224.7	188.3	155.2	113.1
	广州	481.8	395.8	339.7	290.1	233.2	195.4	161.0	117.3
	南海	460.2	378.1	324.4	277.1	222.7	186.7	153.8	112.1
	东莞	534.0	438.7	376.5	321.6	258.5	216.6	178.5	130.0
	增城	487.5	400.5	343.7	293.6	235.9	197.7	162.9	118.7
	顺德	512.4	421.0	361.3	308.6	248.0	207.8	171.3	124.8
	番禺	469.9	386.1	331.3	283.0	227.4	190.6	157.1	114.4
区域	482.6	396.5	340.2	290.6	233.6	195.7	161.3	117.5	
III	区域增长曲线	3.360	2.862	2.517	2.196	1.806	1.531	1.266	0.907
	中山	549.8	468.3	411.9	359.4	295.5	250.5	207.3	148.4
	珠海	708.9	603.8	531.0	463.5	381.1	323.0	267.2	191.3
	深圳	581.5	495.3	435.6	380.1	312.6	264.9	219.2	157.0
	香港	735.3	626.3	550.8	480.7	395.2	335.0	277.2	198.5
	澳门	643.5	548.1	482.1	420.7	345.9	293.2	242.6	173.7
区域	643.8	548.4	482.3	420.9	346.1	293.3	242.7	173.8	
IV	区域增长曲线	2.677	2.421	2.223	2.021	1.744	1.523	1.287	0.925
	鹤山	338.6	306.3	281.3	255.7	220.6	192.7	162.8	117.0
	开平	389.0	351.9	323.1	293.7	253.4	221.4	187.0	134.5
	新会	405.1	366.4	336.5	305.9	263.9	230.5	194.8	140.0
	恩平	529.4	478.9	439.8	399.8	344.9	301.2	254.5	183.0
	台山	427.3	386.5	355.0	322.7	278.4	243.2	205.5	147.7
区域	417.9	378.0	347.1	315.6	272.3	237.8	200.9	144.4	
V	区域增长曲线	3.528	2.957	2.572	2.222	1.807	1.522	1.254	0.901
	龙门	555.0	465.3	404.7	349.6	284.3	239.4	197.3	141.7
	博罗	512.9	430.0	374.0	323.1	262.7	221.2	182.3	131.0
	惠阳	504.7	423.1	368.0	317.9	258.5	217.7	179.4	128.9
	惠东	720.7	604.2	525.5	454.0	369.1	310.9	256.2	184.0
区域	573.3	480.6	418.0	361.2	293.6	247.3	203.8	146.4	

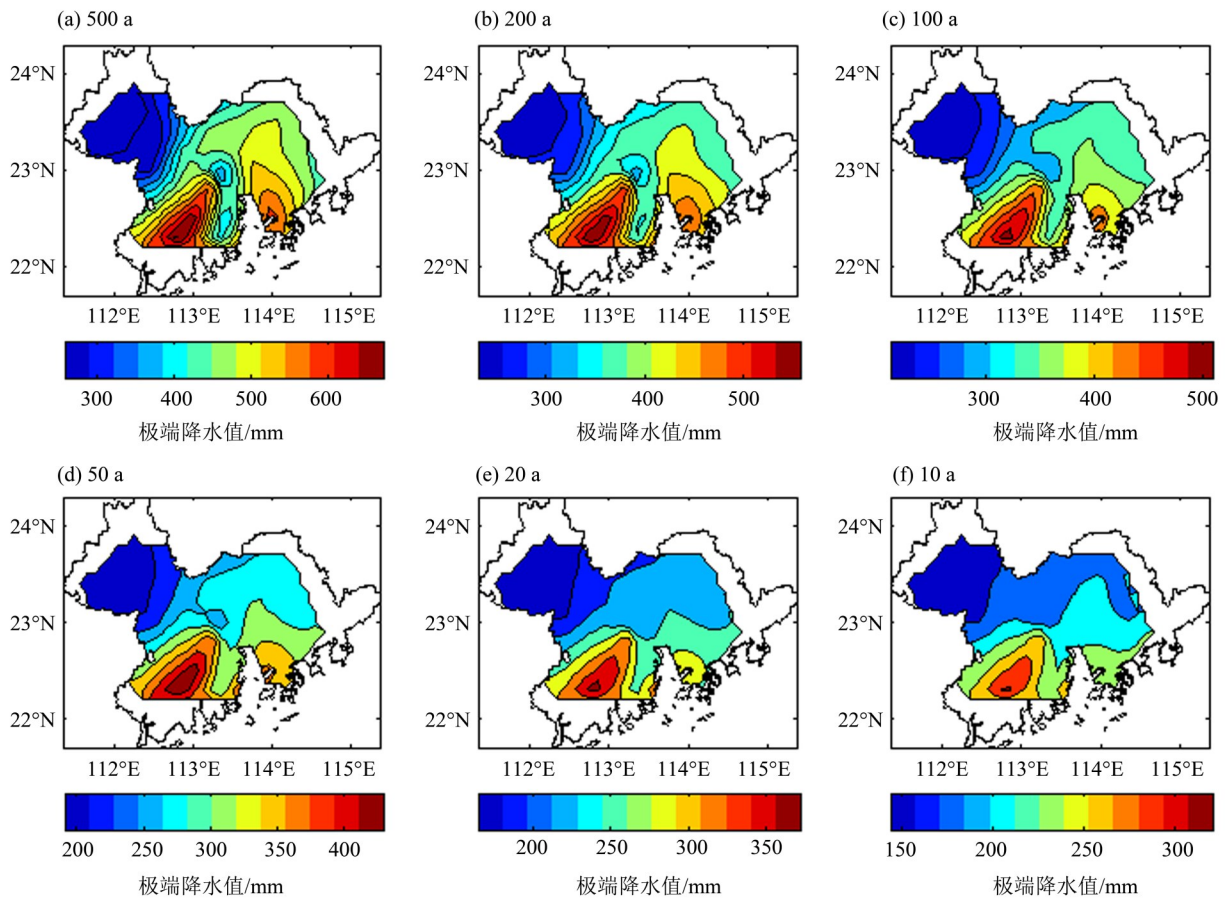


图3 区域最优增长曲线推算的大湾区极端降水设计值

Fig. 3 The design value of extreme precipitation in the Greater Bay Area calculated by the optimal regional growth curve

参考文献:

- 陈元芳,王庆荣,沙志贵等,2002.线性矩法在长江中下游区域水文频率计算中的应用[J].河海大学学报(自然科学版),31(2):207-211.
- 陈永勤,黄国如,2005.基于线性矩法的东江流域区域枯水频率分析[J].应用基础与工程科学学报,13(4):409-416.
- 郭军,任国玉,李明财,2010.环渤海地区极端降水事件概率分布特征[J].气候与环境研究,15(4):425-432.
- 刘瀚博,梁巧倩,2024.2022年5月10-14日广东短时强降雨可预报性分析[J].广东气象,46(3):14-21.
- 吴俊梅,林炳章,邵月红,2015.地区线性矩法在太湖流域暴雨频率分析中的应用[J].水文,5:15-22.
- ADAMOWSKI K, 2000. Regional analysis of annual maximum and partial duration flood data by nonparametric and L-moment methods[J]. J Hydrol, 229: 219-231.
- BONNIN G M, TODD D, LIN B, et al, 2003. Precipitation Frequency Atlas of the United States [M]. NOAA: National Weather Service.
- BONNIN G M, TODD D, LIN B, et al, 2004. Precipitation Frequency Atlas of the United States [M]. NOAA: National Weather Service.
- CUNNANE C, 1988. Methods and merits of regional flood frequency analysis[J]. J Hydrol, 100(1/2/3): 269-290.
- DALRYMPLE T, 1960. Flood frequency analysis [M]. Washington: Water Supply Paper 1953 A.
- GREENWOOD J A, LANDWEHR J M, MATALAS N C, et al, 1979. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form[J]. Water Resour Res, 15(5): 1049-1054.
- GUTTMAN N B, 1993. The use of L-moments in the determination of regional precipitation climates[J]. J Climate, 6: 2309-2325.
- HAYLOCK M, NICHOLLS N, 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998[J]. Int J Climate, 20: 1533-1541.
- HOSKING J R M, 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics[J]. J R Stat Soc Ser B Methodol, 52(1): 105-124.
- HOSKING J R M, 1994. The four parameters kappa distribution

- tion[J].IBM J Res Dev, 38(3):251-258.
- HOSKING J R M, WALLIS J R, 1993. Some statistics useful in regional frequency analysis [J]. Water Resour Res, 29(2):271-281.
- HOSKING J R M, WALLIS J R, 1997. Regional frequency analysis: An approach based on L-moments [M]. UK: Cambridge University Press.
- JAVELLE P, OUARDA T B M, LANG M, et al, 2002. Development of regional flood-duration-frequency curves based on the index flood method [J]. J Hydrol, 258: 249-259.
- KITE G W, 1988. Frequency and risk analysis in hydrology [M]. USA: Water Resources Publications.
- KJELDSSEN T R, JONES D A, 2006. Prediction uncertainty in a median-based index flood method using L moments [J]. Water Resour Res, 42: 1-12.
- LEE S H, MAENG S J, 2003. Frequency analysis of extreme rainfall using L-moment [J]. Irrig Drain, 52: 219-230.
- LIN B, BONNIN G M, MARTIN D L, et al, 2006. Regional frequency studies of annual extreme precipitation in the United States based on regional L-moments analysis [C]// EWRI Proceedings. Omaha, USA.
- NGONGONDO C S, XU C Y, TALLAKSEN L M, 2011. Regional frequency analysis of rainfall extremes in southern Malawi using the index rainfall and L-moments approaches [J]. Stoch Environ Res Risk Assess, 25: 939-955.
- NORBIATO D, BORGA M, SANGATI M, et al, 2007. Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood [J]. J Hydrol, 345: 149-166.
- PARIADA B P, KACHROO R K, SHRESTHA D B, 1998. Regional flood frequency analysis of Mashi-Sabarmati basin (Subzone 3-a) using index flood procedure with L-moments [J]. Water Resour Manag, 12: 1-12.
- PEE M C, WANG Q J, VOGEL R M, et al, 2001. The utility of L-moment ratio diagrams for selecting a regional probability distributions [J]. Hydrol Sci J, 46(1): 147-155.
- ROSSI F, VILLANI P, 1994. Regional flood estimation methods [M]. Netherlands: Springer.
- SAF B, 2009. Regional flood frequency analysis using L-moments for the buyuk and kukuk Menderes River Basins of Turkey [J]. J Hydrol Eng, 14: 783-794.
- SANKARASUBRAMANIAN A, SIRNIVASAN K, 1999. Investigation and comparison of sampling properties of L-moments and conventional moments [J]. J Hydrol, 218: 13-34.
- SATYANARAYANA P, SRINIVAS V V, 2008. Regional frequency analysis of precipitation using large-scale atmospheric variables [J]. Geophys Res, 113: D24110.
- SHI P, CHEN X, QU S, et al, 2010. Regional frequency analysis of low flow based on L moments: case study in karst area, southwest China [J]. J Hydrol Eng, 15: 370-377.
- STEDIGER J R, VOGEL R M, FOUFOULA-GEORGIU E, 1993. Frequency analysis of extreme events [M]. New York: McGraw-Hill.
- SWEET W V, PARK J, 2014. From the extreme to the mean: Acceleration and tipping points of coastal inundation from sea level rise [J]. Earth's Future, 2(12): 579-600.
- VOGEL R M, FENNESSEY N M, 1993. L-moment diagrams should replace product moment diagrams [J]. Water Resour Res, 29(6): 1745-1752.
- YANG T, XU C Y, SHAO Q X, et al, 2010. Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach [J]. Stoch Environ Res Risk Assess, 24: 165-182.
- YANG T, SHAO Q X, HAO Z C, et al, 2010. Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River basin, China [J]. J Hydrol, 380: 386-405.
- ZRINJI Z, BURN D H, 1994. Flood frequency analysis for ungauged sites using a region of influence approach [J]. J Hydrol, 153: 1-21.

(责任编辑 秦社彩)