

# 基于 DFA 方法的珠江流域极端降水阈值研究\*

刘丙军<sup>1</sup>, 伍丽丽<sup>1,2</sup>, 陆文秀<sup>1</sup>

(1. 中山大学地理科学与规划学院水资源与环境系, 广东 广州 510275;  
2. 珠江水利科学研究院, 广东 广州 510611)

**摘要:** 流域极端降水事件发生频率低、样本少, 其阈值定义存在定义方式不统一、方法多样等不确定性和无规则性问题。论文以珠江流域为研究区域, 系统总结了绝对临界值法、百分位值法、极值分布拟合法在确定流域极端降水阈值存在的不确定性问题, 并尝试运用 DFA 方法, 研究探讨了该方法在确定流域极端降水阈值的物理机制及其适用条件。研究表明, DFA 方法以系统动力学为基础, 通过分析极端降水事件对降水时间序列长程相关性的影响, 由此寻求流域极端降水阈值, 能较好反映流域降水序列的统计效应和物理背景, 比较适合流域极端降水阈值分析。

**关键词:** 珠江流域; 极端降水; 阈值; DFA 方法

中图分类号: P426.6 文献标志码: A 文章编号: 0529-6579(2013)01-0136-06

## Researches on DFA-based Extreme Precipitation Threshold in the Pearl River Basin

LIU Bingjun<sup>1</sup>, WU LiLi<sup>1,2</sup>, LU Wenxiu<sup>1</sup>

(1. Department of Water Resources and Environment, Sun Yat-sen University, Guangzhou 510275, China;  
2. Pearl River Hydraulic Research Institute Guangzhou, Guangzhou 510611, China)

**Abstract:** Extreme precipitation events occur rarely and are difficult to collect samples. There are still uncertainty problems with the definition of extreme precipitation threshold, which is variable and not unified. This study summarizes the uncertainties of extreme precipitation threshold based on Fixed Critical Value Method, Special Percentile Value Method and the Goodness of Extreme Value Distribution Fitting Test. In order to overcome these problems, a new method, named as DFA, has been used to discuss the physical mechanism and applicable conditions. The system dynamics-based DFA, tracks extreme precipitation threshold by analyzing the characteristics of long range correlation of the precipitation series and can well reflect the statistical effects and scientific background of precipitation series. It is applicable to define extreme precipitation threshold within a basin.

**Key words:** the Pearl River basin; extreme precipitation; threshold; detrended fluctuation analysis

全球气候持续变化影响下, 水循环更为复杂多变, 极端降水事件频繁发生。极端降水事件引发的洪涝灾害, 越来越影响到人类的正常生活, 成为全人类共同面临的生存问题。极端降水事件具有发生频率低、样本少、无规律等特点, 其阈值定义方式不统一、方法多样, 导致阈值定义具有较强不确定

性和无规则性。科学研究流域极端降水阈值定义及其时空分布特征, 对降低极端降水事件对社会经济的不利影响, 丰富和完善气候变化下水循环要素变异识别理论具有重要理论和实践意义。

国内外定义极端降水事件常规方法包括绝对临界值法、百分位法和极值分布拟合法等。绝对临界

\* 收稿日期: 2012-08-08

基金项目: 国家自然科学基金资助项目(50909106); 教育部青年教师培育基金资助项目(3161399); 广东省科技计划基金资助项目(2011B030800008)

作者简介: 刘丙军(1976年生), 男, 博士, 副教授; E-mail: liubj@mail.sysu.edu.cn

值法是由世界气象组织气候学委员会 (CCl/WMO) 推荐的一类根据绝对物理界限值定义极端气候事件的方法。它通过经验或统计, 采用某一绝对日降水强度作为临界值, 若某日降水超过此临界值, 则被认为是极端降水事件。由于降水分布空间差异大, 各地极端降水的绝对临界值也不统一。我国通常将日降水量超过 50 mm 的降水事件称为暴雨, 若日降水量超过暴雨标准则可认为是极端降水事件<sup>[1]</sup>; 百分位值法采用概率统计方法, 以超过某一百分位点的日降水量为降水极值, 即对降水湿日 (日降水量  $\geq 0.1$  mm) 序列进行频率分析, 将第  $n$  个百分位对应的日降水量定义为极端降水阈值, 当日降水量超过该阈值时, 则被认为是极端降水事件。王鹏翔<sup>[2]</sup>、杨金虎<sup>[3]</sup>、Judit<sup>[4]</sup> 和 Gemmer<sup>[5]</sup> 等分别选用了 99%、97%、95%、90% 等多类百分位值来定义研究区域极端降水阈值; 极值分布拟合法是利用概率统计方法, 根据原始日降水数据样本确定一组门限值, 对超过各门限值的样本集分别进行极值分布拟合, 将拟合效果最优的样本集门限值作为极端降水阈值。Grieser<sup>[6]</sup>、谢敏<sup>[7]</sup> 和 Santiago<sup>[8]</sup> 等分别对各类极值函数的适用性问题进行了有益探讨。

上述研究方法中, 绝对临界值法采用某一绝对日降水强度作为极端降水阈值, 具有较强的经验性和主观性, 无法反映流域降水时空分布差异特征; 百分位值法和极值分布拟合法均以气候要素统计学理论为基础, 认为极端降水事件属于单纯的小概率事件, 可以利用特定的分布函数拟合来确定极端降水阈值, 而较少考虑极端降水的物理背景及形成机理, 计算精度受样本容量大小和拟合分布函数类型影响较大。鉴于此, 本文在总结上述三类方法在确定流域极端降水阈值存在的不确定问题基础上, 尝试运用去趋势波动分析方法 (Detrended Fluctuation Analysis, 简称 DFA), 综合考虑流域极端降水事件的统计特征和物理背景, 通过系统动力学研究极端降水事件与正常降水事件之间的界限, 确定流域极端降水阈值。当前, DFA 方法被广泛应用于水文气候变化趋势及极端事件阈值确定, 并取得了较好研究成果<sup>[9-11]</sup>。

## 1 DFA 方法

### 1.1 极端降水事件的长程相关性特征分析

自记忆特征是气候系统变化的一项物理特征, 在气候时间序列里存在着长程相关性, 即气象系统的演化过程具有长时距相依的特征。对于极端降水

事件, 其演化状态视为水循环系统受到外界扰动而导致的极端异常状态, 其对应的演化轨迹不属于正常的演化轨迹, 而是水循环系统受到外界扰动而导致的极端异常状态。由于极端降水事件属于小概率事件, 所包含水循环系统演化信息极少, 对正常降水序列的长程相关性影响很小, 甚至可以忽略其统计效应。若对原始降水序列进行数据变换, 剔除极端降水数据后的降水序列演变趋势与原始时间序列基本一致, 而剔除正常降水数据的降水序列演变趋势与原始序列差异较大。因此, 可以根据这一现象, 按照由大到小的规则逐步剔除降水序列高值部分, 对剔除高值部分后的降水序列组进行长程相关性分析, 若能找到降水序列长程相关性系数的突变点, 则可认为在此突变点处被剔除的降水高值为极端降水, 由此可以划分降水事件的极端状态和非极端状态。

因此, 可以利用去趋势波动分析方法在计算时间序列长程相关性的优势, 寻求导致降水序列长程相关性发生突变的降水临界值。将该降水临界值作为极端降水阈值, 超过该降水临界值的降水样本作为极端降水, 而小于该临界值的降水样本作为正常降水。

### 1.2 DFA 方法

DFA 是基于随机过程理论和混沌动力学新发展的一种分析方法, 主要用于检测时间序列的长程相关性特征。DFA 方法通过有效滤去时间序列各个阶段的趋势成分, 从而消除趋势中的伪相关现象, 来探测出非平稳时间序列中的相关性。从动力学角度分析, 通过 DFA 方法进行数据变换后的序列仍保留着与原始序列相同的持久性 (或反持久性) 等演变痕迹, “滤除”了原始序列自身演化的各种短时程相依等趋势成分。因此采用 DFA 方法分析非平稳时间序列, 可以避免对相关性的错误判断。

利用 DFA 方法, 确定降水序列极端降水阈值的基本思路是寻找一个日降水量临界值, 保持小于该临界值的降水数据点位置不变时, 对大于该临界值的降水数据点, 无论这些数据点彼此之间的位置如何变化, 对整个序列的 DFA 指数无影响, 可以认为该临界值为降水序列的极端降水阈值。DFA 方法具体确定降水序列极端降水阈值的处理程序如下:

1) 对某一站点湿日 (日降水量大于或等于 0.1 mm) 降水序列  $\{x_i, i = 1, 2, \dots, n\}$ , 确定该降水序列  $\{x_i\}$  的最大值  $x_{\max}$ ; 确定参考点  $x_R$ ,  $x_R$  值可以是降水序列平均值  $x_{ave}$  或者介于最大降水量与

最小降水量之间的某一中值  $x_{med}$  ;

2) 从降水序列  $\{x_i\}$  的最大值  $x_{max}$  开始, 随机化挑选落在区间  $\{x_i; ifx_i \geq x_{max} - d \times k\}$  内的降水数据, 其余数据顺序保持不变, 直到  $x_i = x_R$ , 将依次生成新的降水序列  $Y_j, j = x_{max} - d \times k$ , 其中  $d$  为区间间隔,  $k = 1, 2, \dots, (x_{max} - x_R)/d$  或  $k = (x_{max} - x_R)/d$  ;

3) 计算每个新生成序列  $Y_j$  的长程相关性指数  $DFA_j$ , 分析长程相关性指数随区间  $J$  变化的变化过程。对于某一长度为  $m$  的降水序列  $Y_j = \{x_j, j = 1, 2, \dots, m\}$ , 其  $DFA$  长程相关性指数计算程序如下:

① 计算降水时间序列  $\{x_j, j = 1, 2, \dots, m\}$  的累积离差:

$$Y(j) = \sum_{k=1}^m (x_k - \bar{x}), \bar{x} = \frac{1}{m} \sum_{j=1}^m x_j, \quad j = 1, 2, \dots, m \quad (1)$$

② 将累积离差  $Y(j)$  等分成  $N_s$  个互不重叠时间长度为  $s$  的等时距区间, 其中  $N_s = \text{int}[m/s]$ 。降水序列个数并不总是时间长度  $s$  的整数倍, 可能使得小部分降水数据未被利用。为了不丢失降水数据的信息, 保证降水序列所有数据得到利用, 对  $Y(j)$  的逆序同样进行上述操作, 这样将共有  $2N_s$  个等时距区间。

③ 对于每个区间  $v(v = 1, 2, \dots, 2N_s)$ , 利用多项式回归拟合, 得到降水序列的局部趋势, 滤去该趋势后的序列记为  $Y_s(j)$ , 表示原序列与拟合序列之差, 即

$$Y_s(j) = Y(j) - P_v(j), j = 1, 2, \dots, m \quad (2)$$

式中,  $P_v(j)$  是第  $v$  个子区间的拟合多项式。可见  $DFA$  滤去了累积离差中的  $k$  阶趋势成分及原始序列中的  $k - 1$  阶趋势成分。

④ 计算每个子区间滤去局部趋势后的方差:

$$F^2(v, s) = \frac{1}{s} \sum_{j=1}^s Y_s^2[(v-1)s + j], v = 1, 2, \dots, N_s$$

$$F^2(v, s) = \frac{1}{s} \sum_{j=1}^s Y_s^2[N - (v - N_s)s + j], \quad v = N_s + 1, N_s + 2, \dots, 2N_s. \quad (3)$$

⑤ 计算标准  $DFA$  的波动函数:

$$F(s) = \sqrt{\frac{1}{2N} \sum_{v=1}^{2N_s} F^2(v, s)} \quad (4)$$

若降水序列  $\{x_j\}$  是长程幂律相关的, 则  $F(s)$  与  $s$  应满足幂律关系, 即

$$F(s) \propto s^a \quad (5)$$

对 (5) 式两边取对数, 得

$$\log_{10} F(s) \propto a \log_{10} s \quad (6)$$

⑥ 在双对数坐标系  $(s, F(s))$  中绘制  $a \log_{10} s - \log_{10} F(s)$  函数关系的散点图, 可确定波动函数的  $\alpha$  标度指数, 该指数即为  $DFA$  指数。当  $0.5 < \alpha < 1$  时, 表明此序列各个值之间不是独立的, 具有长程相关性。

4) 不同降水序列  $Y_j$  的  $DFA$  指数序列  $\{DFA_j\}$ , 会逐渐收敛于原始降水序列的  $DFA$  指数值。但指数序列  $\{DFA_j\}$  收敛于原始值的收敛程度并不完全一致, 而是围绕原始  $DFA$  指数值有着非常小幅度的波动。因此, 需要通过计算  $\{DFA_j\}$  的方差序列  $\{\text{var}_j\}$  来消除这一波动对判断收敛点带来的影响。方差定义为:

$$\text{var}^2 = \frac{1}{N-1} \sum_{j=1}^N (DFA_j - E)^2 \quad (7)$$

$$E = \frac{1}{N} \sum_{j=1}^N DFA_j \quad (8)$$

式中,  $DFA_j$  表示第  $j$  个新生序列的  $DFA$  指数,  $N$  表示总的新生成序列  $Y_j$  个数,  $E$  表示指数序列  $\{DFA_j\}$  的平均值。 $\{\text{var}_j\}$  与  $\{DFA_j\}$  的收敛性具有同步性<sup>[12]</sup>, 要确定  $\{DFA_j\}$  的收敛点, 则应先划分出  $\{\text{var}_j\}$  的收敛区。本文采用 BG 算法确定收敛区转折点, 并运用  $\chi^2$  检验检测  $\{\text{var}_j\}$  的收敛点。BG 算法和  $\chi^2$  检验方法的具体原理与步骤可参见文献 [13]。

5) 根据确定  $\{DFA_j\}$  收敛点, 找到导致  $DFA$  指数发生突出的降水临界值, 该临界值即为降水序列的极端降水阈值。

## 2 珠江流域极端降水阈值分析

### 2.1 珠江流域基本概况和数据

珠江流域位于东经  $102^\circ 14' - 115^\circ 53'$ , 北纬  $21^\circ 31' - 26^\circ 49'$  之间, 西起云贵高原、北起南岭山地, 经两广丘陵、珠江三角洲入南海。该流域地处低纬季风地区, 地域广大, 气候复杂, 大部分地区属于亚热带湿润气候, 北部为中亚热带湿润气候, 西部滇南河谷为北热带湿润气候。气候温和多雨, 多年平均温度在  $14 \sim 22^\circ\text{C}$  之间, 流域降水空间分布差异明显, 总体呈东多西少变化趋势, 多年平均年降水量在  $1200 \sim 2200 \text{ mm}$  之间。流域下游地区较上游地区易发生强降水, 最大连续降雨量、最大连续降雨天数以及最大连续降雨强度及三者叠加综合指数在空间分布上呈东大西小的局势, 珠江三角洲、桂林市、百色市等地区是流域极端降雨高发区; 流域极端降水总量无明显变化, 极端降水强



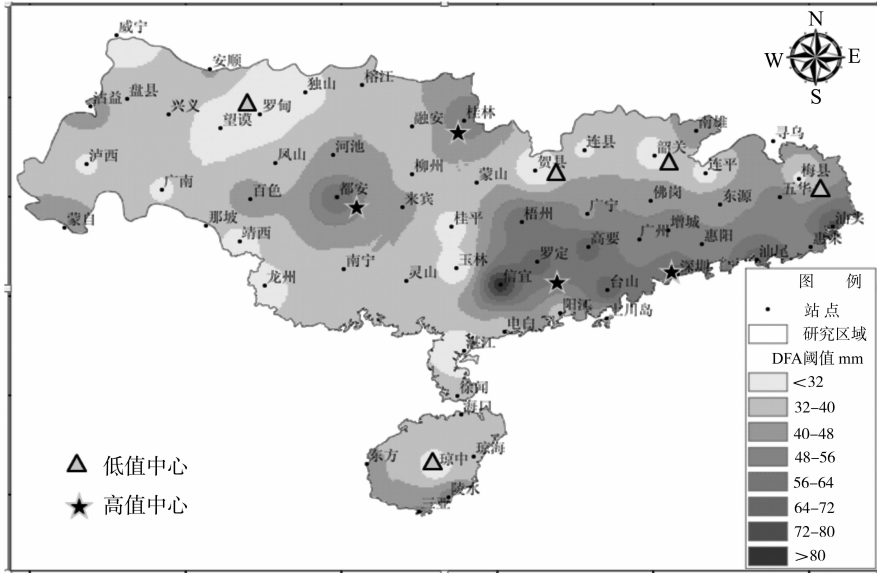


图 2 DFA 方法确定的极端降水阈值结果

Fig. 2 The result of extreme precipitation threshold base on DFA

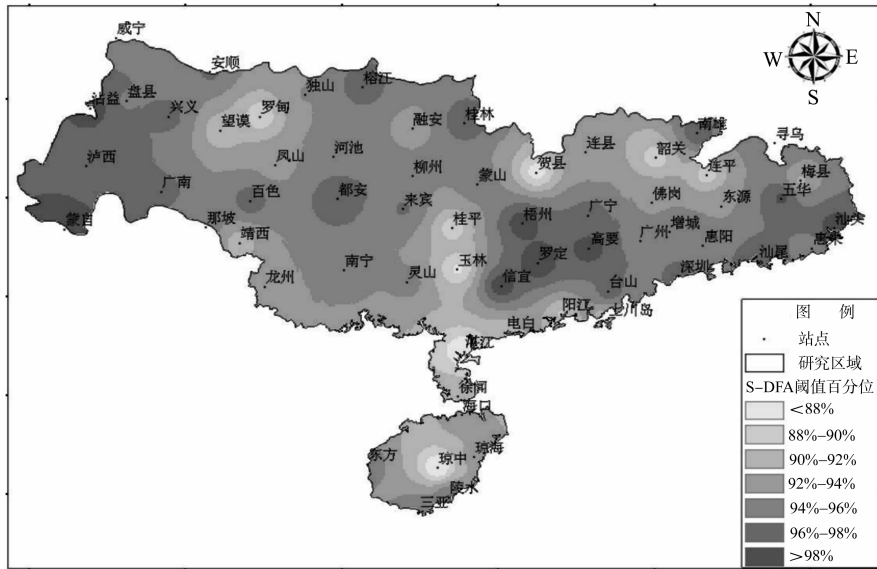


图 3 珠江流域 DFA 极端降水阈值百分位

Fig. 3 Percentile of extreme precipitation threshold base on DFA in the Pearl River Basin

点极端降水阈值与其湿日  $C_w$  值、湿日  $C_s$  值呈正比关系。如流域东南部沿海地区处于降水高值区, 该区域极端降水阈值普遍较大, 达到 55 mm/d 以上; 流域西部广西桂林等地湿日降水  $C_w$  值、 $C_s$  值较高, 表明该区域降水过程变幅较大, 其极端降水阈值也较大。

### 3 结论

流域极端降水事件发生频率低、样本少, 受全球气候变化、下垫面条件等多种内外部因素影响, 其时空演化具有较强不确定性。本文针对流域极端

降水阈值定义方式不统一、方法多样等特点, 总结了绝对临界值法、百分位值法、极值分布拟合法在确定流域极端降水阈值存在的不确定性问题, 并尝试运用 DFA 方法确定流域极端降水阈值, 探讨了该方法的物理机制及其适用性。得出如下结论:

1) 运用绝对临界值法、百分位值法和极值分布拟合法等确定流域极端降水阈值, 三种方法以经验公式或统计公式为主, 较少考虑极端降水的物理背景及形成机理, 且计算结果受样本容量大小和拟合分布函数类型等因素影响较大, 方法存在较强经验性和主观性;

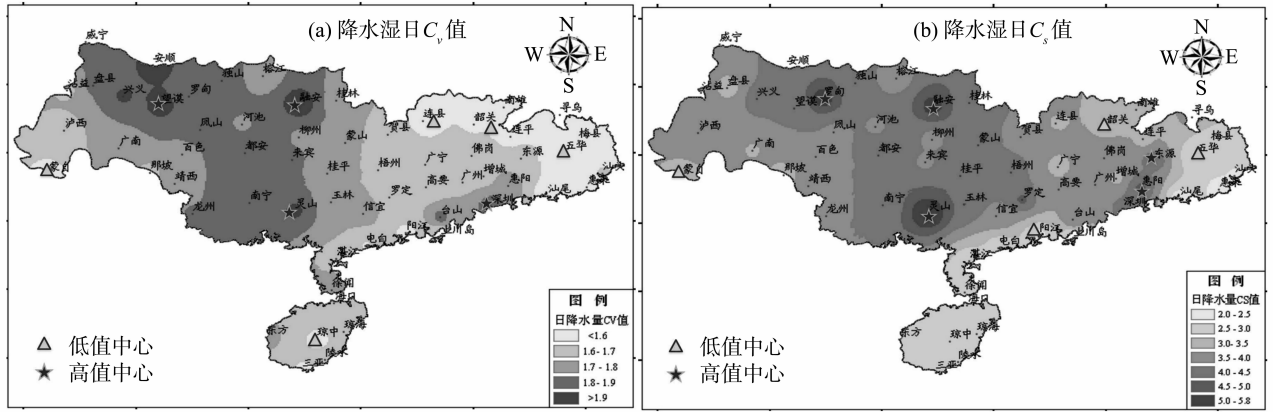


图 4 珠江流域降水湿日  $C_V$  值和  $C_S$  值分布情况

Fig. 4 Distribution of the  $C_V$  and  $C_S$  for the dairy rainfall in the Pearl River Basin

2) DFA 方法以系统动力学为基础，利用变换前后降水序列的长程相关指数变化特征，确定流域各站点极端降水阈值。与常规方法相比，该方法原理正确严谨，计算结果唯一，能较好反映流域降水序列的统计效应和物理背景。但该法要求样本容量大，且计算复杂。

参考文献：

[1] 闵岫,钱永甫. 中国极端降水事件的区域性和持续性研究[J]. 水科学进展,2008,19(6):763-771.

[2] 王鹏翔,杨金虎. 中国西北近 45 年来极端高温时间及其区域性增暖的响应[J]. 中国沙漠,2007,27(7): 649-655.

[3] 杨金虎,江志红,王鹏祥,等. 中国年极端降水事件的时空分布特征. 气候与环境研究[J],2008(1):75-83.

[4] JUDIT B, RITA P. Analysis of precipitation conditions for the Carpathian Basin based on extreme indices in the 20th century and climate simulations for 2050 and 2100 [J]. Physics and Chemistry of the Earth, 2010,35:43-51

[5] GEMMER M, FISCHER T, JIANG T, et al. Trends in Precipitation Extremes in the Zhujiang River Basin, South China[J]. Journal of Climate, 2011,24(3): 750-761.

[6] 谢敏,江志红,丁裕国. 运用 Gumbel-Logistic 模式模拟区域暴雨的试验[J]. 沙漠与绿洲气象,2011,5(1):1-5.

[7] GRIESER J, STAEGER T, SCHOENWIESE C D. Esti-

mates and uncertainties of return periods of extreme daily precipitation in Germany[J]. Meteorologische Zeitschrift, 2007,16(5): 553-564

[8] SANTIAGO B, MARTA A, SERGIO M, et al. Assessing trends in extreme precipitation events intensity and magnitude using non-stationary peaks-over-threshold analysis: a case study in northeast Spain from 1930 to 2006 [J]. International Journal of Climatology, 2011, 31: 2102-2114

[9] 张斌,史凯,刘春琼. 元谋干热河谷近 50 年分季节降水变化的 DFA 分析[J]. 地理科学, 2009,29(4): 561-566.

[10] BUNDE E K, KANTELHARDT J W, BRAUN P, et al. Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies [J]. Journal of Hydrology, 2006,322: 120-137.

[11] LUCIANO T, JORGE O P, BEATRICE S. Long-term persistence and multifractality of river runoff records: Detrended fluctuation studies [J]. Physica A, 2012, 391:1553-1562

[12] 侯威,章大全,周云,等. 一种确定极端事件阈值的新方法:随机重排去趋势波动分析方法[J]. 物理学报, 2011,60(10):1-15.

[13] 封国林,龚志强,董文杰,等. 基于启发式分割算法的气候突变检测研究 [J]. 物理学报,2005,54(11): 5494-5499.

[14] 彭俊台,张强,陈晓宏. 珠江流域极端降雨时空演变特征分析[J]. 灾害学,2011, 26(4): 24-34.