

基于去趋势波动分析方法确定极端事件阈值^{*}

杨 萍^{1,2)} 侯 威^{1,2)†} 封国林^{2,3)}

1) 兰州大学大气科学学院, 兰州 730000)

2) 中国科学院大气物理研究所东亚区域气候-环境重点实验室, 北京 100029)

3) 国家气候中心, 中国气象局气候研究开放实验室, 北京 100081)

(2007 年 11 月 14 日收到, 2008 年 3 月 5 日收到修改稿)

极端事件或者极值事件脱离了自身的正常演化状态, 是系统演化的极端状态或系统受到外界扰动而导致的异常状态. 去趋势波动分析法得到的指数是衡量系统在某一时间尺度内演化的长程相关性的参数, 系统的长程相关性不受极端事件的影响或影响很小. 基于这一思想, 提出了利用去趋势波动分析法确定极端事件的阈值方法, 并验证了该方法的有效性. 使用该方法对北京极端高温事件、极端低温事件和极端降水事件进行了分析和讨论, 确定北京 1951—2004 年极端高温、极端低温事件和极端降水事件的阈值. 50 余年来, 北京极端高温和低温事件在 20 世纪 70 年代之前整体而言属于极端事件发生平缓期, 在 70 年代至 90 年代属于极端事件间歇期. 从 20 世纪 90 年代中期至今, 极端高温事件在短时期内集中爆发, 高温一再突破历史记录, 日最低温度极大值也进入集中发生时期, 属于极端高温事件高发期. 对于极端降水事件, 其发生强度和频次具有明显的周期性, 20 世纪 60 年代中期和 80 年代中期为多发期, 且强度较大, 其余时期为间歇期, 强度较小, 两个完整的周期在 70 年代中期发生更替. 研究表明, 北京地区极端高温、低温事件和极端降水事件均在 70 年代中期发生了一次转折.

关键词: 去趋势波动分析, 极端事件, 阈值

PACC: 9260X

1. 引 言

天气和气候极端事件的频率和强度的变化对人类社会经济和环境很可能产生巨大影响, 因而受到越来越多的关注. 近几年, 不少学者利用多种方法从不同空间和时间角度对极端气温的变化展开了一系列的研究^[1-7]. 气候的定义从其本质上看与某种天气事件的概率分布有关, 当天气的状态严重偏离其平均态时就可以认为是不易发生的事件, 不容易发生的事件在统计意义上就可以称为极端事件. 目前国际上在气候极值变化研究中最多见的是采用某个百分位值作为极端事件的阈值, 超过该阈值的值被认为是极值, 该事件则可以认为是极端事件. 这些方法中阈值的确定受人为因素的影响较大, 且是从数据统计分析的角度出发. 在确定阈值的过程中并没有考虑数据或系统本身的演化特性, 如标度不变性、长程相关性等. 因此, 由传统的百分位方法得到的极

端事件阈值具有不确定的一面.

无标度性广泛存在于自然界系统(包括气候系统)中, 其特征之一是可观测量存在幂律关系. 标度不变性的存在表明, 系统发展是一个内外因素共同作用的过程, 系统过去发生的事件影响现在直到未来, 表现为一种有偏差的随机游走. 在复杂气候系统中, 许多信号是非平稳的, 其均值、标准差或自相关函数等特征量都随时间而变化, 传统的方法(如功率谱分析和相关分析)适合于计算平稳信号的相关特征. 为可靠地分析气候序列中的长程幂律相关性, 有必要辨别数据中内在的长程波动引起的趋势成分, 如果分析时没有滤去趋势成分, 则其中的强趋势成分会对长程相关的分析结果给出虚假的信息^[8]. 近年来, 随着基于分形自相似理论发展起来的去趋势波动分析(detrended fluctuation analysis, 简记为 DFA)是 Peng 等一些生物物理学家于 1994 年探测 DNA(脱氧核糖核酸)内部分子链的相关可能性的程度时首先提出的^[9]. 所谓 DFA 分析就是先除去序列中的

^{*} 国家科技支撑计划(批准号: 2007BAC03A01)资助的课题.

[†] E-mail: hou_w@sohu.com

趋势,再对序列进行记忆特征的研究.气象观测数据通常是非平稳的,采用传统的方法在计算自相关系数时经常受到噪声或某些趋势成分的影响,这会对分析结果的可靠性产生影响.DFA对于分析此类含有趋势成分的序列的标度行为有很大的优势.它可以有效滤去各阶趋势成分,能够探测出非平稳时间序列中的相关性,消除其中的伪相关现象.近几年,DFA已成功应用于噪声分析、心率动力学、物理学、地质学以及经济学等领域,在分析气候模拟结果及变率方面也有成功应用的先例^[10-14].

本文以 DFA 方法为基础,分析了在理想时间序列中极值和非极值数据对 DFA 方法计算结果的影响,并进一步运用 DFA 方法来确定时间序列中的极值,即确定极端事件的阈值.DFA 方法在确定极端事件阈值的过程中考虑了数据或系统本身的长程相关特性,因而具有较好的客观性,适合用来处理气候系统等复杂系统的极值问题.

2. DFA 方法

DFA 方法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,用于检测时间序列的物理特征.从动力学角度看,这种方法中变换的序列仍残留着原序列的痕迹,与原序列保持着相同的持久性(或反持久性).同时,变换可较好地“滤除”其自身演化的趋势成分,剩下的离差序列主要就是波动成分.因此在分析非平稳时间序列时,采用 DFA 方法可以避免对相关性的错误判断.在 DFA 的基础上,Kantelhardt 等^[15]在 2002 年又进一步提出非稳定有限序列的多重分形去趋势波动分析法(multifractal detrended fluctuation analysis,简记为 MF-DFA).MF-DFA 实际是 DFA 思想的进一步广义化,具有更强的功能.MF-DFA 不仅可以检测长程相关性、确定其标度不变性,即分形结构特征,还能判定序列是否具有多重分形属性并确定多分形特征^[15-18].

对长度为 N 的序列 $\{x_k, k = 1, 2, \dots, N\}$, DFA 方法可分为 5 个步骤.

步骤 I 建立一新序列

$$y(i) = \sum_{k=1}^N [x_k - \bar{x}] \quad (i = 1, 2, \dots, N), \quad (1)$$

其中 \bar{x} 为原序列 $\{x_k\}$ 的均值.

步骤 II 将新序列 $y(i)$ 划分为长度为 s 的不重叠等长度子区间,长度为 N 的序列共被分为 $N_s =$

$\text{in}(N/s)$ 个子区间.因序列长度 N 不一定被子区间长度 s 整除,为保证原序列信息不丢失,可以从序列末端开始反向前再划分一次,这样可得到共 $2N_s$ 个子区间.

步骤 III 对每个子区间 v ($v = 1, 2, \dots, 2N_s$) 的数据进行多项式回归拟合,得到局部趋势函数 $y_v(i)$. $y_v(i)$ 可以是一阶、二阶或更高阶的多项式,分别记为 DFA1, DFA2 等.消除各子区间内趋势,计算其方差均值

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s \{y[(v-1)s+i] - y_v(i)\}^2 \quad (i = 1, 2, \dots, N_s), \quad (2)$$

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

步骤 IV 确定全序列的 q 阶波动函数

$$F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(v, s)]^{q/2} \right\}^{1/q}, \quad (4)$$

其中 q 可以取为任何非零实数. $q = 0$ 时 (4) 式变为

$$F_0(s) = \exp \left\{ \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln [F^2(v, s)] \right\}. \quad (5)$$

步骤 V 通过分析双对数坐标图 $F_q(s) \propto s^{h_q}$ 的关系,可以确定波动函数的标度指数 h_q ,即存在幂律关系

$$F_q(s) \propto s^{h_q}. \quad (6)$$

在本文的计算中,对每个子区间 v ($v = 1, 2, \dots, 2N_s$) 的数据进行二阶多项式回归拟合,并取 $q = 2$.此时,对于平稳序列而言, h_q 为赫斯特指数;对于非平稳序列,当 $h_q = 0.5$ 时,该序列为一独立过程;当 $0.5 < h_q \leq 1$ 时,该序列存在长程相关性;当 $h_q < 0.5$ 时,该序列存在负长程相关,即反持久性.本文无意去讨论 DFA 指数(即长程相关性)的物理意义,而将其作为衡量系统物理性质的参数,讨论极端值对系统 DFA 指数的影响,以此为依据来确定极端事件的阈值.

3. 利用 DFA 确定极端事件的阈值

3.1. 方法概述

人们已经认识到,气候系统的变化是具有自记

忆特征的,在气候时间序列里存在着长程相关性,即系统的演化状态具有持续性^[19-21].由于极端事件或者极值事件是系统演化的极端状态或是系统受到外界扰动而导致的异常状态,不属于系统自身正常演化状态的范畴^[22,23],DFA 方法得到的 DFA 指数衡量的是系统在某一时间尺度内演化的长程相关性,而系统整体的长程相关性不受极端事件的影响或影响很小.基于这一思想,可以用 DFA 确定极端事件的阈值.已知系统演化的某一组时间序列为 $\{x_i, i = 1, \dots, n\}$,确定极端事件阈值的具体过程如下:

- 1) 确定 $\{x_i\}$ 的最大值 x_{max} 和最小值 x_{min} .
- 2) 确定序列 $\{x_i\}$ 的中间点 R , R 值可以是平均值 x_{ave} 或者介于最大值与最小值之间的某一中值 x_{med} .
- 3) 从 $\{x_i\}$ 的最大值 x_{max} 开始,依次舍去 $\{x_i, x_i \geq x_{max} - d \times k\}$ 数据区间的 数据点,直到 $x_i = R$,依次得到新序列 $Y_J, J = x_{max} - d \times k$,其中 d 为区间间隔, $k = 1, 2, \dots, (x_{max} - R)/d$.
- 4) 从 $\{x_i\}$ 的最小值 x_{min} 开始,依次舍去 $\{x_i, x_i \leq x_{min} + d \times k\}$ 数据区间的 数据点,直到 $x_i = R$,依次得到新序列 $Y_J, J = x_{min} + dk$,其中 d 为区间间隔, $k = 1, 2, \dots, (R - x_{min})/d$.
- 5) 计算每个新序列 Y_J 的长程相关性指数 D_J , 得到其随舍去区间 J 的变化.

6) 当 D_J 的变化开始趋于平缓且收敛于数据 $\{x_i\}$ 的原始 DFA 指数 D_{origin} 时,取此时的 J 值作为数据 $\{x_i\}$ 极端事件的阈值.

7) 区间间隔 d 代表了本方法的分辨率,其取值越小,确定的阈值分辨率越大,同时计算量也越大;反之则阈值分辨率小,计算量也相对较小.

3.2. 各个区间数据对 DFA 指数的影响

首先验证该方法的有效性.使用混沌系统 Lorenz 方程的 x 分量的 10000 个数据点作为原始序列 $\{x_i\}$,且 $18.57 > \{x_i\} > -18.667$.从其中随机选择任意点 x_i 并改变其值大小,使得 $x_i > 18.57$ 或 $x_i < -18.667$ 并使其成为极端值点,得到序列 $\{x'_i\}$,如图 1(b)所示.按照上述步骤来确定序列的极值事件阈值,计算中局部趋势函数 $y_r(i)$ 使用二阶多项式,即 DFA2;不重叠等长度子区间的长度 s 取值为 $10 \leq s \leq n/2$, n 为序列长度;取区间间隔 $d = 1$,中间点 R 取为零值.

从图 1(a)可以看到:在 $x'_i \geq R$ 时,舍弃序列的极大值 $x'_i \geq 3\sigma$ (此时 $J = 32$) 后得到的 Y_J 的 DFA 指数 D_J 与原始序列的 DFA 指数 D_{origin} (图 1 中点线所示) 相比几乎没有任何变化.随着 J 值的逐渐减小,序列中 $x'_i \geq J$ 的点也越来越多,意味着所舍弃的点也越来越多,直到 $J = 18$ 为止, D_J 值与 D_{origin} 相比仍

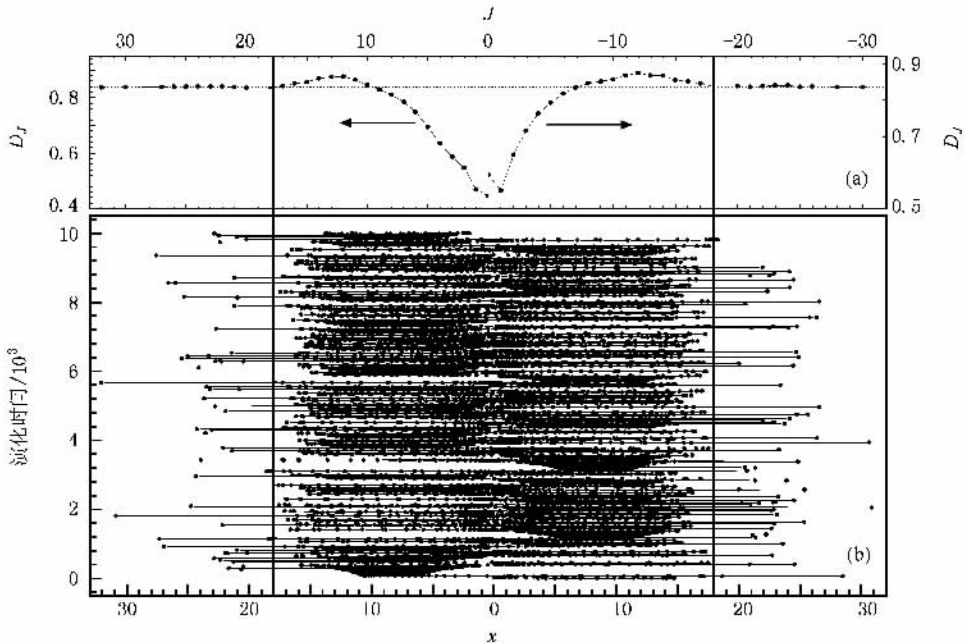


图 1 对含极端值的 x 分量舍去不同区间数据后 DFA 指数的变化 (a) 去除不同区间数据点后序列 Y_J 的 DFA 指数变化 (b) 加入极值点的 Lorenz 系统 x 分量

几乎相等.这说明当 $J \geq 18$ 时舍弃的点所包含系统的整体演化信息非常少,其动力学控制机制远远偏离于系统的整体演化方程.舍弃这些点并没有改变序列本身的演化规律,序列 Y_J 和原始序列 $\{x'_i\}$ 满足相同的演化机制,所以 D_J 几乎等同于 D_{origin} .这些点可以认为是当前序列的极端值, $J = 18$ 也就可以认为是极大值的阈值.同样的情况也出现在 $x'_i \leq R$ 时,在舍去极小值点 $x'_i \geq -30$ (此时 $J = -30$) 后, $D_J = D_{origin}$,这一现象持续到 $J = -18$.可以将 $x'_i \leq -18$ 的数据点均视为极端值, $J = -18$ 也就是极小值的阈值.当 J 值向中间点 R 逼近时, Y_J 的 DFA 指数 D_J 也逐渐偏离原始序列的 DFA 指数 D_{origin} ,当 $J = R$ 时偏离达到最大.并且对于不同的 J 值, D_J 也有较大的差异.这表明当 J 越来越逼近中间点 R 时,所舍弃的数据点服从系统动力学控制方程的控制.舍弃数据点的同时也包含越来越多的系统演化信息,序列 Y_J 与原始序列 $\{x'_i\}$ 的演化机制相比有

了很大改变,故其 DFA 指数 D_J 也偏离原始值 D_{origin} 较大. $18 > x'_i > -18$ 的数据点是当前序列的非极端值.

为了验证该方法的有效性,用该方法对原始序列 $\{x_i\}$ 进行同样的操作计算,如果该算法有效,即极端值对序列的 DFA 指数没有影响,则序列 $\{x_i\}$ 的极端值阈值与 $\{x'_i\}$ 极端值阈值应该相同,并且序列 $\{x_i\}$ 和 $\{x'_i\}$ 的 DFA 指数也应该相同.计算过程中采用相同的参数,结果如图 2 所示.从图 2 同样可以看到,在 $J = 18, -18$ 时,序列 Y_J 的 DFA 指数 D_J 恰好开始收敛于原始序列的 DFA 指数 D_{origin} (图 2 中点线).这表明当 $J \geq 18$ 和 $J \leq -18$ 时,所舍弃的点所包含系统的整体演化信息非常少, $J = 18, -18$ 可以认为是序列 $\{x_i\}$ 极端值的阈值.序列 $\{x_i\}$ 的 DFA 指数为 0.8378,序列 $\{x'_i\}$ 的 DFA 指数为 0.8357,两者仅在小数点后第三位具有微小差异.这说明本文所使用的方法是有效的.

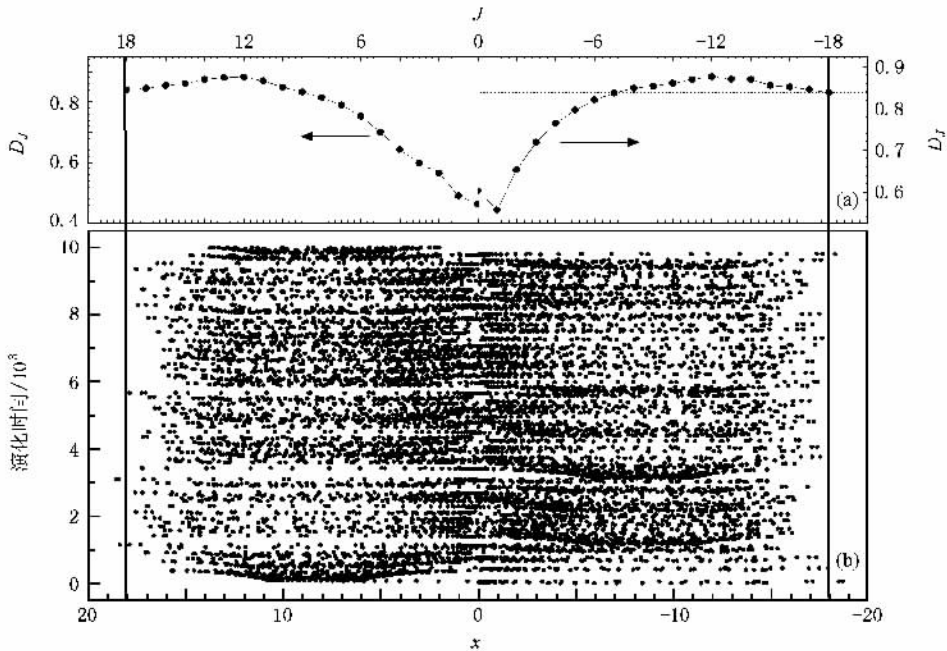


图 2 对原始 x 分量舍去不同区间数据后 DFA 指数的变化 (a) 去除不同区间数据点后序列 Y_J 的 DFA 指数变化 (b) 原始的 Lorenz 系统 x 分量

n 为序列长度, 区间间隔 $d = 1$ 取零值为中间点 R .

4. 实际资料的应用

资料来源于中国气象局 2005 年公布的中国地面国际交换站气候资料日值数据集. 在计算过程中, 同样采用二阶多项局部趋势函数 $y_s(i), 10 \leq s \leq n/2$,

4.1. 极端高温事件

选取北京站(站点编号 54511)1951—2004 年无缺测的逐日最高温度和最低温度观测数据. 由于实际气候资料的极端事件受时间尺度的限制, 不同的

时间尺度可能具有完全不同的极端事件,也就具有不同的极端事件阈值.在这里只讨论序列长度范围

内的极端事件,即只确定 1951—2004 年这一时间尺度内极端事件的阈值.

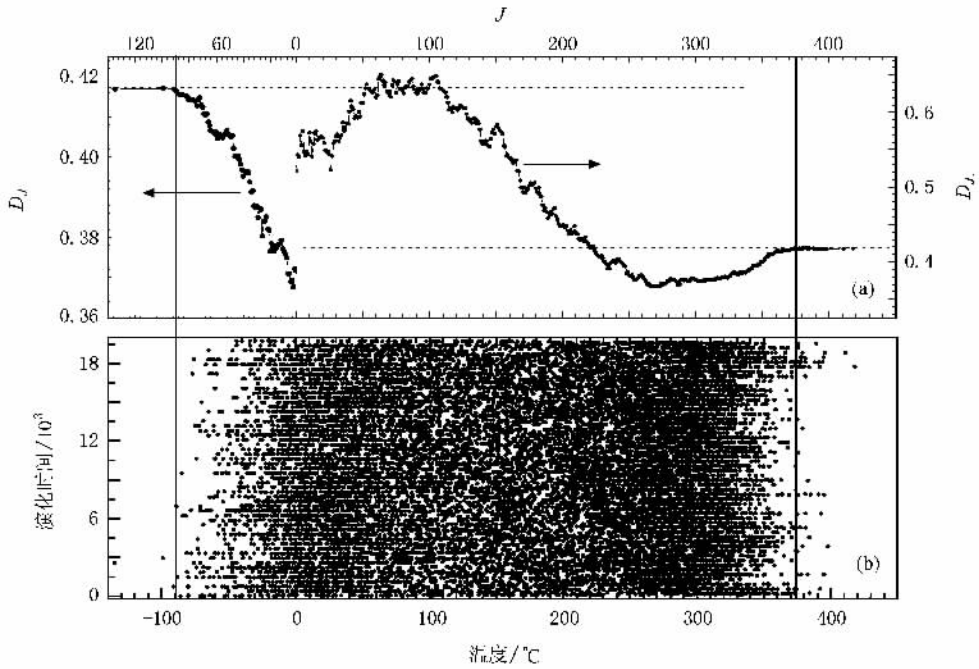


图3 对北京站日最高温度资料舍去不同区间数据后 DFA 指数的变化 (a)去除不同区间数据点后序列 Y_J 的 DFA 指数变化 (b)1951—2004 年日最高温度观测资料

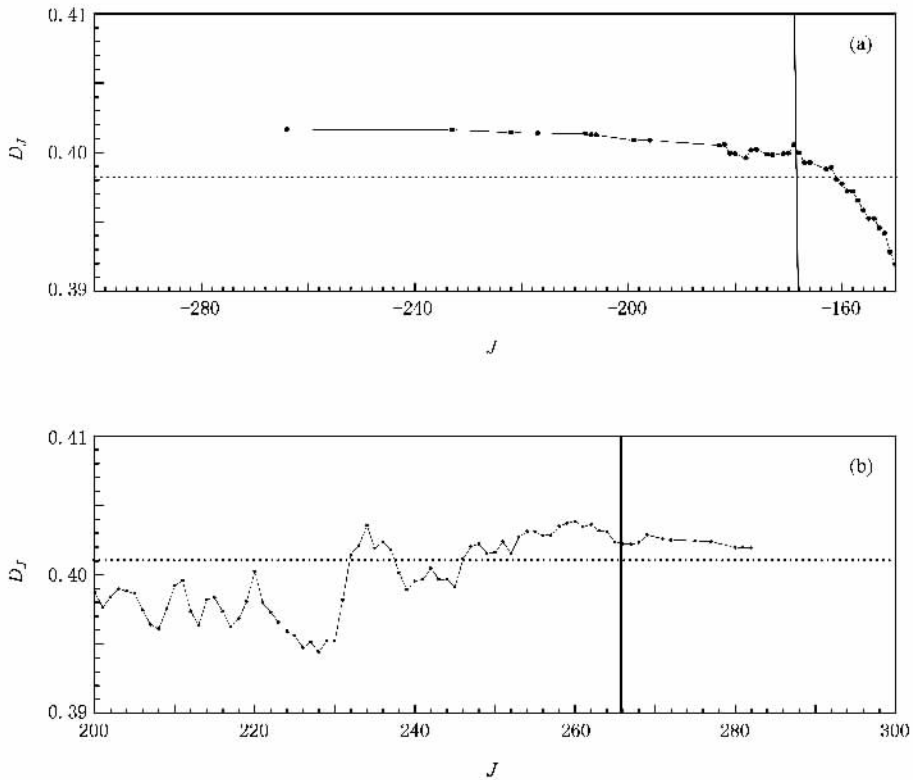


图4 北京站日最高温度的低温区间和高温区间 DFA 指数收敛区域放大图 (a)低温区间 (b)高温区间

对北京地区日最高温度而言,从图 3、图 4 由于原始数据记录格式的标准化原因,故需将图 3、图 4 中横坐标乘以 0.1 代表观测值,点线代表原始 DFA 指数 D_{origin} 可以看出:当 $J = 373$ 时, D_J 开始收敛于 D_{origin} ,并随着 J 值的增加, D_J 值变化不大,而当 $J < 373$ 时, D_J 值之间有着很大差异,同时与原始值 D_{origin} 之间偏差较大.可以认为温度 $T = 37.3\text{ }^\circ\text{C}$ 是北京地区 1951—2004 年间日最高温度极大值,即极端高温事件的阈值.对于低温事件,当 $J = -90$ 时, D_J 收敛于 D_{origin} ,且随着 J 值的减小, D_J 只有微小变化,当 $J > -90$ 时, D_J 值不论是彼此之间还是相对于原始值 D_{origin} 都有很大的偏差.可以确定北京日最高温度极小值的阈值为 $T = -9.0\text{ }^\circ\text{C}$.

根据确定的极端事件阈值,极端高温事件在 1951 年、1952 年、1955 年、1960 年、1961 年、1965 年、1968 年、1972 年和 1973 年共发生 28 次,虽然发生频率和强度都较高,但时间分布比较均匀.在 1974—

1996 年间只有 4 次,分别出现在 1978 年、1981 年、1986 年和 1988 年,频率和强度均较低.随后在 1997 年、1999 年、2000 年、2001 年、2002 年、2003 年和 2004 年共发生 20 次,极端高温事件不仅在短时期内频繁发生,而且高温强度也大幅度上升.对于日最高温度极小值,在 1953 年、1959 年、1958 年和 1970 年共发生 4 次,其强度也逐渐减弱.

综合来看,无论是极端高温事件的发生次数还是强度都有明显的增加趋势.以 20 世纪 70 年代为界,在 1951—1977 年极端高温事件虽然发生次数多且强度较高,但两者随时间变化不大,处于极端高温事件平缓期,同时日最高温度极小值也只在 1953 年、1959 年、1958 年和 1970 年出现 4 次.1978—1996 年极端高温仅出现 4 次且强度都较低,属于间歇期,日最高温度极小值在 1970 年之后未发生.1997—2004 年极端高温事件在短时期内集中爆发,温度也一再突破历史记录,进入极端事件高发期.

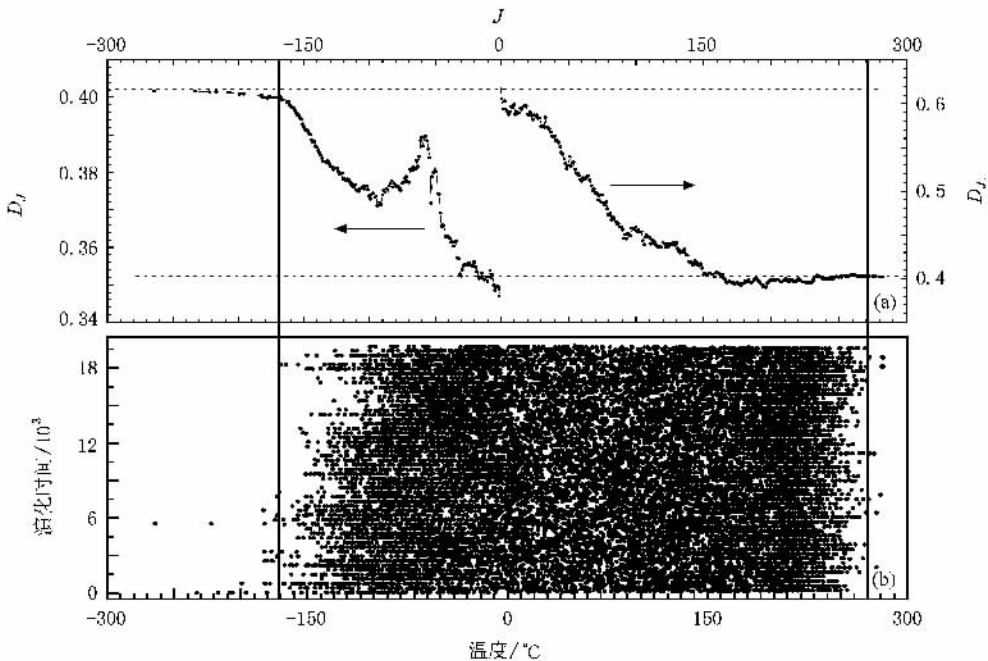


图 5 北京站日最低温度资料舍去不同区间数据后 DFA 指数的变化 (a)舍去不同区间数据点后序列 Y_J 的 DFA 指数变化 (b)1951—2004 年日最低温度观测资料

4.2. 极端低温事件

进一步利用北京地区日最低温度资料来确定极端低温事件的阈值.从图 5 和图 6 (图 5、图 6 中的横坐标同样乘以 0.1 代表观测值,点线代表原始 DFA 指数)可以看出:当 $J = -169$ 时, D_J 开始收敛于原

始值 D_{origin} .随着 J 值的进一步减小, D_J 几乎完全相等;当 $J > -169$ 时, Y_J 的 DFA 指数 D_J 变化明显且偏离原始值 D_{origin} .可以确定 $T = -16.9\text{ }^\circ\text{C}$ 为北京地区 1951—2004 年间极端低温事件的阈值.当 $J = 266$ 且随着 J 值的增大, D_J 均收敛于原始值 D_{origin} ;当 $J < 266$ 时,不同 Y_J 的 DFA 指数 D_J 彼此之间差异较

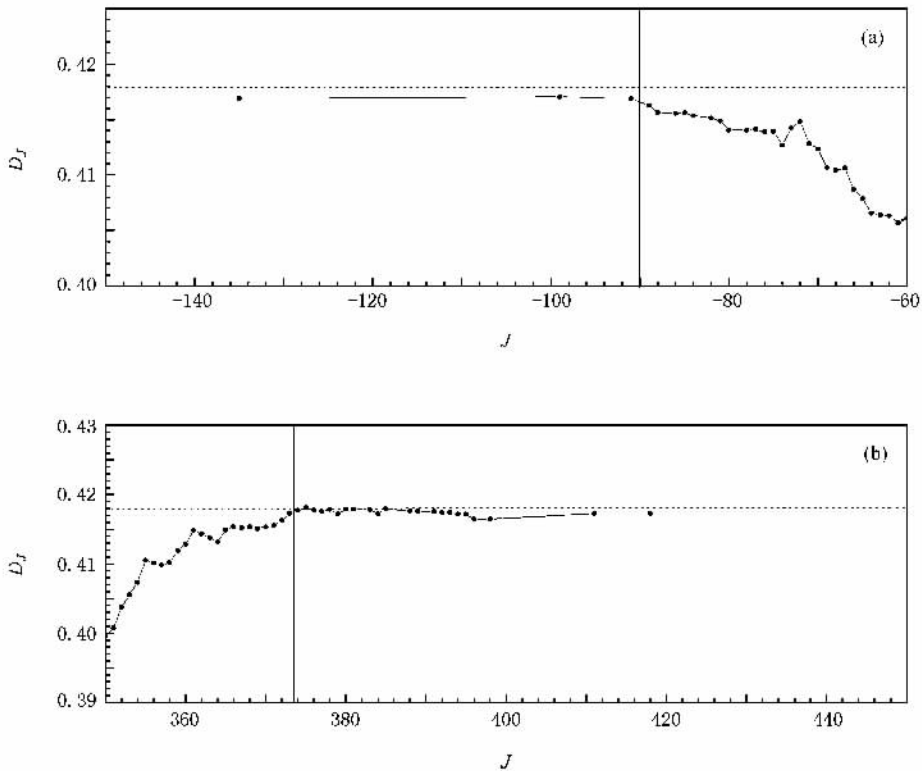


图6 北京站日最低温度低温区间和高温区间 DFA 指数收敛区域放大图 (a)低温区间 (b)高温区间

大,与原始值 D_{origin} 相比也有很大的偏差.可以认为日最低温度极大值的阈值为 $T = 26.6\text{ }^{\circ}\text{C}$.

对于极端低温事件,20 世纪 70 年代之前,在 1951 年、1952 年、1953 年、1958 年、1959 年、1960 年、1966 年、1967 年、1969 年、1972 年和 1973 共发生 26 次,频率和强度均较大.此后直到 2001 年才再一次出现,在 1974—2000 年以及 2001 年之后未发生,无论是发生次数还是低温强度总体呈下降趋势.对于日最低温度极大值,在 1956 年、1958 年、1961 年、1968 年、1971 年、1972 年、1973 年和 1981 年共发生 13 次,其出现频率和强度随时间变化不明显.直到进入 20 世纪 90 年代后半叶再次集中发生,在 1995 年、1997 年、1999 年、2000 年、2002 年和 2004 年共发生 12 次.

由以上所述可知,自 1951 年以来极端低温事件的发生次数有非常明显的减少,且大部分极端低温事件发生在 20 世纪 70 年代前,1974 年之后只发生了一次,且强度刚刚达到阈值标准.日最低温度极大值除在整个 80 年代未出现,其余各个时期均有发生.频率和强度在 70 年代末之前变化不明显,但在 20 世纪后半叶集中爆发.

4.3. 极端降水事件

选取北京站 1960—2000 年无缺测的逐日观测降水资料,讨论 1960—2000 年内的极端事件,确定这一时间尺度内极端降水事件的阈值.从图 7(a), (b) 图 7 中的横坐标同样乘以 0.1 代表观测值,点线代表原始 DFA 指数)可以看出:当 $J \geq 1030$ 时,不同序列 Y_j 的 DFA 指数 D_j 均收敛于原始值 D_{origin} ; 当 $J < 1030$ 时,不同 Y_j 的 DFA 指数 D_j 值差异较大且偏离 D_{origin} ,确定北京 1960—2000 年间极端降水的阈值为 $H = 1030\text{ mm}$.进一步分析发现,极端降水事件在 1960 年、1963 年、1964 年、1975 年、1984 年、1985 年、1986 年、1988 年和 1998 年共发生 10 次.

从图 7(c)可以看出:极端降水事件在 1960—1975 年、1975—2000 年出现两个完整的周期(如图 7(c)中箭头所示).无论是发生强度还是发生次数,在 1960—1964 年和 1984—1988 年为波峰,处于极端事件多发期、高强度期,1964—1984 年和 1988—2000 年为波谷,处于极端事件间歇期、低强度期.周期长度大约为 20 a,两个周期在 1975 年左右发生更替.

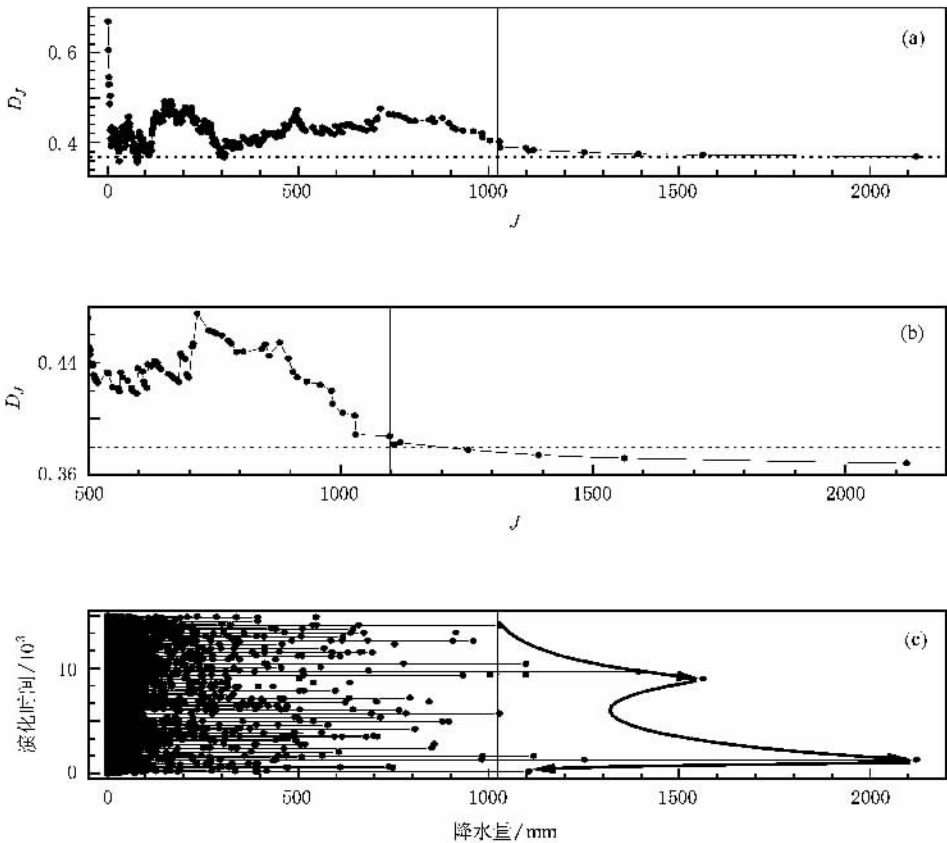


图7 对北京站降水资料舍去不同区间数据后 DFA 指数的变化 (a)去除不同区间数据点后序列 Y_J 的 DFA 指数变化 (b)收敛区域放大图 (c)1952—2004 年日降水观测资料

4.4. 综合分析

综上所述,50 余年来北京地区极端高温和极端低温事件以 20 世纪 70 年代为界.在此之前,极端高温事件的强度较高,发生频率随时间变化不大,分布较均匀,同时日最高温度极小值仅仅出现 4 次.极端低温事件几乎全部集中在 20 世纪 70 年代前,发生频率变化不大但低温强度有非常明显的下降趋势,日最低温度极大值的发生频率和强度同样随时间无明显变化,整体而言属于极端事件发生平缓期.在 70 年代之后,极端高温仅仅在 1978—1996 年间出现 4 次且强度都较低,极端低温事件只在 2001 年出现一次,其强度刚刚达到阈值标准;日最高温度极小值在 70 年代之后未发生,日最低温度极大值在整个 80 年代同样没有出现,属于极端事件间歇期.从 20 世纪 90 年代中期至今,极端高温事件在短时期内集中爆发,高温温度也一再突破历史记录,日最低温度极大值也进入集中发生时期,属于极端高温事件高发期.整体而言,北京地区极端高温事件的频率和强

度都在上升,而极端低温事件则在经历了下降趋势后已经多年未发生,气候系统偏向于高温增暖.

对于极端降水事件,其发生强度和频次具有明显的周期性.20 世纪 60 年代中期和 80 年代中期为多发期,且强度较大,其余时期为间歇期,强度较小.两个完整的周期在 70 年代中期发生交替.由此可以看出,北京地区的极端高温、低温事件和极端降水事件在 70 年代中期发生了一次转折,与华北地区近几十年来的气候转折^[24—27]一致.

5. 结 论

本文讨论了事件极端值对系统 DFA 指数即长程相关性的影响.计算表明:当系统的演化状态处于极端事件时,去除这些极端状态对系统的长程相关性几乎没有任何影响;当系统的演化状态为非极端值,去除不同非极端状态后的 DFA 指数彼此之间有很大差异.基于这一思想给出了利用 DFA 方法确定极端事件阈值的详细方法,并验证了方法的有效性.

使用 DFA 方法对北京极端高温事件、极端低温事件和极端降水事件进行了分析讨论. 确定北京 1951—2004 年极端高温事件的阈值为 $T = 37.3\text{ }^{\circ}\text{C}$, 极端低温事件的阈值为 $T = -16.9\text{ }^{\circ}\text{C}$, 极端降水的阈值为 $H = 1030\text{ mm}$. 在 20 世纪 70 年代之前, 北京地区的极端高温事件的强度较高, 发生频率随时间变化不大, 极端低温事件几乎全部集中在 70 年代之前. 在 70 年代之后, 极端高温事件仅仅在 1978—1996 年间出现了 4 次且强度都较低, 极端低温事件只在 2001 年出现一次; 日最高温度极小值在 1970 年之后未发生. 从 20 世纪 90 年代中期至今, 极端高温事件在短时期内集中爆发, 高温温度也一再突破历史记录, 日最低温度极大值也进入集中发生时期. 北京地区极端高温事件的频率和强度都在上升, 而极端低温事件则在经历了下降趋势后已经多年未发生, 气候系统偏向于高温增暖. 对于极端降水事件,

其发生强度和频次具有明显的周期性. 60 年代中期和 80 年代中期为多发期, 且强度较大, 其余时期为间歇期, 强度较小. 两个完整的周期在 70 年代中期发生交替. 由此可知, 北京地区的极端高温事件、极端低温事件和极端降水事件在 20 世纪 70 年代中期发生了一次转折, 与华北地区近几十年来的气候转折一致.

同时也可以看出, 在全球增暖的大背景下, 极端事件的发生频率和强度经历了从平缓期到过渡间歇期, 最后演化至高发期这一过程, 与百年来全球不断增暖的现象并不完全一致. 这说明极端事件具有特殊的发生发展规律. 始于 20 世纪 90 年代后半叶的高发期是否原本应为平缓期, 但在全球增暖的作用下其出现频率和强度大幅度上升, 还是由于 20 世纪 70 年代中期发生的转折改变了气候系统内在演化规律而导致, 这些问题还有待于进一步研究.

- [1] IPCC 2001 *Climate Change* (New York: Cambridge University Press) p155
- [2] Karl T R, Easterling D R, Quayle R G 1996 *Bull. Am. Meteor. Soc.* **77** 279
- [3] Jones P D, Horton E B, Folland C K, Hulme M, Parker D E, Basnett A 1999 *Climatic Change* **42** 131
- [4] Ren F M, Zhai P M 1998 *Sci. Atmos. Sin.* **22** 217 (in Chinese) [任福民、翟盘茂 1998 大气科学 **22** 217]
- [5] Zhai P, Sun A, Ren F, Liu X, Gao B, Zhang Q 1999 *Clim. Chan.* **42** 203
- [6] Li S G, Wu J T 2002 *Fractal and Wavelet* (Beijing: Science Press) (in Chinese) [李水根、吴纪桃 2002 分形与小波 (北京: 科学出版社)]
- [7] Dai X G, Fu C B, Wang P 2005 *Chin. Phys.* **14** 850
- [8] Zheng Z F, Zhang X L, Cao H X, Xie Z, Pan J H 2007 *Chin. J. Geophys.* **50** 420 (in Chinese) [郑祚芳、张秀丽、曹鸿兴、谢庄、潘家华 2007 地球物理学报 **50** 420]
- [9] Peng C K, Buldyrev S V, Havlin S, Simons M, Stanley H E, Goldberger A L 1994 *Phys. Rev. E* **49** 1685
- [10] Janosi I M, Janecsko B, Kondor I 1999 *Physica A* **269** 111
- [11] Ausloos M 2000 *Physica A* **285** 48
- [12] Koscielny B E, Bunde A, Havlin S 1996 *Physica A* **231** 393
- [13] Pelletier J D 1997 *J. Clim.* **10** 1331
- [14] Fraedrich K 2002 *Stoch. Dynam.* **2** 403
- [15] Kantelhardt J W, Zschiegner S A, Koscielny B E 2002 *Physica A* **316** 87
- [16] He Y L, Yao L K, Su F H 2005 *Syst. Eng. Theor. Prac.* **7** 125 (in Chinese) [何越磊、姚令侃、苏凤环 2005 系统工程理论与实践 **7** 125]
- [17] Zhi R, Gong Z Q, Wang D Y, Feng G L 2006 *Acta Phys. Sin.* **55** 6185 (in Chinese) [支蓉、龚志强、王德英、封国林 2006 物理学报 **55** 6185]
- [18] Zhi R, Lian Y, Feng G L 2007 *Acta Phys. Sin.* **56** 1837 (in Chinese) [支蓉、廉毅、封国林 2007 物理学报 **56** 1837]
- [19] Cao H X 1993 *Sci. China B* **23** 104 (in Chinese) [曹鸿兴 1993 中国科学 B **23** 104]
- [20] Shukla J 1998 *Science* **282** 728
- [21] Feng G L, Dong W J, Jia X J 2004 *Chin. Phys.* **13** 413
- [22] Li J P, Chou J F 2003 *Chin. Sci. Bull.* **48** 703 (in Chinese) [李建平、丑纪范 2003 科学通报 **48** 703]
- [23] Gao L, Li J P 2007 *Adv. Earth Sci.* **22** 486 (in Chinese) [高丽、李建平 2007 地球科学进展 **22** 486]
- [24] Hou W, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 2664 (in Chinese) [侯威、封国林、董文杰、李建平 2006 物理学报 **55** 2664]
- [25] He W P, Feng G L, Gao X Q, Chou J F 2006 *Acta Phys. Sin.* **55** 3196 (in Chinese) [何文平、封国林、高新全、丑纪范 2006 物理学报 **55** 3196]
- [26] Gong Z Q, Feng G L, Dong W J, Li J P 2006 *Acta Phys. Sin.* **55** 3180 (in Chinese) [龚志强、封国林、董文杰、李建平 2006 物理学报 **55** 3180]
- [27] Feng G L, Hou W, Dong W J 2006 *Acta Phys. Sin.* **55** 962 (in Chinese) [封国林、侯威、董文杰 2006 物理学报 **55** 962]

Determining the threshold of extreme events with detrended fluctuation analysis^{*}

Yang Ping^{1,2)} Hou Wei^{1,2)†} Feng Guo-Lin^{2,3)}

¹ *College of Atmospheric Science, Lanzhou University, Lanzhou 730000, China*

² *Key Laboratory of Regional Climate-Environment for Temperature East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China*

³ *Laboratory for Climate Studies of China Meteorological Administration, National Climate Center, Beijing 100081, China*

(Received 14 November 2007; revised manuscript received 5 March 2008)

Abstract

When the dynamic system is in the extreme states, or it is affected by an intensive disturbance, we will get an extreme event. The detrended fluctuation analysis (DFA) measures the long-term correlation of a system, so the extreme events has no effect on system's long-term correlation. Based on this idea, we determined the threshold of extreme events with detrended fluctuation analysis and the results testify its effectiveness. By this method, we analyzed the extreme events in Beijing for the temperature and rain. The results show that before the metaphase of the 1970s, the frequency and intensity of the extreme temperature events is mitigatory, and from 1970s to 1990s, there were very few extreme events. But from the metaphase of the 1990s to now, the extreme temperature events are frequent. The extreme low temperature event has a downtrend from the 1950s to now and never happened after 1978s. The frequency and intensity of extreme precipitations are periodic. The frequency and intensity were high in the metaphase of 1950s and 1960s, and were low in the remaining period. The transition occurred in the 1970s.

Keywords : detrended fluctuation analysis, extreme event, threshold

PACC : 9260X

^{*} Project supported by the State Key Program of Science and Technology of China (Grant No. 2007BAC03A01).

[†] E-mail hou_w@sohu.com