

基于随机重排去趋势波动分析的 全国极端日降水事件综合指标

侯 威¹ 钱忠华² 杨 萍³ 封国林¹

(1. 国家气候中心, 北京 100081; 2. 兰州大学大气科学学院, 甘肃 兰州 730000;

3. 中国气象局气象干部培训学院, 北京 100089)

摘 要: 将去趋势波动分析法(Detrended Fluctuation Analysis, DFA)和替代数据法相结合,同时引入启发式分割算法和卡方检验,提出了一种确定极端气候事件阈值的新方法,称为随机重排去趋势波动分析(Stochastic resort detrended Fluctuation Analysis, S-DFA)方法。同百分位阈值方法相比,S-DFA方法明确指出了极端事件和非极端事件之间的临界值。利用随机重排去趋势波动分析(S-DFA)方法计算并分析了中国极端降水事件阈值的空间分布特征,并对S-DFA方法在实际资料中的应用进行了检验。基于极端降水事件综合指标将中国1961~2006年间极端降水事件分为3个不同等级的地区,进一步发现我国1961~2006年间极端降水的综合指标整体没有表现出明显的上升或下降趋势,保持平稳的波动变化。

关键词: 随机重排去趋势波动分析; 极端降水事件; 阈值; 综合指标

中图分类号: P467

文献标识码: A

引 言

随着非线性动力学在认识天气与气候系统过程和机制方面取得巨大进展^[1-3],“通过何种方式减少对极端天气气候事件认识上的不确定性”正成为目前国内外极端气候事件研究的重点之一。极端天气气候事件研究中最初的或最基本的不确定性来源于极端天气气候事件的定义。本文基于系统的整体动力学特征,将去趋势波动分析法和替代数据法相结合,得到一种具有物理背景和统一定义的确定的极端事件阈值的新方法,称为随机重排去趋势波动分析(S-DFA)方法,并通过实际数据对S-DFA方法进行了检验,验证了S-DFA方法的有效性。

发生频次和强度是极端事件的2个重要属性,通过深入研究这2个属性可以加深对极端事件本身及其背后物理机制和气候背景的认识^[4-7]。极端事件的发生频次和强度彼此之间也不是独立存在的,如果将这二者分隔开来、独立的加以分析研究,也就无法从整体上把握极端事件的变化特征。本文从一

个理想模型出发,通过考虑极端事件阈值、极端事件平均强度与阈值的差值及极端事件的发生次数,给出了极端事件综合指标的定义。在确定各站点极端降水事件阈值的基础上,讨论了极端降水事件的阈值分布、发生次数和综合指标的变化规律。

1 资料和方法

1.1 资料

所用资料为中国气象局公布的中国165个国际交换站1961~2006年无缺测逐日降水资料。为了更加突出图像的变化,本文图、表中涉及降水量的地方均乘以10。

1.2 随机重排去趋势波动分析(S-DFA)方法

趋势波动分析(DFA)方法^[8]是基于随机过程理论和混沌动力学新发展的用来检测时间序列的物理特征的一种分析方法。对某一长度的时间序列,去趋势波动分析法的具体过程及其物理意义可参见文献^[9-10]。Theiler和Prichard提出的替代数据法^[11-12]是为了获得与原始数据具有相同的均值、

收稿日期:2011-12-20; 改回日期:2012-04-06

基金项目: 国家自然科学基金(41005043)、全球变化研究国家重大科学研究计划(2012CB955901)和科技支撑项目(2009BAC51B04)共同资助

作者简介: 侯威(1981-),男,江苏扬州人,博士,高级工程师,主要从事气候变化及极端天气气候事件研究。E-mail: houwei@cma.gov.cn

方差和自相关函数以及概率分布函数的替代数据^[13-14]。

对系统演化的某一组时间序列 $\{x_i, i=1, L, n\}$, 首先得到 $\{x_i\}$ 的最大值 x_{\max} 和最小值 x_{\min} , 确定参考点 R , R 值可以是序列均值 x_{ave} 或者介于 x_{\max} 与 x_{\min} 之间的某一中值 x_{med} ; 然后从 x_{\max} 开始, 对序列中位于第 k 个数据区间内的数据点 $\{x_i, x_i \geq x_{\max} - d \times k\}$ 的顺序进行随机化, 其中 d 为区间间隔, d 的数量级取为序列 $\{x_i\}$ 数量级的 $[\frac{1}{10}, \frac{1}{100}]$ 之间, $\text{Int}[1, L, (x_{\max} - R/d)]$, $\text{Int}()$ 表示取整, 同时保留序列中其余数据的顺序不变, 直到 $x_i = R$, 依次得到新序列 $Y_j, j = x_{\max} - d \times k$; 再从 x_{\min} 开始, 对第 k 个数据区间内的数据点 $\{x_i, x_i \leq x_{\min} + d \times k\}$ 的顺序进行随机化, 同样 d 为区间间隔 $k = \text{Int}[1, L, (R - x_{\min})/d]$, 保留其余数据的顺序不变, 直到 $x_i = R$, 依次得到新序列 $Y_j, j = x_{\min} + d \times k$; 将 j 看作为各个重排区间下限或上限值, 最后计算每个新序列 Y_j 的长程相关性指数 DFA_j , 得到其随 j 的变化。

利用理想随机模型和混沌模型进行数值试验证明, 当重排区间内数据点很少时, 改变这些数据点顺序得到的新序列 Y_j 的 DFA 指数收敛于原始值, 即改变这些数据点的顺序对序列 DFA 指数影响不大或几乎没有影响, 由于这部分数据点的出现概率非常小, 其所包含的系统演化信息极少, 所对应的状态是系统演化的极端状态或是系统受到外界扰动而导致的极端异常状态, 属于小概率事件的范畴; 对于那些出现概率较大的数据点, 改变其顺序对序列 DFA 指数的影响也较大, 其包含了丰富的系统演化信息, 属于系统演化的常规状态, 不属于小概率事件的范畴。由于极端事件或者极值事件属于小概率事件, 同样可以认为此类事件所对应的演化状态是系统的极端状态或异常状态, 不属于系统自身正常演化状态的范畴。由前面的分析可知, 将 DFA 方法和替代数据法相结合, 通过确定序列 Y_j 的 DFA 指数何时开始收敛于原始值, 此时所对应的 j 值就是该序列小概率事件的阈值, 即系统极端事件的阈值, 所以可将上述算法运用于确定序列或系统极端事件或极值事件的阈值。下面给出具体实例及判断收敛点的方法。

1.3 方法示例

将北京 1961~2006 年逐日降水序列 $\{x_i, i=1, L, n\}$ 使用替代数据方法进行重排计算, 计算 DFA 指数时局部趋势函数 $y_e(i)$ 使用二阶多项式, 不重叠等

长度子区间长度 s 取值为 $250 \leq s \leq n/15$, n 为序列长度, 关于 s 取值的详细算法可参见文献^[15]; 取区间间隔 $d=1.0, R=0$ 。计算得到序列 $\{x_i\}$ 的 DFA 指数为 0.55215(图 1 中点线所示)。

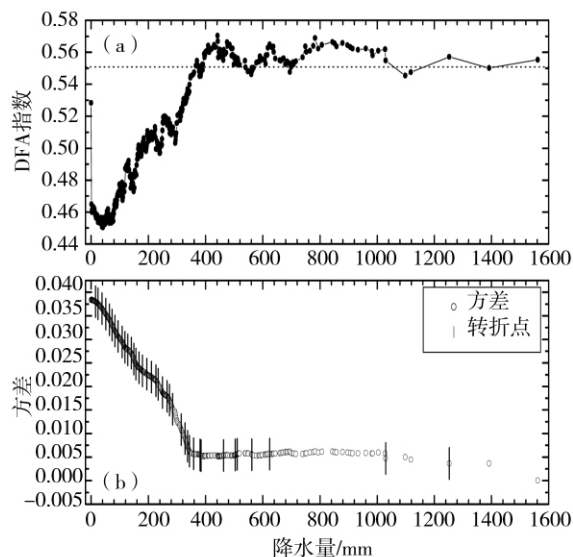


图 1 对逐日降水量序列的不同区间数据进行顺序随机化后 DFA 指数的变化(a), DFA 指数的方差及其突变点(b)

Fig. 1 The changes of DFA exponent by randomizing the sequence of data in different interval of precipitation(a), and variance and abrupt change point of the DFA exponent(b)

图 1 中的点线代表原始序列降水量序列 $\{x_i\}$ 的 DFA 指数。不同 Y_j 的 DFA 指数序列 $\{DFA_j\}$ 随 j 的演化逐渐收敛于原始序列 DFA 指数值, 由于随着 j 值逐渐变大 $\{DFA_j\}$ 的收敛并不是完全一致的逼近收敛形式, 而是围绕原始序列 DFA 指数值有着振幅非常小的波动, 因此通过计算 $\{DFA_j\}$ 的方差序列 $\{var_j\}$ 来消除这一波动对判断收敛点带来的影响。随着 $\{DFA_j\}$ 逐渐收敛于原始序列 DFA 指数值, 不同 DFA_j 值之间的差异逐渐减小, 其方差 $\{var_j\}$ 也逐渐减小最终收敛于 0 值。 var_j 值越小, 说明 $\{DFA_j\}$ 收敛程度越高, 反之亦然, 且 $\{var_j\}$ 与 $\{DFA_j\}$ 同时开始收敛。由于序列 $\{var_j\}$ 向着零值逐级收敛, 随着收敛程度的变化, 在此过程中 $\{var_j\}$ 存在一系列的转折点 $\{x_j\}$, 每个转折点表示 $\{var_j\}$ 的收敛程度在此点前后是不一样的, 同时这些转折点也代表 $\{DFA_j\}$ 向着原始序列 DFA 指数值收敛程度的变化。一种情况是 $\{var_j\}$ 值增大, 此时 $\{var_j\}$ 背离于收敛, 越来越远离收敛值; 另一种情况是 $\{var_j\}$ 值减小, 此时 $\{var_j\}$ 进一步收敛, 越来越接近收敛值。因

为要确定收敛点,所以只考虑序列 $\{var_j\}$ 中愈来愈收敛(即 $\{var_j\}$ 值越来越小)的区域即收敛区(如图1a内虚线框所示),在收敛区域内,所有的转折点均代表了序列 $\{var_j\}$ 向收敛值进一步逼近,收敛程度进一步提高。采用BG算法确定序列 $\{var_j\}$ 中的转折点 $\{x_j\}$ 。Bernaola-Galvan在2001年提出的启发式分割算法是一种有别于传统理论的突变检测方法,其详细算法过程可参见文献^[16]。

由于转折点 $\{x_j\}$ 中有些点所对应的DFA值远远偏离了原始序列DFA指数值,这些点不是收敛点,所以对收敛区域内的转折点采用总体方差显著性检验^[17-18](卡方检验)方法来检验各个转折点是否就是收敛点。如果在收敛区域内存在 $A \rightarrow F$ 一系列转折点,且已知A点不是收敛点和D点为收敛点,则收敛点D之前的转折点因为还没有达到收敛,其收敛程度将继续提高,由于各转折点的收敛情况不一致,所以转折点A、B、C之间应具有显著差异;而D点之后的转折点因为已经达到收敛,其收敛程度已经非常高且可以有余地非常小,各转折点的收敛情况一致,所以转折点D、E、F之间应无

显著差异。采用卡方检验对收敛区域内各个转折点之间的差异进行显著性检验,据此就可以确定出收敛点为D。这样就可以通过确定最后一组总体方差与样本方差有显著差异,也就是收敛情况有显著差异的转折点来确定收敛点。计算DFA指数序列的方差并用BG算法来得到转折点,通过卡方检验来确定收敛点,也就是极端降水事件的阈值。

以最后一个转折点 $x_{j=1252}$ 对应的方差为样本方差 σ_0^2 , $n_{j=1252}=3$;其后一个转折点 $x_{j=1030}$ 对应方差作为总体方差 s^2 ,卡方检验取显著性水平为99%,即 $\alpha=0.01$ 。在表1中,在点 $x_{j=285}$ 之后的各转折点之间也无显著差异;从转折点 $x_{j=337}$ 到 $x_{j=285}$ 之间的各个转折点彼此之间差异显著,收敛程度随J值的增大不断提高;从点 $x_{j=337}$ 开始直至最后一个转折点 $x_{j=1252}$,各点之间无显著差异;根据BG算法可以判断点 $x_{j=337}$ 与其后的数据点的变化情况是一致的,不存在转折,可以认为序列 $\{DFA_j\}$ 的收敛点即为 $x_{j=337}$ 逐日降水量序列 $\{R\}$ 的极端降水事件的阈值即为33.7 mm。

表1 各个转折点之间差异的显著性检验

Tab. 1 The test of significance of difference between turning point

转折点	n	χ^2	$\chi^2_{(a/2)}$	$\chi^2_{(1-a/2)}$	差异是否显著
$x_{j=1252}$	3	$\chi^2_{1030 \ 1252} = n_{j=1252} s^2 / \sigma_0^2 = 1.811$	12.8	0.071	否
$x_{j=1030}$	6	$\chi^2_{624 \ 1030} = n_{j=1030} s^2 / \sigma_0^2 = 4.139$	18.5	0.676	否
$x_{j=624}$	46	$\chi^2_{562 \ 624} = n_{j=251} s^2 / \sigma_0^2 = 48.883$	74.437	25.041	否
$x_{j=562}$	60	$\chi^2_{511 \ 562} = n_{j=562} s^2 / \sigma_0^2 = 57.702$	91.952	35.534	否
$x_{j=511}$	71	$\chi^2_{504 \ 511} = n_{j=511} s^2 / \sigma_0^2 = 75.301$	105.43	44.058	否
$x_{j=504}$	76	$\chi^2_{463 \ 504} = n_{j=504} s^2 / \sigma_0^2 = 79.620$	111.5	47.997	否
$x_{j=463}$	91	$\chi^2_{384 \ 463} = n_{j=463} s^2 / \sigma_0^2 = 91.073$	129.49	60.005	否
$x_{j=384}$	121	$\chi^2_{381 \ 384} = n_{j=384} s^2 / \sigma_0^2 = 114.637$	164.81	84.686	否
$x_{j=381}$	123	$\chi^2_{358 \ 381} = n_{j=381} s^2 / \sigma_0^2 = 113.64$	167.14	86.356	否
$x_{j=358}$	131	$\chi^2_{343 \ 358} = n_{j=358} s^2 / \sigma_0^2 = 106.718$	176.44	93.063	否
$x_{j=343}$	136	$\chi^2_{337 \ 343} = n_{j=343} s^2 / \sigma_0^2 = 101.632$	182.23	97.275	否
$x_{j=337}$	141	$\chi^2_{328 \ 337} = n_{j=337} s^2 / \sigma_0^2 = 93.855$	188	101.5	是
$x_{j=328}$	148	$\chi^2_{315 \ 328} = n_{j=328} s^2 / \sigma_0^2 = 100.596$	196.06	107.44	是
$x_{j=315}$	157	$\chi^2_{285 \ 315} = n_{j=315} s^2 / \sigma_0^2 = 79.977$	206.39	115.11	是
$x_{j=285}$	172	$\chi^2_{274 \ 285} = n_{j=285} s^2 / \sigma_0^2 = 135.487$	223.52	127.98	否

据确定的极端降水事件阈值,消除序列 $\{x_i\}$ 中的极值点或极端事件,也就是去除序列 $\{x_i\}$ 中 $x_i \geq 33.7$ 点,以达到消除极端事件的目的,得到新序列 $\{x'_i\}$,其中 $\{x'_i, x'_i = x_i - 33.7, \text{ if } x_i \geq 33.7\}$,对序列 $\{x'_i\}$ 采用同样的思路和算法进行计算,以检验对于一个不含极小值的序列,本方法是否仍然有效,即能确定该序列中无极端事件。发现收敛区内各个转折点的收敛程度彼此之间均无显著差异,收敛情况一致,序列 $\{DFA_j\}$ 和 $\{VAR_j\}$ 只有收敛的趋势,但未真正到达收敛时刻,不存在收敛点即极端事件的阈值,说明对去除极小值后的序列 $\{x'_i\}$ 使用本方法进行分析时,可以判定其不存在极小值事件。

2 极端降水事件阈值的确定

2.1 极端降水事件阈值的分布

图2给出了利用DFA-S方法得到的中国各地区极端降水事件的阈值,可以看出极端降水事件的阈值具有明显区域变化特征,整体表现出中部和东南部地区高,西北地区小的趋势。在中国的中部和东南部地区,这些地区的极端降水事件的阈值较高,均在100 mm以上;在阈值较高的区域中,一个大值中心出现在长江中下游和江淮地区,另一个大值中心位于华南南部。在广大的西南、东北、西北、河套和华北大部分地区,极端降水事件的阈值较低,均没有超过30 mm;西北地区的阈值最小,基本都<30 mm;其余地区的阈值介于30~100 mm之间。在东北西南部也有一个大值区;在四川的中部地区,极端降水的阈值也比周围地区偏高。

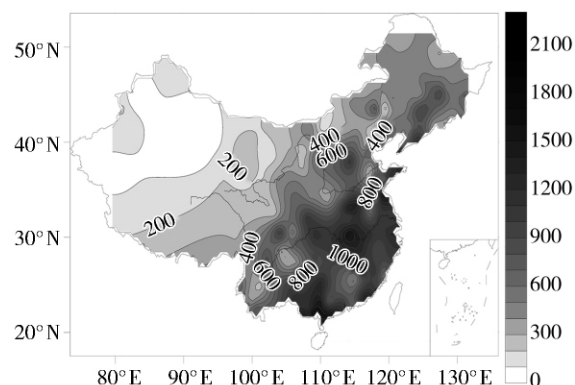


图2 中国日平均降水量极端降水事件阈值(单位: 10^{-1} mm)

Fig. 2 Spatial distribution of the threshold of extreme precipitation in China(Unit: 10^{-1} mm)

2.2 极端降水事件阈值有效性检验

为了验证阈值的有效性,将各站点降水序列 X_i 中 \geq 阈值的观测值全部减去阈值,即做如下处理:

$$\begin{aligned} X'_i &= X_i - J, \text{ if } X_i \geq J, \\ X'_i &= X_i, \text{ if } X_i < J \end{aligned}$$

使得处理后得到的新序列 X'_i 中不再包含有>阈值的数据点,不包含极端值,分别计算序列 X_i 和 X'_i 的DFA指数,如图3所示。可以看出,各站点的DFA指数在去除序列中极端值前后,不论是分布情况还是其大小并无差别,几乎完全一致,说明了使用DFA-S方法得到的阈值是有效的。

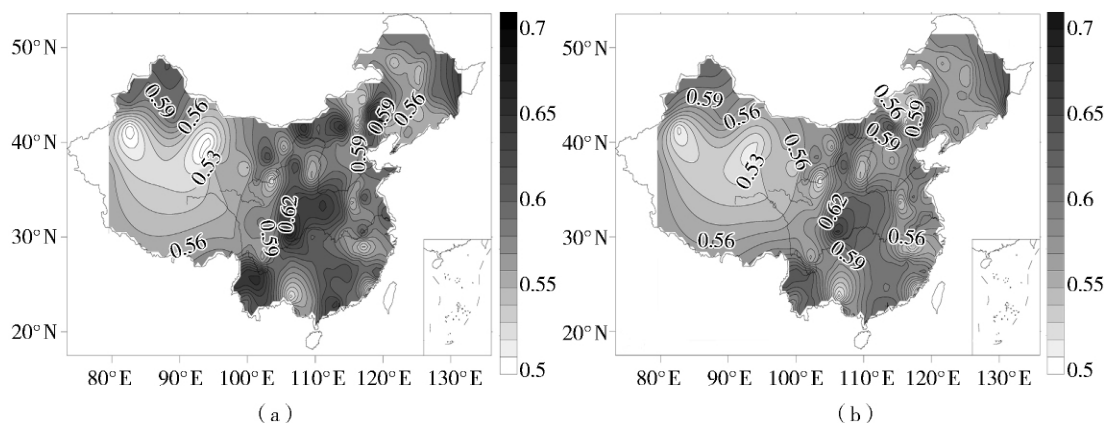


图3 各站点原始DFA指数(a),各站点去除极端值后DFA指数(b)

Fig. 3 The original DFA exponent of all stations(a), the DFA exponent of all stations after removing the extremum(b)

因为此时 X_i' 中已经不再包含极端降水事件, 所以应检测不出极端事件, 即要么不能得到极端降水事件的阈值或者得到的阈值很明显是一个错误值, 要么得到的阈值与使用原始数据得到的阈值一样, 以此来证明算法的合理性、有效性。图 4a 给出了对序列 X_i' 进行分析得到的极端降水事件阈值分布图, 将其同原始序列的极端降

水事件阈值进行比较, 可以看到, 使用序列 X_i' 得到的极端降水事件阈值与使用原始序列 X_i 得到的阈值, 二者的分布和值域大小几乎完全一致。在不包含极端降水值的序列中, 依然会得到与原始数据同样的阈值, 说明使用 S-DFA 算法得到的极端降水事件阈值是合理的、客观的, 也证明了本文算法针对实际资料的有效性。

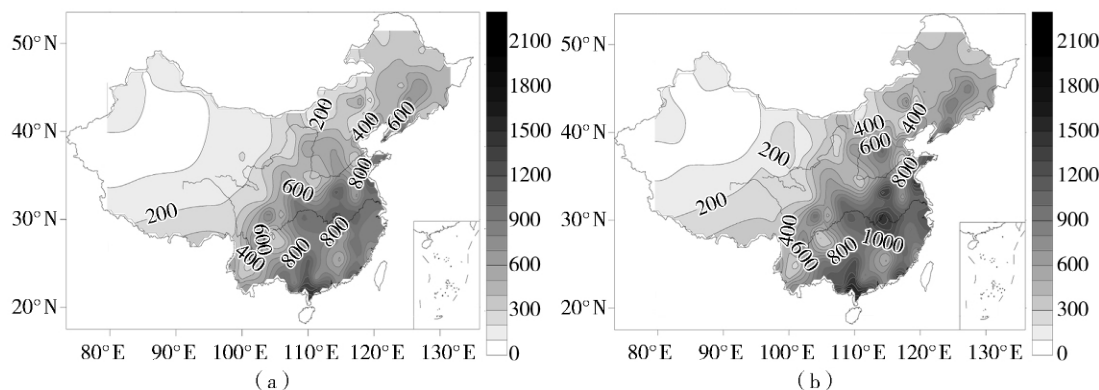


图 4 对序列 X_i' (a) 与原始序列 (b) 进行分析得到的极端降水事件阈值分布图(单位: 10^{-1} mm)

Fig. 4 spatial distribution of the threshold of extreme precipitation in series X_i' (a), Spatial distribution of the average precipitation in 1961–2006 (Unit: 10 mm) (b)

2.3 极端降水事件发生次数

根据确定的极端降水事件的阈值, 图 5 给出了各站点在 1961~2006 年间, 极端降水事件的发生次数。40 a 来中国大部地区极端降水事件的年均发生次数为约 1 次, 约占全部站点总数的 2/3; 其余 1/3 的站点发生次数较多, 年均发生次数超过 1 次。将图 5 和图 2 进行对比, 可以发现极端降水发生次数同极端降水阈值的分布情况并不一致, 阈值大的地区, 其发生次数未必就少, 反之亦然, 二者之间并无必然的联系。进一步对比发现, 极端降水发生次数较多的地区主要集中在 3 个区域内, 西南区、华北区和西北区, 在图 5 中分别以虚线框、实线框和灰色线框标出。从图 6 中可以看出在收敛点之后, 新序列的 DFA 指数序列确实收敛于原始序列的 DFA 指数, 改变小于这些收敛点的数据点在原始序列中的位置, 对系统的长程相关性的影响不大或几乎没有影响。各个站点极端降水事件的阈值的高低、发生次数的多少完全取决于各个站点数据本身, 也就是各个站点所具有的各自独特的区域气候系统。其余站点的图像同图 6 完全类似, 限于篇幅, 就不一一给出。

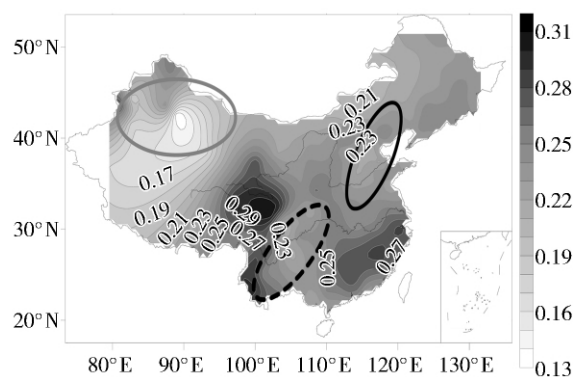


图 5 各站点 1961~2006 年间极端降水事件的发生次数(单位: 次数)

Fig. 5 Spatial distribution of the numbers of extreme precipitation in 1961–2006 of China

3 极端降水事件综合指标

3.1 极端降水事件综合指标的定义及其分布特征

确定用来定义极端降水事件综合指标的变量, 采用各站点的阈值 T_{thres} 、平均极端降水量与阈值的差值 T_{diff} 和极端降水发生次数 T_{NUM} 。这 3 个变量中

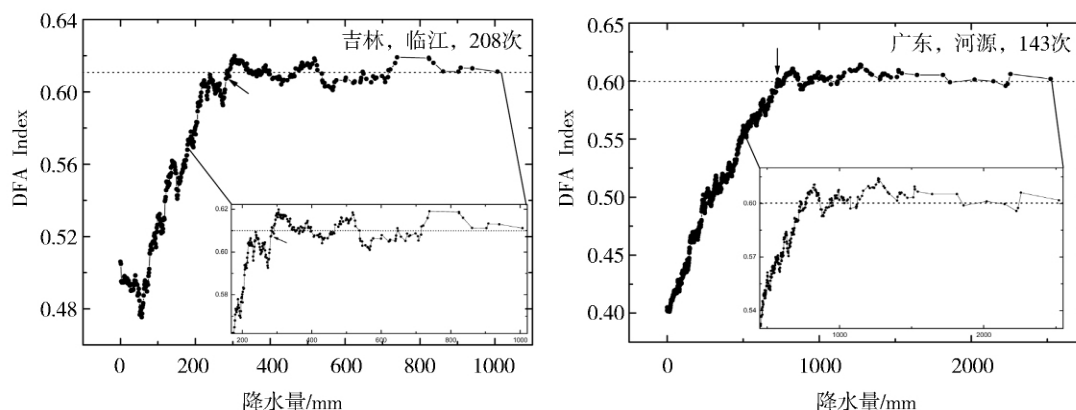


图6 发生次数较多的2个站点的DFA指数随替代数据区间的变化
(箭头所指为阈值处,虚线为原始序列的DFA指数)

Fig. 6 The DFA exponent change with surrogate intervals in Linjiang(a) and Heyuan(b)

的阈值包含了各个站点实际气候条件的信息,平均极端降水量与阈值的差值则代表了在各个站点实际气候条件的基础上,各个站点极端降水事件的大小程度。极端降水发生次数反映了极端降水事件的发生频率。为便于各站点之间的对比分析,将165站点的某个变量组成一个序列,采用粗粒化的方法将其转化成符号形式的等级序列,使每个站点都有一个属于自己的“等级数”,使用“熵最大化”^[19-20]方法来确定各个变量的“可分等级数”,取熵最大时所对应的划分的等级数目对3种变量进行粗粒化处理。粗粒化程度也就是“可分等级数”,确定每个站点各个变量的对应等级 RT_{thres} 、 RT_{diff} 和 RT_{NUM} 。确定等级之后,进一步进行归一化处理,使其值介于[0, 1]之间,得到这3种变量相对于其它站点的等级指数,等级指数越高则极端程度越大。进行粗粒化的等级处理并归一化以后,阈值和发生次数的等级指数的分布同处理前相比没有发生任何变化,且更有利于观察各变量的全国分布特征(图略)。

$$\dot{x}(t) = -\sigma x(t) + \sigma y(t)$$

Lorenz 方程组 $\dot{y}(t) = rx(t) - y(t) - x(t)z(t)$ 中的 $r = R_a/R_c$ 称为雷诺数; x 和对流 $\dot{z}(t) = x(t)y(t) - wz(t)$ 运动的强度成正比,称为速度模; y 正比于上升气流与下沉气流之间的温度差,称为温度模; z 正比于温度垂直廓线同线性分布的偏离程度,称为温度梯度模。如果实际极端降水事件可以用一个Lorenz方程组来描述,则可将极端降水事件的阈值看作下界面的温度,平均极端降水量看作上界面的温度,平均极端降水量与阈值的差值作为上升气流与下沉气流之间的差。阈值和对流运动成正比,对

流弱则加热弱,阈值也较低,反之亦然,阈值就相当于速度模;平均极端降水量与阈值的差值相当于温度模;极端降水发生次数本身代表了气候系统的降水分量偏离正常状态的多少,可以视为梯度模。这样就可以将阈值、平均极端降水量与阈值的差值和极端降水发生次数,这三者的等级指数看作是Lorenz方程组的3个分量。这样就可以将阈值、平均极端降水量与阈值的差值和极端降水发生次数的等级指数求和,给出极端降水事件综合指标的定义:

$$IDE = RT_{thres} + RT_{diff} + RT_{NUM},$$

IDE(Index of Extreme Events)代表了极端降水事件综合强度,其值越大,极端降水事件的综合强度越高,反之亦然。

图7为由上述定义得到的中国1961~2006年间,极端降水事件综合强度的空间分布情况,依据其综合强度的大小和区域分布情况,可以将中国1961~2006年来的极端降水事件分为3个不同等级的地区,最强的1级极端降水事件发生在华南南部、长江中下游、东南沿海和华北东南部以及四川盆地地区(图7中灰色线区域内,以字母A标出);3级强度的极端降水事件发生地主要集中在中国西北地区,覆盖了中国西北和北方的部分区域,包括新疆大部、青海、甘肃、内蒙西部和中部和东北北部(图7中黑色线区域内,以字母C标出);其余地区则属于极端降水的第2级地区(图7中以字母B标出),新疆西北部也有部分地区属于2级。总体而言,极端降水事件综合强度指标较好地体现了中国降水的特征,其中1级极端降水事件基本位于中国降水的2个主雨带上。

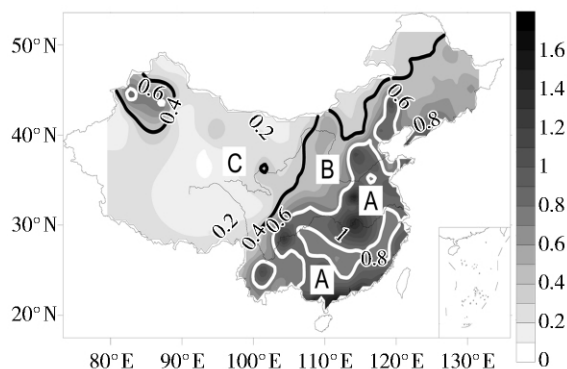


图7 极端降水事件综合强度等级的空间分布

Fig. 7 Spatial distribution of the composite index rank for extreme precipitation

3.2 极端降水事件综合指标的年际变化

图8给出了1961~2006年中国极端降水事件综合指标的年变化曲线。极端降水事件的年综合指标是将各年内中国165个国际交换站极端降水事件的综合指标求和,从而得到整个中国40 a来各年的综合指标。可以看出极端降水的综合指标整体没有表现出明显的上升或下降趋势,保持平稳的波动变化。在1970年代初期之前,其值偏高,总体大于综合指标的多年平均值;在1970年代初至1980年代末偏低,小于多年平均值;在1980年代末之后,综合

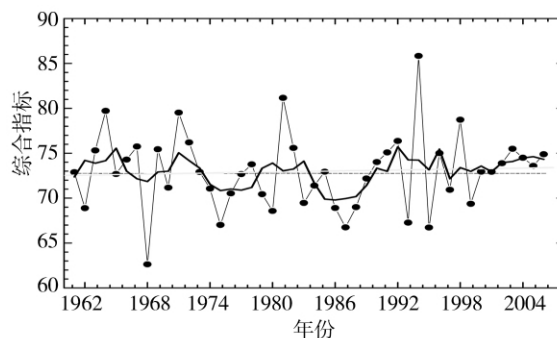


图8 极端降水事件综合指标的逐年变化
(虚线为平均值,点线为逐年变化,实线为5点平滑值)

Fig. 8 The changes of the composite index for extreme precipitation in 1961–2006 (The dashed line for the average, the dot line for yearly variation, the real line for the smoothing by five dots)

指标再次偏高,高于多年平均值。

3.3 极端降水事件综合指标的合理性

极端降水事件的综合指标将发生频次和事件强度综合起来,并兼顾了不同地区各自特有的气候状态或气候基准态。对于不同的地区而言,同一类极端事件的发生概率并不一样,同样次数、强度的某种极端事件发生在不同的地区,其“极端程度”也是大不相同的。从图9中可以看出,1967年极端降水事

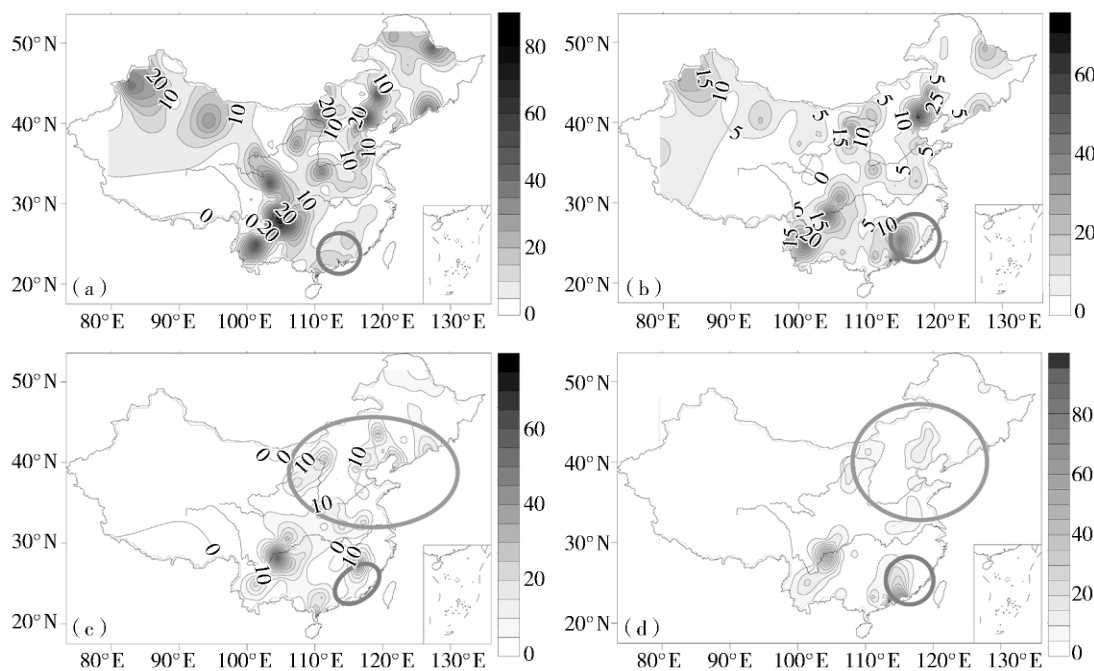


图9 1967年(a)和1968年(b)极端降水事件的发生次数分布(单位:次),
1967(c)和1968(d)极端降水事件的累计极端降水量分布(单位:mm)

Fig. 9 Spatial distribution of the numbers for extreme precipitation in 1967(a) and 1968(b) (Unit: times);
Spatial distribution of the accumulated extreme precipitation in 1967(c) and 1968(d) (Unit: 10^{-1} mm)

件的综合指标要远大于 1968 年。在 1967 年极端降水事件在 84 个站点出现, 共计发生 215 次, 累计极端降水量为 3 629.8 mm; 而在 1968 年极端降水事件在 76 个站点出现, 共计发生 218 次, 累计极端降水量为 4 020.9 mm, 可以看出, 1968 年中极端降水事件的发生站点数和发生次数都与 1967 年相差不多, 1967 年的累计极端降水量要小于 1968 年, 但 1967 年的综合指标却还是大大高于 1968 年。从图 9 中可以看出, 1967 年与 1968 年的极端降水事件发生次数的差别主要在华南南部, 其在 1968 年发生次数要远多于 1967 年, 在中国其余地区则相差不多, 区域分布也基本类似; 累计极端降水量的差别主要集中在华南南部和江淮、河套和华北北部地区, 其中华南南部在 1967 年和 1968 年累计极端降水量相差不多, 只是地区有所不同, 而在江淮、河套和华北北部地区在 1968 年非常小, 1967 年情况恰恰相反, 在该地区的累计极端降水量明显偏大。在前面已经分析, 中国南方的极端降水事件的发生次数和累计极端降水量在整个中国范围内属于中等偏上水平, 而中国北方则属于中等偏下的水平。江淮、河套和华北北部, 尤其是河套和华北北部地区由于其整体水平较低, 其次数和强度相对于整个中国区域内都偏强的年份显然更加“极端”, 导致了 1967 年极端降水综合指标要高于 1968 年, 也再次说明了本文极端事件综合指标的合理性、客观性。

4 结 论

基于中国气象局公布的中国 165 个国际交换站 1961~2006 年无缺测的逐日降水资料, 利用随机重排去趋势波动分析(S-DFA)方法计算并分析了中国极端降水事件阈值的空间分布特征, 并对 S-DFA 方法在实际资料中的应用进行了检验。中国极端降水事件阈值整体表现出中部和东南部地区高, 西北地区小的趋势。在阈值较高的区域中, 一个大值中心出现在长江中下游和江淮地区, 另一个大值中心位于华南南部。在广大的西南、东北、西北、河套和华北大部分地区, 极端降水事件的阈值较低; 西北地区的阈值最小。

基于“熵最大化”方法, 通过考虑极端降水事件的阈值、发生次数及极端降水事件平均强度与阈值的差值, 给出了极端降水事件强度综合指标的定义。这一综合指标将极端降水事件的发生频次和强度综合起来且兼顾了不同地区各自特有的区域气候背

景, 并以 1967 年和 1968 年的极端降水事件为例进一步说明了综合指标定义的合理性。综合指标代表了极端事件综合强度, 其值越大, 极端事件的综合强度越高, 其值越小则综合强度越弱。

基于极端降水事件综合指标的空间分布规律, 将中国 1961~2006 年间极端降水事件分为 3 个不同等级的地区, 最强的 1 级极端降水事件发生在华南南部、长江中下游、东南沿海和华北东南部以及四川盆地地区; 3 级强度的极端降水事件发生地主要集中在西北北部地区, 覆盖了中国西北和北方的部分区域, 包括新疆大部、青海、甘肃、内蒙西部和中部和东北北部; 其余地区则属于极端降水的第 2 级地区, 新疆西北部也有部分地区属于 2 级。1961~2006 年中国极端降水的综合指标整体没有表现出明显的上升或下降趋势, 保持平稳的波动变化。

参考文献:

- [1] Easterling D R, Evans J L, Groisman P Y, et al. Observed variability and trends in extreme climate events: a brief review [J]. Bull. of the Amer Mete Soci, 2006, 81(3): 417-425.
- [2] Redner S, Petersen M R. Role of global warming on the statistics of record-breaking temperatures [J]. Phys. Rev. E, 2006, 74, 061114, doi: 10.1103/PhysRevE.74.061114.
- [3] 齐冬梅, 李跃清. 高原季风研究主要进展及其科学意义 [J]. 干旱气象, 2007, 25(4): 74-79.
- [4] 胡宜昌, 董文杰, 何勇. 21 世纪初极端天气气候事件研究进展 [J]. 地球科学进展, 2007, 22(10): 1066-1075.
- [5] 钱维宏, 符娇兰, 张玮玮, 等. 近 40 年中国平均气候与极值气候变化的概述 [J]. 地球科学进展, 2007, 22(7): 673-684.
- [6] 王文, 张薇, 蔡晓军. 近 50a 来北京市气温和降水的变化 [J]. 干旱气象, 2009, 27(4): 350-353.
- [7] 韩晖, 孟丽霞. 甘肃省近 46a 年降水结构空间变化分析 [J]. 干旱气象, 2010, 28(1): 30-34.
- [8] Peng C K, Buldyrev S V, Havlin S. Mosaic Organization of DNA Nucleotides [J]. Phys. Rev. E, 1994(49): 1685-1689.
- [9] 杨萍, 侯威, 封国林. 基于去趋势波动分析方法确定极端事件阈值 [J]. 物理学报, 2008, 57(8): 5333-5342.
- [10] He W P, Feng G L, Wu Q, et al. A new method for abrupt change detection in dynamic Structures [J]. Non Proc. Geophys, 2008, 15(4): 601-606.
- [11] Theiler J, Linsay P S, Rubin D M. Time Series Prediction: Forecasting the Future and Understanding the Past [M]. Addison-Wesley: Reading Mass Press, 1993. 429.
- [12] Theiler J, Eubank S, Longtin A, et al. Testing for nonlinearity in time series: the method of surrogate data [J]. Physica D, 1992, 58: 77-94.
- [13] Kugiumtzis D. Surrogate data test for nonlinearity including non-monotonic transforms [J]. Phys. Rev. E, 2006, 62: 25-50.
- [14] Timmer J. What Can Be Inferred from Surrogate Data Testing [J].

- PhysRevLett, 2006, 85: 2647, doi: 10.1103/PhysRevLett.85.2647.
- [15] 侯威, 章大全, 杨萍, 等. 去趋势波动分析方法中不重叠等长度子区间长度的确定[J]. 物理学报, 2010, 59(12): 8986–8993.
- [16] Bernaola G P. Scale invariance in the nonstationarity of human heart rate[J]. Phys Rev Lett, 2001, 87: 168–105.
- [17] Plackett R L. Karl Pearson and the Chi-Squared Test[J]. International Statistical Review, 1983, 51(1): 59–72.
- [18] Greenwood P E, Nikulin M S. A guide to chi-squared testing [M]. Wiley, New York, 1996. ISBN047155779X
- [19] Wong A K C, Sahoo P K. A gray-level threshold selection method based on maximum entropy principle [J]. IEEE Trans SMC, 1989, 19(4): 866–871.
- [20] Pal N R, Pal SK. Entropy: a new definition and its application [J]. IEEE Trans SMC, 1991, 21(5): 1260–1270.

A Composite Index of Daily Extreme Precipitation in China Based on Stochastic Resort Detrended Fluctuation Analysis

HOU Wei¹, QIAN Zhonghua², YANG Ping³, FENG Guolin¹

(1. National Climate Center, Beijing 100081, China; 2. College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China; 3. Training Center of China Meteorological Administration, Beijing 100089, China)

Abstract: By combining detrended fluctuation analysis (DFA) method with surrogate data method, and using the heuristic segmentation algorithm as well as Chi-Square statistics, the stochastically re-sorting detrended fluctuation analysis (S-DFA) method was developed to define the threshold of extreme events. By using S-DFA method, we obtained the thresholds of extreme precipitation events from 1961 to 2006 in China and analyzed its spatial-temporal distribution characteristics. We also validated the effectiveness of S-DFA method through extreme events detection by using precipitation series. The composite index of extreme precipitation events was given in this paper, which integrated the information about frequency and strength of extreme precipitation events, considering the characteristic of regional climate system. Based on the composite index, we divided into three zones according to different precipitation rank of extreme precipitation events from 1961 to 2006 in China. The composite index of extreme precipitation maintained smooth fluctuation with no obvious increasing or decreasing trend on the whole during 1961–2006 in China.

Key words: stochastic resort detrended fluctuation analysis; extreme precipitation event; threshold; composite index