

利用中值检测方法研究近 50 年中国极端 气温变化趋势*

章大全^{1)†} 钱忠华^{1)‡}

1) 扬州大学物理科学与技术学院, 扬州 225009)

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) 国家气候中心气候研究开放实验室, 北京 100081)

(2007 年 11 月 19 日收到, 2008 年 1 月 12 日收到修改稿)

利用中国 194 个站点 1957—2001 年的逐日温度观测资料和月平均温度资料, 分别采用中值和均值两种极端事件检测方法, 检测并比较了中国极端高温事件和极端低温事件的变化趋势和空间分布特征. 检测表明, 年均极端高温的次数在近 50 年中趋于上升, 而年均极端低温的次数有所减少, 这与目前观测结果一致, 符合全球变暖的特点. 在空间分布上, 除西南地区部分站点外, 近 50 年中国大部分地区极端低温事件的年均发生次数趋于减少, 而极端高温事件发生频率的变化则呈现出东南沿海地区减少、西北内陆地区增加分布特点.

关键词: 极端事件, 全球增暖, 变化趋势, 中值

PACC: 9260X

1. 引 言

近百年来, 全球气候正经历一次以变暖为主要特征的显著变化. 其中, 全球变暖背景下极端气候事件的频率、强度的变化尤其引人注目. 已有的研究表明, 全球陆面温度的升高过程中多数地区的最低温度升高明显, 其变化幅度高于最高温度的升高, 因而表现出一种日夜增暖的不对称性, 使得日较差变小^[1, 2]. 政府间气候变化委员会第四次评估报告(IPCC, AR4)进一步指出, 自从 1951 年以来, 热日(hot days)和暖夜(warm night)显著增多, 而冷日和冷夜却减少了^[3]. 我国近几十年的日最高温度略有增加, 最低温度显著增加, 日较差明显变小. 在最近 40—50 年中, 部分区域极端最低温度和平均最低温度有明显上升, 尤其以北方冬季更为突出^[4, 5].

当天气的状态严重偏离其平均态时被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端变化研究中最常见的是采用某个百分位值作为极端值的阈值, 超过这个阈值的值被认为是极值, 并称为极端事件. 百分位值的确定一般需要了解气候要素的概率分布函数, 或对该气候要素的概率分布函数做出某种假设, 然后根据统计学和概率论相关知识给出便于实际操作的公式. 以温度为例, 研究表明温度时间序列的概率密度分布接近正态分布^[6], 通常将某一时间段内温度的平均值加上(或减去)标准偏差乘以一个系数作为极端气候事件的阈值, 一旦温度高于(或低于)这个阈值, 则认为发生了极端高温(低温)事件. 这种做法的缺陷在于计算均值时将极端事件即极端值也包括在内, 某种程度上掩盖了系统背景的真实信息, 从而可能导致某些极端事件无法检测到^[7].

月平均资料的分析可以提供极端气候事件强度和频率方面的低频信息. 通过对观测数据的分析和理论研究来揭示平均值和极端事件之间的关系, 由此可以从月平均资料分析得到的低频信息中推断出

* 国家自然科学基金(批准号 40675044)、国家重点基础研究项目特别基金(批准号 2006CB400503)和中国气象局气候研究开放实验室开放项目(批准号 LCS-2006-04)资助的课题.

† E-mail: z_daquan@126.com

极端气候事件发生频率和强度等的变化规律^[8]。但目前对月平均资料中的极端气候事件的强度、频率和持续时间等性质仍然研究较少^[9,10]。

本文采用中值和均值两种极端事件检测方法,分别对中国 194 个站点 1957—2001 年的日温度资料和月平均温度资料进行分析,并对两种不同检测方法得出的结果进行比较,以反映我国近 50 年极端高温事件和极端低温事件的变化规律及其空间分布情况。

2. 方法介绍

极端值包括极大值和极小值,下面以极大值为例分别介绍基于均值的极端事件检测方法和基于中值的极端事件检测方法,并对极小值的检测采取类似操作。以传统的极端事件均值检测方法为例,如果

$$x(i) > \bar{x} + z\sigma, \quad (1)$$

$i = 1, \dots, n$, 则 $x(i)$ 被认为是极端值。 \bar{x} 为样本均值, σ 为标准偏差。根据概率理论^[11], 取 $z = 1.28$, 当 $x(i)$ 符合正态分布时, 超过这个阈值的极端事件发生概率小于 0.1。

Hampel 提出一种新的极端值检测方法^[12], 如果

$$x(i) > \text{MED}(x) + z\text{MAD}(x), \quad (2)$$

则 $x(i)$ 被认为是极端值。样本 $x(i)$ 长度为 n , $\text{MED}(x)$ 和 $\text{MAD}(x)$ 分别是原始序列 $x(i)$ 和序列 $\{ |x(i) - \text{MED}(x)|; i = 1, \dots, n \}$ 的中值。中值的定义如下^[12]。如果将观测值按递增的顺序排列, 中值 M 就是这些数据的中心点, 一半的观测值在中值之下, 一半在中值之上。将数据排列为 $x_1 \leq x_2 \leq \dots \leq x_n$, 那么中值就可以用下式计算:

$$M = \begin{cases} \frac{x_{\frac{n+1}{2}}}{2}, & n \text{ 是奇数,} \\ (\frac{x_{\frac{n}{2}}}{2} + \frac{x_{\frac{n}{2}+1}}{2})/2, & n \text{ 是偶数.} \end{cases} \quad (3)$$

均值和中值都用来度量序列分布的中心位置。均值对于一些特别高的数据值很敏感, 但中值却不受这一变化的影响, 因为中值只取决于多少数据高于或低于它, 而不取决于高多少或低多少。因此当序列中的极端值较多时, 中值能够更好地反映系统的背景状态。 z 为选定的参数, 如果 z 取值过小检测结果包含过多的虚假信息, 反之 z 过大则将导致检测得到的极端值过少。Hampel 使用蒙特卡罗模拟实验得到^[13], 当原始序列符合正态分布时 z 取 $1.92 \approx 1.28 \cdot 0.67$, 超过该阈值的极端事件发生概率小

于 0.1。

为了消除时间序列 $x(i)$ 中趋势项的影响, 本文采用滑动窗口中检测极端事件的方法^[14]。时间序列 $\{x(i), t(i); i = 1, \dots, n\}$, $t(i)$ 单调递增, 如果

$$x(i) > \text{MED}_{j=i-k}^{j=i+k}(x(j)) + z\text{MAD}_{j=i-k}^{j=i+k}(x(j)), \quad (4)$$

$i = k + 1, \dots, n - k$ 则 $x(i)$ 为检测到的发生在时间 $t(i)$ 的极端值。 $\text{MED}_{j=i-k}^{j=i+k}(x(j))$ 是滑动窗口 $x(j)$ 的中值 (窗口长度为 $2k + 1$, $k \leq (n - 1)/2$), 作为原始序列趋势项的估计值; $\text{MAD}_{j=i-k}^{j=i+k}(x(j))$ 是窗口中原序列减去趋势项结果的中值。对极端事件的均值检测方法采取同样的处理, 如果

$$x(i) > \frac{1}{2k + 1} \sum_{j=i-k}^{j=i+k} x(j) + z\sigma_{j=i-k}^{j=i+k} x(j), \quad (5)$$

$i = k + 1, \dots, n - k$ 其中 $\sigma_{j=i-k}^{j=i+k} x(j)$ 表示窗口内子序列的标准偏差, $x(i)$ 为检测到的发生在时间 $t(i)$ 的极端值。

3. 数值实验

为了检验极端事件中值检测方法的有效性, 并对基于中值和均值的两种极端事件检测方法进行比较, 我们构建理想时间序列 $x(i)$, 并分别应用这两种方法进行检测。时间序列 $t(i)$, $x(i)$ 图 1(a)(b) 中的细实线。生成规则如下: $t(i) = i, i = 1, \dots, 300$; $x(i) = 5 + \epsilon(i) + 3s(i)$, 其中 $\epsilon(i)$ 为区间在 0 到 1 之间均匀分布的随机数, $s(j) = \sin(2\pi(t(j) - 100)/100), j = 100, \dots, 200$, 当 j 小于 100 或者大于 200 时, $s(j) = 0$ 。在时间序列中 $x(i)$ 加入 18 个极端值, 分别是 $x(20) = 10, x(22) = 12, x(24) = 9, x(50) = 15, x(55) = 8, x(60) = 10, x(100) = 16, x(120) = 19, x(130) = 15, x(140) = 20, x(145) = 18, x(175) = 17, x(180) = 13, x(185) = 15, x(200) = 14, x(220) = 15, x(240) = 13, x(260) = 18$ 。

从图 1 可以看出, 中值检测方法能检测出所有 18 个极端事件, 而均值检测方法仅检测出 13 个, 尤其当几个极端事件前后间隔很短时, 均值检测方法往往只能检测出其值最大的极端事件。当某一时段内有多极极端值连续出现, 即极端事件存在群发现象时, 均值检测方法受极极端值影响较大, 无法准确反映气候的平均状态, 从而造成部分极极端值的缺测现象。换言之, 当气候要素序列中的极端事件存在较

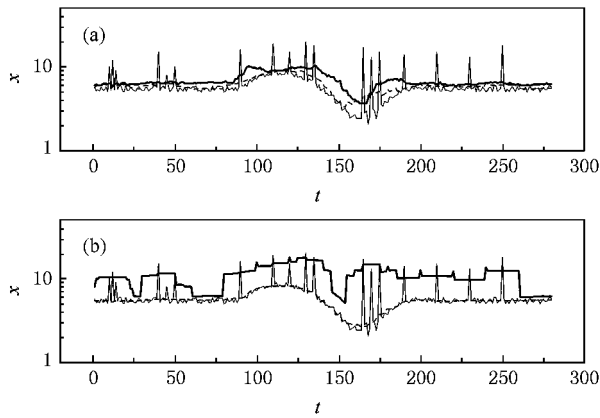


图 1 理想时间序列中极端值的检测(细实线为理想序列 $x(i)$, 虚线分别为滑动窗口的中值序列和原序列的滑动平均值 ($k=10$)粗实线为使用上述两种方法确定的极端事件阈值) (a)(b)分别为极端事件的中值和均值检测方法得到的结果

强的群发现象时,均值方法检测到的极端气候事件的次数往往偏小,中值方法则弥补了这一方面的不足,因此对中值方法和均值方法结果差异的分析一定程度上可以反映极端事件的群发现象的显著程度.

4. 实际资料分析

所用资料为中国 194 测站逐日和逐月的最高气温、最低气温.资料来源为国家气象信息中心发布的中国均一化历史气温数据集,该数据集为中国 194 个国际交换站观测统计资料,时间跨度为 1957 年 1 月至 2001 年 12 月.对缺测较多或者不满 45 年连续观测的站点加以剔除,经检验有 182 个站点满足上述要求,对仍存在的个别缺测使用插值方法补足.

4.1. 全国平均极端高温和极端低温变化趋势

图 2 是分别使用极端事件的中值和均值检测方法检测全国平均年极端高温和极端低温事件发生次数变化趋势.使用中值方法分析逐日资料的结果表明,年均极端高温和极端低温次数无显著变化,近 40 年上升或下降幅度不超过 1 天.逐月资料的结果则显示,年均极端高温的次数在过去几十年中趋于上升,上升幅度约为 0.1 次/10 a,而年均极端低温的次数有所减少.

均值方法的分析结果显示出不同的变化趋势.逐日资料的结果显示,年均极端高温和极端低温次

数都有所减少,下降幅度约为 0.5 d/10 a.逐月资料得到的结果则显示年均极端高温和极端低温事件的发生次数都在上升,而在全球变暖背景下,全球大部分地区呈现出极端高温增多、极端低温减少的变化事实^[15].这可能是由于均值方法在极端事件检测中未能真实地反映系统的背景信息,检测结果包含了过多的虚假分量所致.

从中值和均值方法的结果对比来看,分析逐日资料时,无论年均极端高温次数还是极端低温次数,中值方法检测结果都比均值方法高 3 次左右,而在处理逐月资料时两种方法得到的极端事件发生次数无明显差异.这说明中值方法具有较强的检测和识别极端事件的能力,同时由于月资料是对日资料的平均,平滑了一些高频信息,因此对于低频信息两种方法的检测结果等价.

4.2. 中值方法和均值方法结果的对比分析

为进一步分析中值和均值两种不同的极端事件检测方法检测结果的异同性,这里对每个站点使用中值方法分析逐日资料得到的年平均极端事件发生次数减去均值方法的结果,并给出两种方法得到的年均极端事件发生频率差异的空间分布情况.

从年均极端高温发生次数来看,较之传统均值方法,使用中值方法得到的年均极端高温次数普遍要多 3 至 4 天,但不同地区的差异较大.两种方法得到的年均极端高温发生次数差异最大的地区为四川盆地,达到 6 天以上,说明该地区的极端高温事件可能存在较强的群发性.

年均极端低温次数差异最大的区域为东北和华北北部地区,而西南地区两种方法得到的结果几乎相同.从图 3(b)可以看出,差异较大地区主要为受冷空气南下影响显著的地区,而西南地区由于受到青藏高原的阻挡,受冷空气影响较弱,极端低温事件的群发性并不明显.

4.3. 极端高温变化趋势的空间分布

从空间分布来看,西北地区年均极端高温次数普遍上升,在河西走廊和新疆南部表现为大于 1 次/10 a 的明显增加趋势(图 4(a)),在华东、华南及四川盆地则表现为减少趋势,大部分测站年均极端高温次数减少幅度为 0.5 次/10 a;均值方法的检测结果与中值方法的检测结果在空间分布上基本一致,差

异仅仅是均值方法检测的年均极端高温次数上升的幅度相对较小(图 4(b)).除东北地区外,极端高温变化趋势的空间分布与任国玉等得到的年平均气温变化趋势的空间分布^[16]有着相近的分布特征,从侧面说明东北地区的增温可能主要是温度平均值的上升造成的.

中值检测结果分析表明,逐月资料得到的年均极端高温次数上升区域主要分布在河套和江淮地区,上升幅度在河套地区超过 0.2 次/10 a,年均极端高温次数减少的区域集中在华北北部和东北北部地区,新疆、四川盆地和东南沿海地区年均极端高温次数变化趋势并不明显(图 4(c)).

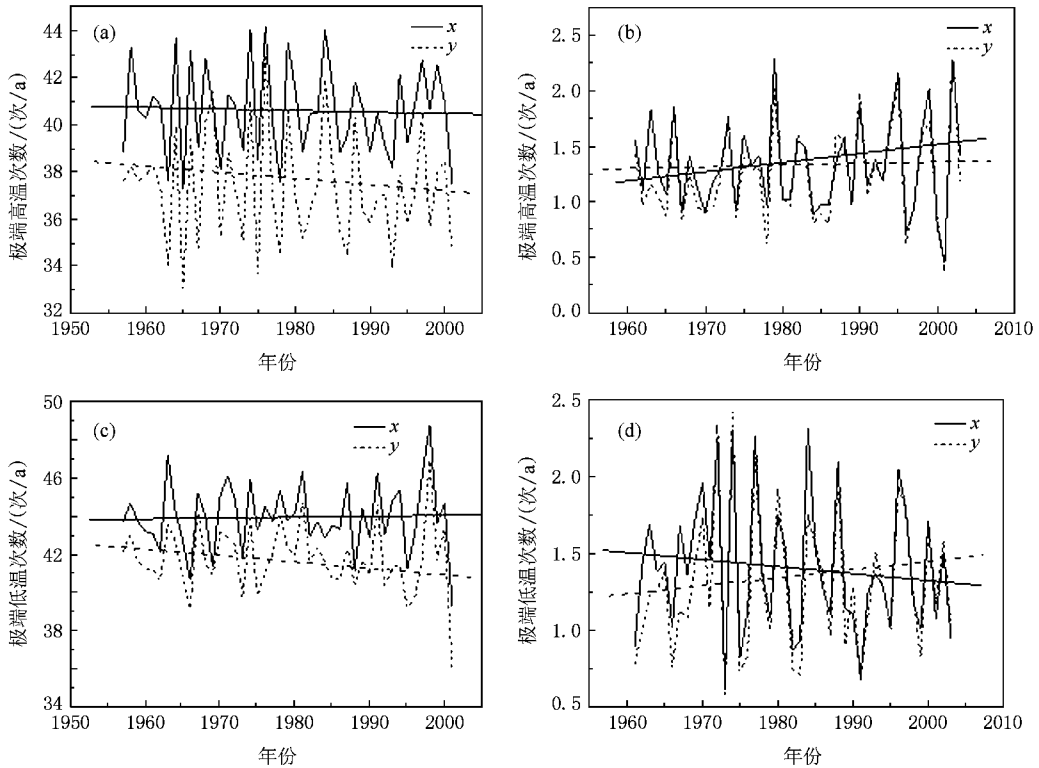


图 2 全国平均极端高温和极端低温事件变化趋势(实线 x 为中值方法的检测结果,虚线 y 为均值方法的检测结果,斜线为线性趋势)(a)分析逐日资料得到的年平均极端高温发生频率变化(b)分析逐月资料得到的年平均极端高温发生频率变化(c)分析逐日资料得到的年平均极端低温发生频率变化(d)分析逐月资料得到的年平均极端低温事件发生频率变化

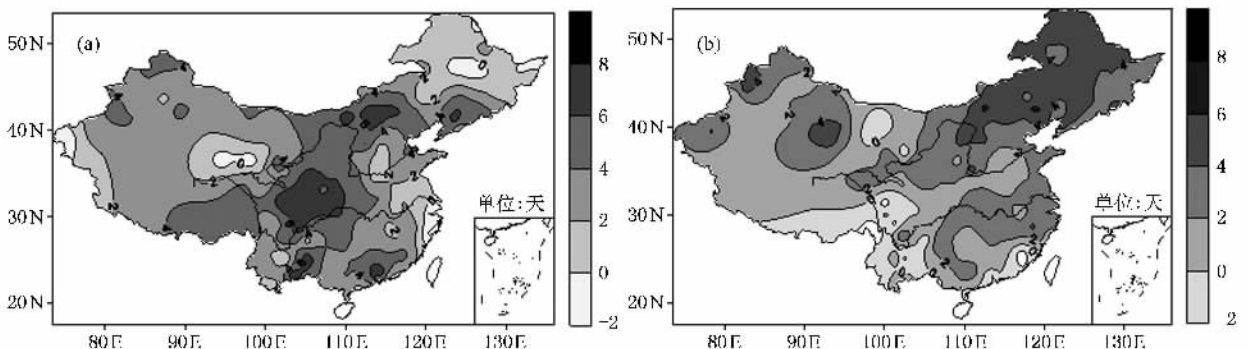


图 3 中值方法和均值方法结果对比(a)中值方法与均值方法分析逐日资料得到的年均极端高温事件发生次数的差异(b)中值方法与均值方法分析逐日资料得到的年均极端低温事件发生次数的差异

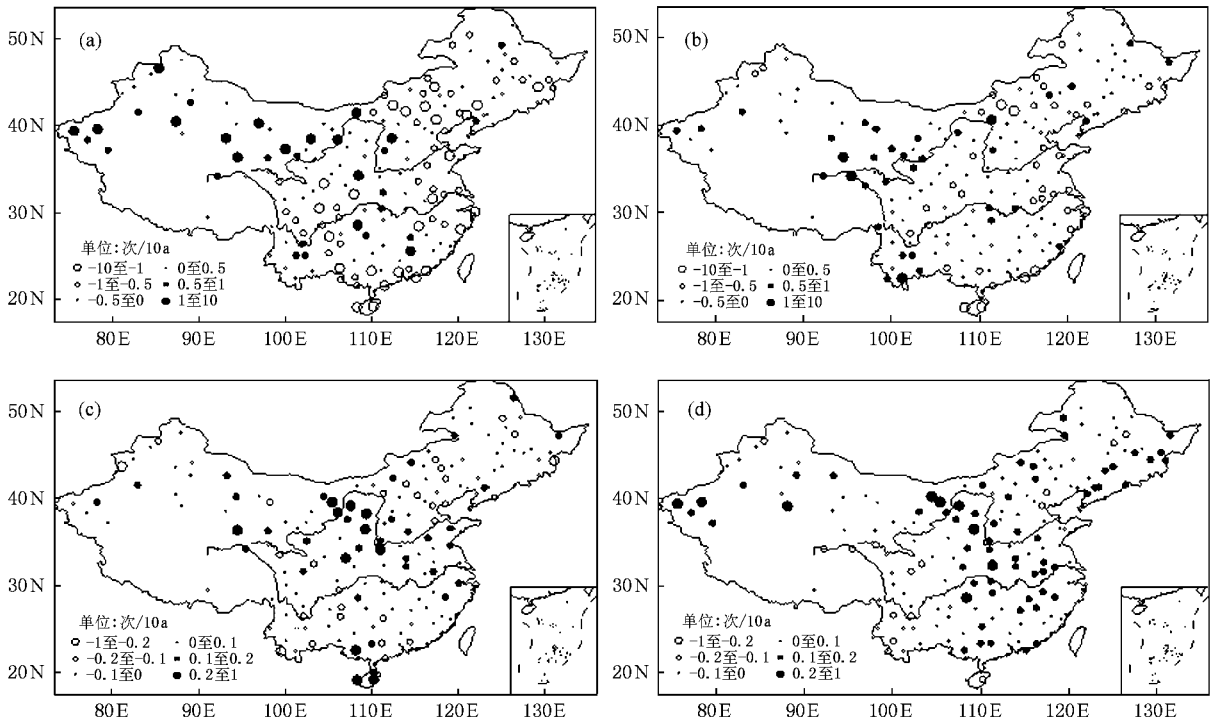


图 4 极端高温变化趋势的空间分布 (a) 中值方法分析逐日资料得到的年平均极端高温事件发生频率变化的空间分布 (b) 均值方法分析逐日资料得到的年平均极端高温事件发生频率变化的空间分布 (c) 中值方法分析逐月资料得到的年平均极端高温事件发生频率变化的空间分布 (d) 均值方法分析逐月资料得到的年平均极端高温事件发生频率变化的空间分布

均值方法的检测结果显示(图 4(d)), 全国大部分地区近几十年的年均极端高温次数都增加了, 这种增加趋势在河套地区和长江中下游地区最为显著. 与逐日资料检测结果相比, 逐月资料检测结果中极端高温事件分布比较集中, 可能是由于逐月资料反映的是气候变化的低频信息, 序列的平稳性较好所致.

4.4. 极端低温变化趋势的空间分布

从逐日资料的中值检测结果来看(图 5(a)), 总体上年均极端低温次数呈现出东南沿海地区减少、西北内陆地区增加的趋势, 均值方法得到的结果则显示出南方地区年均极端低温次数增加, 北方减少的区域分布特征. 长江以南除东南沿海部分地区外, 大部分站点的年均极端低温次数增加幅度超过 0.5 次/10 a (图 5(b)).

逐月资料的中值检测结果(图 5(c))显示, 除西南地区部分站点外, 近几十年中国大部分地区极端低温事件的年均发生次数趋于减少, 其中华北北部和东北南部减少趋势最为明显, 部分站点的减少幅度超过 0.2 次/10 a.

使用均值方法分析逐月资料的结果表明(图 5(d)), 华北、西北和四川盆地部分地区年均极端低温

次数呈较弱的减少趋势, 而华南中部、东北西部和内蒙古北部地区年均极端低温次数上升趋势明显, 其中华南中部部分站点上升幅度超过 0.2 次/10 a.

5. 结 论

本文采用中值和均值两种极端事件检测方法, 分别对中国 194 个站点 1957—2001 年的日温度资料和月平均温度资料进行分析, 并对两种检测方法得出的结果进行比较, 得到我国近 50 年极端高温事件和极端低温事件的变化及其分布情况, 具体结论如下.

1. 从全国平均来看, 中值方法分析逐日资料的结果显示, 近几十年年均极端高温次数和极端低温次数无明显变化趋势, 而均值方法检测到极端高温和极端低温的次数都有所下降, 但总体而言下降趋势并不明显. 中值方法分析逐月资料的结果表明, 年均极端高温的次数在过去几十年中趋于上升, 而年均极端低温的次数有所减少. 均值方法的结果则显示出年均极端高温和极端低温的次数都在上升. 通过与观测事实的对比分析发现, 中值检测方法更能有效地反映极端气候事件的真实变化信息.

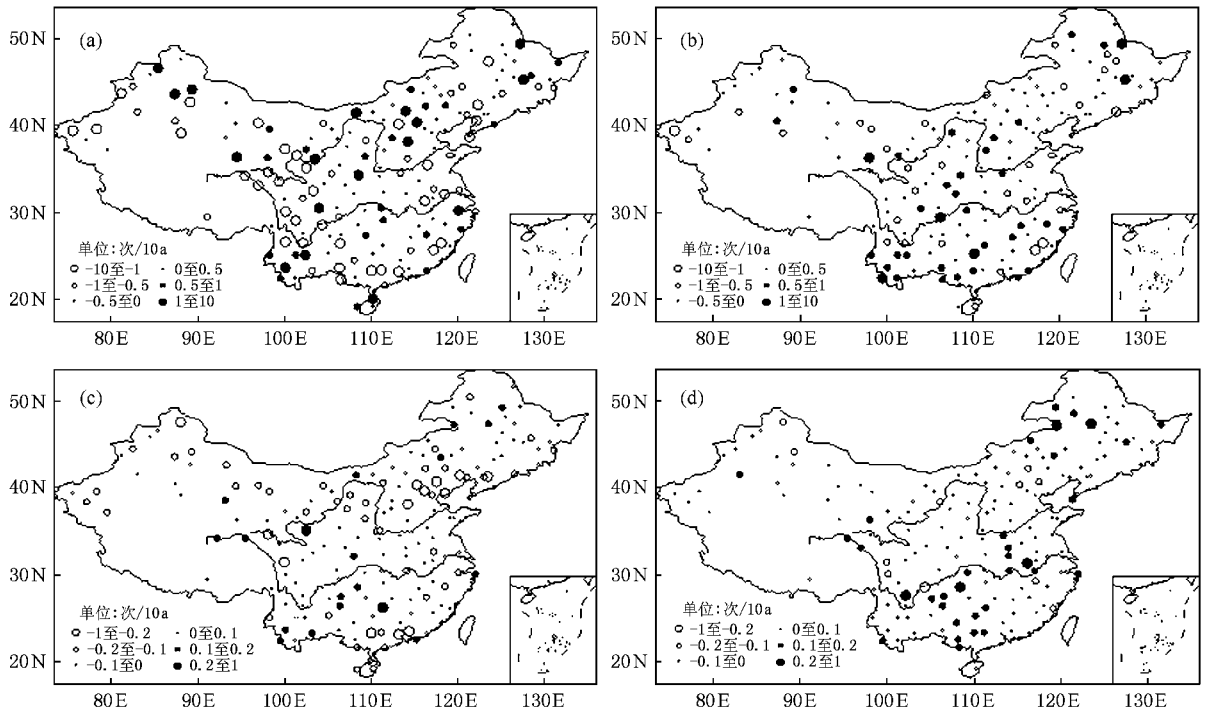


图 5 极端低温变化趋势的空间分布 (a)中值方法分析逐日资料得到的年平均极端低温事件发生频率变化的空间分布 (b)均值方法分析逐日资料得到的年平均极端低温事件发生频率变化的空间分布 (c)中值方法分析逐月资料得到的年平均极端低温事件发生频率变化的空间分布 (d)均值方法分析逐月资料得到的年平均极端低温事件发生频率变化的空间分布

2. 两种极端事件检测方法得到的结果对比表明,较之传统均值方法的结果,使用中值方法得到的年均极端高温次数普遍要多 3—4 天,但不同地区的差异较大。两种方法得到的年均极端高温发生次数差异最大的地区为四川盆地,年均极端低温次数差异最大的区域为东北和华北北部地区,而西南地区两种方法得到的结果几乎相同。

3. 西北地区年均极端高温次数普遍上升,在河西走廊和新疆南部表现出超过 1 次/10 a 的明显增加趋势,在华东、华南及西川盆地则表现为减少趋势,幅度约为 0.5 次/10 a。均值方法的结果在空间分布上与中值方法基本一致,而年均极端高温次数上升的幅度相对较小。逐月资料的分析结果显示,年均

极端高温次数上升区域主要分布在河套和江淮地区,年均极端高温次数减少的区域集中在华北北部和东北北部,新疆、四川盆地和东南沿海地区年均极端高温次数变化趋势不明显。

4. 整体上年均极端低温次数呈现出东南沿海地区减少、西北内陆地区增加的趋势,均值方法得到的结果则显示出南方地区年均极端低温次数增加,北方减少的空间分布特征。长江以南除东南沿海部分地区外,大部分站点的年均极端低温次数增加幅度超过 0.5 次/10 a。逐月资料的分析结果显示,除西南地区部分站点外,近几十年中国大部分地区极端低温事件的年均发生次数趋于减少,其中华北北部和东北南部减少趋势最为明显。

- [1] Karl T R, Kukla G 1991 *Geophys. Res. Lett.* **18** 2253
 [2] Karl T R, Jones P D, Knight R W 1993 *Bull. Amer. Meteor. Soc.* **74** 1007
 [3] Solomon S, Qin D H, Manning M, Alley R B, Bertsen T 2007 *Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the IPCC Fourth Assessment Report* (Cambridge, United Kingdom and New York, NY, USA: Cambridge University

Press) 316

- [4] 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]
 [5] Feng G L, Gong Z Q, Dong W J, Li J P 2005 *Acta Phys. Sin.* **54** 5494 (in Chinese) [封国林,龚志强,董文杰,李建平 2005 物理学报 **54** 5494]
 [6] Render S, Mark R P 2006 *Phys. Rev. E* **74** 1539

- [7] Lanzante J R 1996 *Inter. J. Clim.* **16** 1197 (Beijing : Science Press) p5 (in Chinese) [张仁铎 2005 空间变异理论及应用 (北京 : 科学出版社) 第 5 页]
- [8] Thomas C P , Francis Z , Albert K T 2002 IPCC *Workshop on Changes in Extreme Weather and Climate Events Breakout Group 1 : Temperature* (Beijing : China Meteorological Press) p9
- [9] Feng G L , Dong W J 2003 *Chin. Phys.* **12** 1076
- [10] Feng G L , Dong W J , Li J P 2004 *Chin. Phys.* **13** 1582
- [11] Tian Y H , Yang N , Leng Z K 2002 *Probability and Statistics* (Beijing : High Education Press) p43 (in Chinese) [田应辉、阳妮、冷志魁 2002 概率论与数理统计 (北京 : 高等教育出版社) 第 43 页]
- [12] Zhang R D 2005 *Theory and Application of Spatial Variability* (Beijing : Science Press) p5 (in Chinese) [张仁铎 2005 空间变异理论及应用 (北京 : 科学出版社) 第 5 页]
- [13] Hampel F R 1985 *Technometrics* **27** 95
- [14] Manfred M 2006 *Computers & Geosciences* **32** 141
- [15] Ding Y H , Reng G Y 2007 *National Assessment Report of Climate Change* (Beijing : Science Press) p33 (in Chinese) [丁一汇、任国玉 2007 气候变化国家评估报告 (北京 : 科学出版社) 第 33 页]
- [16] Ren G Y , Xu M Z , Chu Z Y , Guo J , Li Q X , Liu X N , Wang Y 2005 *Acta Meteor. Sin.* **63** 943 (in Chinese) [任国玉、徐铭志、初子莹、郭军、李庆祥、刘小宁、王颖 2005 气象学报 **63** 943]

Analysis of extreme events in China's temperature in recent 50 years using detecting method based on median ^{*}

Zhang Da-Quan^{1,2,†} Qian Zhong-Hua^{1,2,3}

¹ *College of Physics Science and Technology , Yangzhou University , Yangzhou 225009 , China*

² *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 19 November 2007 , revised manuscript received 12 January 2008)

Abstract

Based on daily and monthly temperature records from 1957 to 2001 of China , using climate extreme detecting method based on average and median , we have analyzed the transition and distribution of frequency of temperature extremes . The results indicate that while the average number of extreme high temperature events per year has been increased in the last 50 years , the extreme low temperature events decreased . Except for some stations located in southwest China , the frequency of extreme low temperature events in most regions over China decreased . The spatial pattern of transition of frequency of extreme high temperature events showed decreased in southeast coastal regions and increase in inland regions .

Keywords : extreme events , global warming , trend , median

PACC : 9260X

^{*} Project supported by the National Natural Science Foundation of China (Grant No. 40675044) , the State Key Development Program for Basic Research of China (Grant No. 2006CB400503) , and the Laboratory for Climate Studies of China Meteorological Administration Climate research Program (Grant No. LCS-2006-04) .

[†] E-mail : z. daquan@126.com