

doi:10.3969/j.issn.1672-0504.2019.03.008

时空聚集性探测方法在极端高温事件 聚集分析中的应用研究

张 婷^{1,2,3,4}, 程昌秀^{1,2,3,4*}, 杨山力^{1,2,3,4}, 宋长青^{1,3,4}

(1. 北京师范大学地表过程与资源生态国家重点实验室, 北京 100875; 2. 北京师范大学环境演变与自然灾害教育部重点实验室, 北京 100875; 3. 北京师范大学地理科学学部, 北京 100875; 4. 北京师范大学地理数据与应用分析中心, 北京 100875)

摘要: 极端高温是灾害性气候的研究热点之一, 进行其时空聚集性探测对极端高温灾害的深入研究与危险性评价至关重要。传统温度阈值法从超越概率角度进行极端事件提取, 无法实现时空一体化分析, 缺少对时空聚集程度的量化评价。该文引入扫描统计量, 构造时空柱体扫描窗口对研究区进行时空一体化粗探测, 通过相对危险度定量评价聚集程度, 最后对探测结果进行局部 Moran's *I* 热点分析, 结合热点置信度优化聚集区的空间定位。以中国大陆地区 2013 年 7 月的温度数据为例, 采用上述方法, 探测出研究区的 6 个显著时空聚集区, 相对危险度最高为 3.99, 最低为 2.60, 范围覆盖华北、华中、华东、西北的部分地区, 最后以当年的气象资料及报道验证探测结果。结果表明, 时空扫描和局部 Moran's *I* 分析的集成是一种有效探测极端高温事件时空聚集区的方法, 不仅能实现极高温事件准确可靠的时空定位, 还能给出极高温区域的时空聚集指标。

关键词: 极端高温事件; 时空扫描统计量; 局部 Moran 指数; 时空聚集性; 温度阈值

中图分类号: P423 文献标识码: A 文章编号: 1672-0504(2019)03-0051-07

0 引言

科技进步、工业发展在给人类生活提供极大便利的同时, 也引发了一系列严重的气候环境问题, 其中, 极端高温事件对社会的影响尤为显著^[1], 且未来极端高温事件发生的频率和强度可能会持续增加^[2,3]。围绕这类灾害事件的研究逐渐成为灾害性气候研究的热点之一。

极端高温事件是指在一定区域内长时间持续性的温度超过限制温度的极端天气事件^[4]。目前, 国内外学者基本采用以下 3 种方法对事件进行提取: 1) 固定阈值法, 即人为给定一个温度值作为极端温度的判断标准^[5]。陈敏等^[6]认为连续 3 天以上温度超过 35 °C 的事件为极端高温事件, 张可慧等^[7]、肖嗣荣等^[8]也选用 35 °C 为高温日的判别标准, 但该方法没有考虑极端值的空间差异, 只适用于范围小且环境差异小的研究区。2) 百分比阈值法, 即以温度序列的某一百分位值为温度阈值^[9,10]。杨萍等^[11]以 90% 设定阈值, 分析极端温度事件频次的年际变化趋势; 游庆龙等^[12]以 90% 设定阈值, 分析极端温度、极端降水日数的年际变化趋势; 严晓瑜等^[13]以 95% 设定阈值, 认为东北大部分地区年极端高温随时间

呈递增趋势。该方法虽然表征出极端值的区域差异, 但其分析集中于高温事件在时间尺度上的演变问题。3) 基于基准温度与变异值的极端温度事件检测法, 即通过比较事件温度数据与基准温度的变异值进行极端温度事件检测^[14]。Chen 等^[4]认为基于异常值分析的极端温度事件检测方法可通过提取瞬态异常显著改善事件检测结果, 但该方法仍是基于不同区域温度时间序列进行高温事件探测, 仍未打破时空独立探测的常规思路。目前的提取方法虽可实现较大时空尺度上的极端高温定位, 但在时间演变上人为划分整体时间范围(1 a, 5 a, 10 a 等)、割裂连续时间, 且传统统计学中各探测单元相互独立, 无法对事件进行时空一体化探测, 只由温度阈值进行极端温度事件判别, 无额外指标定量评价聚集程度, 显著性分析仍停留在定性层面。

时空扫描统计量在解决上述问题中有优势。该方法中各探测单元不独立, 可以实现扫描窗口在整个研究区时空一体化探测, 同时定量给出聚集性指标, 但该方法目前主要集中应用在流行病学^[15,16]、社会经济学^[17,18]等领域。因此, 本研究将时空扫描统计量引入高温的时空相关性研究中, 试图在传统温度阈值法提取结果的基础上, 通过构建时空柱体的

收稿日期: 2018-04-01; 修回日期: 2018-06-04

基金项目: 国家自然科学基金面上项目“基于历史事件的全球海啸灾害时空分异规律及风险评估”(41771537)

作者简介: 张婷(1994-), 女, 硕士研究生, 研究方向为灾害类数据分析。* 通讯作者 E-mail: chengcx@bnu.edu.cn

扫描窗口,将事件的时间属性和空间属性放在同一坐标系下进行时空一体化粗探测,计算最大似然比、相对危险度;分别用于确定聚集窗口、定量评价窗口聚集程度;再对探测结果进行空间局部自相关性热点分析,根据以往研究经验,选用适用范围更广、频率更高^[19,20]的局部 Moran's *I* 指数进行局部热点探测,结合热点置信度优化聚集区的空间定位。该方法可以实现时空一体化的极端高温事件探测,并对聚集区的聚集程度进行定量评价。

1 实验数据与研究方法

本研究的数据主要包括:1)中国省级行政区划,用于界定研究区域(18.16°~53.56°N,73.45°~135.09°E);2)1979—2013年研究区温度网格数据,即从网站 <http://apps.ecmwf.int/datasets> 获取的 ECMWF Re-Analysis Interim (ERA-Interim) 全球温度数据产品,空间分辨率为 0.5°×0.5°,时间分辨率为 6 h。2013 年高温现象显著,截至 8 月上旬,中国高温天气日数已达 55 d,为 60 年来之最^[21]。因此,本研究以 2013 年 7 月中国大陆地区的温度数据为例,进行极端高温事件的时空聚集性探测。

时空聚集性探测方法由 Naus 提出,之后广泛用于二、三维数据聚集性探测^[22],通过构建时空三维柱体的探测窗口进行时空统计分析,并构建统计探测量(对数似然比、相对危险度)对柱体窗口内外的聚集程度进行计算,可以实现高温事件时空聚类的分布分析及定量评价。该方法需预先确定极端高温点,并将其作为聚集性探测的数据输入。为体现空间差异对极端值的影响,本文选用百分比阈值法确定高温阈值,提取极端高温点,进而使用时空扫描分析法对高温点进行时空聚集区粗探测;由于时空扫描提取出的空间聚集区以一定半径的圆形区域表示,对事件空间位置描述的弹性度小,因此,需对上述粗探测结果进行局部 Moran's *I* 分析^[23,24],优化空间位置。该方法基于探测单元之间的相互关联提取高温区域,结果以探测单元的集合表示,而加入局部 Moran's *I* 分析,可以精确化极端高温事件在某一时间累积区间内的空间定位,完善时空聚集性探测结果。最后以当年的气象资料及报道对探测结果进行验证。技术路线见图 1。

2 极端高温时空聚集区探测

2.1 高温阈值确定与极端高温点提取

高温阈值确定:以 1979—2013 年的逐日最高温

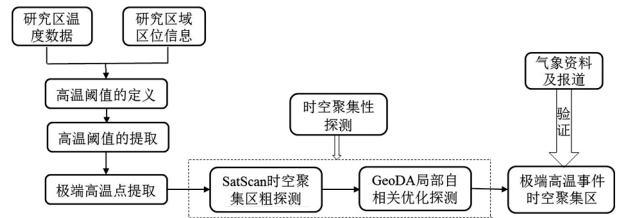


图 1 极端高温时空聚集性探测流程
Fig. 1 Flow chart of extreme high temperature spatio-temporal integration scanning

度数据作为高温阈值的累积数据,阈值的百分位确定直接影响极端高温事件的危险水平及提取数量。本研究涉及范围大、数据累积年份长,为保证极端高温事件提取的有效性,参考以往长时间、大范围的温度阈值提取方法^[13,25],对各单元点日最高温度序列递增排序,将第 95 个百分位的数值定义为该点的高温阈值,将阈值低于 25 °C 的单元点修正为 25 °C,保证高温阈值均不低于 25 °C^[26](图 2,彩图见附录 2)。

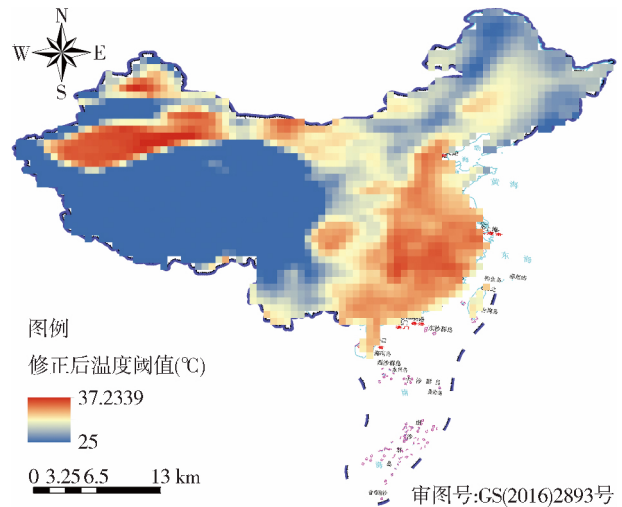


图 2 修正后研究区高温阈值
Fig. 2 Threshold of extreme high temperature in the study area after correction

极端高温点提取:选用研究区 2013 年 7 月逐日最高温度数据,当待测单元点 (*i, j*) 的日最高温度等于或超过该点高温阈值,则 *k* 日的该单元点被记录为极端高温。极端高温点 T_M 满足式(1),当 $T_M \geq 0$,该点发生极端高温事件,否则不发生极端高温事件。

$$T_{M,i,j,k} = t_{i,j,k} - T_{i,j} \quad (1)$$

式中: $t_{i,j,k}$ 为 (*i, j*) 处、*k* 日最高温度; $T_{i,j}$ 为 (*i, j*) 处高温阈值; $T_{M,i,j,k}$ 为 (*i, j*) 处、*k* 日的极端高温点数据。

2.2 极端高温时空聚集区粗探测

对上述极端高温点用时空扫描的方法进行时空聚集区粗探测。首先,确定扫描区域并在区域内设定一个大小、位置可变的聚集性探测窗口。扫描窗口为底面代表空间聚集区、高代表时间聚集段的时空柱体,且随时间、空间不断扩大,直至窗口设定的

上限(图 3)。其中,空间扫描窗口内的事件数不得超过探测事件总数的 50%;时间扫描窗口的最大值不得超过总时间序列长度的 90%,时间窗口下限设定为 3 d(高温事件需至少累积发生 3 天)。最后,基于假设检验的思想,构建检验扫描的统计量:对数似然比(Log Likelihood Ratio, LLR)、相对危险度(Relative Risk, RR),利用 LLR 对聚集区进行判定,RR 对聚集区进行评价。

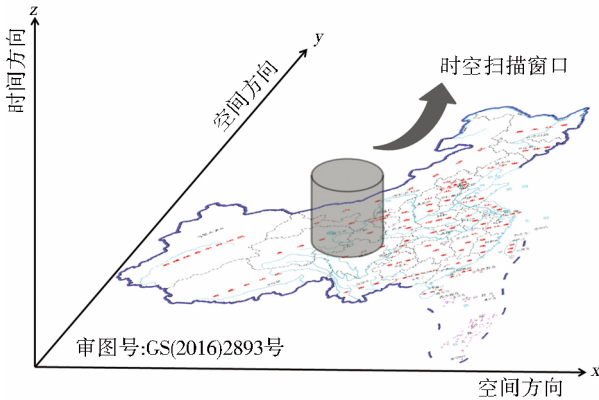


图 3 时空扫描探测方法示意
Fig. 3 Schematic diagram of spatio-temporal scanning method

扫描统计量的构建基于事件的概率分布,本研究中各单元极端高温点只有高温与否两种情况,且整个研究区的单元点构成事件全集,故选用离散型伯努利分布构建扫描统计量。提出无效假设 H_0 : 极端高温事件在时间、空间上呈现完全随机分布,有似然函数 L_0 :

$$L_0 = \prod p(x_i) = (N/M)^N (1 - N/M)^{M-N} \quad (2)$$

式中: $p(x_i)$ 为事件点的概率; N 为研究区极端高温点总数; M 为研究区温度记录总数。

备择假设 H_1 : 极端高温事件在探测窗口内存在时空聚集性分布,有似然函数 L_1 :

$$L_1 = \prod p(x_i) = p^n (1-p)^{m-n} q^{N-n} (1-q)^{(M-m)-(N-n)} \quad (3)$$

式中: n 为扫描窗口中极端高温点数; m 为扫描窗口中温度记录总数; $p = n/m$ 为窗口内极端高温点比例; $q = (N-n)/(M-m)$ 为窗口外极端高温点比例。

根据式(4)得到 LLR, 表征窗口内事件聚集的可能性, 即在该空间位置(圆心坐标及圆的半径)和时间区间的 LLR 越大, 发生高温聚集的可能性越高。存在时空聚集的窗口需满足 $LLR > 0$ 且窗口内极端高温点比例 ($p = n/m$) 大于窗口外极端高温点比例 ($q = (N-n)/(M-m)$), LLR 最大值对应的窗口为聚集可能性最大的窗口。最后, 根据式(5)对各时空聚集窗口计算 RR, 用以衡量时空窗口内相较于时空窗口外的相对高温事件数, 评价高温事件的持续聚集性强弱。

$$LLR = \log\left(\frac{L_1}{L_0}\right) = \ln[p^n (1-p)^{m-n} q^{N-n} (1-q)^{(M-m)-(N-n)}] - \ln[(N/M)^N (1-N/M)^{M-N}] \quad (4)$$

$$E(m) = \frac{N}{M} \times m, RR = \frac{n/E[m]}{(N-n)/(E[M]-E[m])} \quad (5)$$

式中: $E[m]$ 为基于无效假设的窗口内极端高温点数; $n/E[m]$ 表征窗口内事件比例; $(N-n)/(E[M]-E[m])$ 表征窗口外事件比例。

为保证时空扫描结果更具统计意义, 本文基于扫描统计的模型假设(伯努利模型), 使用蒙特卡罗法进行 999 次模拟数据集实验, 计算置信度 P 以评价 LLR 的置信度水平。即使用处理真实数据集的方法, 对 999 个模拟数据集进行 LLR 值的计算, 然后对真实数据集的 LLR 和模拟数据集的 LLR 进行从小到大排序, 如果真实数据集的 LLR 排第一位, 则 $P = 1/(1+999) = 0.001$ 。本研究认为, 当 $P \leq 0.001$ 时具有统计意义。

2.3 极端高温时空聚集区优化探测

对粗探测结果中置信度 P 符合要求的聚类加入局部 Moran's I 分析, 改善极端高温事件空间定位, 此过程的核心在于进行分析的温度数据必须为粗探测聚类各时间累积区内的高温点积温数据。空间位置提取结果与空间权重矩阵的设定有关^[27], 积温数据中事件点的空间位置分布规律、温度事件数量多, 选用基于空间邻接概念的 Rock 权重矩阵, 即仅有公共边界的邻接关系才定义为邻接。

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nm} \end{bmatrix} \quad (6)$$

式中: n 为积温事件个数; w_{ij} 表示积温事件 i 与 j 的邻接关系, 当且仅当积温事件 i, j 符合 Rock 临界条件时 $w_{ij} = 1$, 否则 $w_{ij} = 0$ 。

对高温点积温数据计算局部空间自相关指标(LISA)。根据式(7)对第 i 个积温点计算局部自相关指标 I_i , 反映一个积温点与邻近区域积温点的空间聚集程度。

$$\begin{cases} S^2 = \frac{1}{n} \sum_{i=1}^n (t_i - t)^2 \\ I_i = \frac{t_i - t}{S^2 \sum_{j=1}^n w_{ij}} \sum_{j=1}^n w_{ij} (t_j - t) \\ Z_{I_i} = (I_i - E[I_i]) / \sqrt{V[I_i]} \end{cases} \quad (7)$$

式中: S 为积温标准差; t_i, t_j 分别表示 i, j 处高温点积温值; t 为研究区高温点积温均值; Z_{I_i} 服从标准正态分布。

根据各积温点的 LISA 绘制散点图,分析积温事件的局部空间不稳定性。Moran 散点图分为 4 个象限,分别表示 4 种局部空间相关关系。其中,第一象限(HH)表示积温的高值区域单元被同是高值的积温区所包围,空间差异程度小;同时,落在散点图 HH 与对应粗探测圆形窗口内的热点为该积温时段精确化空间聚集区,形成新的时空聚类。

通过将局部自相关指标 I_i 标准化计算 Z_{I_i} ,对应得到每个高温积温点局部自相关指标 I_i 的显著性 P ,本文认为当 P 值不超过 0.001 时具有统计意义。至此完成了极端高温事件时空聚集性探测。该方法与传统极端温度事件统计方法有本质上的差异:传统方法从超越概率角度对各探测单元进行相互独立的探测;空间统计学在分析这类带有时间、空间属性

的数据时,考虑到了数据的时空关联作用,因而可以对时间、空间进行统一坐标系下的分析,同时构建特定指标对结果进行定量评价。

3 结果分析

3.1 时空扫描粗探测结果

以空间窗口半径(0~700 km)、时间窗口长度(3~27 d)为时空扫描窗口进行粗探测,采用横轴墨卡托投影提取研究区 2013 年 7 月极端高温事件的时空聚类分布(图 4)。可以看出,研究区内极端高温事件并非随机分布,而是呈现出明显的时空聚集特征,有 7 个具有统计意义的极端高温事件时空聚集区,可以根据其时间累积区段划分为 6 类。

由具体的扫描结果(表 1)可知,聚类 1 的极端高

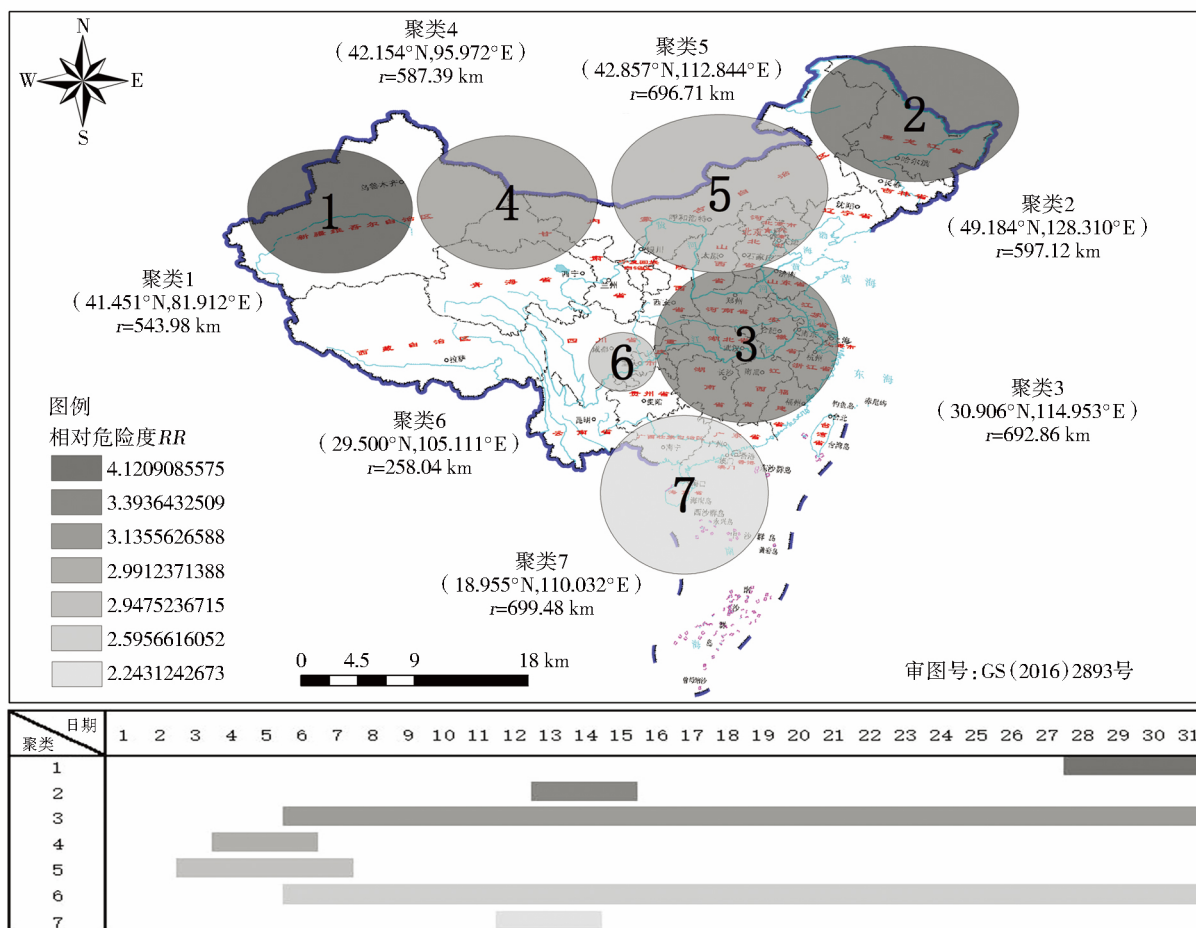


图 4 2013 年 7 月 1—31 日的 SatScan 时空扫描结果
Fig. 4 SatScan spatio-temporal scanning results on 1-31 July, 2013

温聚集风险是其他地区的 4.12 倍,相对危险度最高,分布于 28—31 日的新疆西部地区,聚类中心为(41.451°N,81.912°E),半径为 543.98 km;聚类 2 的极端高温聚集风险是其他地区的 3.39 倍,相对危险度较高,分布于 13—15 日的黑龙江省、内蒙古西北部等地区,聚类中心为(49.184°N,128.310°E),半

径为 597.12 km;聚类 3 的聚集风险是其他地区的 3.14 倍,相对危险度较高,分布于 6—31 日的华中、华东大部分地区,涵盖了 11 个省及上海市,聚类中心为(30.906°N,114.953°E),半径为 692.86 km,是探测中的一级聚类区。聚类 1 和聚类 2 虽属于二级聚类区,但其相对危险度更高,属于聚集性最为显著

(聚类 7, $RR=2.24$) 未探测出结果。由此, 时空聚集性探测方法给出 6 个具有统计意义的时空聚集区(聚类 1-6)。根据其时间累积区划分为 5 类, 其中 3-7 日的聚集高温主要发生在北京、天津、河北和山西的大部以及内蒙古的部分地区, 相对危险度中等(2.95); 4-6 日的聚集高温发生在甘肃与内蒙古和新疆的交汇处, 相对危险度中等(2.99); 6-31 日的聚集高温分为两个聚类, 上海、江苏、浙江、湖南和湖北等地区相对危险度较高(3.14), 四川、重庆及贵州的交汇处极端高温的相对危险度较低(2.60); 13-15 日的聚集高温主要发生在黑龙江省中部, 相对危险度较高(3.39); 28-31 日的聚集高温主要发生在新疆西南部, 相对危险度最高(4.12)。时空聚集性探测方法实现了极端高温事件时空一体化分析, 且对各聚类区均有置信度和显著性指标定量评价聚集程度。

由中国气象局网站(<http://www.cma.gov.cn/>) 2013 年 7 月发布的数据统计 7 月各省会城市高温情况, 结果显示, 重庆高温发生于 24-31 日, 长沙于 1-31 日, 杭州于 14-31 日, 上海于 20-31 日, 福州于 28-31 日, 武汉、合肥、南京于 23-31 日, 郑州于 25-31 日, 与聚类 1、聚类 5 结果基本吻合。南疆盆地南部部分地区于 30 日、31 日持续高温, 与聚类 2 结果基本吻合。天津于 3 日、24 日、28 日、北京于 2-5 日、石家庄于 2-7 日持续高温, 与聚类 3 结果基本吻合。由于温度数据来源不同, 且气象局统计资料只包括持续超过 35°C 的高温数据, 只能对提取范围做到相对验证, 但也反映出本研究方法在探测高温事件的优势。

4 讨论与展望

本文时空扫描中利用离散型伯努利分布计算似然函数值, 因气温存在时间和空间的相关性, 使得本方法存在理论缺陷。目前尚未提出一种顾及时空聚集性的概率计算方法, 因而模型中暂用传统方法。时空扫描模型的处理过程可以在一定程度上消除这种偏差。例如, 本文的研究对象为极端高温事件, 不直接使用相关性较大的温度数据可以在一定程度上消除研究对象的时空相关性。因而, 本文将特定时间、特定位置的极端高温点是否发生极端高温事件看作是一次伯努利试验 E , 研究区在整个研究时段的所有极端高温点可看作是将伯努利试验 E 独立重复地进行 n 次的集合, 认为不同时间、不同位置的极端高温点是否发生极端高温是相互独立的。实验结果证明, 时空扫描和局部 Moran's I 分析的集成是一

种有效探测极端高温事件时空聚集性的方法。其中, 时空扫描统计量对确定极端高温的时间累积区有明显优势, 局部 Moran's I 分析对确定空间聚集区的精确划分也有明显优势, 二者有机结合不仅能对高温事件进行准确可靠的时空定位, 还能定量评价高温事件区域的时空聚集程度。

该方法中置信度是很重要的评价指标: 在提取极端高温点时, 本文设定置信度为 95% (即温度阈值设定标准); 对探测结果进行筛选和过滤时, 本文设定显著性 P 均不超过 0.001。此外, 有些时空参数的设置也值得重点关注: 在时空扫描阶段, 如何设置合理的扫描时空窗口; 在进行局部 Moran's I 分析阶段, 空间权重矩阵的设置也很关键。

上述研究方法不仅适用于极端高温的回顾性探测, 同时可以根据研究区近日的气象数据以及次日的气象预报数据, 大致判断次日极端高温危险区域, 并给出其时空聚集强度指数, 从而进行高温预警。该方法还可以推广至极端低温事件, 甚至其他点数据的时空模式分析中, 如自然资源、地震等灾害类数据等。本研究基于伯努利(0-1)分布对日尺度数据进行探测, 采用不同的概率分布(如泊松分布)可将研究推广到季尺度、年尺度等更大的时间尺度。

参考文献:

- [1] 秦大河, STOCKER T. IPCC 第五次评估报告第一工作组报告的亮点结论[J]. 气候变化研究进展, 2014, 10(1): 1-6.
- [2] WANG H J, CHEN Y N, CHEN Z S. Spatial distribution and temporal trends of mean precipitation and extremes in the arid region, north-west of China, during 1960-2010[J]. Hydrological Processes, 2013(12): 1807-1818.
- [3] DU H B, WU Z F, LI M, et al. Characteristics of extreme daily minimum and maximum temperature over Northeast China, 1961-2009[J]. Theoretical and Applied Climatology, 2013(1-2): 161-171.
- [4] CHEN Y, HU Q, YANG Y M, et al. Anomaly based analysis of extreme heat waves in Eastern China during 1981-2013[J]. International Journal of Climatology, 2016(1): 509-523.
- [5] 徐金芳, 邓振镛, 陈敏. 中国高温热浪危害特征的研究综述[J]. 干旱气象, 2009, 27(2): 163-167.
- [6] 陈敏, 耿福海, 马雷鸣, 等. 近 138 年上海地区高温热浪事件分析[J]. 高原气象, 2013, 32(2): 597-607.
- [7] 张可慧, 李正涛, 刘剑锋, 等. 河北地区高温热浪时空特征及其对工业、交通的影响研究[J]. 地理与地理信息科学, 2011, 27(6): 90-95.
- [8] 肖嗣荣, 张可慧, 刘芳圆, 等. 石家庄市高温热浪与“三大火炉”城市的对比研究[J]. 地理与地理信息科学, 2010, 26(5): 87-92.
- [9] 齐月, 陈海燕, 房世波, 等. 1961~2010 年西北地区极端气候事件变化特征[J]. 干旱气象, 2015, 33(6): 963-969.

- [10] 杨萍,刘伟东,王启光,等. 近 40 年我国极端温度变化趋势和季节特征[J]. 应用气象学报,2010,21(1):29-36.
- [11] 杨萍,刘伟东,侯威. 北京地区城郊极端温度事件的变化趋势及差异分析[J]. 气候与环境研究,2013,18(1):80-86.
- [12] 游庆龙,康世昌,闫宇平,等. 近 45 年雅鲁藏布江流域极端气候事件趋势分析[J]. 地理学报,2009,64(5):592-600.
- [13] 严晓瑜,赵春雨,王颖,等. 近 50 年东北地区极端温度变化趋势[J]. 干旱区资源与环境,2012,26(1):81-87.
- [14] 钱维宏. 天气尺度瞬变扰动的物理解析原理[J]. 地球物理学报,2012,55(5):1439-1448.
- [15] CHEN C C, TENG Y C, LIN B C, et al. Online platform for applying space-time scan statistics for prospectively detecting emerging hot spots of dengue fever[J]. International Journal of Health Geographics,2016,15(1):40-43.
- [16] ALOE C, KULLDORFF M, BLOOM B R. Geospatial analysis of non medical vaccine exemptions and pertussis outbreaks in the United States[J]. Proceedings of the National Academy of Sciences of the United States of America,2017,114(27):7101-7105.
- [17] 李德辉,方兆本,余雁. 扫描统计量—检测基金业绩持续性的新方法[J]. 运筹与管理,2006,15(1):82-87.
- [18] TUIA D, KAISER C, DACUNHA A, et al. Socio-Economic Cluster Detection with Spatial Scan Statistics. Case Study: Services at Intra-Urban Scale[Z]. Geo-computation, National University of Ireland,2007.
- [19] CLIFF A D, ORD J K. Model building and the analysis of spatial pattern in human geography[J]. Journal of the Royal Statistical Society,1975,37(3):297-348.
- [20] 昌晶亮,余洪,罗伟伟,等. 珠三角地区 PM_{2.5} 浓度空间自相关分析[J]. 生态与农村环境学报,2015,31(6):853-858.
- [21] 杨涵洧,封国林. 2013 年盛夏中国持续性高温事件诊断分析[J]. 高原气象,2016,35(2):484-494.
- [22] GLAZ J, NAUS J, WALLENSTEIN S. Scan Statistics[M]. Berlin:Springer,2001. 755-788.
- [23] 徐国鑫,金晓斌,周寅康. 基于 DEA 和空间自相关的我国土地市场化程度分析[J]. 地理与地理信息科学,2011,27(5):64-68.
- [24] 于文丽,蒲英霞,陈刚,等. 基于空间自相关的中国省际人口迁移模式与机制分析[J]. 地理与地理信息科学,2012,28(2):44-49.
- [25] 翟盘茂,潘晓华. 中国北方近 50 年温度和降水极端事件变化[J]. 地理学报,2003,58(S1):1-10.
- [26] SHI P J, KASPERSON R. World Atlas of Natural Disaster Risk [M]. Berlin Heidelberg:Springer,2015. 193-194.
- [27] 翟卫欣,程承旗. 一种空间权重矩阵的优化方法[J]. 武汉大学学报(信息科学版),2017,42(6):731-736.

Application of Spatio-temporal Aggregation Detection Method in Extreme High Temperature Aggregation Analysis

ZHANG Ting^{1,2,3,4}, CHENG Chang-xiu^{1,2,3,4}, YANG Shan-li^{1,2,3,4}, SONG Chang-qing^{1,3,4}

(1. State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875;
2. Key Laboratory of Environmental Change and Natural Disaster, MOE, Beijing Normal University, Beijing 100875;
3. Faculty of Geographical Science, Beijing Normal University, Beijing 100875;
4. Center for Geodata and Analysis, Beijing Normal University, Beijing 100875, China)

Abstract: Extreme high temperature is one of the hottest research topics of disastrous climate. The spatio-temporal aggregation detection is the basis of the deep analysis of this kind of disaster, and plays an important role for risk assessments of extreme high temperature events. The traditional temperature threshold method extracts extreme events from the perspective of transcendental probability. Therefore, it cannot realize the spatio-temporal integration scanning and lacks quantitative evaluation of the spatio-temporal aggregation degree for clusters. In this paper, the scanning statistics is utilized to construct the spatio-temporal cylinder scanning window, and realize the spatio-temporal integration detection. Relative risk (RR) is used to quantitatively assess the aggregation degree of clusters. Then, local Moran's I analysis is used to optimize the spatial location of the clusters according to the confidence of the hotspots. Taking the temperature data of July 2013 in mainland China as an example, 6 significant spatio-temporal clusters have been detected in the study area, with maximum RR (3.99) and minimum RR (2.60). These clusters cover parts of North China, central China, East China and Northwest China. Finally, the meteorological data and reports are used to verify the detection clusters. The results show that the integration of spatio-temporal scanning and local Moran's I analysis is an effective method to detect spatio-temporal clusters of extreme high temperature events. This method can not only locate space position and time position of the high temperature events accurately and reliably, but also evaluate the aggregation degree of the high temperature spatio-temporal clusters quantitatively.

Key words: extreme high temperature events; spatio-temporal scanning statistics; local Moran index; spatio-temporal clustering; temperature threshold