

文章编号: 2096-1618(2023)02-0221-06

气温日最低值空间差值序列的中位值解析

陈乐¹, 贺南¹, 陈岚², 文斌²

(1. 成都市温江区气象局, 四川 成都 611130; 2. 成都信息工程大学通信工程学院, 四川 成都 610225)

摘要:为了深入开展气象观测数据的精细化分析,提升观测资料的应用价值,使用1970–2017年成都地区的气象资料,以数据序列的频数特征为起点,对最低气温数据空间差值序列的中位值进行解析。提出序列中位值的提取方案;揭示这类序列中位值所关联的客观现象;分析这类序列的中位值与其统计平均值之间的联系和差异;阐述使用这种中位值进行解析运算的一些优势。结果表明,所提出的方案能够可靠提取到0.01℃精度的中位值;成都地区15 km间距以上样本中,有49%~88%的最低气温两站空间差值序列的中位值更具序列代表意义;与最低气温空间差的平均值序列的标准差相比,98%的对应的中位值序列的标准差偏小,通常偏小量为11%~19%,这种特点更有利于短序列使用,以提升观测数据使用时效;所有样本中,最低气温空间差的平均值序列与中位值序列同步演化,两者时间序列的相关系数为0.87~0.99,这两个不同统计方式的数据存在这种稳定的关联说明还有内在机理有待深入发掘。

关键词:气象学;气候变化;气象数据分析;空间差值;日最低值;中位值

中图分类号:P423.3⁺6

文献标志码:A

doi:10.16836/j.cnki.jcuit.2023.02.014

0 引言

长期缜密的气象观测已积累了海量宝贵的科学数据^[1-2],为推进气象事业高质量发展,发挥这些科学数据的应用价值,有必要利用现代信息工程技术对这些数据进行再分析与发掘^[3-6]。实践中多种数据分析工具都可以方便开展气象数据序列的频数乃至频谱分析^[7-11],但对现象、规律的认识是一项基础性工作,从新的分析角度去发掘、验证现象无疑是一种挑战,进一步系统地发掘应用价值更需要不懈的努力。

气温是现代气象观测精密的核心基础数据^[12],两个相邻区域的气温差值是热动力分析^[13-15]的基础量,其中气温日最低值作为气温日周期循环^[16]的节点表征量,包含来自日照、下垫面及大气运行的信息,并且形成于气温日周期循环的放热末端,客观背景较单一,有利于入手开展其频数特征与客观规律的关联分析。许多学者对气温开展了诸多研究,王秀文等^[17]将30年累年的气候平均值资料纳入中期预报业务系统中,建立了一套规范、标准、完整的适合中期预报业务使用的气候资料数据库和显示系统。任芝花等^[18-19]绘制了全国基准站建站至2000年气温4次与24次定时累年月平均值差值的平均值分布图,并做了初步分析,结果表明,采用两种不同时间分辨率统计气象要素月平均值会出现统计误差,气压、相对湿度、风速4次与24次定时观测资料月平均值偏差在观测精度之内,气温

的差值超过了观测精度。李兰等^[20-21]用新疆99个气象站年、季、月气温、降水要素的1981–2010年平均值和1971–2000年平均值资料,从区域平均和空间分布两方面对旧2个气候平均值进行对比分析,结果表明气温平均值有较大改变,一些过去的气候评价结果在长序列评价中需要重新考虑。刘小宁等^[22-23]利用1961–2002年699个国家基本(基准)站的时值气温观测资料,分析了中国不同方法计算的气温平均值的差异。这些都说明在使用气温平均值时,要充分重视平均气温值的计算方法,避免误差。

本文以成都平原区域相邻站点间的日最低气温的差值序列为样本,从站间日最低气温差序列的中位值入手,分析站点间日最低气温差序列的中位值与平均值的差异,了解站点间日最低气温差序列的中位值所呈现的客观现象的成因与其限制条件,为检验和发掘相关实际应用搭建技术线路。

1 资料来源与数据处理

1.1 资料来源

成都平原西部为青藏高原东缘,东部为海拔900 m左右的龙泉山脉,样本站点中13个位于成都平原,另一个简阳站位于龙泉山脉东侧。成都地形封闭,湿润寡照,通常按气候分为西部沿山、中部平原和东部山区。研究资料来源于成都市气象局,使用成都地区温江、都江堰、彭州、郫都、崇州、大邑、邛崃、蒲江、新

津、双流、龙泉驿、新都、金堂、简阳 14 个国家站 1970–2016 年的逐日最低气温资料。

1.2 数据处理

1.2.1 最低气温站间差值分析序列的建立

将站点 A 与站点 B 的日最低气温的差值 $T_{\min A} - T_{\min B}$ 记为 D_T_{\min} , $D_T_{\min} \in \{\dots, -0.2\text{ }^{\circ}\text{C}, -0.1\text{ }^{\circ}\text{C}, 0\text{ }^{\circ}\text{C}, +0.1\text{ }^{\circ}\text{C}, +0.2\text{ }^{\circ}\text{C}, \dots\}$ 。本地站点 D_T_{\min} 的值域通常在 $\pm 5.0\text{ }^{\circ}\text{C}$ 以内,在此范围以 $0.1\text{ }^{\circ}\text{C}$ 为步进顺序排列温差 $X_i, X_i = 0.1i\text{ }^{\circ}\text{C} (i = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5)$ 。一个 D_T_{\min} 值会对应某个档位的 X_i 值一次,即 $D_T_{\min} = X_i$ 时,计数 $P_i = 1$; M 个 D_T_{\min} 的序列 $\{D_T_{\min j}\}, (j = 1, 2, 3, \dots, M)$ 中,统计各档 X_i 出现的频数 $f(X_i)$:

$$f(X_i) = \sum_{j=1}^M P_i, D_T_{\min j} = X_i, P_i = 1; D_T_{\min j} \neq X_i, P_i = 0 \quad (1)$$

可生成反映某个温差位 X_i 拥有多少样本量的频次分析序列 $\{X_i, f(X_i)\}$ 。

1.2.2 分析序列的中位值的定义与提取

中位值 (Mo) 指序列 $\{X_i, f(X_i)\}$ 的频次量的中位数 (Me) 对应的特殊 X_i 。具体描述: 设序列时长度 M (天), 则有频次量中位数 $Me = M/2$, 进而中位值就是指 Me 对应的 X_i 。Mo 与 Me 联系如下:

$$\sum_{-\infty}^{Mo} f(X_i) = \sum_{Mo}^{+\infty} f(X_i) = Me = M/2 \quad (2)$$

文中主要使用季序列, $M \approx 92$, $Me \approx 46$, 这时 Mo 就是从 $\{D_T_{\min}\}$ 中的正值或负值最远端, 顺序计数至第 46 个频次量时对应的 X_i 。

为使 Mo 能够以 $0.01\text{ }^{\circ}\text{C}$ 的精度与平均值进行平行对比运算, 还需要使用插值公式提高 Mo 的取值分辨率, 运算公式如下:

$$Mo = X_a - \left(\frac{0.1}{f_{X_a} - f_{X_b}} \right) \times (f_{X_a} - Me) \quad (f_{X_b} < Me < f_{X_a}) \quad (3)$$

式中: Mo 为中位值; Me 为频率中位数; Mo 位于两个精度档位 X_a, X_b 之间 ($|X_a - X_b| = 0.1$); X_a 拥有频率为 f_{X_a} ; X_b 拥有频率为 f_{X_b} 。

2 $\{D_T_{\min}\}$ 序列中位值的客观规律

2.1 $\{D_T_{\min}\}$ 序列的平均值、中位值的数学意义及观察方法

根据数学原理, 数据序列的平均值、中位值及众数值都属于序列的代表数, 其中众数稳定性较弱, 后文主要考察序列的中位值与平均值。平均值是通过计算得到, 因此它会随单个数据变化而变化, 容易受极限值影

响; 中位值通过排序取得, 不受序列中的极限值影响, 为考察 $\{D_T_{\min}\}$ 序列的中位值和平均值对于序列的代表性的强弱, 后续内容将通过大量检查 $\{D_T_{\min}\}$ 序列中, 每个具体 D_T_{\min} 值与其序列平均值、序列中位值的距离 (绝对离差), 统计比较每个 D_T_{\min} 样本更接近平均值的发生率, 和更接近中位值的发生率, 从而区分代表性强弱, 进而观察数据序列的客观背景, 分析中位值的适用条件。

D_T_{\min} 的 30 a 序列概率分布方式接近正态, 这种序列长度下, 平均值与中位值基本等量。将序列的长度减小到季, D_T_{\min} 概率分布方式很少表现为 30 a 长序列的分布特点, 更多为显著的偏态分布; 进一步, 用连续多年 (n 年) 的同季的 $\{D_T_{\min}\}$ 的中位值, 和 $\{D_T_{\min}\}$ 的平均值, 形成 $\{Mo\}$ 和 $\{\overline{D_T_{\min}}\}$ 两个时间序列, 在本地站点样本 182 种 48 a 时长的对比组合中, $\{Mo\}$ 和 $\{\overline{D_T_{\min}}\}$ 的相关系数为 $0.87 \sim 0.99$, 平均水平为 0.95 , 证明中位值与平均值的时间变化量高度一致性, 同时也证明包括插值方式在内的中位值采样运算过程具有很好的重复稳定性和解析分辨率, 可以用于开展高精度的分析。图 1 是季样本的 $\{\overline{D_T_{\min}}\}$ 和 $\{Mo\}$ 历年同季变化的示例。

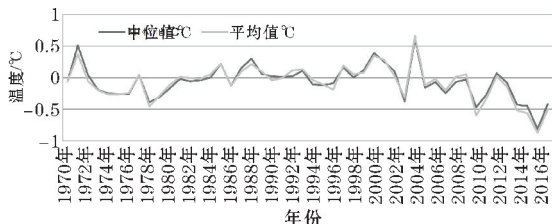
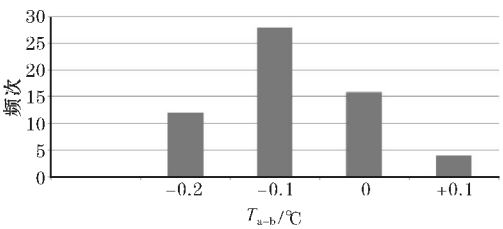


图1 1970–2016 年温江站与郫都站 $\{Mo\}$ 和 $\{\overline{D_T_{\min}}\}$ 的同季比较

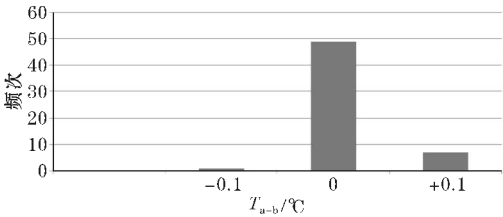
2.2 $\{D_T_{\min}\}$ 分布特点的分析

$\{D_T_{\min}\}$ 内的 D_T_{\min} 值统计各个 X_i 的频次分布, 总是呈现趋向某段位置集中的趋势, 这种稳定存在的现象必然有客观原因。

首先从两个站点距离为 0 开始, 分析测量读数误差对 $\{D_T_{\min}\}$ 的影响, 距离为 0 时两个气温传感器的差值就是读数误差, 读数误差分布是以 $0\text{ }^{\circ}\text{C}$ 为中心的正态分布, 其中包括读数时间不同步产生的误差。图 2 是国家基准站主备气温传感器温差值样本片段, 其中包含有 3 分钟读数时差的样本差值分布范围为 $\pm 0.2\text{ }^{\circ}\text{C}$, 读数时间订正后的差值分布范围为 $\pm 0.1\text{ }^{\circ}\text{C}$ 。历史资料中, 有较长时段为人工观测数据, 其中的读数和读数时间误差更大, 由此可知, 在站间日最低气温差值序列中, 总是会附加有 $\pm 0.1\text{ }^{\circ}\text{C} \sim \pm 0.2\text{ }^{\circ}\text{C}$ 的正态随机分布误差。两个站点的气温差异越小, 读数随机误差的影响比重越大。



(a) 分钟数据的频次分布(60 min),备份站慢了3 min的样本



(b) 更正了3 min后的分钟数据 T_{a-b} 的频次分布

图2 双气温传感器读数误差样本

因此, X_i 的频次集中分布在某段位置这种现象中, 有两个站点比对计算时被突出的测量误差这一因素, 它弱化 D_T_{min} 中的有用信息, 附加随机性, 使 $\{D_T_{min}\}$ 的平均值代表性加强。

另一个差值的常见集中分布点为 0 温差位。这是由于大尺度天气系统的侵入, 或盛行风对本地空气团的置换, 相邻两个站点有较大概率处于同一等温线控制, 阴雨天气也会减小 D_T_{min} 。这些因素使 $\{D_T_{min}\}$ 中一定会有在 $X_i=0$ 的位置附近出现频次量, 这个特点说明 $\{D_T_{min}\}$ 对 X_i 的概率分布并不完全随机。

2.3 $\{D_T_{min}\}$ 的优势代表数统计

对实际样本进行统计观察的目的, 是为了解 $\{D_T_{min}\}$ 的 Mo 和 $\{\overline{D_T_{min}}\}$ 以及 0 °C 在什么条件下成为多数 D_T_{min} 的趋向值, 即挑选 $\{D_T_{min}\}$ 的优势代表数。

2.3.1 统计方法

分别检查 $\{D_T_{min}\}$ 中每个日值 D_T_{minj} 与 Mo 、 $\{\overline{D_T_{min}}\}$ 、0 °C 之间的距离, 检查以日数据为基础, 以季为比较单元, 采用直接穷举比较。

如果某一季有 M 天, 统计计算时, 假设:

$$\frac{\sum_{j=1}^M |D_{Tminj} - Mo|}{M} = A \tag{4}$$

$$\frac{\sum_{j=1}^M |D_{Tminj} - \{\overline{D_{Tmin}}\}|}{M} = B \tag{5}$$

$$\frac{\sum_{j=1}^M |D_{Tminj}|}{M} = C \tag{6}$$

即 A 为 $\{D_T_{min}\}$ 相对于 Mo 的平均绝对离差; B 为 $\{D_T_{min}\}$ 相对于 $\{\overline{D_T_{min}}\}$ 的平均绝对离差; C 为 $\{D_T_{min}\}$ 相对于 0 °C 位的平均绝对离差。

统计使用 31 年资料 124 个季, 每个季值对应上述 3 个离差值记为 A_i 、 B_i 、 C_i , 同时对应条件计数量 P_{a_i} 、

P_{b_i} 、 P_{c_i} :

$A_i < B_i$ 并且 $A_i < C_i$, 则令 $P_{a_i} = 1$, 否则 $P_{a_i} = 0$

$B_i < A_i$ 并且 $B_i < C_i$, 则令 $P_{b_i} = 1$, 否则 $P_{b_i} = 0$

$C_i < A_i$ 并且 $C_i < B_i$, 则令 $P_{c_i} = 1$, 否则 $P_{c_i} = 0$

Mo 的优势占比 = $\sum_{i=1}^{124} \frac{P_{a_i}}{124} \times 100$ (7)

$\{\overline{D_T_{min}}\}$ 的优势占比 = $\sum_{i=1}^{124} \frac{P_{b_i}}{124} \times 100$ (8)

趋向 0 °C 的优势占比 = $\sum_{i=1}^{124} \frac{P_{c_i}}{124} \times 100$ (9)

2.3.2 中位值占优势的典型样本数据及其背景条件

经统计, 在本地站点组合中, 基于季 $\{D_T_{min}\}$ 的

Mo 和 $\{\overline{D_T_{min}}\}$ 以及 $X_i=0$ 温差位进行比较时, D_T_{minj} 更接近 Mo 的情况占多数, 以都江堰站与彭州站的组合为例, 在 1970–2000 年的连续 11325 组日数据中, D_T_{minj} 最接近 Mo 的情况有 5871 日次, 占比 52%; D_T_{minj} 最接近 0 的情况有 3280 日次, 占比 29%; D_T_{minj} 最接近 $\{\overline{D_T_{min}}\}$ 的情况有 2048 日次, 占比 18%; 另有倾向不明朗的(等距离)样本 126 日次, 占比 1%。统计结果见表 1。

表 1 中位值占优势的样本数据案例 单位: %			
参数	金堂-简阳组	都江堰-彭州组	温江-崇州组
中位值的样本占比	73	88	49
平均值的样本占比	23	10	31
0 差值的样本占比	4	2	20

样本的背景条件: 金堂-简阳间距 50 km, D_T_{min} 绝对值平均量为 0.84 °C; 都江堰-彭州间距 28 km, D_T_{min} 绝对值平均量为 0.73 °C; 温江-崇州间距 15 km, D_T_{min} 绝对值平均量为 0.46 °C。

2.3.3 平均值占优势的典型样本数据及其背景条件

表 2 平均值占优势的样本数据 单位: %	
占比	温江-郫县组
中位值的样本占比	26
平均值的样本占比	54
0 差值的样本占比	20

样本的背景条件: 站点间距为 11 km; 地形是平原; D_T_{min} 的绝对值平均量为 0.31 °C。

2.3.4 0 温差值占优势的典型样本数据及其背景条件

0 温差值占优势的样本数据见表 3。

表 3 0 温差值占优势的样本数据 单位: %	
参数	金堂-简阳阴天的样本组
中位值的样本占比	19
平均值的样本占比	8
0 差值的样本占比	73

样本的背景条件: 取自两个站点同为无日照的阴天的记录(3884 组日数据); 站点间距 50 km; 间隔龙泉山脉; D_T_{min} 的绝对值平均量为 0.84 °C。

2.3.5 现象的综合与分析

现象1:本地样本 $D_{T_{minj}}$ 更接近其季值 $\{D_{T_{min}}\}$ 的 Mo 的情形最为常见。

分析出现这种现象的原因,源于 Mo 的频次对等的取样方式,这种取样方式下,如果将 Mo 看作一个动态的 0 值参考点,则两个站点日最低气温差值的矢量方向(正负值)出现的概率将一致;由于日最低气温会跟随气温的日周期振荡更新,因此,两个站间日最低气温差值的矢量方向出现概率一致的情形,符合两个站点基于日周期作用下的,在有限距离上的热力平衡趋势,并且这种平衡可以建立在时空量更大的气温梯度演化趋势之上。

现象2: $D_{T_{min}}$ 的绝对值平均量越小, $D_{T_{minj}}$ 越容易接近其季平均值。

出现这种现象的原因,源于气温传感器随机读数误差的影响,总体偏小的气温差数据序列受到读数随机误差干扰更明显,其平均值的代表性会增强。这种现象会使 $D_{T_{min}}$ 在实际应用分析中存在精度限制,即 $|D_{T_{min}}|$ 偏小随机性增强,应用分析的可靠性将减弱,除非进一步提高采样准确度。

现象3:在两个站点日照时数都为 0 的条件下, $D_{T_{minj}}$ 更趋向 0℃。

与这种现象关联的已知背景因素有两个,首先是两个站点同为阴天时,大概率受到天气系统影响,本地气团被入侵冷空气整体置换,使两个相邻站点的气温值接近;另一个原因是日照弱,并且两个站点受热均匀成因单一,呈现熵最大化倾向。

3 中位值应用的优势

气温差值序列中位值是气温差值分布特征中的一个重要的参数。对于两个站点的日最低气温差值序列,序列中位值更具有客观代表性,因此,在观察和提取气温差值序列的相关客观规律时,使用中位值作为序列代表数,会比使用平均值作为代表数更稳定可靠;在涉及两站气温差值分布参数的运算中,还需要充分考虑测量误差和温差值趋 0 的不同客观背景条件的影响。中位值结合客观规律可以设计出某些特殊的技术线路,在贺南等^[24]的国家发明专利中就是利用了温差序列中位值存在确定客观意义这种优点,通过跟踪中位值对应的频率值的变化规律,设计出归类分型的运算方法,从而实现对大量区域自动站探测环境的动态监控。在成都地区站点管理实践中,使用该方法对 256 个区域自动站开展自动化动态分类,精准地识别出了“城市型”“公园湿地型”“乡村型”“山地型”等不同下垫面条件的站点。识别出的不同类型站点之间,气象要素的演化规律呈现一致的明确差异。其中不同

类型站点之间,湿度值的差异在 6%~15%,日较差的差异值在 0.5℃~3℃,了解和修正这些差异将有利于提高精细化预报的准确度。另外,方法还自动判断并经实际验证排查出其中 4 套环境不良站点,显示出这一技术具有的独特潜力。

气温差值序列中位值比平均值稳定这种优点,也可直接提升一些简单运算的效果,如应用分析中常见的平行序列比较法,由于中位值的时间序列的标准差,多数情况都小于平均值的时间序列的标准差,这时使用中位值序列进行平行比对分析可以取得更高的精确度或稳定性。以成都地区 14 个国家站为样本,各站的 48 a 逐日最低气温(T_{min})的季均值 $\overline{T_{min}}$,可生成 A 站、B 站两两组合互为参照的观察序列 $\{\Delta \overline{T_{min}}\}$ 共 91 种,其中 $\Delta \overline{T_{min}} = \overline{T_{minA}} - \overline{T_{minB}}$;同时每季数据还可统计出 $\{T_{minA} - T_{minB}\}$ 的 Mo ,构成 48 a 同期中位值序列 $\{Mo\}$;将这 91 种 $\{\Delta \overline{T_{min}}\}$ 结合 4 个季节的 364 组 $\{\Delta \overline{T_{min}}\}$ 的标准差(记为 $STDEV\{\Delta \overline{T_{min}}\}$)与 $\{Mo\}$ 的标准差(记为 $STDEV\{Mo\}$)进行比较。其中 $STDEV\{Mo\} < STDEV\{\Delta \overline{T_{min}}\}$ 的占比高达 98%,平均水平上, $STDEV\{Mo\}$ 比 $STDEV\{\Delta \overline{T_{min}}\}$ 小 11%,对于成都地区,冬春季 $STDEV\{Mo\}$ 比 $STDEV\{\Delta \overline{T_{min}}\}$ 小 19%,华西秋雨季两者差别较小。 Mo 时间序列的这种标准差小的特点,使其在序列的平行比时有更大优势。这是因为对于气候变化、观测环境变化、传感器计量基点漂移这类缓变量而言,整体的缓变量在中位值序列与平均值序列中的变化幅度是一致的,但中位值序列的标准差比平均值小 10%~20%,意味着提取缓变量时被视为噪声的瞬变量幅度显著减小,这将十分有利于各类缓变量信息的算法提取。同时, Mo 的时间序列的标准差相对较小这种特点在较短的序列中更具优势,可以提高资料的使用时效,如针对成都地区台站受城市化影响较弱时期的气温记录,大部分 T_{min} 季均值,只需要 8 a 时长平滑(平均), Mo 序列的自然波动量就可以位于 30 a 标准值的 ± 0.1 ℃ 以内,而同样条件下,多数平均值序列波动范围为 30 a 标准值的 ± 0.2 ℃ 范围。这时如果都以 8 a 时长去对比 30 a 历史同期标准值从而提取缓慢变化量,使用 Mo 序列的潜在的最大分析误差为 0.2℃,使用平均值序列的潜在最大的分析误差为 0.4℃,两者可靠的分析精确度相差很大, Mo 序列的这种使用效果,使其在较短的时间序列应用中具有潜力,有望改善气象记录的使用时效。

4 中位值序列与平均值序列的比较

以都江堰站与郫都站日最低气温差的 10 月、11 月、

12 月的季度样本为例,其季序列中位值与平均值的30 a 经年变化一致(相关系数 0.96),但不等值,见图 3。

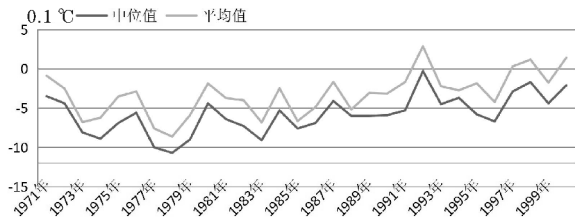


图3 都江堰站与郫都站的中位值与平均值

图 3 中平均值始终高于中位值,从平均值易受突变量影响这种计算特点分析,都江堰站日最低气温减去郫都站日最低气温的形成的差值中,始终会存在更多的正向的突变量。

实际数据支持这种解释,逐日检查 1971–2000 年 30 a 间第四季度都江堰站与郫都站间最低气温的差值数据:第四季度的两站间温差 30 a 的平均量为 $-0.32\text{ }^{\circ}\text{C}$,在逐日数据中,以该平均量 $-0.32\text{ }^{\circ}\text{C}$ 的 ± 10 倍值作为突出量判定阈值,在正值突出方向阈值为 $+2.9\text{ }^{\circ}\text{C}$,30 a 中超过此正向阈值的个值分散出现 43 次,这是都江堰站日最低气温大幅度超过郫都站的次数;温差 $-3.5\text{ }^{\circ}\text{C}$ 是以平均量 $-0.32\text{ }^{\circ}\text{C}$ 为对称轴的负向 10 倍突出量判断阈值,对应都江堰站日最低气温异常的大幅度低于郫县郫都站的情况,30 a 中只出现了 2 次,很明显,以平均量为参照,正负突变量严重不对称,见图 4。

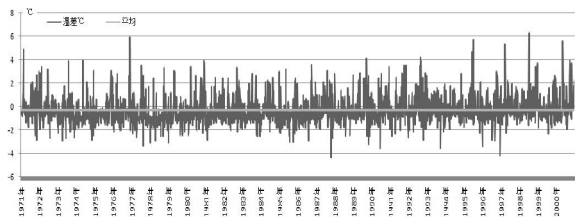


图4 都江堰站与郫都站的正负突变量

这种温差数据分布意味着:尽管 30 a 平均值上郫都站最低温高过都江堰 $0.32\text{ }^{\circ}\text{C}$,但都江堰站的日最低气温“大幅度”超过郫都站的概率却要高很多,这种现象的成因有待进一步研究,其中都江堰与郫都站的海拔高度差为 159 m,该样本 30 a 序列平均值为 $-0.32\text{ }^{\circ}\text{C}$,中位值的 30 a 均值为 -0.58 。类似样本广泛存在,如图 5 中福建省福州站与闽侯站 30 a 的数据(中位值序列与平均值序列相关系数为 0.9),这说明虽然中位值序列与平均值序列比较相似,但客观物理意义存在差异。本地资料中,气温差值序列平均值中可以观察到中位值、随机测量误差量、突出值影响量三个组成部分,其中中位值是基础量。利用温差序列中位值与平均值的这种关系,可以追踪分析测量误差影响或突出值影响的背景信息,中位值的这种客观性和基础性,使其在应用时具有一定意义和潜力。

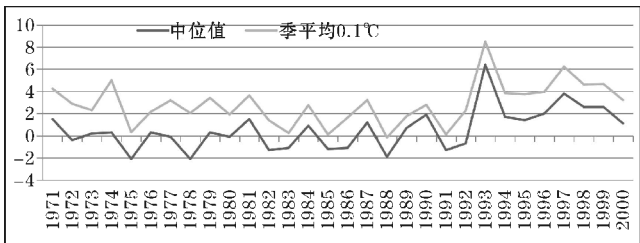


图5 福州站与闽侯站的中位值与平均值

5 结论

研究基于 1970–2017 年成都地区的气象资料,以数据序列的频数特征为起点,对最低气温数据空间差值序列的中位值进行解析,主要结论如下:

(1) 两个相邻站点日气温的空间差值序列的中位值,与局部区域热力平衡机制密切相关,是重要的分析参数,是比平均值更好的序列代表数。在 $15\sim 50\text{ km}$ 网格间距中,中位值作为优势代表数的 $\{T_{\min A}-T_{\min B}\}$ 序列占比可以达到 $49\%\sim 88\%$ 。

(2) 通过与平均值序列的比对分析,以及中位值序列具体的应用实践表明,提出的中位值统计方案能够在 $0.01\text{ }^{\circ}\text{C}$ 精度上生成可靠的中位值分析序列。

(3) 当两个站点的平均温差很小时,气温传感器的读数随机误差会形成显著干扰,因而影响气温差值序列中客观规律的表达与应用,样本中两站距离小于 11 km 、两站平均温差的绝对值小于 $0.31\text{ }^{\circ}\text{C}$ 时,温差序列的平均值已经显示出更强的代表性,这时读数随机误差会干涉到分析结论。

(4) 值得注意的是,两个相邻站点间气温传感器的读数误差服从正态分布,但现实中包含天气背景信息的相邻站点气温空间差值并不服从正态分布,此时序列中位值与平均值不等量,长序列中两者往往存在 0.9 以上的极强相关性,通过检查中位值与平均值存在的这种稳定差异量,有可能进一步找出具备应用价值的有用信息。

参考文献:

[1] 赵洪杰,闫景东,陈广瑞,等. 积温对比法在鸭梨始花期预报中的应用[J]. 气象科技,2019,47(5):872–878.

[2] 傅玮东,李兰,顾雅文. 库车日平均气温 $\geq 3\text{ }^{\circ}\text{C}$ 初日与杏树开花期的关系[J]. 沙漠与绿洲气象,2017,11(4):86–89.

[3] 裴浩,Alex Cannon,Paul Whitfield,等. 近 40 年内蒙古候平均气温变化趋势[J]. 应用气象学报,2009,20(4):8.

[4] 周国华,王盘兴,罗小莉,等. 基于 160 站资料的我国表面气温异常特征[J]. 应用气象学报,

- 2011,22(3):9.
- [5] 蔡文华,潘卫华,张辉,等. 2004/2005 年连江县冬季沿坡地地面气温观测和分析研究[J]. 应用气象学报,2006,17(4):483-487.
- [6] 薛湛彬,陈炯,张瑛,等. ECMWF 高分辨率模式 2 m 温度预报误差订正方法研究[J]. 气象,2019,45(6):831-842.
- [7] 张成军,雷学锋,李娜,等. 基于动态最优 PP 法的乡镇温度预报技术[J]. 干旱气象,2019,37(3):508-514.
- [8] 徐丽娇,胡泽勇,赵亚楠,等. 1961-2010 年青藏高原气候变化特征分析[J]. 高原气象,2019,38(5):911-919.
- [9] 孙健,李栋梁,邵鹏程,等. 中国冬季气温月际变化特征及其对大气环流异常的响应[J]. 气象学报,2019,77(5):885-897.
- [10] 陈百硕,李守伟,何建敏,等. 天气衍生品中时变均值回复的气温预测模型研究[J]. 管理工程学报,2014(2):6.
- [11] Luo H, Bao W, Wang Y. Spatial-Temporal Variation and the Tendency of Droughts and Floods in Henan Province[J]. Yellow River,2016(3).
- [12] 王效瑞,徐敏. 气温基准平均值的差异及其统计检验[J]. 气象科技,2003,31(1):33-37.
- [13] 杨倩,刘登峰,孟宪萌,等. 1960-2017 年山西省降水和气温的时空变化特征分析[J]. 人民珠江,2019,40(6):7.
- [14] 刘小宁,张洪政,李庆祥. 不同方法计算的气温平均值差异分析[J]. 应用气象学报,2005,16(3):12.
- [15] 郑斯中,冯丽文. 中国月平均气温标准差的研究[J]. 大气科学学报,1987(2):235-241.
- [16] 王秀文,李月安. 新气候平均值在中期预报业务中的应用[J]. 气象,2003,29(1):43-45.
- [17] 孙化南. 不同时间分辨率对气象要素月平均值统计的影响[J]. 应用气象学报,2004,15(B12):8.
- [18] 任芝花,许松,孙化南,等. 全球地面天气报历史资料质量检查与分析[J]. 应用气象学报,2006,17(4):9.
- [19] 白素琴. 不同气候平均值对气候评价业务的影响[J]. 沙漠与绿洲气象,2016,10(1):88-94.
- [20] 李兰,杜军,白素琴,等. 莫索湾地区农业气候资源的变化特征分析[J]. 中国农业气象,2009(S1):5.
- [21] 刘小宁,张洪政,李庆祥. 不同方法计算的气温平均值差异分析[J]. 应用气象学报,2005,16(3):12.
- [22] 刘小宁,鞠晓慧,范邵华. 空间回归检验方法在气象资料质量检验中的应用[J]. 应用气象学报,2006(1):37-43.
- [23] 刘小宁,任芝花. 地面气象资料质量控制方法研究概述[J]. 气象科技,2005,33(3):5.
- [24] 贺南,夏昕,杨东,等. 一种区域站的客观分群方法[P]. 中国专利:1253654,2021-06-29.

Median Value Analysis of Spatial Difference Series of Daily Minimum of Air Temperature

CHEN Le¹, HE Nan¹, CHEN Lan², WEN Bin²

(1. Chengdu Wenjiang District Meteorological Bureau, Chengdu 611130, China; 2. College of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: In order to further carry out the fine analysis of meteorological observation data and improve the application value of observation data, the meteorological data in Chengdu from 1970 to 2017 are used to analyze the median value of spatial difference series of minimum temperature data from the frequency characteristics of data series. The extraction scheme of median value of this sequence is proposed and the objective phenomenon associated with the median value of this kind of sequence is revealed. The relationship and difference between the median value and its statistical mean value of this kind of series are analyzed and some advantages of using this median value for analytical operations are described. The results show that the proposed scheme can reliably extract the median value with an accuracy of 0.01 °C. Among the samples with a distance of more than 15 km in Chengdu, the median value of the spatial difference series of the two stations with 49% to 88% of the lowest temperature is more representative. Compared with the standard deviation of the mean value series of the spatial difference of the lowest temperature, the standard deviation of 98% of the corresponding median value series is small, usually 11% to 19%. This feature is more conducive to the use of short series to improve the timeliness of observation data. In all samples, the average value series and median value series of the spatial difference of the lowest temperature evolve synchronously, and the correlation coefficient of the two time-series is 0.87 to 0.99. The stable correlation between the data of these two different statistical methods shows that there is still an internal mechanism to be further explored.

Keywords: meteorology; climate change; meteorological data analysis; spatial difference; daily minimum; median value