

# 江西九江气象站百年均一化月、年平均气温数据集（1924–2023）研发

占龙飞<sup>1,2</sup>, 董保华<sup>1,2</sup>, 许彬<sup>1,2\*</sup>, 李勇<sup>1,2</sup>, 辛佳洁<sup>1,2</sup>, 王丽燕<sup>3</sup>

1. 江西省气候中心, 南昌 330096; 2. 南昌国家气候观象台, 南昌 330200;  
3. 江西信息应用职业技术学院, 南昌 330043

**摘要:** 长期的均一化观测序列对于气候变化的准确评估和归因至关重要。然而, 我国多数气象台站受到了台站迁址、仪器换型、环境变迁等非自然因素的影响, 造成观测序列中存在非均一性。本研究基于多源月平均气温资料, 在数据整合和质量控制基础上, 以武汉气象站为参照站, 采用标准化序列法对江西九江气象站 1924–2023 年逐月平均气温资料进行了插补; 利用惩罚最大 F 检验 (Penalized Maximal F Test, PMF) 方法对插补后数据的均一性进行了检验, 同时采用分位数匹配方法对其进行了订正, 构建了江西九江气象站百年均一化月、年平均气温数据集 (1924–2023)。通过与邻近气象站的均一化百年气温数据对比分析, 构建的九江气象站百年逐月平均气温数据与邻近站的相关性达到 0.9 以上, 一定程度上验证了构建方法的科学性及其数据集的可靠性。本数据集为江西九江气象站 1924–2023 年逐月/年平均气温 (单位:  $^{\circ}\text{C}$ ), 数据存储格式为 .txt, 由 2 个数据文件组成, 数据量为 16.5 KB。

**关键词:** 九江; 气温; 百年序列; 插补; 均一化

**DOI:** <https://doi.org/10.3974/geodp.2025.01.10>

**CSTR:** <https://cstr.escience.org.cn/CSTR:20146.14.2025.01.10>

**数据可用性声明:**

本文关联实体数据集已在《全球变化数据仓储电子杂志 (中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2024.12.03.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2024.12.03.V1>.

## 1 前言

在全球气候变化背景下, 区域气候变异的研究对理解气候变化的空间特征及其潜在影响具有重要意义<sup>[1,2]</sup>。东亚季风区是全球气候变化的敏感区域之一, 其气候变化既受全球变暖的直接驱动, 也受到区域大气环流、地表覆盖变化等多种因素的影响<sup>[3]</sup>。九江作为长江

收稿日期: 2024-12-30; 修订日期: 2025-02-20; 出版日期: 2025-03-25

基金项目: 中国气象局 (CMA2024QN15); 江西省科学技术厅 (20223BBG71019, 2023KYG01001); 上海市气象局 (QYHZ202106); 广东省气象局 (ZJLY202312); 南昌国家气候观象台 (JX2023Z09); 江西省气象局 (JX2022ZHFXPC06, JX2023Q07)

\*通讯作者: 许彬, 江西省气候中心, 1176325432@qq.com

数据引用方式: [1] 占龙飞, 董保华, 许彬等. 江西九江气象站百年均一化月平均气温数据集 (1924–2023) 研发[J]. 全球变化数据学报, 2025, 9(1): 87–95. <https://doi.org/10.3974/geodp.2025.01.10>. <https://cstr.escience.org.cn/CSTR:20146.14.2025.01.10>.

[2] 占龙飞, 许彬, 董保华等. 江西九江气象站百年均一化月平均气温数据集 (1924–2023) [J/DB/OL]. 全球变化数据仓储电子杂志, 2024. <https://doi.org/10.3974/geodb.2024.12.03.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2024.12.03.V1>.

中下游的重要城市,具有典型的东亚季风气候特征,其长时间序列的气候记录为研究区域气候变化特征及其与全球气候变化的联系提供了宝贵的数据基础。然而,由于历史气象观测数据来源多样、缺测较多以及时间序列非均一性等问题,准确反映区域气候长期变化趋势的高质量数据集仍然稀缺<sup>[4]</sup>。

构建连续、可靠的长时间序列均一化气象数据是气候变化研究的基础工作<sup>[5,6]</sup>。均一化处理是将多源气象观测数据经过系统校正与一致化处理,以消除数据间的系统性差异和非均一性特征,使其能够反映真实的气候变化特征。近年来,许多研究利用均一化技术修正历史气象数据,取得了显著成果<sup>[7-10]</sup>,但在具体的应用上,仍需结合区域特征及历史数据特点,进行更为细致和针对性的处理。

本研究以九江气象站为研究对象,整合 1924–1938 年、1951–2023 年的多源气象数据及 1924–2016 年的邻近站均一化数据,采用标准化序列法、均一性检验和断点订正技术,构建了九江站 1924–2023 年的均一化逐月气温数据集,并揭示了九江站近百年来气温的变化特征。本研究可为九江地区的气候变化研究提供重要的数据支撑,也可与其他区域气象数据的均一化处理和序列构建提供借鉴。

## 2 数据集元数据简介

《江西九江气象站百年均一化月、年平均气温数据集(1924–2023)》<sup>[11]</sup>的名称、作者、地理区域、数据年代、时间分辨率、空间分辨率、数据集组成、数据出版与共享服务平台、数据共享政策等信息见表 1。

## 3 数据研发方法

### 3.1 资料来源

九江站 1924 年 1 月–1938 年 3 月的数据,来自中央气象局和中国科学院地球物理研究所联合资料室编印并于 1954 年出版的《中国气温资料》<sup>[13]</sup>;1951–2023 年的数据,来自于国家气象信息中心研发的“天擎”气象大数据云平台<sup>[14]</sup>;百年均一化邻近参照站数据取自中国科学院大气物理研究所修正百年来中国 32 站均一化逐月气温序列集后建立的数据集,该数据集能很好地代表中国过去百年大尺度气候变化特征<sup>[15]</sup>。九江站地理位置如图 1 所示。

### 3.2 算法原理

#### 3.2.1 数据插补

基于《中国气温资料》<sup>[13]</sup>所记载的逐月平均气温数据,构建九江站 1951 年以前的气温序列。然而,初步整合后的九江站气温数据在 1938 年 4 月至 1950 年 12 月期间出现了连续的缺测情况,缺测比例高达 12.8%。为了恢复数据的完整性,选取长序列完整的邻近站作为参考站,采用标准化序列法对九江站缺测资料进行插补。

表 1 《江西九江气象站百年均一化月、年平均气温数据集（1924–2023）》元数据简表

条 目	描 述
数据集名称	江西九江气象站百年均一化月、年平均气温数据集（1924–2023）
数据集短名	MeanTempJiujiang1924–2023
作者信息	占龙飞，江西省气候中心，lf.zhan@foxmail.com 许彬，江西省气候中心，1176325432@qq.com 董保华，江西省气候中心，dongbaohua_jx@163.com 李勇，江西省气候中心，908791309@qq.com
地理区域	九江
数据年代	1924–2023 年
时间分辨率	月、年
空间分辨率	气象站点
数据格式	.txt
数据量	16.5 KB
数据集组成	江西九江气象站 1924–2023 年逐月/年平均气温
基金项目	中国气象局（CMA2024QN15）；江西省科学技术厅（20223BBG71019，2023KYG01001）；上海市气象局（QYHZ202106）；广东省气象局（ZJLY202312）；南昌国家气候观象台（JX2023Z09）；江西省气象局（JX2022ZHFPC06）
数据计算环境	Python, R
出版与共享服务平台	全球变化科学研究数据出版系统 <a href="http://www.geodoi.ac.cn">http://www.geodoi.ac.cn</a>
地址	北京市朝阳区大屯路甲 11 号 100101，中国科学院地理科学与资源研究所
数据共享政策	（1）“数据”以最便利的方式通过互联网系统免费向全社会开放，用户免费浏览、免费下载；（2）最终用户使用“数据”需要按照引用格式在参考文献或适当的位置标注数据来源；（3）增值服务用户或以任何形式散发和传播（包括通过计算机服务器）“数据”的用户需要与《全球变化数据学报（中英文）》编辑部签署书面协议，获得许可；（4）摘取“数据”中的部分记录创作新数据的作者需要遵循 10%引用原则，即从本数据集中摘取的数据记录少于新数据集总记录量的 10%，同时需要对摘取的数据记录标注数据来源 <sup>[12]</sup>
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS, GEOSS, PubScholar, CKRSC

标准化序列法计算方法参考文献<sup>[16]</sup>，该方法可表示为以下公式：

$$Z_j = \frac{X_j - \bar{X}_j}{S_j} \tag{1}$$

$$Z_{avg} = \frac{1}{n} \sum_{j=1}^n Z_j \tag{2}$$

$$X_i = Z_{avg} S_i + \bar{X}_i \tag{3}$$

式中， $Z$  表示标准化序列， $Z_{avg}$  为参考站平均标准化序列， $X_i$  表示第  $i$  月待插补数据， $j$  代表第  $j$  个参考站， $X_j$  为  $j$  站第  $i$  月数据， $\bar{X}_j$  和  $S_i$  分别为  $j$  站第  $i$  月数据的多年平均值和标准差， $n$  表示参考站站数， $\bar{X}_i$  和  $S_i$  分别为待插补站第  $i$  月数据多年的平均值和标准差。

根据文献<sup>[17]</sup>的要求，选取邻近站时应以九江站为基准，选择水平距离在 300 km 范围内的邻近气象站数据。所选气象站的起始观测年份需早于 1924 年，且数据完整性要高，同时海拔高度应与九江站相近。经综合考量，最终确定武汉站作为邻近参考站。

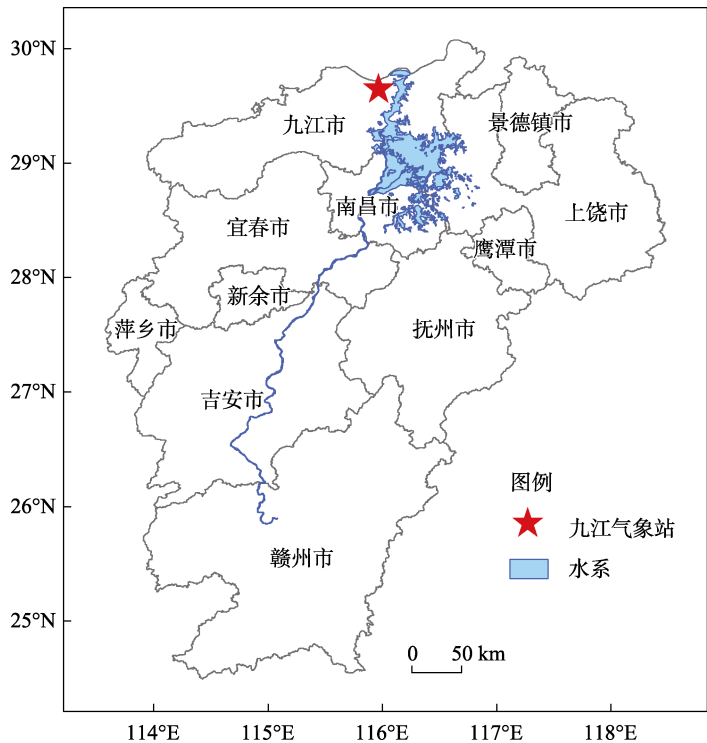


图 1 九江气象站地理位置图

3.2.2 均一性检验和订正

(1) 断点的检验方法

从 20 世纪 50 年代初开始，我国才有了较为完整和系统的观测资料；1951 年之前由于我国的气象观测站记录资料来源多样，时间序列的连续性、一致性和规范性均存在不足，缺乏可靠的依据来评估 1951 年之前非均一性检验结果的合理性，增加了我们构建真实反映当地气候变迁的参考序列的难度。

RHtest V4 软件包是一个用于气候数据均一性检验和调整的统计工具，主要用于检测和修正气候时间序列数据（如温度、降水等）中的非气候因素（如站点迁移、仪器变更、观测方法变化等）导致的偏差，从而提高数据的可靠性和一致性。其中的惩罚最大 F 检验（Peirce’s Maximum F-test, PMF）是一种无需参考序列即可开展的检验方法，其优势在于能够显著降低因非均一性参考序列以及元数据信息不详尽所引发的检验偏差。过往研究已充分证明该方法的有效性与可靠性。因此，本研究基于月平均气温序列，借助 PMF 方法，对九江站 1924 年至 2023 年的月平均气温序列进行了均一性检验。检验遵循了文献<sup>[18]</sup>中的方法。

(2) 断点的订正方法

采用分位数匹配（Quantile Matching, QM）方法<sup>[19]</sup>订正月值序列，该方法可在消除线性趋势后，确保待检序列中的各个片段在经验分布上能够相互匹配。本研究对以下两类断点进行了订正：

- 1) 第一类断点是 1954 年之前的断点，这些断点在年和月序列中同时被检测出，并且

是显著的间断点，但由于缺乏详尽的元数据支持，因此订正位置依据月值序列中检测出的断点时间来确定。

2) 第二类断点为 1954 年之后所检测到的断点，这些断点在年序列或月序列中被识别出来，并且有相应的元数据作为支撑。若这些断点的出现时间与台站元数据记录的信息相差不超过 1 年，那么便依据元数据记录的具体时间，将该断点的位置进行替换调整<sup>[6]</sup>。

检验采用 95% 的信度水平，以确保检验结果的可靠性。这种方法有助于减少因非均一性参考序列和缺乏详细元数据信息所带来的检验误差，从而提高序列订正的准确性。

3.3 技术路线

《江西九江气象站百年均一化月、年平均气温数据集（1924–2023）》研发的技术路线如图 2 所示。整合 1924–1938 年、1951–2023 年气象观测数据及 1924–2016 年统计均一化校正数据，通过标准化序列法插补了九江站缺测数据，再基于均一化检验和订正，构建了九江站 1924 年至 2023 年的逐年均一化气象数据序列。

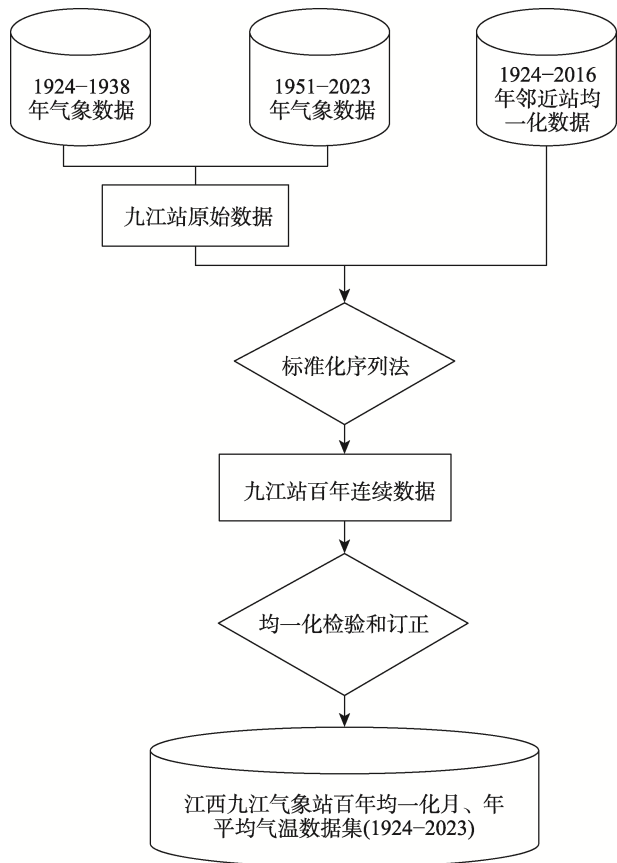


图 2 数据集研发技术路线图

4 数据结果与验证

4.1 数据集组成

本数据集为 1924–2023 年九江站逐月、年平均气温资料。月平均气温列出了年、月和

对应的月平均气温（℃）数据；年平均气温列出了年份和对应的年平均气温（℃）数据。数据格式为.txt，大小为 16.5 KB。

4.2 数据结果

本研究构建了一套完整的百年时间序列气温数据集，由图 3 可知，50 年代之前，九江站年平均气温呈现下降趋势，50 年代至 90 年代初，无明显变化趋势，90 年代开始呈现明显上升的趋势，1924–2023 年变化趋势约为 1.0℃/100a。由图 4 可知，九江站逐月平均气温分布存在明显的季节性变化：夏季（6 月至 8 月）温度显著较高，冬季（12 月至 2 月）温度显著较低，呈现出典型的季风气候特征。此外，20 世纪末至 21 世纪初，夏季的高温持续时间和强度有所增加，反映了气候变暖的趋势。这一变化与全球变暖背景下的区域气候变化特征一致。

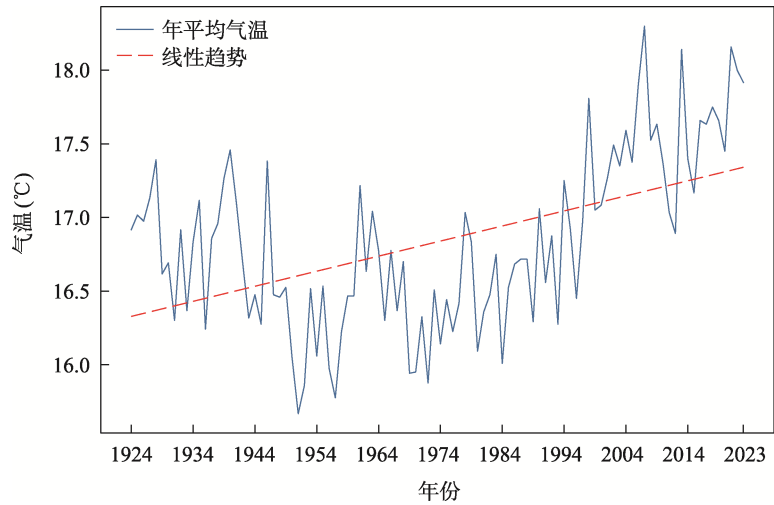


图 3 九江站年平均气温统计分析图（1924–2023）

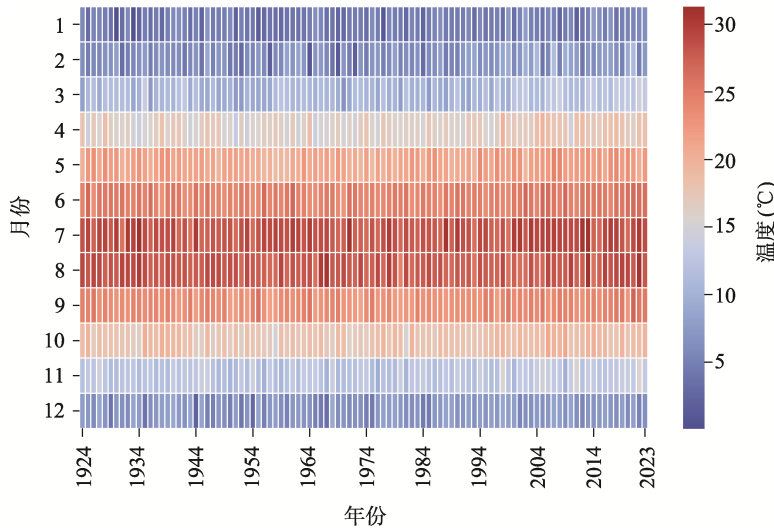


图 4 九江站均一化逐月平均气温统计图（1924–2023）

图 5 为不同时段气温变化的趋势及显著性检验结果。可以看出，气温在 1–6 月、10–11 月及年呈现显著上升趋势，尤其 2–4 月增幅最明显，这可能与气候变暖趋势加剧有关，冬春升温加快，秋季热量释放延迟，导致这些时段气温上升显著。而在 7–9 月和 12 月期间，气温变化较小，甚至在 8 月出现下降趋势，这些趋势均未达到显著水平。这可能是夏季受副热带高压等稳定天气系统控制，气温年际变化幅度相对平缓；12 月为初冬，冷空气活动频繁但强度和持续时间存在年际波动，使得气温变化未达到显著水平。

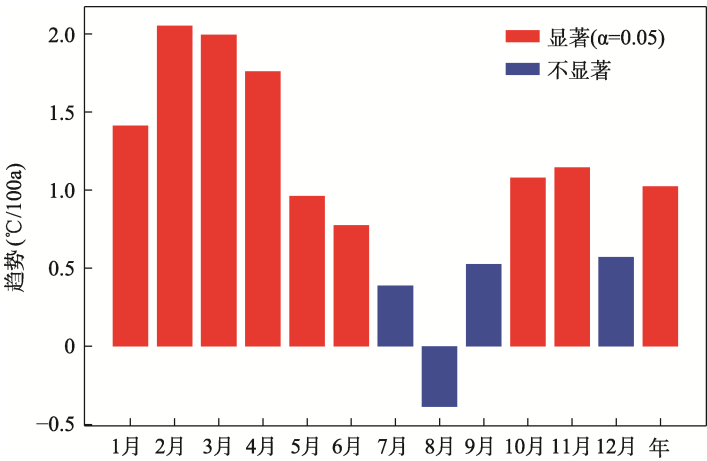


图 5 九江站气温变化趋势统计分析图（1924–2023）

4.3 数据结果验证

相关系数、均方根误差和标准差可以反映本数据集在不同时间段的表现。由表 2 可知，本数据集在 1924–2023 年平均数据上的相关系数为 0.94，表明本数据集与参照站之间具有很高的相关性。全年均方根误差为 0.27，显示出本数据集与参照站之间的平均偏差较小；标准差为 0.59，说明全年数据的波动性较低。

表 2 1924–2023 年九江站各月及年平均气温与参照站对比检验表

月/年	相关系数	均方根误差（℃）	标准差（℃）
1	0.97	1.09	1.42
2	0.98	0.77	1.90
3	0.97	0.46	1.60
4	0.94	0.50	1.30
5	0.93	0.50	1.18
6	0.91	0.82	0.97
7	0.92	0.53	1.22
8	0.90	0.63	1.11
9	0.88	0.55	1.09
10	0.91	0.63	1.09
11	0.95	0.89	1.24
12	0.97	1.13	1.39
年平均	0.94	0.27	0.59

从月度数据来看,相关系数范围为 0.88–0.98,且冬半年相关系数高于夏半年,表明本数据集在冬半年与参照站有更好的相关性表现。均方根误差范围为 0.46–1.13,春季均方根误差最小,冬季最大,表明春季与参照站的值最接近,而冬季误差最大。这是由于春季气温逐渐回暖,变化相对平缓,与参照站的均方根误差较小;冬季气温较低,且受到冷空气的影响,气温波动较大,与参照站的均方根误差较大。标准差范围为 0.97–1.9,其中 6 月的标准差最低,表明数据波动较小,而 2 月的标准差最高,数据波动较大。

## 5 讨论和总结

本研究构建了江西九江气象站 1924 年至 2023 年的均一化逐月、年平均气温数据集。研究综合利用《中国气温资料》、国家气象信息中心“天擎”气象大数据云平台,以及邻近站百年均一化数据,通过标准化序列法对缺测数据进行插补,并采用 RHtest V4 软件对序列的均一性进行检验与订正,确保了时间序列的连续性和可靠性。验证结果表明,本数据集与参照站相关性较高(全年相关系数达 0.94),均方根误差较小(0.27),能够较好地反映九江地区百年气候变化特征。

研究发现,九江站的气温变化具有明显的季节性特征,夏季(6月至8月)温度显著较高,冬季(12月至2月)显著较低,反映了典型的东亚季风气候特征。从长期趋势看,除8月外,九江站各月平均气温均呈现上升趋势,尤其在2月至4月的增温最为显著,这与全球变暖背景下春季和冬季升温幅度较大的区域特征一致。然而,7月至9月气温变化较小,尤其8月呈现下降趋势,未达到显著水平,这可能反映了该区域在夏季受局地气候变化或季风变率影响的复杂性。此外,20世纪末至21世纪初夏季的高温持续时间和强度有所增加,进一步表明区域气候正受到全球变暖的深远影响。

尽管取得了显著进展,但本研究仍存在一定局限性。由于1951年前气象数据来源多样且元数据信息不足,断点的检验与订正依赖于序列的统计特征,这可能会对结果的精确性造成一定影响。此外,极端气候事件对数据均一性可能带来的潜在影响需要进一步探讨。未来研究中,应结合更高分辨率的气候模拟与观测数据,进一步探讨区域气候变化的驱动机制与响应特征,为制定应对气候变化的政策提供更为精确的科学支持。

**作者分工:** 许彬、董保华对数据集的开发做了总体设计; 占龙飞采集和处理了九江气温数据,并设计了模型和算法; 占龙飞、李勇做了数据验证; 占龙飞、辛佳洁、王丽燕撰写了数据论文等。

**利益冲突声明:** 本研究不存在研究者以及与公开研究成果有关的利益冲突。

## 参考文献

- [1] Arnella, N. W., Lowe, J. A., Challinor, A. J., *et al.* Global and regional impacts of climate change at different levels of global temperature increase [J]. *Climatic Change*, 2019, 155(3): 377–391.
- [2] 潘小多. 基于区域气候模式未来气候变化研究综述[J]. *高原气象*, 2018, 37(5): 1440–1448.
- [3] Ren, G., Chan, J. C. L., Kubota, H., *et al.* Historical and recent change in extreme climate over East Asia [J].



- Climatic Change*, 2021, 168(3): 1–19.
- [4] Zhang, J. L., Liu, B., Ren, S. Q., *et al.* A 4 km daily gridded meteorological dataset for China from 2000 to 2020 [J]. *Scientific Data*, 2024, 11(1): 1230.
- [5] 唐国利, 丁一汇, 王绍武等. 中国近百年温度曲线的对比分析[J]. 气候变化研究进展, 2009, 5(2): 71–78.
- [6] 司鹏, 郝立生, 罗传军等. 河北保定气象站长序列气温资料缺测记录插补和非均一性订正[J]. 气候变化研究进展, 2017, 13(1): 41–51.
- [7] 李庆祥. 我国气候资料均一性研究现状与展望[J]. 气象科技进展, 2016, 6(3): 67–74.
- [8] 申燕玲. 青海省长序列高分辨率气象要素栅格数据集构建与评估[J]. 气象科技进展, 2023, 13(4): 29–33.
- [9] 王海军, 涂诗玉, 陈正洪. 日气温数据缺测的插补方法试验与误差分析[J]. 气象, 2008, 34(7): 83–91.
- [10] 王雅萍. 甘肃省新旧气象站址观测资料对比及均一性分析[J]. 高原气象, 2023, 42(2): 506–514.
- [11] 占龙飞, 许彬, 董保华等. 江西九江气象站百年均一化月、年平均气温数据集(1924–2023)[J/DB/OL]. 全球变化数据仓储电子杂志, 2024. <https://doi.org/10.3974/geodb.2024.12.03.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2024.12.03.V1>.
- [12] 全球变化科学研究数据出版系统. 全球变化科学研究数据共享政策[OL]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [13] 中央气象局. 中国气温资料[M]. 北京: 中央气象局, 1954.
- [14] 中国气象局. 气象大数据云平台[OL]. <http://www.weatherdt.com.cn/datasource.html>.
- [15] 中国科学院大气物理研究所. 百年来中国 32 站均一化逐月气温序列集[OL]. <http://www.weatherdt.com.cn/datasource.html>.
- [16] Li, Z., Yan, Z. W., Cao, L. J., *et al.* Further-adjusted long-term temperature series in China based on MASH [J]. *Advances in Atmospheric Sciences*, 2018, 35(8): 909–917.
- [17] 余予, 李俊, 任芝花等. 标准序列法在日平均气温缺测数据插补中的应用[J]. 气象, 2012, 38(9): 1135–1139.
- [18] 司鹏, 解以扬. 天津太阳总辐射资料的均一性分析[J]. 气候与环境研究, 2015, 20(3): 269–276.
- [19] Wang, X. L., Chen, H., Wu, Y., *et al.* New techniques for the detection and adjustment of shifts in daily precipitation data series [J]. *Journal of Applied Meteorology and Climatology*, 2010, 49(12): 2416–2436.