

近半个世纪中国区域历史气温网格数据集的建立^{*}

李庆祥^{1,2} 李 伟³

1 中国气象局国家气象信息中心, 北京, 100081

2 中国气象局气候变化开放实验室, 北京, 100081

3 中国气象局国家气象中心, 北京, 100081

摘 要

气候序列长度不一致以及由此导致的空间抽样误差在气候变化检测研究中常常产生影响, 因此往往需要把全球或者区域尺度气候数据转化为格点气候数据集。文中分别采用修改后的一级差分方法和普通 Kriging 方法, 把中国大陆(共约 728 个站, 不含港、澳、台地区)气象台站 1951 年 1 月—2004 年 12 月经过质量控制和均一化的历史气温数据转化为 $2.5^{\circ} \times 2.5^{\circ}$ 经纬度网格化数据集。对比分析表明: 应用上述两种方法, (1) 格点化带来的空间抽样误差较小; (2) 格点化后的格点气温序列和站点序列相关程度很高; (3) 不同的方法建立的格点数据集在序列的相关性、距平场的相似性描述方面均非常一致。二者的相关性和相似性也表明文中所建立的格点数据集是比较合理的。最后利用格点化后的气温数据集, 分别采取距平平均方法和一级差分方法, 对近 54 年中国气温变化趋势进行了更为准确的估计。不同的计算方法对中国区域气温增暖幅度的研究表明, 整个中国区域内近 50 年气温表现出一致的增暖特点, 其中以北方地区(西北、东北)气温增暖幅度最大, 个别地区达到近 $0.6^{\circ}\text{C}/(10\text{ a})$, 而最低的长江中、上游地区, 年均气温增暖幅度较小, 还不到 $0.1^{\circ}\text{C}/(10\text{ a})$, 全国平均的年平均气温近 54 年来增暖速率约为 $0.23\text{—}0.25^{\circ}\text{C}/(10\text{ a})$, 这和其他专家得出的结论是基本一致的, 而由于文中对数据集进行了较为系统的质量控制, 使得估计结果更为精确、可信。

关键词: 格点数据集, 气温, 增暖, 误差, 均一性。

1 引 言

近年来全球、半球尺度的气温变化研究中, 以 Jones 等^[1-4]的全球历史气候数据集、美国 NCDC 的 Peterson 等^[5]的 GHCN 数据集和 Hansen^[6-7]的全球数据集为最有代表性的几个成果。他们的结果也已经被 IPCC 作为主要的参考依据^[8], 为人们和科学界认识气候变化提供了很好的数据基础来源。中国地理位置处于东亚季风气候区, 其气候变化对全球变化的贡献和影响相当大。但在研究中国气候变化时, 由于资料本身的质量问题和空间抽样误差所导致的随机性和科学不确定性问题, 则讨论得还很不够。

在研究全球或区域大尺度气候变化序列时, 往

往将气候序列首先格点化, 以确保各个格点序列基本代表相同的局域面积上的气候变化, 这样计算的区域平均序列更加有代表性。上述提到的 3 个数据集也不例外, 不过采取的数学方法各有不同, Jones 等^[2]采取的方法为 CAM(Climate Anomaly Method), 其基本思想为计算气候距平的平均; Hansen 等^[6]的方法 RSM(Reference Station Method)是在较大的格点内选择一个时间最长的台站, 作为参考站, 然后计算每个时次计算所有站的距离权重平均, 形成该格点的序列; 而 Peterson 等^[5]则提出一个尽可能利用所有台站信息的一级差分方法(FDM, First Difference Method)。因此, 本文拟建立中国区域格点数据, 从而更精确地描述气候变化的趋势

^{*} 初稿时间: 2006 年 1 月 16 日, 修改稿时间: 2006 年 3 月 7 日。

资助课题: 国家自然科学基金项目(40605021), 中国科技部专项基金项目(2004DKA20170-01-02), 中国气象局气候变化专项基金项目(CCSF2005-2-QH17)。

作者简介: 李庆祥, 从事气候资料和气候变化研究。Email: liqx@cma.gov.cn。

和幅度。

本文首先从已通过质量控制和均一化的历史气温资料序列入手,采用略作修改后的一级差分方法(FDM)和 Kriging 插值方法分别从不同角度对中国大陆地区近 54 年的历史气温数据进行空间插值,分别建立了近 54 年中国历史气温网格化数据集。最后,分别利用两种格点数据集采取两种方法(FDM 方法和距平平均方法)对近 54 年中国气温变化幅度进行检测,得出了 1951—2004 年中国气温变暖更为准确的变化趋势。

2 数据情况及格点化数据集建立方法

2.1 资料及处理情况简述

本文选用 1951 年 1 月—2004 年 12 月中国所有基本、基准气象观测站的逐月平均气温数据(平均气温选取北京时间 02,08,14,20 时 4 个时次的算术平均)^[9]。自 1951 年以来,许多台站发生过业务改变、台站迁移、更迭等,有的已经取消,有的为新建站,因此共约牵涉到总的站点序列数为 728 个(图 1),目前尚在业务运行的还有约 670 个,从站点变化曲线(图 2)可以明显看到这一变化特点。

历史资料分析的一个重要方面是质量控制问题,这当中包括了一般的质量检查和均一性检查以及相应的调整和订正。我们借鉴了 Peterson 等^[10]的手段对历史气温数据进行了包括气候极值检查、一致性检查和空间检查在内的一系列检查,剔除了一些明显与气候事实不符的数据;进一步研究认为约 40% 的平均气温序列存在着非均一性问题,并采用统计方法和台站元数据结合的手段,对其中明显

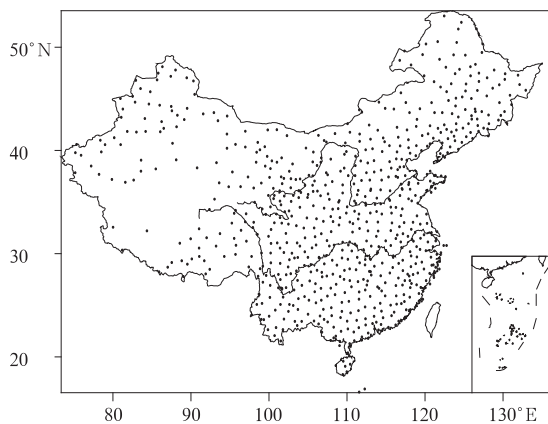


图 1 中国基本、基准气象站的地理分布

Fig. 1 Geographic locations of all reference and principal weather stations over Mainland China

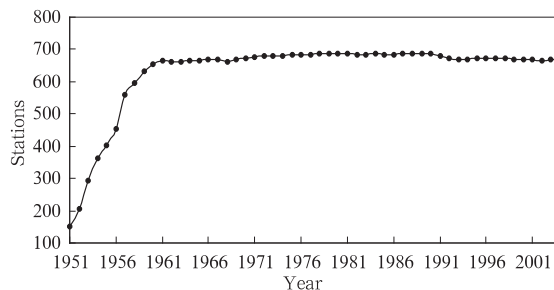


图 2 1951—2004 年观测站点个数的变化曲线

Fig. 2 Temporal variations of the station numbers over Mainland China during 1951–2004

的不连续序列进行均一化订正。订正结果表明,均一化数据集在描述区域气温变化尤其是局地气温变化时具有明显的优势^[11]。

2.2 数据格点化方法

数据网格化方法很多,各种方法本身由于其侧重点和应用对象的不同而有所区别。大体上可以分为两类:一类如反距离加权插值, Kriging 插值^[12], Cressman 插值等,其思路是根据某个时次的气象场离散点的值进行外推,预测(估计)网格点所在位置的值;另一类如上面提到的 CAM、RSM 和 FDM 等,他们的思路则是选取一定的网格范围,而将该范围内的所有离散点的时间序列进行加权平均,从而形成网格化的数据集。两种思路的出发点是不同的,前一种是针对某个时次(日、月或年)数据的客观分析,第二种则是针对某一个格点区域内的变化特征,从而形成格点序列。Kriging 方法在权重的选取、插值误差的估计方面有独特的优势^[13-14],因此本文选取该方法作为第一种思路的代表;而 FDM 方法则是第二种思路中能够尽可能将所有站点序列应用来进行网格化,充分利用数据信息,可以在一定程度上避免格点序列的非均一性等优势。如果单纯采用第一种思路,在站点数目变化明显的地区往往容易产生时间序列的非均一性问题;而第二种思路则由于对格点内站点数、序列长度等要求比较高,出现缺测数据格点比较多。因此本文选取上述两种方法对中国历史气温数据集进行网格化,一方面通过比较插值结果的异同可以互相比较插值效果,另一方面可以减少由于单一方法导致的随机性误差^[5]。

2.2.1 一级差分方法

Peterson 等^[5]的 FDM 方法的主要思路不再赘述,本文对该方法略作修改,将向前差分改为向后差

分式(1),这样可以避免由于序列起始年份站点过少,导致估计误差较大。

$$\delta_i = x_{t-1} - x_t \quad (1)$$

根据 Jones 等^[15] 对大尺度气温平均的空间抽样误差的估计结果,定义每个网格内的标准误差方差:

$$\sigma^2 = s^2(1 - r_m) / [1 + (n - 1)r_m] \quad (2)$$

作为一级差分方法插值误差的度量,显然,插值误差主要依赖于几个因素:(1) 网格范围内站点个数 n ; (2) 网格范围内各站点的相关性(r_m 为格点内各站点的平均相关系数); (3) 格点序列的时间变率(s^2 为格点序列的标准差)。显然,这种误差随时间的变化只与网格内的站点数量有关。

2.2.2 Kriging 插值方法

使用 Kriging 方法进行插值的第一步就是要确定变异函数模型,研究表明在对气象要素场进行插值的时候使用球面模型效果比较好。第二步就是在确定好的变异函数模型基础上进行插值计算,首先对空间场进行结构分析,包括采样点距离待插点的距离和采样点的空间分布,每个采样点的权重取决于半变异图、该点距待插点的距离和待插点周围的观测值的空间分布关系,特别是块 Kriging 方差的大小以及样本点和内插点的设置,接近内插点的样本点上具有较大的权系数,远离内插点的样本点的权系数较小甚至为 0,即内插点的估计主要决定于邻近的样本点。在对各点预测的时候,总希望各点预测的值之和与待插点的真实值之差为 0,这就要求误差方差越小越好。本文采用球面变异函数(其球面半径取 10),首先将各站点序列转化为所有时段上的距平序列,以避免站点变化对插值结果序列的影响^①。

3 数据格点化结果和效果讨论

考虑到如果网格尺度选的太细,插值出来的格点数据缺测太多,因此本文选取 $2.5^\circ \times 2.5^\circ$ 经纬度网格点进行 FDM 方法插值。在利用式(1)求取各网格内平均时,如果资料序列开始于一个极端冷(暖)年份时,最后一个点将出现异常,由此可能会较为敏感地影响求平均结果^[5]。因此在对最后一个点处理时,采取了3倍均方差的裁减方案:即如果最

后一个点超过了序列的 3 倍均方差,则去除这个点。

相关性往往用来表征序列之间的相似程度,本文计算了 FDM 计算的各个格点序列和原属该格点内的站点序列的相关性(图 3)(取 $n=50$ a, 0.1% 的显著性水平下的相关系数为 0.44),从图中来看,超过显著水平的地区几乎覆盖了全国,相关相对小的区域出现在西部边境、西南及长江流域部分地区,但一般也在 0.45—0.7;同样,各站点与相应站点的线性趋势差别也较小(图略)。因此,可以看出,利用一级差分方法得到的格点数据是比较忠实于站点数据集的。特别是全国平均的相关系数超过 0.90,说明格点数据场解释了站点气温场的方差比重非常大。同样,Kriging 方法插值前站点序列和格点序列也反映了类似的特点,更详细的比较另文详述^①。考察 FDM 方法得到的格点序列和 Kriging 方法得到的格点序列的线性相关性和空间的相似性。从年平均气温序列的结果来看,54 年来,两种插值方法得到的格点场之间的相关非常大,两种插值方法的效果均不错,可以认为插值方法的选取对网格化的结果影响是较小的。

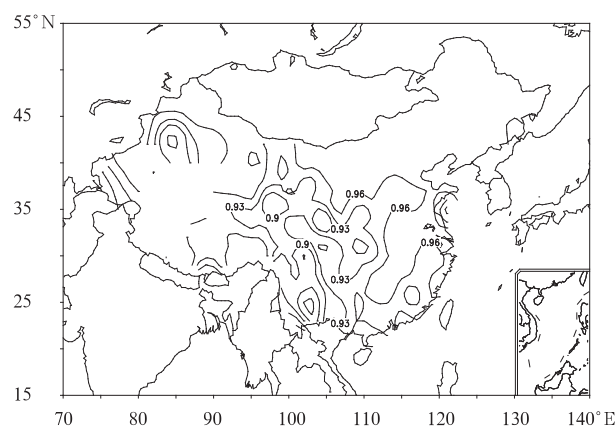


图 3 台站年平均气温和 FDM 方法得到的格点年平均气温的相关分布

Fig. 3 Correlation coefficients of station and grid (FDM) point annual mean temperature series

根据式(2)计算了从 1951—2004 年空间插值的平均抽样误差分布(图略),在中国东部站点较为密集的地区,插值抽样误差方差是较小的,约为 0.1,而在西部特别是青藏高原等站点稀疏的地区,抽样

① 李伟,李庆祥,江志红. 用 Kriging 方法对中国历史气温数据插值可行性讨论. 南京气象学院学报,待刊。

误差方差略大。但从全国平均来看,20 世纪 50 年代初期的抽样误差相对比较大,而从 60 年代以来,均处于一个较低的水平,由此可以看出,本文采用 FDM 方法和 Kriging 方法插值后的误差方差均比

较小(图 4a,b)。从二者的比较也可以看到,Kriging 方法插值的格点数据的误差方差绝对值比一级差分方法略大,但二者绝对值都很小。这也正反映了一级差分方法在 50 年代初的效果好于 Kriging 方法。

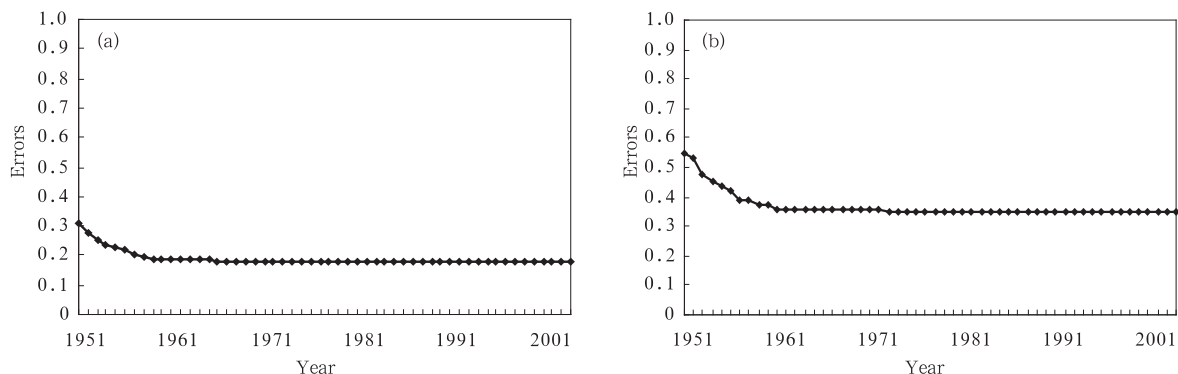


图 4 用一级差分方法(a)和 Kriging 方法(b)建立中国年均气温序列的抽样误差方差随时间变化

Fig. 4 Temporal variations of the spatially averaged sampling standard errors of the annual temperature series over Mainland China constructed using (a) FDM and (b) Kriging method

4 近 50 年年、季平均气温变化序列对比分析

4.1 年平均气温变化曲线

对于中国近百年尺度的气温变化幅度问题已经有很多的论述^[16-24],对近 100 年的气温增暖幅度和主要特征已经比较明朗。但过去许多研究大都是基于较少站点(如 160 个站点)的气温数据得到的。近年来,气候数据的质量控制技术得以不断发展并应用到各国的数据管理业务当中^[23-24],数据集的质量不断提高,使用加密以后的气候数据对于中国区域和全国气温变化的幅度问题也得到了一些研究者的关注^[25-26]。因此用均一化数据对近 50 年气温变化幅度(特别是近 10 年以来的异常气温增暖)进一步认识具有一定的科学意义。

图 5 给出了各个格点的年均气温趋势分布,总体来看,反映出的特点和其他工作^[21]类似。但我们分析的趋势分布在区域性上更加平缓,这主要是因为所用的气候数据经过了均一化的缘故。

为了避免由于计算方法差异带来的随机性影响,本文计算中国气温变化序列时,选取了两种思路:一种仍然是 Peterson 等^[5]的思路,即用 FDM;另一种是采用 Jones 等^[2]的思路,即计算距平序列的中国平均的方法(CAM)。

图 6 分别给出了用两种方法对两种格点数据集

计算得到的中国年平均气温距平变化曲线(考虑到在用 FDM 计算时,第一年的数值如超过 3 倍均方差时需要被舍弃,因此只用了 1952—2004 年的统计结果)。明显看出:4 条序列基本都较一致地反映了 53 年中国区域气温变化的类似情况,1998 年气温达到了最高值,2002 年为第二极端高值,自 1997 年以来,中国区域平均气温已经持续多年处于较高水平。

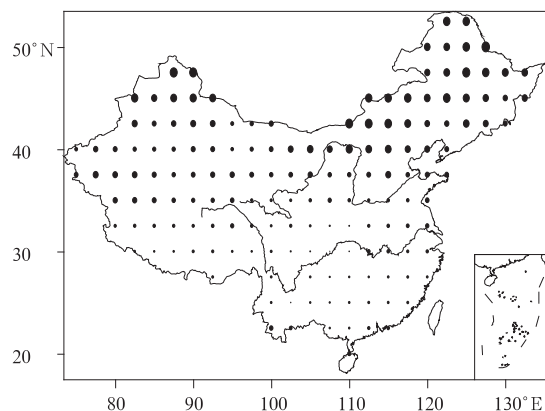


图 5 1952—2004 年中国年均气温变化趋势分布(最小值 0,最大值 0.588,黑圆点大小为线性变化关系,单位 $0.1^{\circ}\text{C}/(10\text{ a})$)

Fig. 5 Linear trends of annual temperatures for each grid box over 1952–2004 (the sizes of the black dots are linearly proportional to the trend value from 0 (smallest dot) to 0.588 (largest dot); unit: $0.1^{\circ}\text{C}/(10\text{ a})$)

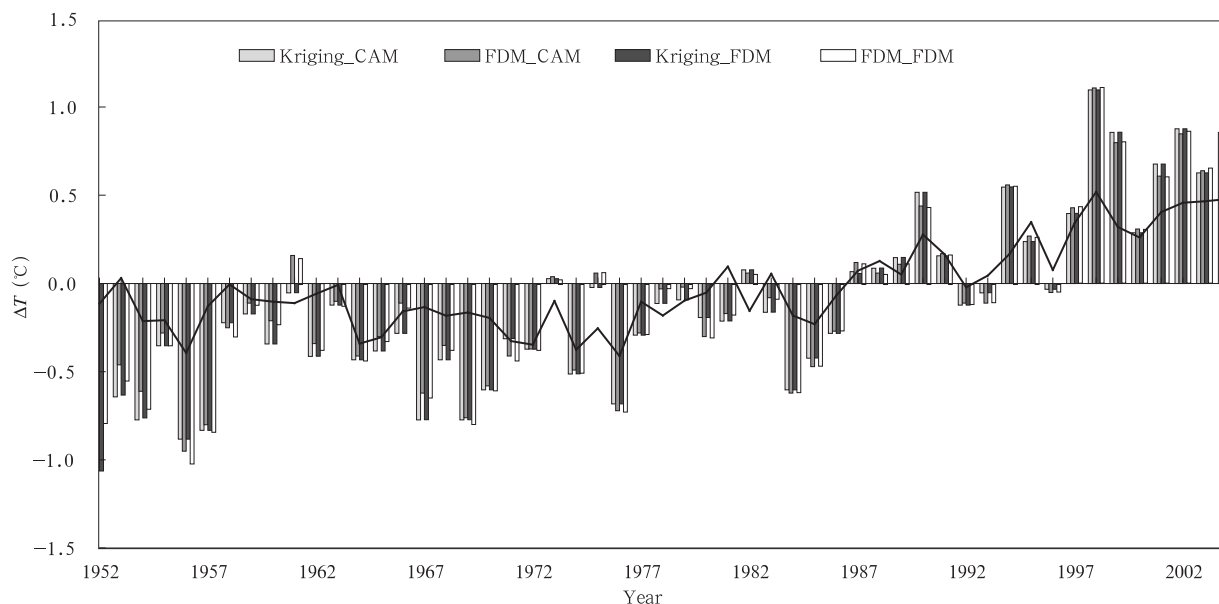


图 6 1952—2004 年中国年平均气温距平($\Delta T(^{\circ}\text{C})$)变化曲线
(粗实线为 Jones P D 的北半球序列)

Fig. 6 Time series of the annual surface air temperature anomaly ($^{\circ}\text{C}$) over Mainland China in 1952–2004
(the solid curve is the Northern hemispheric surface air temperature series by Jones P D)

这样看来,本文用 FDM 方法和 Kriging 方法建立的区域气温变化序列均是比较合理可信的;另一方面,在 20 世纪 50 年代初的 3 年(1952、1953、1954 年),用 Kriging 插值得到的结果要明显低于 FDM 方法的结果,主要是因为这几年站点资料过少,Kriging 方法仅仅通过东部地区的距平数据进行全国范围内的插值,因此不如 FDM 方法更可信。

在此基础上,还计算了 1952—2004 年年平均和各个季节气温变化序列的线性趋势(表 1)和均方差(表 2),以反映所建序列同以前类似研究的差异。结果表明:春、秋季气温序列的线性趋势和年均气温的趋势比较一致,在 $0.23\text{--}0.25\text{ }^{\circ}\text{C}/(10\text{ a})$,夏季序列的趋势仍然有 $0.11\text{--}0.14\text{ }^{\circ}\text{C}/(10\text{ a})$,冬季变暖最为明显,达到了 $0.39\text{--}0.41\text{ }^{\circ}\text{C}/(10\text{ a})$,近 50 年来,中国冬季平均气温升高了近 $2\text{ }^{\circ}\text{C}$,这个结论和其他专家的结论是较一致的。类似地,对两种方法建立的序列均方差进行了统计对比,不同的方法反映了类似的规律,即按冬季—春季—秋季—夏季的顺序依次变小;并且在数值上也比较一致,以一级差分的格点数据集作区域平均时方差略小。

从图 6 中国气温序列和 Jones 等同期的北半球序列进行比较,可以看到:中国和北半球平均气温的年代际变化特点基本类似。但很明显中国的气温变

表 1 1952—2004 年季平均与年平均气温
序列变化趋势($^{\circ}\text{C}/(10\text{ a})$)

Table 1 Trends of annual and seasonal temperature series in 1952–2004

区域平均	格点化	春季	夏季	秋季	冬季	年
一级差分	一级差分	0.26	0.11	0.21	0.41	0.24
	Kriging	0.24	0.14	0.22	0.39	0.25
距平平均	一级差分	0.24	0.11	0.20	0.39	0.23
	Kriging	0.25	0.14	0.22	0.39	0.25

表 2 1952—2004 年季平均与年平均气温序列均方差($^{\circ}\text{C}$)
Table 2 mean square deviations($^{\circ}\text{C}$)of annual and seasonal temperature series in 1952–2004

区域平均	格点化	春季	夏季	秋季	冬季	年
一级差分	一级差分	0.67	0.36	0.56	0.96	0.47
	Kriging	0.59	0.39	0.57	0.93	0.49
距平平均	一级差分	0.61	0.36	0.54	0.93	0.46
	Kriging	0.62	0.39	0.57	0.93	0.49

化幅度较北半球为大,计算 Jones 的北半球序列,1952—2004 年的线性趋势为 $0.13\text{ }^{\circ}\text{C}/(10\text{ a})$,大约为中国气温趋势的一半,出现这个结果的原因从图中可以看出主要是因为 20 世纪 50—80 年代中国气温低于北半球气温,而 80 年代后期特别是 90 年代后期以来中国气温明显高于北半球平均气温,这个现象值得进一步深入研究。

4.2 季节平均气温变化曲线

根据上述可知,在 20 世纪 50 年代后期以来,两种格点数据集的效果差别不大,结合前面的误差方差分析结果,认为 FDM 方法建立的格点数据在 50 年代初更加合理,且用 FDM 方法进行格点化的物理含义更加清楚。图 7 给出用 FDM 插值后的格点序列得到的 1952—2004 年中国季节气温变化。显

然,各个季节的气温变化和年均气温变化特点既有相当的联系,但也存在着一些区别。例如,从 54 年气温最高的 1998 年来说,春、夏季并不是处于最高值的水平,而是因为秋季、冬季气温异常偏高所致。由此可以看出,50 年气温偏高、偏低的年份出现的原因也不尽相同,但总体趋势都是偏暖年份越来越多,偏冷年份越来越少。

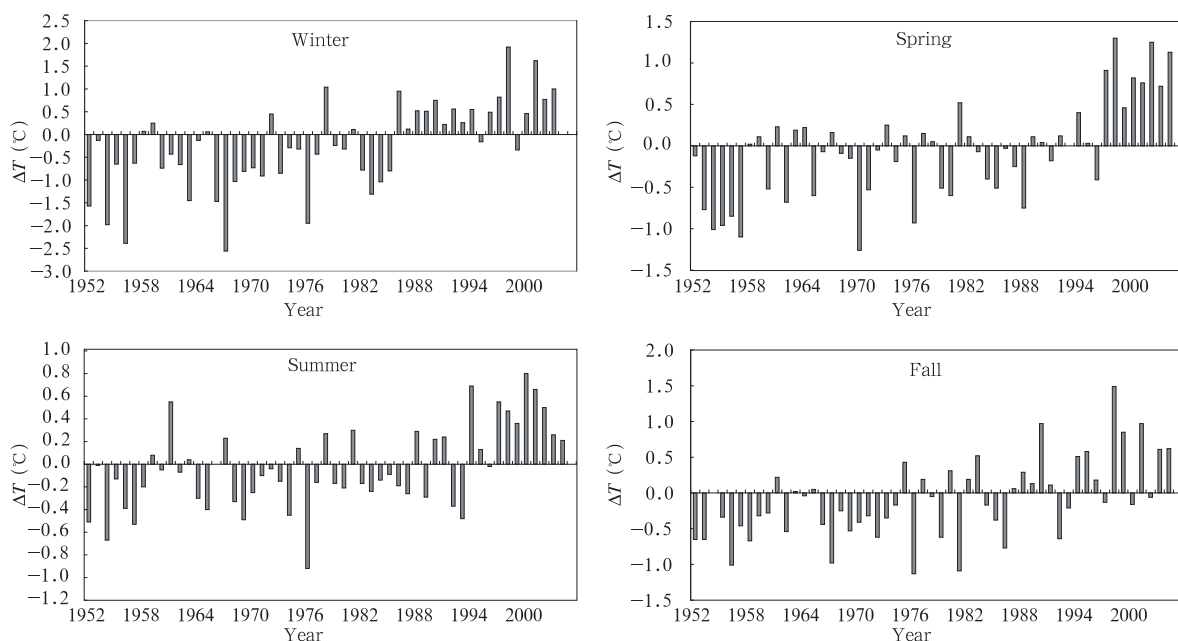


图 7 1952—2004 年 FDM 插值计算的中国各季节平均气温变化序列

Fig. 7 Time series of the seasonal surface air temperatures (FDM interpolation) over Mainland China during 1952—2004

5 结论与讨论

与以前的一些类似工作^[27-28]相比,本文的特色在于:

(1) 基础数据——站点气温序列的均一化过程结合了台站元数据信息和统计方法,使得数据均一化过程更加科学,数据质量更好;

(2) 通过采取不同方法的对比论证,认为格点化方法对于这种站点密度下中国区域气温格点数据集研制的影响是比较小的,从而也更加证明了格点化数据集的可靠性;

(3) 本文的数据在空间覆盖上更加全面。这种格点化的思路对于其他气候要素(如气压、高空要素等)也比较适用,因此具有一定的推广和应用前景。

由于格点化方法和区域平均方法的选择,导致

区域气候变化检测的计算结果有一定的差别,采取多种方法结合的手段可以尽可能降低这种计算的随机性^[5]。本文利用两种格点化的数据集,分别用一级差分方法(FDM)和距平平均(CAM)方法建立了近 50 年中国区域(不含台湾、香港、澳门和西沙群岛)年、季气温变化序列。通过对比分析认为:不同格点化方法、不同区域平均方法对于中国近 50 年年、季气温增暖特点的描述是较一致的,这正说明本文通过将站点数据进行格点化后建立的区域平均序列是较准确的。

由于近 50 年来中国区域城市热岛对区域增暖的贡献水平比实际气温变化幅度低一个量级左右^[25,29-30],因此本文在对数据集格点化时并没有去除城市热岛的影响。

致谢:本文完成过程中,曾多次和美国国家气候资料中

心(NCDC)的 Thomas C. Peterson 博士讨论一级差分方法
和应用,谨致谢意。感谢审稿专家的建设性意见。

参考文献

- [1] Jones P D. Hemispheric Surface air temperature variations: Recent trends update to 1987. *J Climate*, 1988, 1: 654-660
- [2] Jones P D. Hemispheric Surface air temperature variations: A reanalysis and an update to 1993. *J Climate*, 1994, 7: 1794-1802
- [3] Jones P D, Raper S C B, Bradley R S, et al. Northern hemisphere surface air temperature variations, 1851-1984. *J Climate Appl Meteor*, 1986, 25: 161-179
- [4] Jones P D, Raper S C B, Bradley R S, et al. Southern hemisphere surface air temperature variations, 1851-1984. *J Climate Appl Meteor*, 1986, 25: 1213-1230
- [5] Peterson T C, Karl T R, Jamason P F, et al. First difference method: Maximum station density for the calculation of long-term global temperature change. *J Geophys Res*, 1998, 103: 25967-25974
- [6] Hansen J, Lebedeff S. Global trends of measured surface air temperatures 1987. *J Geophys Res*, 1987, 92: 13345-13372
- [7] Hansen J, Lebedeff S. Global surface air temperatures: update through 1987. *Geophys Res Lett*, 1988, 15: 323-326
- [8] PCC, 2001: Climate Change 2001: The Science of Climate Change // Houghton J T, Ding Y, et al. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate
- [9] 中国气象局. 地面气象观测规范. 北京:气象出版社, 1979: 186pp
- [10] Peterson T C, Russell S V. An overview of the Global Historical Climatology Network temperature data base. *Bull Amer Meteor Soc*, 1997, 78: 2837-2849
- [11] Li Q, Liu X, Zhang H, et al. Detecting and adjusting on temporal inhomogeneity in Chinese mean surface air temperature dataset. *Adv Atmos Sci*, 2004, 21(3):260-268
- [12] Oliver M, Webster A, Kriging R. Method of interpolation for geographical information system. *Int J Geogr Inf Syst*, 1990, 4:313-332
- [13] 冯锦明,赵天保,张英娟. 基于台站降水资料对不同空间内插方法的比较. *气候与环境研究*, 2004, 9(2): 261-277
- [14] 刘峰. 应用 Kriging 算法实现气象资料空间内插. *气象科技*, 2004, 32(2): 110-115
- [15] Jones P D, Osborn T J, Briffa K R. Estimating sampling errors in large-scale temperature averages. *J Climate*, 1997, 10: 2548-2568
- [16] 屠其璞. 近百年来我国气温变化的趋势和周期. *南京气象学院学报*, 1984, 7(2):151-161
- [17] 丁一汇,戴晓苏. 中国近百年来的气温变化. *气象*, 1994, 20(12):19-26
- [18] 林学椿,于淑秋. 近 40 年我国气候趋势. *气象*, 1990, 16(10): 45-53
- [19] 翟盘茂,任福民. 中国近 40 年来最高最低温度变化. *气象学报*, 1997, 55(4):418-429
- [20] 王绍武,叶瑾琳,龚道溢等. 近百年中国年气温序列的建立. *应用气象学报*, 1998, 9(4):396-401
- [21] 秦大河. 中国西部环境演变评估(综合卷). 北京:科学出版社, 2002:80pp
- [22] 陈隆勋,周秀骥,李维亮等. 中国近 80 年来气候变化特征及其形成机制. *气象学报*, 2004, 62(5):634-646
- [23] 李庆祥,刘小宁,张洪政等. 定点观测气候序列的均一性研究. *气象科技*, 2003, 31(1):2-11
- [24] Peterson T C, Easterling D R, et al. Homogeneity adjustments of in situ atmospheric climate data: a review. *Int J Climat*, 1998, 18:1493-1517
- [25] Li Q, Zhang H, Liu X, et al. UHI effect on annual mean temperature during recent 50 years in China. *Theor Appl Climatol*, 2004, 79: 165-174
- [26] 任国玉,郭军,徐铭志等. 近 50 年中国地面气候变化基本特征. *气象学报*, 2005, 63(6): 942-956
- [27] Song F, Qi H, Qian W. Quality control of daily meteorological data in China, 1951-2000: A new dataset. *Int J Climat*, 2004, 24: 853-870
- [28] Zhai P, Chao Q, Zou X. Progress in China's climate change study in the 20th century. *J Geogr Sci*, 2004, 21(z1): 3-11
- [29] 黄嘉佑,刘小宁,李庆祥. 中国南方沿海地区城市热岛效应与人口的关系研究. *热带气象学报*, 2004, 20(6):713-722
- [30] Zhou L, Dickinson R E, Tian Y, et al. Evidence for a significant urbanization effect on climate in China. *Proceedings National Academy of Sciences USA*, 2004. 101: 9540-9544

CONSTRUCTION OF THE GRIDDED HISTORIC TEMPERATURE DATASET OVER CHINA DURING THE RECENT HALF CENTURY

Li Qingxiang^{1,2} Li Wei³

1 National Meteorological Information Center, CMA, Beijing 100081

2 Laboratory for Climate Studies, CMA, Beijing 100081

3 National Meteorological Center, CMA, Beijing 100081

Abstract

The detection of large-area averaged temperature changes using long-term in situ observations is always hampered by the spatial sampling errors caused by the different lengths of series. Usually, people need interpolate these stations' temperature series onto grid boxes/points on regional or global scales. Based on the quality controlled and homogenized observational monthly mean surface air temperature data series over Mainland China (the number of the stations added up to 728, and the stations of HongKong, Taiwan and Macao were excluded in this study), traditional Kriging methods and modified First Difference Method (FDM) were used in this paper to grid the monthly historic temperature dataset during Jan 1951—Dec 2004 over Mainland China by $2.5^{\circ} \times 2.5^{\circ}$ resolutions, and much validation work was performed later to confirm the quality of the grid dataset. Comparison analysis shows no matter by which method the grid dataset is constructed: (1) the spatial sampling errors of each grid box/point are relatively small, which means the interpolation is relatively accurate for a local region; (2) the correlations between gridded series and corresponding station series are very high, which means the gridded series did not change the original inter-annual and decadal regional climate change in station series; and (3) the anomaly distribution shows very high consistencies even by using different interpolating methods. Of course, these high correlations and consistencies between two different methods can also prove the rationality of the two versions of gridded datasets. In addition, the linear trend of surface air temperature changes over China during recent 54 years were also analyzed with the two versions of gridded datasets constructed by First Difference Method (FDM) and widely used Climate Anomaly Method (CAM), similar results were achieved. During recent 50 years (1952—2004), the annual temperature over China shows consistently warming trends, with the most significant warming trend in Northern China, especially in Northeast China and Northwest China. It was near $0.6^{\circ}\text{C}/10\text{a}$ in the maximum trend grid box in Northwest China, but much weaker in the upper-middle reaches of the Changjiang River Basin (less than $0.1^{\circ}\text{C}/10\text{a}$ in the lowest trend region). The whole country-averaged linear trend during the 54 years was about $0.23\text{--}0.25^{\circ}\text{C}/10\text{a}$, and for seasonal means, the wintertime temperature's warming trend was the largest, while the summer's was the smallest. All these results are coincident with many earlier researches, but our evaluation in this study seems more reliable and accurate due to the facts that the dataset we used is systematically quality controlled.

Key words: Gridded dataset, Surface air temperature, Warming, Error, Homogeneity.