

近 57 年我国气温格点数据集的建立和质量评估^{*}

张 强 阮 新 熊安元

(国家气象信息中心,北京 100081)

摘 要

在普通克里金插值方法的基础上,引入高程因子并充分考虑插值的边界效应,对 1951—2007 年我国气温站点资料进行空间结构分析和插值,得到我国地面气温日、月、年平均值 $1^{\circ} \times 1^{\circ}$ 格点数据集。数据集的质量评估结果表明:高程在我国区域气温空间结构分析和插值中起着重要作用,高程资料的引入有效提高了大部分高山地区的插值效果;相比站点资料,所建立的格点数据集在描述我国年平均气温以及季节平均气温分布时更为合理,突出了温度场的大尺度特征;数据集反映出了我国年平均气温变化趋势主要的空间差异;数据集较好地反映了我国年平均气温变化状况,1951—2007 年气温变暖幅度约为 1.6°C , 增温速率 $0.28^{\circ}\text{C}/10\text{a}$, 比全球或半球同期平均增温速率明显偏高,且气温增暖主要发生在最近的 20 余年之内。另外,格点数据集显示,1998—2007 年是 1951 年以来最暖的 10 年,其中 2006 年全国平均气温距平接近 2000 年之前的历史最高年份 1998 年,而 2007 年全国平均气温距平值超过 1998 年,达到 1951 年以来的历史最高值 1.3°C , 为最暖的一年。

关键词: 气温; 空间插值; 格点数据集; 质量评估

引 言

基于定点持续观测积累的气象要素序列,由于站点空间分布不均、序列长短不一、观测台站环境变迁等等缺陷,在气候分析和数值模拟等研究中,不能完全真实代表区域气候变化特征,在实际应用中面临诸多限制。利用空间插值技术将离散的站点资料转换成规则的网格点序列,可以有效地反映气象要素的空间信息,大大提高了序列在对应网格范围的气候代表性。

国外科研机构非常重视站点资料的网格化,进行了大量的研究工作^[1-7],根据不同研究方向和使用目的,建立了不同版本的全球或区域格点数据集,这些数据集中较有影响力的包括:美国国家海洋大气局(NOAA)研制的在降水网格化研究中最具代表性的全球综合分析降水数据集(CMAP),英国 East Anglia 大学的 Climatic Research Unit (简称 CRU) 研制的高分辨率地表气候变量数据集(CRU TS 2.1),美国研制的全球海洋大气综合数据集(COADS),以及 Jones 等^[1]全球气温格点数据集,Han-

son 等^[2]气温数据集,Hulme^[3]降水数据集等。这些数据集作为认识气候变化的基础数据已经被 IPCC 作为主要的参考依据^[8]。对于中国区域而言,这些格点数据集所包含的站点观测信息有限(通常只有大约 200 个台站的观测资料),且局地地理信息相对欠缺,没有对历史资料的均一性进行检验,资料的质量及其在中国区域的代表性存在一定问题,在中国部分地区误差较大。随着气象业务发展的需要,近几年国内也开展了站点资料网格化的研究,但都比较零散^[9-17],并没有建立一个完整的、长时间序列的、大家普遍认可的地面要素格点数据集。

本文利用国家气象信息中心最新研制的均一化气温数据集,首先对常规空间插值方法的插值效果进行对比检验,根据检验结果选取普通克里金插值法(ordinary Kriging)作为我国地面温度场空间插值的基本方法。同时,引入数字高程资料尽可能消除我国独特地形条件下高程对气温空间插值精度的影响,并利用全球历史气候网数据集资料对我国站点进行了外围扩充,以减小由于边界资料稀少造成的空间插值边界效应。在此基础上,建立了近 57 年我国 $1^{\circ} \times 1^{\circ}$ 气温格点数据集,本文重点介绍了对数据

* 国家科技基础条件平台建设项目“气象科学数据共享中心”(2005DKA31700-02)资助。

2008-11-11 收到,2009-05-20 收到再改稿。

集从天气学和气候学两方面进行质量检验和评估的主要结果。

1 资料与方法

1.1 资料来源和处理

用于空间插值的资料主要包括3个部分:国家气象信息中心气象资料室最新研制的1951—2007年我国731站均一化日值和月值气温数据集,美国国家气候资料中心(NCDC)研制的1951—2007年全球历史气候网数据集(GHCN_V2.0),利用全球30"分辨率高程数据集(Global 30 Arc-Second Elevation Data Set,简称GTOPO30)生成的全国范围 $1^{\circ} \times 1^{\circ}$ 的数字高程模型资料(DEM)。

由于插值过程中客观存在的边界效应,对于中国区域,如果只使用国内的观测资料,边界地区的空间结构信息只能依靠边界内部资料确定的空间结构信息,通过数学外推方法获得,这将导致空间分析误差,对插值结果造成较大影响。为了消除边界效应,将资料的选取范围扩大是最直接和有效的方法。本文在全国731个站点资料的基础上,引入NCDC研制的GHCN_V2.0对中国区域以外的资料进行补充,以减小边界效应造成的误差。GHCN数据集作为WMO监测和评估当今全球气候变化的核心参考数据^[18-19],具有较高的可信度。资料选取的具体范围为 $10^{\circ} \sim 60^{\circ} \text{N}$, $70^{\circ} \sim 150^{\circ} \text{E}$ 。图1是用GHCN资料进行中国区域外围资料补充前后的台站数量对比图。图1显示:经过GHCN资料的补充,20世纪90年代之前参加空间插值的站点数有大幅度提高,特

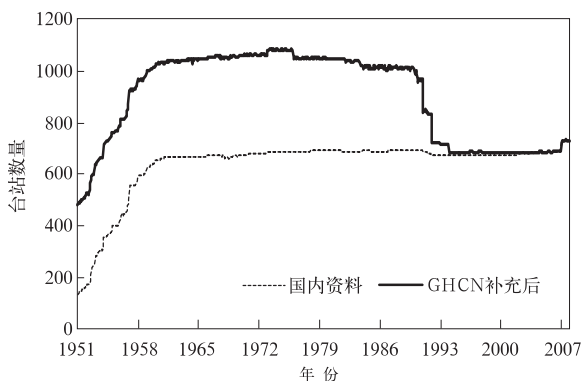


图1 1951年1月—2007年12月GHCN资料补充前后台站数量对比

Fig. 1 The station number variation after using GHCN in China from January 1951 to December 2007

别是60—90年代,参加空间插值的站点数由600个左右提高到1000个左右。90年代以后,由于GHCN资料包含中国区域的站点数较少,站点补充不明显,外围资料对减小边界地区的插值误差作用有限,但90年代以后的台站相对稳定、且密度较大,边界效应不如之前显著。

用GHCN资料补充后,参加实际空间插值的站点空间分布显示:除我国东南沿海地区观测台站较少外,中国区域其他边界周围的台站补充较为充分,特别是东部沿海地区,日本和朝鲜半岛的外围资料补充十分明显(图略)。

1.2 格点化方法

由于各种空间分析方法基于不同的地理假设和不同的数据建模方法,对于众多的空间数据分析方法而言,没有绝对最优的空间插值方法,只有特定条件下的最优方法^[20]。研究证明,并不是所有方法都适应气象要素场的要求,本文从常用于气象要素的空间插值方法中,重点选取反距离加权插值法、普通克里金插值法、最近临点法、线性插值三角网法、多元回归法以及局部多项式法6种插值方法进行温度场插值效果检验。

1.2.1 检验标准和结果

采用交叉验证法(cross validation)验证以上6种不同插值方法对实际温度场的插值效果。交叉验证假设每个站点的气象要素值未知,用周围站点的值来估算,然后根据所有站点实际观测值与估算值的误差大小评判插值方法的优劣。具体采用平均误差(mean error,简称ME)、平均绝对误差(mean absolute error,简称MAE)、均方根预测误差(root-mean-square prediction error,简称RMSPE)作为评估插值效果的标准。对1998—2007年近10年全国7月温度场(统一订正到3000 m拔海高度)分别使用6种空间数据分析方法进行空间插值,并计算出每种方法ME、MAE和RMSPE的10年平均值,结果如表1所示。

由表1可知,在气温的空间插值中,普通克里金方法相对其他几种方法各项误差较小,对温度场的插值具有较明显的优势。

1.2.2 普通克里金插值方法的改进

普通克里金插值方法以空间结构分析为基础进行估值,它利用了数据空间场的性质,在插值过程中可以反映空间场的各向异性,并且充分利用数据点之间的空间相关性。从表1的对比中,可以看出普通克里金方法在诸多气象要素场的空间分析方法中

表 1 6 种插值方法对近 10 年全国 7 月温度场的空间插值交叉验证结果

Table 1 The spatial interpolation effect of ordinary Kriging and five other methods over China in July during 1998—2007

误差	反距离加权	普通克里金	最近邻点	线性/三角网	多元回归	局部多项式
ME	-0.01407	0.0075	0.02804	-0.03156	0.00008	-0.11671
MAE	0.65751	0.54506	0.74918	0.56089	1.51776	0.73674
RMSPE	0.92832	0.78538	1.06555	0.7858	1.99304	1.00241

具有一定的优势,但我国地形复杂,特别是青藏高原大地形作为特有的地理因素,对温度场的分布影响巨大^[13]。如何在复杂地形条件下利用有限的观测数据插值得到气温的合理空间分布始终是一个难题。普遍认为,为了减小误差就必须将高程对气温的影响作为一个重要的因素来考虑。本文在普通克里金方法的基础上,引入高程资料对其进行了改进,以适应我国特殊地形的要求。

假设空间变量由背景值和分析值增量组成,那么,某一网格上的气温可以被表示为:

$$T(X_0) = T^*(X_0) + \Delta T(X_0) \quad (1)$$

式(1)中, $T^*(X_0)$ 是气温背景值,由式(2)计算:

$$\begin{cases} T^*(X_0) = \sum_{i=1}^n \lambda_i T^*(X_i) \\ T^*(X_i) = T(X_i) - \gamma[H(X_i) - H^*] \end{cases} \quad (2)$$

式(2)中, $T^*(X_i)$, $T(X_i)$ 和 $H(X_i)$ 分别是第 i 个气象站的气温背景值、气温观测值和数字高程; H^* 是进行气温订正的高程; γ 是气温直减率(本文取值 $0.65\text{ }^\circ\text{C}/100\text{ m}$)。当气温被订正到同一高程(本文取值 1000 m)后,可以被认为是一个随机变量,采用普通克里金方法对其进行估值。 $\Delta T(X_0)$ 是分析增量,

$$\Delta T(X_0) = \gamma[H(X_0) - H^*] \quad (3)$$

式(3)中, $H(X_0)$ 是待计算网格的高程,取值于DEM。

1.2.3 我国地面温度场空间结构分析

利用改进的普通克里金方法进行中国区域温度场的插值,其插值精度取决于对温度场空间结构的准确分析。按照空间变异理论,在连续的空间介质范围内,变量性质随距离具有某种相关关系,这种关系可以通过已知的基本变异函数模型进行有效拟合^[21]。对于气温场而言,这种相关性体现在空间任意两个样本点随着距离 h 的增加(图2),其温度相关性呈现某种趋势的变小(对应变异函数值 $r(h)$ 逐渐增大,如图2所示),当样本间的分离距离超过一定范围(变程)后,样本点在空间上就不再相互关联。对于气象要素的月平均值,温度场的变程通常为 1200 km ,而降水只有 450 km 。本文中温度场变异函数的变程控制在 $20\sim 22$ 格距(约 1000 km 左右)之间。

本文通过试验证明,温度场的空间相关性受高程的影响较为明显,图2显示了这种影响和变化。由图2可见,原始站点温度场和经过高程订正到同

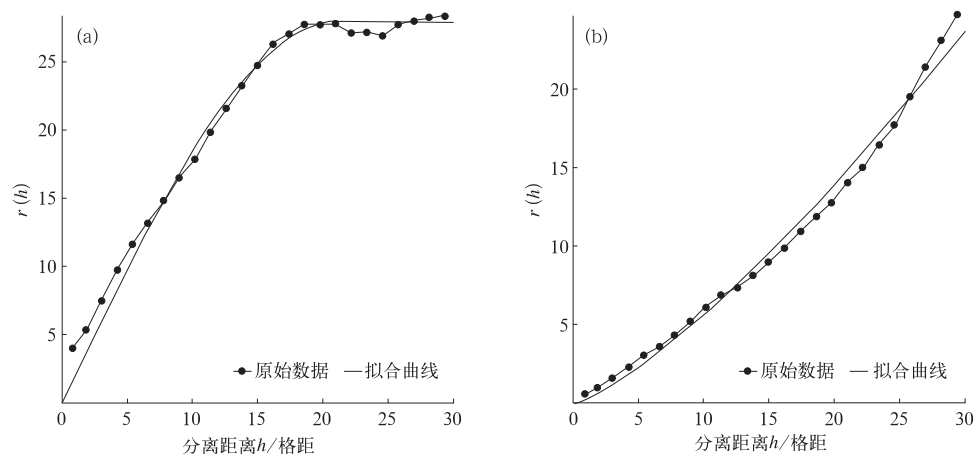


图 2 高程订正前后 1980 年 7 月我国地面月平均气温空间结构及试验变异函数拟合效果对比 (a)原始站点温度,(b)高程订正后的温度
Fig. 2 Original(a) and adjusted(b) spatial structure of the monthly mean temperatures over China in July 1980

一海平面的温度场在空间结构上有较大变化,这种变化一方面体现在空间变异函数模型的不同,另一方面,也可以看出,原始站点温度场的变异函数(图2a)的样本点并未通过坐标原点,这和空间变异理论相矛盾,其主要原因是高程本身产生的温度递减率削弱了空间两个临近点在水平空间上的相关性,而经过高程订正后的图2b没有出现这种现象。由此可见,通过高程订正,可以消除高程在温度场空间结构分析中的这种不利影响,得到的空间结构能比较真实反映温度场样本点之间的相关性,最终提高空间插值的精度。

对1951年1月—2007年12月我国温度场的空间结构分析表明:中国区域地面气温变异函数的试验曲线基本可以被拟合为幂函数模型或是高斯模型两种。

2 数据集质量评估

鉴于本文建立的格点数据集是对历史气候数据进行空间插值,而不是对某一个时次的要素场进行插值。因此,对数据集的质量评估重点从格点序列和原始站点序列的相关性以及所描述的中国区域累年气温空间分布、趋势分布等几个方面进行,以检验格点数据集能否应用于气候变化等方面的研究。

2.1 相关性检验

序列的相关性可以检验插值后的格点数据集能否准确反映原来观测序列的变化特征。选择中国区域1961—2007年有完整观测资料的560个台站,逐站计算了这些台站的年平均气温,以及1,4,7,10月的月平均气温观测资料序列与其临近格点(空间上与站点距离最近的格点)序列的相关。

结果表明:94%的台站年平均气温观测资料序列与其临近格点序列的相关系数超过0.9,即使在地形复杂的西南地区,序列相关也普遍超过0.9,仅个别站为0.7~0.8(图3)。在1961—2007年有完整观测资料海拔在3000 m以上的34个台站中,仅天山南脉的吐尔尕特和喜马拉雅山脉东部的隆子两个站观测序列与插值序列的相关低于0.6,喜马拉雅山脉中部的定日和横断山区的稻城相关分别为0.6和0.7,其他30个高海拔台站的序列相关都在0.8以上,24个站相关系数超过0.9(表略)。这与插值过程中引入高程资料以减少地形对气温空间插值的影响有直接关系。进一步对1,4,7,10月的月平均气温观测资料序列与其临近格点序列的统计表明,相关系数超过0.9的台站分别为96%,98%,95%和97%,而各月序列相关低于0.6的台站比例均不超过2%,且集中在边境和沿海岛屿,如中俄边境上的满州里和呼玛、中蒙边境上的满都拉、东部沿海的成山头、大陈岛等地(图略)。

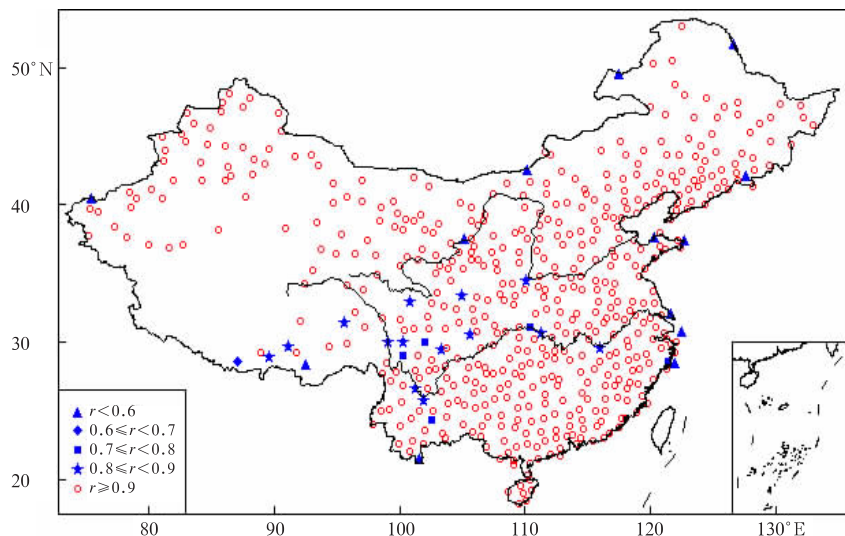


图3 1961—2007年560个台站年平均气温观测资料序列与其临近格点序列的相关系数空间分布

Fig. 3 Correlation coefficients of the annual mean temperature in 560 stations with that in its adjacent grid point during 1961—2007

总体上,本文建立的格点数据集与原始站点序列在全国绝大部分地区的相关程度很高,反映了我国大陆年平均气温和季平均气温的基本特征。对于全国而言,由于使用了国外气象资料消除了边界效应,在边界上的插值误差总体减小,相关性提高。另外,数字高程资料的引入,提高了大部分高山地区,特别是青藏高原大部分地区的插值效果,但个别边界站点仍然受站点稀少的影响导致插值效果不理想,内陆个别站点由于处于寒温带交界处或极值区插值效果相对较差。总体而言,在全国有台站分布的大部分地区,站点和临近格点序列具有很好的相关性,插值效果较好。在台站分布稀疏的西部地区,序列相关性检验具有一定的空间局限性,可以通过对比气温的空间分布进行插值效果检验。

2.2 累年空间分布

从 1971—2000 年的年平均气温分布来看(图 4),本文建立的格点数据集比较准确地描述了我国大陆温度场主要特征。在东部地区,格点数据集平

滑掉了东南和西南部高山站造成的范围很小的极值中心,对温度的纬度地带性变化反映很好。而在西部地区,数据集很好地反映了天山南北麓、塔里木盆地和柴达木盆地等大地形下温度随地形高度的变化,塔里木盆地南缘的温度梯度、等值线的走向更为合理,并准确反映出了青藏高原上不同海拔地区的温度差异,比如昆仑山脉到喀喇昆仑山脉、祁连山和巴颜喀喇山附近的低值区。而站点资料(图 4a)在西部,特别是青藏高原上过于稀少,存在等值线走向以及对温度随海拔高度的变化描述不合理等问题。总体来讲,本数据集描述的我国年平均气温分布更为合理,突出了温度场的大尺度特征,但与站点资料所描述的温度场相比数值偏低,在东部更为明显。

进一步分析了 1971—2000 年温度场累年季节空间分布的变化,结果表明:格点数据集在描述 1 月和 7 月的月平均气温分布上的优越性及与站点资料的差异和年平均气温一致(图略)。

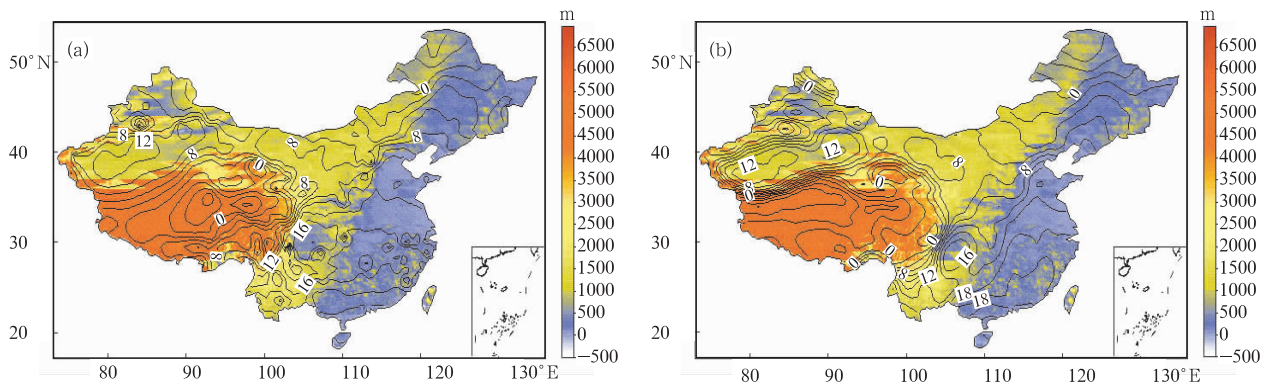


图 4 1971—2000 年我国年平均气温空间分布(单位: $^{\circ}\text{C}$; 底图不同颜色代表数字高程)

(a) 站点资料二维插值, (b) 格点数据集

Fig. 4 Spatial distribution of the annual mean temperature over China during 1971—2000

(unit: $^{\circ}\text{C}$; the different colors denote DEM) (a) station data, (b) grid data sets

2.3 累年趋势分布

对比我国 1961—2005 年年平均气温趋势分布(图 5)可以看出,格点数据集准确反映出了我国年平均气温变化主要的空间差异:北方增温趋势比南方更明显,幅度普遍超过 $0.3\text{ }^{\circ}\text{C}/10\text{ a}$,新疆北部、内蒙古大部分地区和青藏高原广大地区增温幅度超过 $0.4\text{ }^{\circ}\text{C}/10\text{ a}$;而南方大部分地区增温幅度不超过 $0.2\text{ }^{\circ}\text{C}/10\text{ a}$,西南地区增温幅度最小,其中四川盆地东部和云贵高原北部地区还呈现弱的降温趋势。

这与国内多年前和最近的研究结果一致^[22-26]。

相比站点序列,格点数据集在描述我国年平均气温趋势分布时有以下几个方面的差异:格点数据集对巫山附近的负趋势没有反映出来;对内蒙古西部趋势在 $0.4\text{ }^{\circ}\text{C}/10\text{ a}$ 以上的高值区强度和范围的描述都偏大;东北地区的高值区范围一致,但强度偏大。此外,格点数据集和站点资料对青藏高原南部和横断山区气温趋势的反映差异较大。

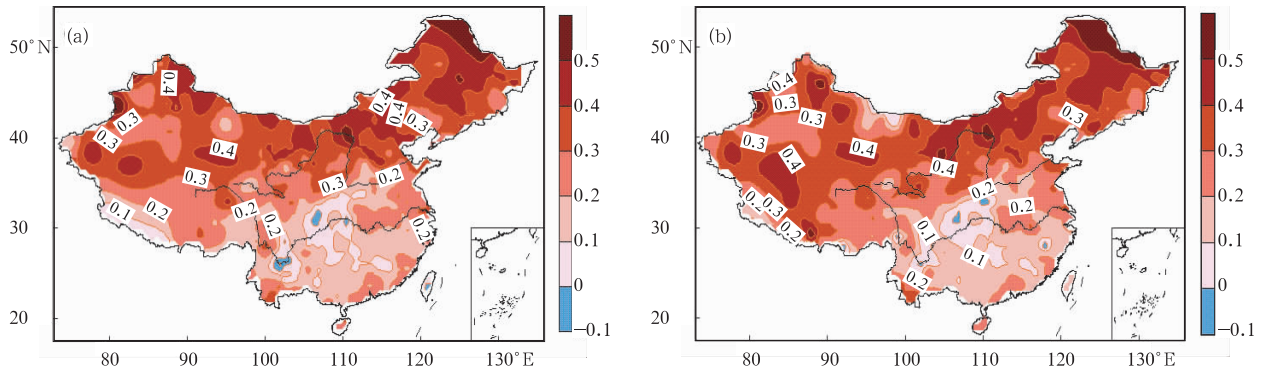


图5 1961—2005年我国年平均气温变化趋势的空间分布(单位: $^{\circ}\text{C}/10\text{a}$)

(a) 站点资料二维插值, (b) 格点数据集

Fig. 5 Trends in annual mean temperature over China during 1961—2005(unit: $^{\circ}\text{C}/10\text{a}$)

(a) station data, (b) grid data sets

2.4 累年季节趋势分布

对比 1961—2005 年 1 月平均气温趋势分布(图 6a, 6b)可以看出, 就全国范围来讲, 格点数据集和站点资料增温趋势超过 $0.3\text{ }^{\circ}\text{C}/10\text{a}$ 以及低于 $0.2\text{ }^{\circ}\text{C}/10\text{a}$ 的覆盖范围均非常一致; 我国西部增温趋势为 $0.5\text{ }^{\circ}\text{C}/10\text{a}$ 以上的 3 个高值区和内蒙古西部的趋势高值区, 长白山南部的高值区, 海拉尔附近以及西

南到华中地区增温趋势小于 $0.1\text{ }^{\circ}\text{C}/10\text{a}$ 的 4 个低值区从强度到范围均吻合很好。但站点资料阿尔泰山东面和二连浩特以北弱的正趋势, 格点数据集描述为负趋势, 格点数据集反映的黑龙江东北部高值区的强度和范围都超过站点资料所描述的情况。此外, 格点数据集描述的攀枝花附近的负趋势中心强度弱、范围小。

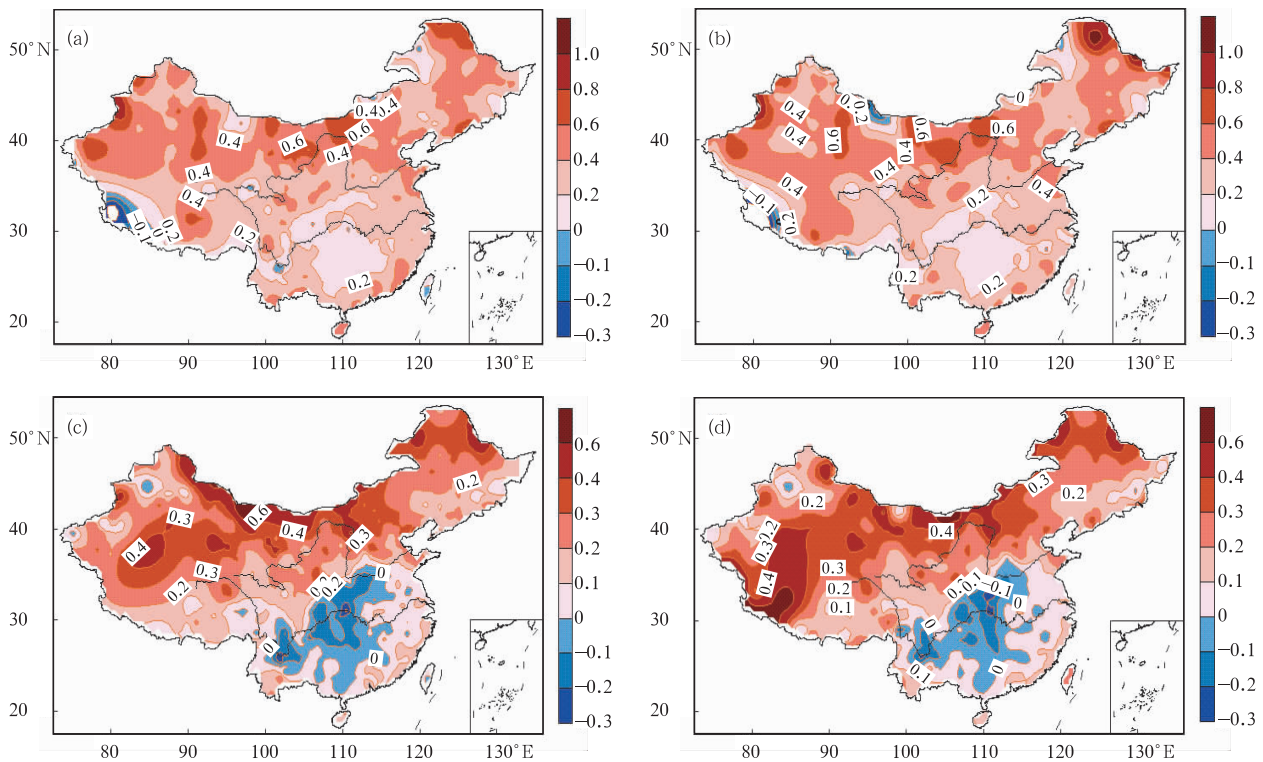


图6 1961—2005年我国1月和7月平均气温趋势(单位: $^{\circ}\text{C}/10\text{a}$)

(a) 1月站点资料二维插值, (b) 1月格点数据集, (c) 7月站点资料二维插值, (d) 7月格点数据集

Fig. 6 Trends in monthly mean temperature over China in January and July during 1961—2005(unit: $^{\circ}\text{C}/10\text{a}$)

(a) station data in January, (b) grid data sets in January, (c) station data in July, (d) grid data sets in July

格点数据集与站点资料对 7 月平均气温趋势的表现大部分地区都非常一致(图 6c, 6d): 北方各增温趋势对应的区域、南方呈降温趋势区域的位置、范围及强度都非常一致。但格点数据集对青藏高原 7 月气温趋势的表现与站点资料有较大差异, 对横断山区以及甘肃北部等地趋势的反映几乎与站点资料相反。

2.5 年平均气温对比

以 1971—2000 年作为全国气温距平值的标准气候期, 分别利用站点资料和格点数据集计算了 1951—2007 年我国年平均气温距平(图 7a)。从图 7a 可以看出, 站点资料和格点数据集反映出我国 20 世纪 50 年代年平均气温变化差异很大, 站点资料表明: 50 年代我国气温明显偏高, 而格点数据集反映出的同期气温则明显偏低。这是因为 1951—1960 年期间, 我国境内基本、基准观测站稀少, 分布很不均匀, 且主要集中在气温相对较高的东部地区, 因此

用站点资料算术平均得到的我国平均气温偏高; 随着我国气象台站的建设, 我国境内基本、基准站迅速从 1951 年的 133 个增加到 1960 年 12 月的 656 个, 站点东西分布也相对均衡, 因此站点资料与格点数据集反映出的我国年平均气温变化状况差异也随之迅速减小; 1961 年之后境内观测站稳定在 660~690 个之间, 来自 GHCN 的境外资料主要消除了边界效应对插值的影响, 所以 60 年代中期以后, 无论是整个中国区域还是我国东部地区, 格点数据集与站点资料序列都比较吻合, 年际变化的趋势和程度都非常一致。图 7b 比较了站点资料和格点数据集对我国东部地区(103°E 以东)气温变化状况的反映。可以看出, 由于 50 年代东部地区观测站点相对丰富, 格点数据集和站点资料反映出的东部地区年平均气温尽管数值上存在一定差异, 但变化趋势是一致的, 而且这种差异随着站点的增加迅速减小。60 年代中期以后, 两个序列吻合得很好。

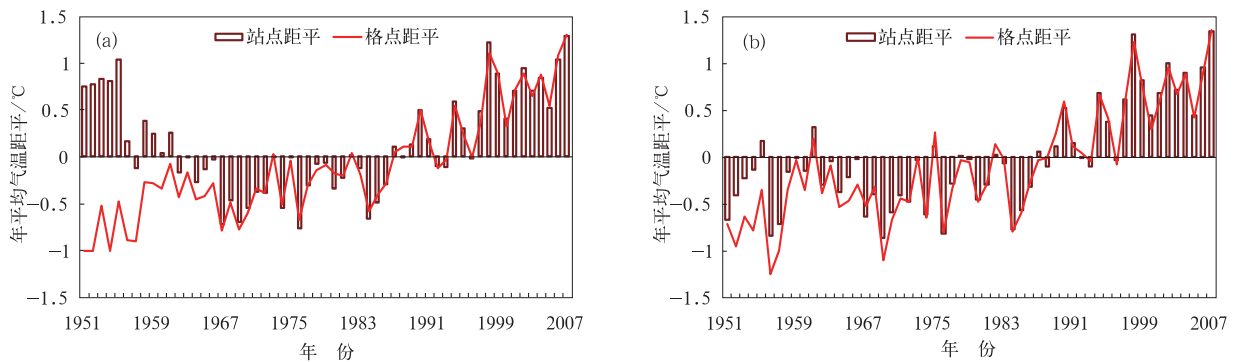


图 7 1951—2007 年全国(a)及我国东部地区(b)年平均气温距平时间序列

Fig. 7 The time series of the annual mean temperature anomaly over China (a) and the east part of China (b) during 1951—2007

图 7a 中格点数据集反映出 1951—2007 年我国年平均气温整体上升趋势非常明显, 温度变化速率达 $0.28\text{ }^{\circ}\text{C}/10\text{ a}$, 57 年平均气温上升了约 $1.6\text{ }^{\circ}\text{C}$, 比全球或半球同期平均增温速率明显偏高。气温增暖主要发生在近 20 年, 20 世纪 80 年代中期之前, 全国气温基本在较小的范围内波动, 之后气温呈明显上升趋势。从偏暖年份看, 20 世纪 80 年代中期之前, 只有 1973 年和 1982 年全国平均温度距平值为正, 且绝对值很小(分别为 $0.027\text{ }^{\circ}\text{C}$, $0.039\text{ }^{\circ}\text{C}$)。80 年代中期之后, 出现了 18 个偏暖年份, 且偏暖的程度越来越大。这与国内近几年已有研究成果^[26-30]是一致的。另外, 图 7a 中格点数据集显示, 1998—2007 年是 1951 年以来最暖的 10 年, 其中 2006 年

全国平均温度距平接近 2000 年之前的历史最高年份(1998 年), 而 2007 年全国平均温度距平值超过 1998 年, 达到 1951 年以来的历史最高值 $1.3\text{ }^{\circ}\text{C}$, 为最暖的一年。

3 结论与讨论

分析表明:

1) 在对温度场空间结构准确分析的基础上, 普通克里金方法对气温空间插值中的各项误差均较小, 对温度场插值具有一定优势。

2) 格点数据集与原始站点序列在全国绝大部分地区的相关程度很高。数字高程资料的引入以及

使用国外资料消除了边界效应,使得大部分高山地区和边界上的插值误差总体减小,相关性明显提高。

3) 总体来讲,格点数据集描述的我国年平均气温及季节平均气温分布更为合理,突出了温度场的大尺度特征,同时格点数据集准确反映了我国北方和南方年平均气温变化的主要空间差异。

4) 格点数据集较好地反映出 1951—2007 年我国气温变化状况。同时,格点数据集显示出 2007 年全国平均温度距平值超过 2000 年之前的历史最高年份 1998 年,达到 1951 年以来历史最高值 1.3 °C,为最暖的一年。

空间插值本身是一个比较复杂的问题,插值误差也是客观存在的。本格点数据集的研制一方面利用了普通克里金方法对温度场插值的相对优越性,同时充分考虑到了高程对气温插值的重要影响和插值过程中的边界效应,尽可能消除影响插值效果的不利因素来减小插值误差。检验结果表明,本格点数据集达到了预期的精度要求。

参考文献

- [1] Jones P D, Raper R, Bradley R S, et al. Northern hemisphere surface air temperature variations: 1851—1984. *J Climate Appl Meteor*, 1986,25(2):161-179.
- [2] Hanson J, Lebedeff S. Global surface air temperature: Update through 1987. *Geophys Res Lett*, 1988,15(4):323-326.
- [3] Hulme M. A 1951—80 global land precipitation climatology for the evaluation of general circulation models. *Climate Dyn*, 1992,7(1):57-72.
- [4] Jones P D. Hemispheric surface air temperature variations: A reanalysis and an update to 1993. *J Climate*, 1994,7(11):1794-1802.
- [5] Huffman G J, Adler R F, Rudolf B R, et al. The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull Amer Meteor Soc*, 1997,78(1):5-20.
- [6] Xie P, Arkin P A. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates and numerical model outputs. *Bull Amer Meteor Soc*, 1997,78(11):2539-2558.
- [7] Adler R F, Kidd C, Petty G, et al. Intercomparison of Global Precipitation Products: The Third Precipitation Intercomparison Project (PIP-3). *Bull Amer Meteor Soc*, 2001,82(7):1377-1396.
- [8] IPCC. Climate Change 2001: The Science of Climate Change // Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. 2001.
- [9] 林忠辉,莫兴国,李宏轩,等. 中国陆地区域气象要素的空间插值. *地理学报*, 2002,57(1):47-56.
- [10] 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.
- [11] 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.
- [12] 苏志侠,程麟生. 两种客观分析方法的比较: 逐步订正和最优内插. *高原气象*, 1994,13(2):194-205.
- [13] 李新. 青藏高原气温分布的空间插值方法比较. *高原气象*, 2003,22(6):565-573.
- [14] 封志明,杨艳昭,丁晓强,等. 气象要素空间插值方法优化. *地理研究*, 2004,23(3):357-364.
- [15] 朱芮芮,李兰,王浩,等. 降水量的空间变异性和空间插值方法的比较研究. *中国农村水利水电*, 2004,7:25-28.
- [16] 庄立伟,王石立. 东北地区逐日气温要素的空间插值方法应用研究. *应用气象学报*, 2003,14(5):605-615.
- [17] 魏凤英,曹鸿兴. 我国月降水和气温网格点资料的处理和分析. *气象*, 1994,20(10):26-30.
- [18] Vose R S, Schmoyer R L, Steurer P M, et al. The Global Historical Climatology Network: Long-term Monthly Temperature, Precipitation, Sea-level Pressure, and Station Pressure Data. ORNL/CDIAC-53, NDP-041, 1992:325.
- [19] Peterson T C, Vose R S. An overview of the global historical climatology network temperature database. *Bull Amer Meteor Soc*, 1997,78(12):2837-2849.
- [20] 李新,程国栋,卢玲. 空间内插方法比较. *地球科学进展*, 2000,15(3):260-265.
- [21] 张仁铎. 空间变异理论及应用. 北京:科学出版社, 2005.
- [22] 张兰生,方修琦. 中国气温变化的区域分布规律. *北京师范大学学报(自然科学版)*, 1988(3):78-85.
- [23] 宋连春. 近 40 年我国气温时空变化特征. *应用气象学报*, 1994,5(1):119-124.
- [24] 陈隆勋,邵永宁,张清芬,等. 近四十年我国气候变化的初步分析. *应用气象学报*, 1991,2(2):164-173.
- [25] 陈隆勋,朱文琴. 中国近 45 年来气候变化的研究. *气象学报*, 1998,56(3):257-271.
- [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):942-955.
- [27] 沙万英,邵雪梅,黄攻. 20 世纪 80 年代以来中国气候变暖及其对自然区域界线的影响. *中国科学*, 2002,32(4):317-326.
- [28] 魏凤英,曹洪兴,王丽萍. 20 世纪 80~90 年代我国气候增暖进程的统计事实. *应用气象学报*, 2003,14(1):79-86.
- [29] 曾昭美,严中伟. 本世纪全球增暖的显著性分析. *应用气象学报*, 1999,10(增刊):23-33.
- [30] 王绍武,叶瑾琳,龚道溢. 近百年中国气温序列的建立. *应用气象学报*, 1998,9(4):392-401.

Establishment and Assessment of the Grid Air Temperature Data Sets in China for the Past 57 Years

Zhang Qiang Ruan Xin Xiong Anyuan

(*National Meteorological Information Center, Beijing 100081*)

Abstract

Temperature data from meteorological stations in China for the past 57 years are interpolated by introducing digital elevation model (DEM) and taking the edge effect of interpolations into consideration on the basis of the ordinary Kriging method. Using this method, the daily, monthly and annual mean data sets of the grid-based temperature in China with the resolution of 1° latitude by 1° longitude are obtained.

The results of the grid data sets quality assessment show that DEM has an apparent effect on the spatial structure of temperatures and plays an important role on interpolations. Great improvements on the spatial interpolation in alpine spots are achieved after exploiting DEM. Compared with station data, grid data sets are more plausible to depict the annual and seasonal mean temperature distribution. The grid data sets can well simulate the spatial difference of the annual mean temperature trends in China. The grid data sets can also well demonstrate the change of annual mean temperature in China which as a whole rises by about 1.6°C for 1951—2007, with a warming rate of about $0.28^{\circ}\text{C}/10\text{ a}$. The warming in the later half 20th century is more rapid than the average values of the world and the Northern Hemisphere. The most evident warming occurs in the past over 20 years.

The grid data sets also indicate that: Period from 1998 to 2007 is a decade when the annual mean temperature is the highest since 1951, amongst which the annual mean temperature anomaly in 2006 is close to that in 1998 with the highest value before 2000, and the annual mean temperature anomaly in 2007 with a value of 1.3°C beyond that in 1998, is the highest since 1951.

Key words: surface air temperature; spatial interpolation; grid data set; quality assessment