DOI:10.19789/j.1004-9398.2016.06.017

局部回归克里格在气温栅格化中的应用*

郭春霞[1,2] 栗忠魁[3] 诸云强[1,4**] 孙 伟[5]

(1. 中国科学院地理科学与资源研究所,北京 100101; 2. 中国科学院大学,北京 100049;
3. 首都师范大学资源环境与旅游学院,北京 100048; 4. 江苏省地理信息资源开发与利用协同创新中心, 江苏 南京 210023;
5. 中国农业科学院农业信息研究所,北京 100081)

摘 要

气温是全球环境变化、农业、生态等研究领域的重要输入变量,气温栅格数据能有效地与其他空间数据进行叠加分析,目前主要通过空间插值获得. 本文采用局部回归克里格(Local Regression Kriging)构建全国 1981—2010 年 30 年累年年均气温空间分布数据集,并与常规插值方法进行精度对比. 插值结果与我国整体的气温分布趋势一致,同时也反映了典型区域气温的局部特征. 西北地区受海拔高低起伏的影响,温度呈现出以青藏高原地区为中心,向外围逐渐升高的特点. 交叉验证结果表明,充分考虑了气温空间异质性以及残差空间自相关性的邻域为 50 的局部回归克里格法插值精度相对最高,其均方根误差(RMSE)为 1.788 ℃,平均绝对误差(MAE)为 1.127℃,拟合优度 R^2 为 0.916 9.

关键词: 局部回归克里格, 累年年均气温, 气温栅格, 局部建模, 精度分析.
中图分类号: P208

0 引 言

气温是反映地球表层系统热量状态的一项基本气象要素,不仅是在自然区域界限划分中需要考虑的关键因素,更是地表陆面过程各类型模拟中的重要参数[1],广泛地应用于全球环境变化、农业、生态等研究领域. 常用的气温数据主要是通过气象台站获取,每个站点仅代表该站点处的气温,由于气象站点个数有限,很难反映全国气温的整体分布特征,而高空间分辨率、网格化的地表气温数据能够更好的表达其连续分布的空间特征,并且可以和其他基于栅格的环境、遥感等数据进行叠加分析,是模型计算的重要输入变量[2-7].

气温栅格数据通常是通过空间插值获得,常用的空间插值方法包括趋势面分析、回归分析、地理加权回归、泰森多边形法、距离倒数权重法、样条函数法以及克里格系列插值法等[8-13]. 廖顺宝等人在分析几个气温栅格化中的具体问题时,基于 1951—1980 年 30 年累年年均气温对比分析了反距离权重插值、克里格插值、样条插值和趋势面插值四种方法的插值精度,发现克里格插值法精度最高[14];李新等人利用反距离平方、趋势面、克里格以及协克里格插值法对青藏高原 1961—1990 年月平均气温进行栅格化,发现四种方法的插值结果均不太理想,是由于气象台站个数太少且分布不均匀[15];Nynke 等人为了获得欧洲高精度、长序列的日均气候数据集,对比了六种空间插值方法,最终选择了全局克里格插值法作为数据集制作方法[16];Fabio 等人在区域尺度上对比了去趋势反距离权重法、泛克里格法以及多层神经网络法的气象要素插值精度,结果表明泛克里格插值法精度较高,这说明引入气温与环境要素相关性分析有助于提高插值精度[17]. 张唯等人在估算日降水量空间分布的过程中发现日降水量监测数据呈现非正态分布,通过非线性克里格插值可获得较反距离加权法与普通克里格精度更高的日降水量空间分布数据集[18]. Mariusz 等人在研究城市热岛效应的空间插值时发现地理加权回归优于普通线性回归,是由于地理加权回归基于局部建模,充分考虑了插值要素的空间非平稳性[19].

分析现有研究成果发现在克里格插值中,引入环境要素的混合插值方法[20-22]或者局部建模的插值方

收稿日期: 2015-03-13
* "十二五"国家科技支撑计划项目(2012BAB11B00);科技基础性工作专项重点项目(2013FY110900).
**通讯作者

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

法均能取得较好的插值结果,但两者结合的建模思想在气温插值的研究中在国内并不多. 本文采用作者孙伟最新提出的基于滑动窗口的局部回归克里格插值法[23-24]对我国 1981—2010 年 30 年的累年年均气温进行栅格化,并与普通克里格、局部克里格、普通线性回归、局部线性回归、地理加权回归以及回归克里格六种插值方法进行对比分析.

1 数据处理及研究方法

1.1 数据来源

1981—2010 年 30 年累年年均气温数据(站点分布如图 1 所示)来源于中国气象科学数据共享服务网(http://cdc.cma.gov.cn). 全国 1 km 分辨率的 DEM 数据集以及全国 1∶25 万行政区划数据均来源于地球系统科学数据共享平台(http://www.geodata.cn). 站点气温源数据包含 922 条数据记录,剔除 29 条冗余数据记录以及 47 条高程值异常的数据记录,得到本文插值所采用的 846 条累年年均气温数据. 全国 1 km 分辨率的坡度、坡向和地表粗糙度数据集分别采用 ArcGIS 中的 Surface、Neighborhood、Raster Calculator 工具,基于 DEM 数据计算所得. 所有空间数据的投影方式均为 Albers 投影;参考椭球体为 Krasovsky;第一标准纬线与第二标准纬线分别为 25°0′0″和 47°0′0″;中央经线为 105°0′0″E;坐标原点为 0°0′0″. 图 1 显示我国气象站点集中分布在东部地区,西部地区气象站点分布稀疏,尤其是青藏高原海拔超过 5 000 m 的地区,几乎没有气象站点分布.

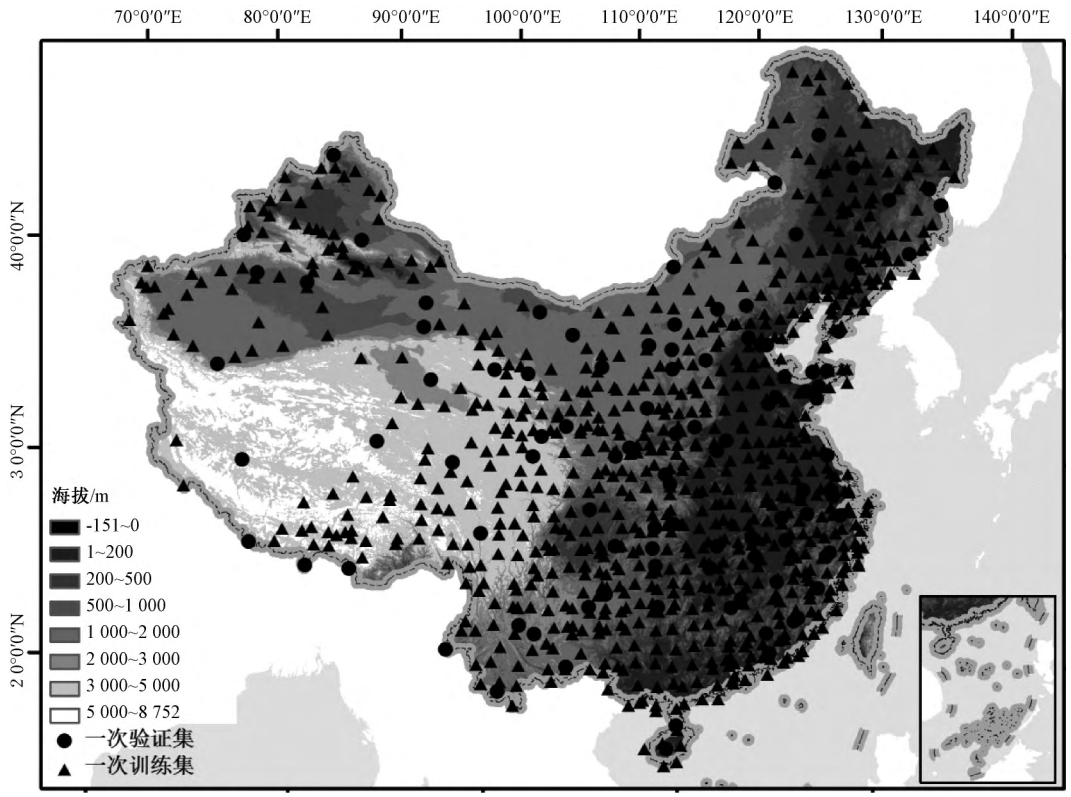


图 1 全国气象站点分布图

1.2 气温及协变量的统计分析

1981—2010 年 30 年累年年均气温自 -5.1 ℃变化到 27 ℃,其数学均值和中值分别是 11.88 ℃和 12.9 ℃(表 1 所示),观测值的空间分布较正态分布左偏低峰,然而峰度系数和偏度系数分别为 -0.69 和 -0.26,绝对值均小于 1,因此累年年均气温值的分布接近于正态分布. 通过相关性分析及逐步回归,最终参与气温插值模型构建的参数主要有经度、纬度、高程和地表粗糙度,全局回归 R^2 为 0.906 8. 地表粗糙度计算方法见公式(1):

$$\mathrm{TRI} = \sqrt{\left|(3\times3\mathrm{MAX})^2-(3\times3\mathrm{MIN})^2\right|}, \quad (1)$$

其中,3×3MAX 为以 3×3 的矩形为邻域统计各个 DEM 栅格的最大值,3×3MIN 为以 3×3 的矩形为邻域统计各个 DEM 栅格的最小值.

1.3 局部回归克里格算法实现

局部回归克里格是在回归克里格的基础上考虑

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

表 1　气温观测值及地形因子描述性统计值

变量	气温/℃	高程/m	坡度/°	坡向/°	地表粗糙度
平均值	11.88	1 062.16	3.02	162.45	515.04
中值	12.9	612	1.64	156.46	343.33
最小值	-5.1	0	0	0	0
最大值	27	5 461	25.63	359.80	5 461
标准差	6.44	1 278.56	3.80	110.58	567.71
峰度系数	-0.69	1.97	6.00	-1.21	2.38
偏度系数	-0.26	1.65	2.16	0.15	1.64

数据空间异质性而提出的优化方法,回归克里格是多元线性回归与普通克里格的结合,首先选择合适的气温影响因子参与气温回归方程模型的构建,其次对回归残差进行全局克里格插值,最后将气温回归结果与残差插值结果叠加得到全国气温栅格数据集. 式(2)给出了回归克里格的算法实现:

$$Z(x_0) = a_0 + \sum_{k=1}^{m} a_k \times V(x_0)_k + \sum_{i=1}^{n} \omega_i \times \varepsilon(x_i), \quad (2)$$

其中,a_0 为回归部分的截距,$a_k(k=1,2,3,\cdots,m)$ 分别为回归方程中 m 个自变量的回归系数,$V(x_0)_k(k=1,2,3,\cdots,m)$ 分别为待估点 x_0 处的 m 个自变量的取值,ω_i 为参与 x_0 属性值估算的周围点的权重系数,$\varepsilon(x_i)$ 为参与 x_0 属性值估算周围点的回归残差.

残差估计采用普通克里格插值法,半变异函数模型通过式(3)构建:

$$\gamma^*(h) = \frac{1}{2N_h}\sum_{i=1}^{N_h}[z(x_i+h) - z(x_i)]^2, \quad (3)$$

其中,$z(x_0)$ 为待估点 x_0 处的属性值,$z(x_i)(i=1,2,3,\cdots,n)$ 为待估点周围点的观测值,ω_i 为各个点所对应的权重系数,h 为变异距离,N_h 为计算距离 h 的变异函数的点对个数.

已有的常用的局部回归主要是通过分大区实现,即将研究区域根据空间化要素的整体特点划分若干区域,每个区域建立各自的回归模型并进行插值计算,缺点是难以体现区域内部差异. 基于滑动窗口的局部回归克里格是针对每一个待估栅格点选择其周围一定邻域内(或一定数量)的观测样点进行线性回归拟合及残差半变异函数拟合并进行估值计算(每个未知点采用的是相同性质的回归函数,即回归模型采用的拟合参数一致,而回归系数随着未知点变化).

式(4)给出了局部回归克里格的算法实现:

$$z(x_0) = a_0(x_0) + \sum_{k=1}^{m} a_k(x_0) \times V(x_0)_k + \sum_{i=1}^{U(x_0)} \omega_i(x_0) \times \varepsilon(x_i), \quad (4)$$

其中,$a_0(x_0)$ 为待估点 x_0 处的回归截距,$a_k(x_0)(k=1,2,3,\cdots,m)$ 为待估点 x_0 处回归方程中各个自变量的回归系数,$V(x_0)_k(k=1,2,3,\cdots,m)$ 为各个自变量在点 x_0 处的取值,$\varepsilon(x_i)$ 为待估点 x_0 在邻域 $U(x_0)$ 内所有已知观测点的回归残差,$\omega_i(x_0)$ 是邻域内各个已知观测点对待估点 x_0 的贡献值即权重.

综合前人研究[1,5,11,25],本文选择了经度、纬度,高程,坡度,坡向及地表粗糙度 6 个地形因子参与累年年均气温逐步回归模型的参数筛选.

1.4　精度验证方法

本文通过十次交叉验证对各个插值方法进行精度验证,将观测数据集随机分为 10 份,每份数据集大约 9/10 的数据作为建模集,剩余 1/10 作为验证集,分别计算每份数据集的均方根误差 RMSE(Root Mean Squared Error)和平均绝对误差 MAE(Mean Absolute Error)两个指标,最后求取 10 份数据的平均 RMSE 和 MAE,在此基础上评价局部回归克里格的插值精度,并与普通克里格、局部克里格、普通线性回归、局部线性回归、地理加权回归、以及回归克里格六种方法进行插值精度的对比分析.

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum[z(x_i) - \hat{z}(x_i)]^2}, \quad (5)$$

$$\mathrm{MAE} = \frac{1}{n}|z(x_i) - \hat{z}(x_i)|, \quad (6)$$

其中,n 为验证数据集站点的个数,$z(x_i)$ 为点 x_i 处的观测值,$\hat{z}(x_i)$ 为点 x_i 处的预测值. RMSE 与 MAE 的值越小,插值精度越高.

局部回归克里格算法主要采用 RKGuider[23] 实现,精度验证则在 R 统计软件中调用 gstat,spgwr,SDMTools,sp 四个包进行了实现.

2　结果及讨论

2.1　邻域大小对局部回归克里格插值结果的影响

局部回归克里格的插值精度及结果与插值邻域大小密切相关,为探索研究获取精度相对最高,插值效果相对最佳的气温栅格,本文随机选择了邻域大小分别为 30,50,80,100,150,200 的局部回归克里格对我国 1981—2010 年 30 年累年年均气温进行空间化(图 2),结果显示 6 种插值结果的空间分布规

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

律基本一致,主要表现为:全国气温整体上自西北至东南逐渐升高,然而在新疆自治区内及青藏高原地区存在温度反常.新疆的喀什地区、和田地区、巴音郭勒蒙古自治州交界处与青藏高原的中部地区是两个低温中心,其温度不仅低于同纬度的其他地区,同时也低于周边各区域.这一现象证明地形是影响气候变化的重要因子,气温随着海拔的升高而降低.对于不同邻域对结果的影响则主要体现在一些区域内部的细节变化,例如青藏高原内部及西南角落的气温分布差异.

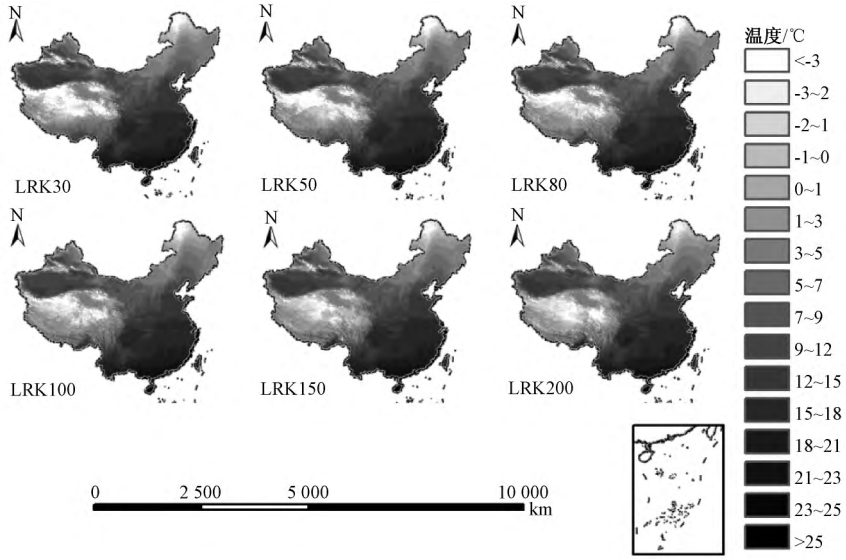


图 2 基于不同滑动窗口的局部回归克里格气温插值结果分布图

2.2 常用插值方法气温栅格化讨论

为验证局部回归克里格插值的优越性,本文也选取了六种常规插值方法进行对比,六种插值方法的气温空间化结果如图 3 所示.与全局模型相比,局部模型更能表达区域内部的气温分布特征,而且能很好地保留一些特征极值,例如局部克里格以及局部线性回归在青藏高原分别对普通克里格和普通线性回归的优化;地理加权回归本质上也是一种基于局部思想的地统计方法,然而气温分布显示低温区的气温变差较小,对于青藏高原内部以及东北漠河一带的极低气温

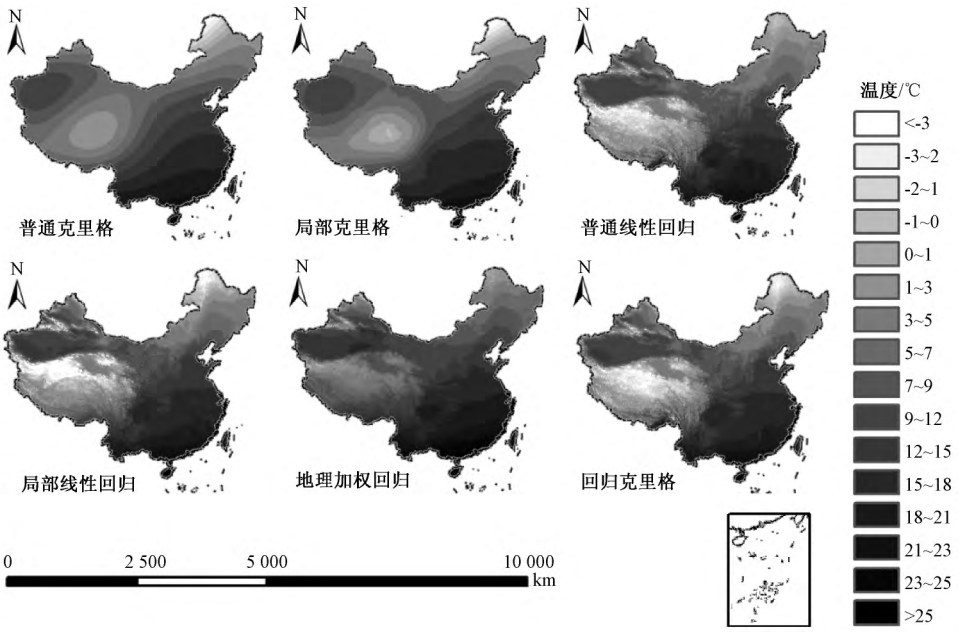


图 3 六种插值方法的气温插值结果分布图

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

未能体现,这一现象说明地理加权回归在插值过程中平滑了一些极值信息; 回归克里格基于方法综合的思想,其插值结果优于普通克里格和普通线性回归. 局部建模及方法综合的分别不同程度地提高了气温插值结果的准确度,因此基于两种思想融合的局部回归克里格具有其独特的优越性.

2.3　局部回归克里格及常规插值方法的精度分析

不同邻域大小的局部回归克里格,其插值精度不同(表 2) ,均方根误差(RMSE) 以及平均绝对误差(MAE) 随着邻域的不断增大呈现出先降后升的趋势. 邻域从 30 增加到 50 时,RMSE 降低了 6.4%,MAE 降低了 5.2%; 邻域从 50 增加到 80 时,RMSE 升高了 1.7%,MAE 升高了 1.1%; 邻域从 80 增加到 100 时,RMSE 升高了 1%,MAE 升高了 0.2%; 邻域从 100 增加到 150 时,RMSE 升高了 0.1%,MAE 升高了 1.5%; 邻域从 150 增加到 200 时,RMSE 升高了 0.4%,MAE 升高了 1.1%. 其中,LRK50 插值精度相对最高,邻域减小,精度大幅度下降,主要是由于小邻域包含的空间信息相对较少甚至不足,这将直接导致模型对已知观测数据的拟合优度下降; 而大邻域虽弱化了区域内部差异,但模型拟合效果不会受到影响,因此精度呈小幅度下降.

如表 2 所示,除却 LRK30,其他所有邻域的局部回归克里格均优于常规插值方法. 局部克里格较普通克里格 RMSE 降低了 8.4%,MAE 降低了 12.3%; 局部线性回归较普通线性回归 RMSE 降低了 4.4%,MAE 降低了 10.6%; 地理加权回归较普通线性回归 RMSE 降低了 4.7%,MAE 降低了 14.2%. 上述数据事实证明局部模型确实有助于提高插值精度,这与图 3 揭示的气温分布特征相吻合. 回归克里格较普通线性回归 RMSE 降低了 7.8%,MAE 降低了 18.4%; 较普通克里格 RMSE 降低了 26%,MAE 降低了 35%. 因此引入环境要素的线性回归与残差克里格的结合优于其中任意一个方法.

表 2　不同插值方法精度验证结果

插值方法	均方根误差 (RMSE/℃)	平均绝对误差 (MAE/℃)
LRK30	1.910	1.189
LRK50	1.788	1.127
LRK80	1.818	1.139
LRK100	1.836	1.137
LRK150	1.838	1.154
LRK200	1.845	1.167
OK	2.428	1.686
LOK	2.224	1.478
OLR	1.948	1.345
LOLR	1.862	1.203
GWR	1.857	1.154
RK	1.797	1.097

2.4　LRK50 在全国七大地理区的拟合优度讨论

由于全国 846 个气象台站在中国内陆地区并非均匀分布,虽然气温观测值与 LRK50 气温估值的全局 R^2 达 0.916 9,事实上,在不同地理区域范围内 R^2 截然不同. 表 3 依据全国七大地理区划: Ⅰ东北地区; Ⅱ华北地区; Ⅲ华东地区; Ⅳ华中地区; Ⅴ华南地区; Ⅵ西南地区; Ⅶ西北地区,分别对比了观测值与 LRK50 的估计值和插值精度. 其中东北地区的模型拟合度最高 R^2 为 0.961 4,高于全国平均水平,其次是华北与西南地区,R^2 分别为 0.892 7 和 0.818 7,西北与华南地区的 R^2 均达到 0.7 以上,华中地区的模型拟合度最差 R^2 为 0.316 1,华东地区的模型拟合度高于华中地区,R^2 为 0.642 9. 华中主要包括河南省、湖北省及湖南省,处于平原地带,DEM 和 TRI 不再是影响该区气温的主要因子,因此该区的拟合度最差; 而华东包括了中国大部分沿海省份,海洋大环境在气温调节中的作用远远高于本研究选取的几个因子, LRK50 的拟合度受到限制. 影响气温的因素很多,包括位置、大气、地形、洋流、植被、水文等,而影响不同区域气温的主要因素不尽相同,因此单一模型的全局插值效果并不能代表全国的插值效果.

表 3　全国七大地理区域气温观测值与 LRK50 气温估值对比

区域编号	观测站点数	气温观测值/℃			LRK50 气温估计值/℃			LRK50 精度结果	
		最小值	均值	最大值	最小值	均值	最大值	RMSE/℃	R^2
Ⅰ	93	-3.9	5.24	10.5	-4.03	5.23	10.40	0.593	0.961 4
Ⅱ	104	-4.0	7.59	14.3	-1.99	7.63	13.73	1.350	0.892 7
Ⅲ	151	5.9	16.15	21.7	11.26	16.23	21.02	1.474	0.642 9
Ⅳ	88	7.8	16.12	18.8	10.97	16.14	19.24	1.487	0.316 1
Ⅴ	52	19.1	21.50	24.8	19.18	21.55	25.56	0.723	0.739 8
Ⅵ	173	-2.4	12.62	23.9	0.57	12.76	22.66	2.583	0.818 7
Ⅶ	185	-5.1	7.40	15.7	-4.53	7.35	14.47	2.108	0.790 0

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

2.5 LRK50 的方差分析及与发布数据的对比

为了验证 LRK50 插值所得气温栅格的科学性及合理性,根据中国气象科学数据共享服务网发布的分辨率为 1°×1°的中国气温年值格点数据集,选择与本文时间段最为接近的 1981－2007 年各年平均气温数据,栅格计算得到 1981—2007 年 27 年累年年均气温分布图(图 4 c),通过对比发现,LRK50 插值结果的值域与发布数据的基本一致,且气温空间分布的整体趋势和区域局部特征基本相同,然 LRK50 插值结果的空间分辨率较发布数据提高了约 10 倍左右. LRK50 的估值方差在西南地区最大(图 4 b),表明该地区区域内部气温变化剧烈,主要受该地区地形起伏变化的影响,而广大东部地区由于其地形平坦,并不存在气温突变,因此方差相对较小.

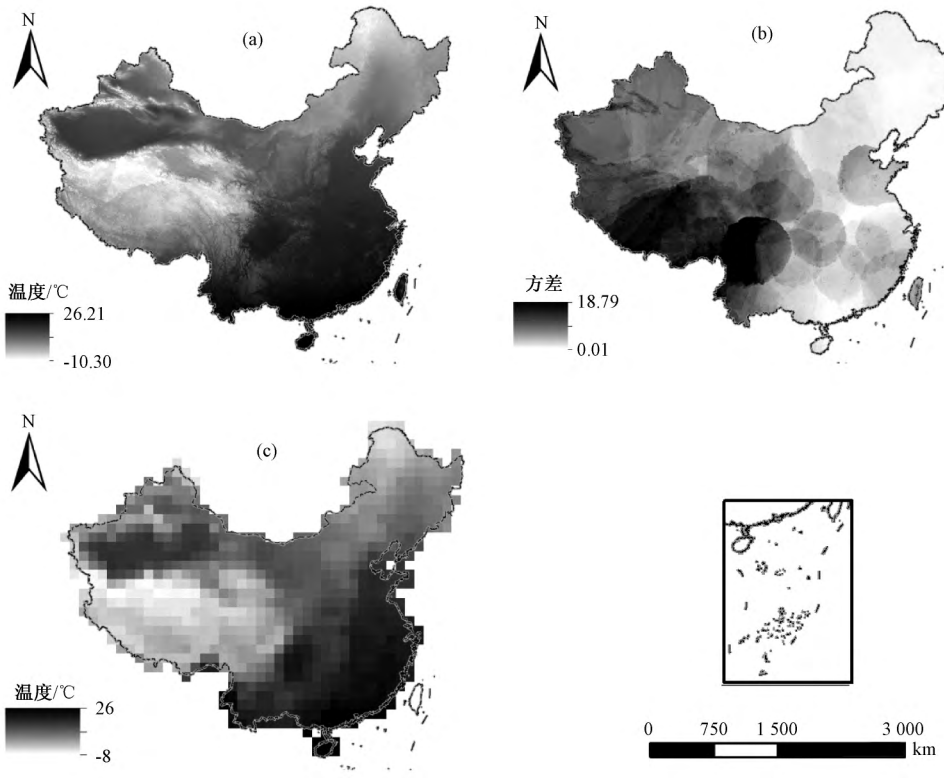


图 4 LRK50 气温插值结果及方差与发布数据气温分布图

3 结 论

气温是全球环境变化、农业、生态等多个研究领域重要参数. 常用气温数据主要通过全国气象监测台站监测获取,然而由于国家级气象台站数量有限,分布不均,因此全国气温分布的准确表达受到限制. 空间插值方法将站点气温插值气温栅格,可以很好地解决上述问题,因此气象要素插值是一项重要的科学研究任务. 本文主要采用局部回归克里格对我国 1981—2010 年 30 年累年年均气温进行了栅格化处理,并就插值结果以及与其他六种插值方法的精度对比进行了分析,得到以下结论:

1) 局部回归克里格的插值精度受插值邻域大小影响. 本研究选取了大小分别为 30、50、80、100、150、200 的邻域进行局部回归克里格的精度验证,发现随着邻域的增大,插值精度先升后降,其中 LRK50 的插值精度相对最高.

2) 插值结果表明全国气温自西北至东北逐渐升高,而在青藏高原中部以及新疆喀什、和田地区和巴音郭勒蒙古自治区的交界处呈现出两个低温中心,气温普遍低于同纬度的其他地区以及周边各个地区.

3) LRK50 在中国东北地区的模型拟合最优,其 R^2 达到 0.9614,高于全国平均水平,其次是华北与西南地区,R^2 分别为 0.892 7 和 0.818 7,西北与华南地区的 R^2 均达到 0.7 以上,华中地区的模型拟合度最差 R^2 为 0.316 1,华东地区的模型拟合度高于华中地区,R^2 为 0.642 9. 因此全国范围的模型在区域研究中需要进一步改进以提高插值结果的精度.

4) LRK50 在西南地区的估值方差最大,区域内

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

部气温变化剧烈,是由于区域内部的地形起伏变化大. 1981—2010 年 30 年累年年均气温的 LRK50 插值结果与气象共享服务网站上发布的气温栅格数据基本一致,证明局部回归克里格科学合理且具有一定的适用性.

5) 局部回归克里格同时考虑了气温的空间自相关性以及受地形地貌等因素影响呈现的空间异质性,插值结果不仅能反映全局气温分布特点,部分区域内部复杂的气温分布特征也得到了充分体现. 然而局部回归克里格算法实现是针对每一个栅格点构建回归模型及半变异函数,因此随着空间分辨率及邻域的增大,算法速度受到限制,因此如何将算法并行化处理是后续研究的重点内容.

参 考 文 献

[1] 蔡迪花, 郭钜,李崇伟. 基于 DEM 的气温插值方法研究 [J]. 干旱气象, 2009, 27(1): 10 – 17.

[2] 潘耀忠, 龚道溢, 邓磊, 等. 基于 DEM 的中国陆地多年平均温度插值方法 [J]. 地理学报, 2004, 59(3): 366 – 374.

[3] 朱会义, 刘述林,贾绍凤. 自然地理要素空间插值的几个问题 [J]. 地理研究, 2004, 23(4): 425 – 432.

[4] 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 [J]. 地理学报, 2002, 57(1): 47 – 56.

[5] 廖顺宝, 李泽辉,游松财. 气温数据栅格化的方法及其比较 [J]. 资源科学, 2003, 25(6): 83 – 88.

[6] 庄立伟,王石立. 东北地区逐日气象要素的空间插值方法应用研究 [J]. 应用气象学报, 2003, 14(5): 605 – 615.

[7] Hijmans R J, Cameron S E, Parra J L, et al. Very high resolution interpolated climate surfaces for global land areas [J]. International Journal of Climatology, 2005, 25(15): 1965 – 1978.

[8] Vicente-Serrano S M, Saz-Sánchez M A,Cuadrat J M. Comparative analysis of interpolation methods in the middle Ebro Valley(Spain): application to annual precipitation and temperature [J]. Climate Research, 2003, 24(161 – 180.

[9] 封志明, 杨艳昭, 丁晓强, 等. 气象要素空间插值方法优化 [J]. 地理研究, 2004, 23(3): 357 – 364.

[10] 李军, 游松财,黄敬峰. 中国 1961—2000 年月平均气温空间插值方法与空间分布 [J]. 生态环境, 2006, 15(1): 109 – 114.

[11] 彭彬, 周艳莲, 高苹, 等. 气温插值中不同空间插值方法的适用性分析—以江苏省为例 [J]. 地球信息科学学报, 2011, 13(4): 539 – 548.

[12] 岳文泽, 徐建华,徐丽华. 基于地统计方法的气候要素空间插值研究 [J]. 高原气象, 2005, 24(6): 974 – 980.

[13] 季青, 余明. 基于协同克里格插值法的年均温空间插值的参数选择研究 [J]. 首都师范大学学报(自然科学版), 2010, 31(4): 81 – 87.

[14] 廖顺宝,李泽辉. 气温数据栅格化中的几个具体问题 [J]. 气象科技, 2004, 32(5): 352 – 356

[15] 李新,程国栋,卢玲. 青藏高原气温分布的空间插值方法比较 [J]. 高原气象, 2003, 22(6): 565 – 573.

[16] Hofstra N, Haylock M, New M, et al. Comparison of six methods for the interpolation of daily, European climate data [J]. Journal of Geophysical Research, 2008, 113(D21):6089 – 6098.

[17] Attorre F, Alfo M, De Sanctis M, et al. Comparison of interpolation methods for mapping climatic and bioclimatic variables at regional scale [J]. International Journal of Climatology, 2007, 27(13): 1825 – 1843.

[18] 张唯, 郑海波, 张剑波. 非线性克里格方法在日降水量分布估算中的应用 [J]. 中国农业气象, 2014, 35(6): 690 – 699.

[19] Szymanowski M, Kryza M. Local regression models for spatial interpolation of urban heat island—an example from Wrocław, SW Poland [J]. Theoretical and Applied Climatology, 2011, 108(1 – 2): 53 – 71.

[20] 陈鹏翔, 毛炜峄. 基于 GIS 的新疆气温数据栅格化方法研究 [J]. 干旱区地理, 2012, 35(3): 438 – 445.

[21] 张莉莉, 陶忠良, 张京红. 基于 DEM 的海南岛平均气温空间插值研究 [J]. 热带作物学报, 2012, 33(4): 699 – 703.

[22] 蒋友严, 黄进. 基于 GIS 技术的甘肃省气温空间分布特征 [J]. 干旱气象, 2013, 31(1): 206 – 211.

[23] Sun W, Minasny B, McBratney A. Analysis and prediction of soil properties using local regression-kriging [J]. Geoderma, 2012, 171 – 172: 16 – 23.

[24] Sun W, Zhu Y Q, Huang S L, et al. Mapping the mean annual precipitation of China using local interpolation techniques

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[J]. Theoretical and Applied Climatology, 2014, 1 - 10.

[25] Perry M, Hollis D. The generation of monthly gridded datasets for a range of climatic variables over the UK [J]. International Journal of Climatology, 2005, 25(8): 1041 - 1054.

Application of Local Regression Kriging in Mean Annual Air Temperature Estimation

Guo Chunxia[1,2] Li Zhongkui[3] Zhu Yunqiang[1,4] Sun Wei[5]

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101; 2. University of Chinese Academy of Sciences, Beijing 100049; 3. College of Resource Environment and Tourism, Capital Normal University, Beijing 100048; 4. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing Jiangsu 210023; 5. Agricultural Information Institute of CAAS, Beijing 100081)

Abstract

Temperature is an important input variable in many research fields, such as global environmental change, agriculture, ecology and so on. Temperature grid can be overlaid to other spatial data efficiently. At present, air temperature raster can be obtained mainly by interpolation techniques. In this paper, mean annual air temperature raster of 30 years from 1981 to 2010 was obtained by using local regression kriging, which was also compared with the other six methods. Interpolation results indicate: temperature interpolation agrees with distribution of Chinese air temperature overall, at the same time reflects local features of air temperature in typical regions. Influenced by topography of the Tibetan Plateau, temperature in the Tibetan Plateau is lowest, which rises continually to the periphery. Cross validation results indicate: local regression kriging with 50 neighbors considering spatial heterogeneity and spatial autocorrelation presents relatively highest accuracy, whose RMSE and MAE are 1.788℃ and 1.127℃ respectively. The goodness of fit R^2 is 0.9169.

Key words: local regression kriging, mean annual air temperature, temperature raster, local modeling, accuracy analysis.

(C)1994-2020 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net