

doi: 10.7621/cjarrp.1005-9121.20190913

· 农业区划 ·

四川省多年平均气温数据空间插值方法与影响因素研究*

何鹏^{1,2}, 张媛^{2,3}, 高文波^{1,2}, 蹇东南², 林正雨^{4*}

(1. 四川省农业科学院农业信息与农村经济研究所, 成都 610066; 2. 四川省农业科学院大数据中心, 成都 610066;
3. 成都理工大学旅游与城乡规划学院, 四川成都 610059; 4. 四川农业大学资源学院, 成都 611130)

摘要 [目的] 探索复杂地形、气候背景下适合当地特点的多年平均气温数据空间插值方法。[方法] 在考虑海拔高度、数据源密度和空间化栅格大小等条件影响下, 选择5种插值方法对1981—2010年四川省144个气象观测站点多年平均气温观测数据进行了空间插值, 并采用预留检验站点和交叉验证两种方法对结果进行评估。[结果] 空间化模型插值精度 IDWE > OKE > MRM > OK > IDW; 不考虑海拔因素时中等密度站点插值精度最高, 考虑海拔因素后高密度站点插值精度最高; 随着空间化尺度的不断增大, 不考虑海拔因素的插值方法误差变化不大, 考虑海拔因素的插值方法误差逐渐增大, 且在空间栅格尺度大于4 000m后误差快速增加; 不同海拔数据源对插值结果有影响, 采用3种方法插值的温度差在-0.98~3.05℃之间, 但绝对温差大于1℃的区域仅占全省面积的0.76%~1.52%。[结论] (1) 空间化模型方法的选择对气温空间化精度影响较大, 考虑了海拔等环境影响因素的模型比单纯的空间自相关模型精度更高; (2) 数据源(气象站点)数量的减少会导致空间化误差增加, 但并不明显; (3) 随着栅格尺度逐渐变大, 空间化误差逐渐增大, 适宜尺度在500~2 000m; (4) 气象站点海拔误差会影响气温空间化精度, 但从占全省面积来看影响并不显著; (5) 复杂地形会影响气温空间化效果, 分区拟合精度优于不分区拟合。

关键词 年平均气温 空间化误差 海拔 影响因素 四川省

中图分类号: S162.2 **文献标识码**: A **文章编号**: 1005-9121[2019]09114-11

0 引言

随着生态学、农业科学、资源环境科学等相关学科的发展, 气象信息在区域和全球尺度变化研究中发挥着重要作用, 特别是各类生态系统模型、作物模型都需要空间化的气象要素, 并且所要求的时间和空间分辨率也越来越高。例如, 在全球变化研究中, MT-CLIM、FOREST-BGC等景观、区域、全球尺度的生态系统模型, 都需要空间化的气温、降水、太阳辐射等环境因子作为输入参数^[1]。因此, 利用现有的地面气象观测资料和相关技术, 开发空间化的气候数据产品成为近年来生态学、农业区划、资源和环境科学研究的重点任务之一^[2-5]。

然而有限的气象站点仅代表一定区域内气象要素的分布情况, 其观测获取的只是局部、离散、有限的空间点数据, 对于没有气象站点分布的广大区域, 只能通过以点代面或者空间内插等方法推算求得。国内外学者已经在气象要素推算方面做了大量研究。国外, Eleanor等人引入地形, 运用多元回归法对气温和降水进行空间化处理^[6]; Marquinez等人引入高程、坡度、坡向、离海岸线的距离和离相对西边的距离等

收稿日期: 2018-11-28

作者简介: 何鹏(1978—), 男, 四川雅安人, 硕士、副研究员。研究方向: 农业资源与环境信息化研究

*通讯作者: 林正雨(1982—), 男, 四川自贡人, 硕士、副研究员。研究方向: 区域农业发展研究。1456875524@qq.com

*资助项目: 四川省财政创新能力提升工程“大数据背景下主要农产品质量安全预警分析技术研究”(2016GYSH-004); 四川省科技支撑计划软科学项目“基于精细区划的农产品气候品质认证评价指标体系研究”(2017ZR0045); 四川省科技支撑计划项目“智慧果园智能化技术及装备引进与示范”(2016NZ0114)

因子,利用多元回归方法和GIS技术拟合降水空间分布^[7];20世纪90年代以来,大量区域性乃至全球性栅格气象数据库不断建立和完善,如基于PRISM建立的美国、加拿大、中国、蒙古、欧洲等国家和地区的空间气象数据信息系统^[8];基于Anusplin插值方法的澳大利亚、南非等国家的气象数据信息系统^[9,10];基于DAYMET的美国生物气象数据信息系统^[11]。国内,20世纪80年代,以傅抱璞、沈国权等为代表的学者通过数理统计模拟方法对气温、降水等要素进行推算,获得了非气象站点所在区域的气象要素空间分布情况^[12-14];20世纪90年代以来,随着GIS技术在国内的流行,于贵瑞、陈晓峰等利用GIS软件并结合地形等因素,对区域气象要素空间分布进行了深入研究^[3-5,15,16],廖顺宝等、蔡福等还对气温空间化的插值方法、精度进行研究^[1,17,18],廖玉芳等以湖南省为例,进行了精细化农业气候区划研究,建立相关的气候空间化数据集^[19,20]。

影响气温空间分布的因素很多,主要以海拔高度和地形条件最为显著。一般情况下,随着海拔高度的增加,气温下降,但其变化速率因山地性质和气候条件而不同。在对气温素进行空间插值时,由于海拔高度的影响非常显著,海拔高度的误差对气温空间插值精度的影响不能忽略^[21-24]。四川省处于我国东部季风区与西南青藏高寒区的交接地带,气候垂直变化十分显著,加之地形复杂,导致气候的区域差异很大。江蕾采用全国633个站点数据,主要考虑地貌拟合插值模型,生成1km空间分辨率的气温数据,该研究对于四川省这一地貌复杂区的气温空间插值误差研究不够^[25]。为了对比和分析海拔高度对四川省多年平均气温空间分布的影响,文章利用5种插值方法,同时考虑数据源密度(3种密度)和空间化栅格大小(10个空间化栅格尺度)对四川省144个气象观测站点的多年平均气温进行了空间插值研究,并利用预留检验站点法对结果进行评估。

1 数据与方法

1.1 气温数据

气温数据来源于中国气象数据网(<http://data.cma.cn/>)申请下载的“中国地面累年值月值数据集(1981—2010年)”,包括气温、降水、气压、湿度、风向和风速共五大类12小类的年均气象数据,数据制作时间2012年,共包括四川省145个国家级地面站点(图1)。参考四川省气候区划(自然资源区划),并考虑原有站点分布疏密现状,采用分层随机抽样法,全省预留检验站13个(约10%),剔除数据缺失较多的站点1个,实际131个站点参与插值运算。

气象站点密度设定。数据源密度变化分析以预留的13个检验站点为基础,仍然采用分层随机抽样方法,并考虑原有气象站点分布的疏密度,131个站点按照1/3的比例递减,得到低密度(43个站点),中密度(87个站点),高密度(131个站点)共3个级别。

1.2 数字高程模型

数字高程模型(DEM)来源为国际科学数据服务平台(<http://datamirror.csdb.cn/>)申请下载的ASTER GDEM第一版本(V1)数据,数据时间2009年,数据投影为UTM/WGS84,空间分辨率30m,空间范围为四川省全域。

空间化栅格大小设定。空间化尺度变化分析结合四川省幅员面积和已有研究文献,依次为100、250、500、1000、2000、3000、4000、5000、7500、10000m共10个尺度等级。

1.3 插值模型

①反距离权重法(IDW)和加入海拔影响因子的反距离权重法(IDWE)。反距离权重法是对采样点进行线性加权决定输出的栅格值,加权值与距离成反比,采样点离输出栅格越远,它对输出栅格的影响越小。反距离权重法的表达式为:

$$T_{i,j} = \frac{\sum_{k=1}^n T_k d_k^{-2}}{\sum_{k=1}^n d_k^{-2}} \quad (1)$$

式(1)中, $T_{i,j}$ 为待插值的网格点要素值; T_k 为第 k 个点的要素值; k 为第 k 个被引用的插值点; d_k^{-2}

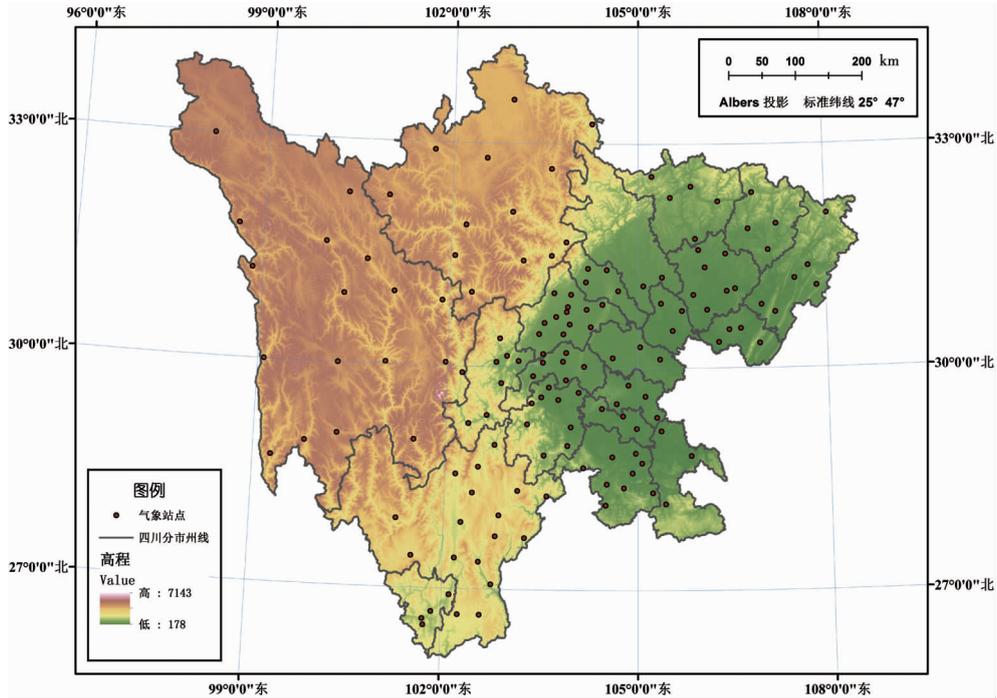


图1 四川省气象站点分布

为第 k 个点对待插值点距离平方的倒数。

考虑了海拔高度影响后, 由于气温随着海拔高度增高而降低, 进行气温空间插值时, 海拔高度是附加的重要变量, 其计算公式为

$$T_{i,j} = \sum_{k=1}^n [T_k + (H_{i,j} - H_k) \times \beta] d_k^{-2} / \sum_{k=1}^n d_k^{-2} \quad (2)$$

式(2)中, H_k 为第 k 个点的海拔高度, β 为气温随海拔高度的递减率, 其余参数与 IDW 同。

②普通克里格法(OK)和加入海拔影响因子的普通克里格法(OKE)。普通克里格法又称局部最优线性无偏估计法, 建立在变异函数理论结构分析基础之上。所谓线性是指估计值是样本值的线性组合, 即加权线性平均; 无偏是指理论上估计值的平均值等于实际样本值的平均值, 即估计的平均误差为 0, 最优是指估计的误差平方差最小。其计算公式为:

$$Z(t_{i,j}) = \sum_{k=1}^n \lambda_k Z(t_k) \quad (3)$$

式(3)中, $Z(t_{i,j})$ 为待插值的栅格点要素值; $Z(t_k)$ 为第 k 个点的要素值; λ_k 为权重系数, 其和等于 1。

加入海拔影响因子的普通克里格法基于普通克里格法并充分考虑了海拔因子的影响, 其计算公式为:

$$Z(t_{i,j}) = \sum_{k=1}^n \lambda_k [Z(t_k) + (H_{i,j} - H_k) \times \beta] \quad (4)$$

③多元回归分析法(MRM)。多元回归分析模型是应用数理统计方法, 建立气象要素与影响气象要素空间插值因子的多元回归模型。该文是基于经度、纬度和海拔高度的多年平均气温多元回归模型, 并计算残差。

多元回归分析模型表达式为:

$$T_{(\lambda, \varphi, h)} = b_0 + b_1 \lambda + b_2 \varphi + b_3 h + \varepsilon \quad (5)$$

式(5)中, b_0 为常数, b_1, b_2, b_3 为待定系数, λ 为经度, φ 为纬度, h 为海拔高度, ε 为残差。

1.4 数据处理

利用地理信息系统软件 ArcGIS 10.1 的数据管理模块(Data Management Tools)、转换模块

(Conversion Tools)、地统计模块 (Geostatistical Analyst Tools) 和空间分析模块 (Spatial Analyst Tools) 进行数据预处理和插值, 并使用预留站点法检验插值精度。属性数据的处理采用 SPSS 20。主要工作如下:

(1) 将气象站点的经度、纬度和多年平均气温在 ArcGIS 10.1 中生成点文件, 分别使用两种常规空间插值方法 (逆距离权重法和普通克里格法) 生成多年平均气温栅格图。

(2) 参考游松财^[21]提出的气温垂直递减法。首先, 在考虑了海拔高度的影响下, 根据气温垂直递减率和各气象站的海拔高度将各气象站的年平均气温订正到虚拟海平面上; 其次, 运用两种常规插值方法 (逆距离权重法和普通克里格法) 对虚拟海平面上的月平均气温进行空间插值, 得到虚拟海平面上的气温空间分布; 最后, 在 GIS 环境下, 利用 DEM 数据和气温垂直递减率进行栅格图层的代数运算得到平均气温的空间分布。

(3) 利用气象站点的海拔高度、经度、纬度和多年平均气温数据, 在 SPSS 20 中进行多元回归分析, 拟合出回归模型。将四川省 DEM 数据在 ArcGIS 10.1 中生成 value 值分别为经度 (x) 和纬度 (y) 的栅格图层, 使用栅格计算器根据回归模型生成多年平均气温栅格图。

1.5 模型验证

在使用气象站点数据进行气候指标插值结果的精度检验时, 交叉检验 (cross-validation) 是最常用的一种方法。该方法假设某一个气象站点的气温未知, 用其他所有气象站点来估算该气象站点的气温, 通过计算所有气象站点的实际气温值与估算值之间的误差来评判空间插值方法的精度。对不同的插值方法, 交叉验证可以准确的验证不同插值方法之间的相对精度, 使用时一般用平均绝对误差 (MAE) 和均方根误差 (RMSE) 作为检验精度的标准。MAE 反映样本数据估值的总体误差或精度水平, RMSE 则反映样本数据的估值灵敏度和极值。计算公式分别为:

$$MAE = \frac{1}{n} \sum_{i=1}^n abs(T'_i - T_i) \quad (6)$$

$$RSME = \sqrt{(T'_i - T_i)^2} \quad (7)$$

式 (6) 中, T_i 表示第 i 个气象站点的实际多年平均气温, T'_i 表示空间化后第 i 个气象站点所在栅格单元的气温, $abs()$ 为绝对值。

2 结果与分析

2.1 空间化模型精度分析

采用 $1\text{km} \times 1\text{km}$ 栅格尺度, 分别采用 IDW、IDWE、OK、OKE 和 MRM 5 种方法, 以 144 个气象站点进行多年平均气温数据的空间化 (图 2), 并对 144 个站点逐一做交叉验证, 统计平均绝对误差和均方根误差 (表 2), 其中 IDWE 精度最高, OKE 精度次之, 其中未引入海拔因子的 IDW、OK 模型, 精度明显低于有海拔因子参与的模型。

表 2 不同模型空间化误差

℃

	IDW	IDWE	OK	OKE	MRM
MAE	2.03	0.88	1.76	0.93	1.20
RSME	3.05	1.28	2.52	1.29	1.48

表 1 不同密度气象站点海拔高度、经纬度与平均气温的相关系数

	经度	纬度	海拔高度
高密度	0.594 **	-0.404 **	-0.912 **
中密度	0.622 **	-0.447 **	-0.920 **
低密度	0.681 **	-0.477 **	-0.931 **

注: * 表示相关系数显著水平为 0.05, ** 表示相关系数显著水平为 0.01

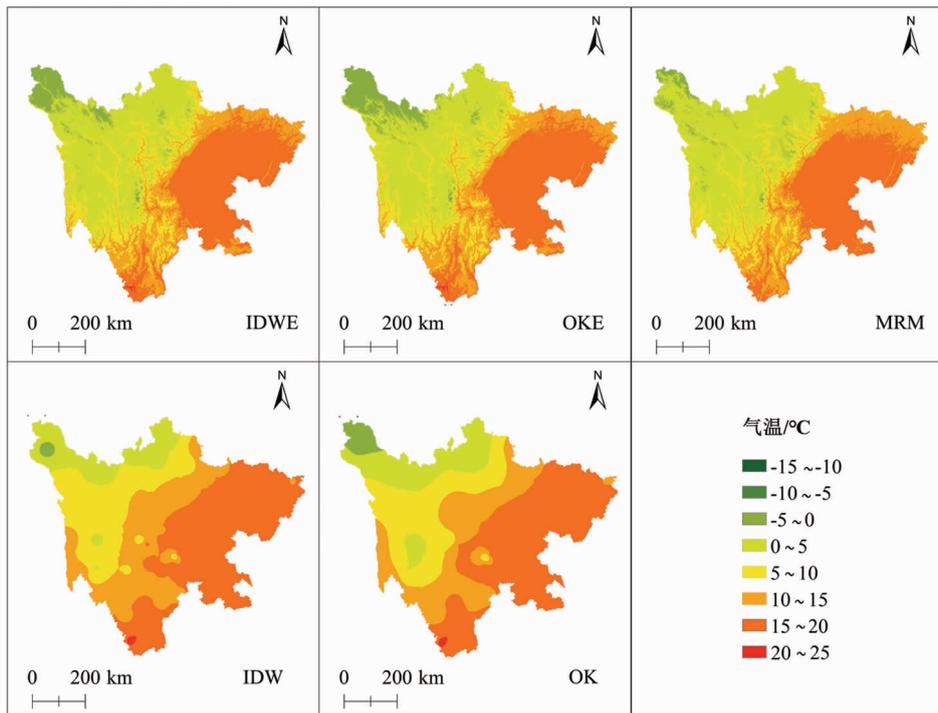


图2 5种方法144个站点空间化结果

2.2 空间化误差与数据源密度的关系

采用 $1\text{km} \times 1\text{km}$ 栅格尺度, 分别采用 IDW、IDWE、OK、OKE 和 MRM 5 种方法, 按照低、中、高 3 个气象站点密度等级进行多年平均气温数据的空间化, 并统计出 13 个检验站点的 MAE、RSME (表 3)。(1) 应用 IDW 和 OK 插值模型, 采用中等密度站点对年平均气温插值精度最好, 高密度站点插值精度次之, 但两者差异不显著低密度站点插值精度最低, 3 种数据源密度之间的插值精度差异较大;(2) 应用 IDWE、OKE、MRM 插值模型, 随着站点密度的增大, 多年平均气温数据空间化误差总体上都保持减小的趋势, 3 种数据源密度插值精度之间的差异较小, 采用高密度站点插值年平均气温误差最小;(3) 3 种插值模型中, IDWE 方法插值精度最好, 低、中、高 3 种密度的年平均气温误差分别为 0.74°C 、 0.76°C 、 0.67°C ; OKE 模型次之, 其 MAE 值分别为 0.87°C 、 0.82°C 、 0.80°C 。

表3 不同站点密度插值精度检验结果

 $^\circ\text{C}$

空间化误差 ($^\circ\text{C}$)		IDW		IDWE		OK		OKE		MRM	
		MAE	RSME								
区域	低 (132)	1.21	1.82	0.74	0.92	1.51	2.40	0.87	1.23	1.08	1.20
站点数	中 (87)	0.76	1.39	0.76	0.97	1.24	2.00	0.82	1.10	1.01	1.17
(个)	高 (43)	0.92	1.53	0.67	0.91	1.27	1.95	0.80	1.04	0.99	1.16

2.3 空间化误差与空间化尺度的关系

针对 131 个气象站点 (即高密度), 分别采用上文中 5 种方法, 按照 10 个尺度等级进行多年平均气温数据的空间化, 并统计出 13 个检验站点的 MAE、RSME 值, 结果见图 4。在空间化尺度由 100m 增大到 10km 的过程中:(1) IDW 和 OK 两种插值模型随着空间化栅格尺度的增大, 其空间化误差变化不大。应用 IDW 和 OK 插值模型, 平均气温误差分别为 $0.92 \sim 0.99^\circ\text{C}$ 和 $1.26 \sim 1.34^\circ\text{C}$;(2) IDWE、OKE 模型下, 随着空间化尺度的增大, 其空间化误差逐渐增大。空间化尺度小于 2km 时, 其平均绝对误差小于 1°C 。

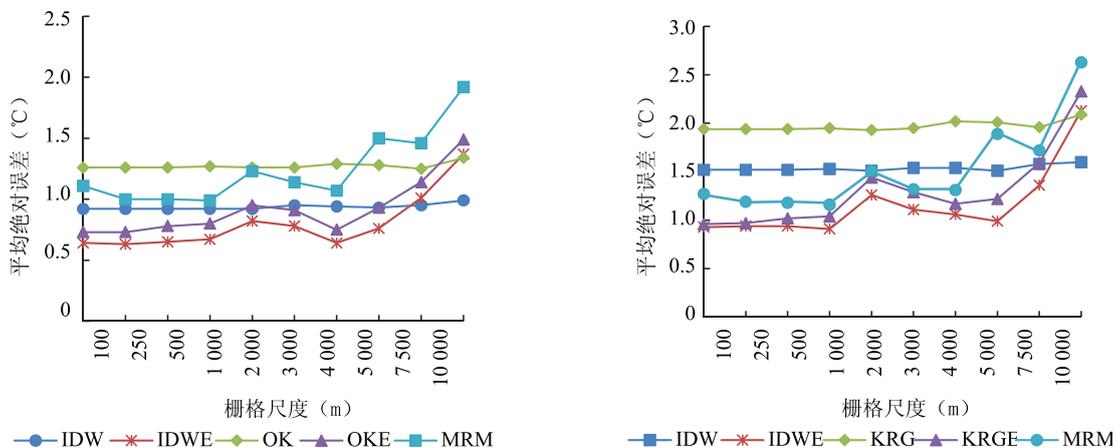


图 3 5 种模型方法下 10 种栅格尺度插值精度的 MAE 和 RSME

(3) 应用 MRM 插值模型下, 总体来说, 随着空间化尺度的增大, 其空间化误差逐渐增大。除采用 1km 栅格尺度外, 其余栅格尺度平均气温误差均超过 1℃; (4) 总体来说, 加入海拔因子的 IDWE、OKE、MRM 插值模型比 IDW 和 OK 插值模型精度更高, 其中以 IDWE 插值精度最好, OK 模型精度最差。

2.4 空间化误差与数据源密度和空间化尺度的定量关系

上述分析表明, 5 种插值方法中, 以 IDWE 插值精度最好。该部分采用插值最好的 IDWE 模型, 分别采用高、中、低 3 种站点密度, 按照 10 个尺度等级进行多年平均气温数据的空间化, 定量统计 13 个检验站点的平均绝对误差 (表 4)。(1) 随着空间化栅格尺度的增加, 空间化误差逐渐增加; 栅格尺度小于等于 2 000m 时, 3 种站点密度的平均气温误差小于 1℃。其中高密度站点插值精度最好; (2) 同样的空间化栅格尺度下, 高密度站点的空间化误差最小。高、中、低 3 种密度之间的空间化误差差距不明显。100m × 100m 尺度下 3 种密度的 MAE 在 0.64 ~ 0.76℃, 10km × 10km 尺度下 3 种密度 MAE 值在 1.37 ~ 1.69℃。

2.5 空间化误差与海拔数据源的关系

游松财等研究表明^[21], 由于国家气象局提供的气象站点的经纬坐标只精确到分, 其空间误差大约 ± 30" (约 1km), 因此在气温内插时考虑到海拔高度的影响, 并且使用 DEM 数据就必须考虑两者的海拔高度差异^[13]。利用四川省 144 个气象站点的经度及纬度资料, 从 ASTER GDEM 中提取相应的海拔高度, 发现有 20 个站点 (约占气象站点总数的 13.9%) 具有 100m 以上的差异, 其中有 9 个站点 (约占气象站点总数的 6.3%) 具有 200m 以上的差异, 已有的研究表明全国基础气象站点 (约 635 个) 中仅有 8% 左右站点海拔高度与 DEM 海拔高度有

表 4 不同数据源密度和不同空间化栅格大小的空间化误差分布
℃

空间化栅格 尺度 (m)	高 (131 个)		中 (87 个)		低 (43 个)	
	MAE	RSME	MAE	RSME	MAE	RSME
100	0.64	0.93	0.76	0.97	0.75	0.92
250	0.63	0.94	0.76	0.97	0.75	0.92
500	0.65	0.94	0.76	0.97	0.75	0.92
1 000	0.67	0.91	0.76	0.97	0.74	0.92
2 000	0.82	1.26	0.99	1.61	1.00	1.66
3 000	0.78	1.11	0.87	1.25	0.81	1.22
4 000	0.64	1.06	1.16	1.79	1.14	1.89
5 000	0.76	0.99	1.21	1.89	1.17	1.96
7 500	1.01	1.36	1.49	2.22	1.64	2.34
10 000	1.37	2.13	1.52	2.24	1.69	2.47

100m 以上的差距, 说明我国西南地区由于海拔较高, 地貌类型复杂, 给气象数据空间化增加了不确定因素。该部分采用 IDWE、OKE、MRM 3 种插值方法, 分析了气象站海拔与 DEM 海拔插值之间的气温绝对误差。气象站海拔与 DEM 海拔插值之间的气温绝对误差的空间分布见图 4。

图 4 表明, 采用 IDWE、OKE、MRM 3 种插值模型, 采取气象站海拔数据插值的气温 (T_{ms}) 与 DEM 海拔数据插值的气温 (T_{DEM}) 均存在差异, 3 种插值模型下的气温差空间分布相似。小部分区域 T_{DEM} 小于

表5 海拔高度差异及其温度差异对比

站号	站名	经度 (°E)	纬度 (°N)	海拔高度 (m)			垂直温度差 (°C)
				DEM	气象站	误差	
56097	九寨沟	104.25	33.27	1 904	1 437.70	-466.30	-1.75
56144	德格	98.58	31.80	3 618	3 185.00	-433.00	-1.62
—	—	—	—	—	—	—	—
56178	小金	102.35	31.00	2 277	2 439.20	162.20	0.61
57315	仪陇	106.40	31.53	453	656.80	203.80	0.76

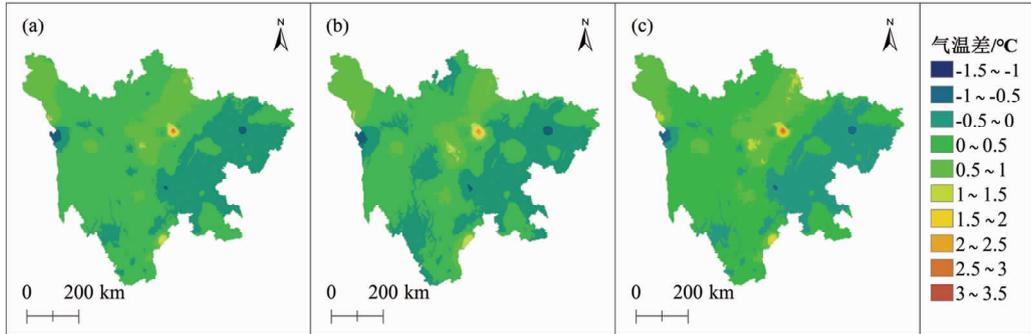


图4 IDWE (a)、OKE (b)、MRM (c) 3种插值模型下气象站海拔与DEM海拔插值的气温差

T_{ms} , 主要位于四川的东部区域; 大部分区域 T_{DEM} 大于 T_{ms} ; IDWE、OKE、MRM 模型下, T_{DEM} 与 T_{ms} 之间的气温差分别在 $-0.98 \sim 2.77^\circ\text{C}$ 、 $-0.98 \sim 2.69^\circ\text{C}$ 、 $-1.06 \sim 3.05^\circ\text{C}$ 范围内。3种插值模型下, 气温差异最大的地方主要位于四川的西北部、中北部以及东南部。以栅格面积统计, 3种插值模型空间化结果中, 气温差绝对值大于 1°C 的区域分别占全省面积的 0.76%、1.13%、1.52%, 其中气温差绝对值大于 1.5°C 的区域仅分别占全省面积的 0.21%、0.26%、0.31%, 所以两种海拔数据插值结果的差异性并不显著。

按 $500\text{m} \times 500\text{m}$ 的栅格密度分别以气象站海拔和DEM海拔进行气温插值, 将海拔分为10个级别, 分别统计不同方法下各海拔范围内两者之间的气温差(表6)。结果表明: (1) 采用3种方法对不同海拔插值的气温存在差异, 在 $0 \sim 500\text{m}$ 范围内两者之间差异较小, 气温差小于 0.1°C 。(2) 海拔大于 1500m 时, 气温差普遍大于 0.3°C ; 海拔在 $2000 \sim 2500\text{m}$ 范围内时, 气温差平均值为 0.709°C ; 海拔在 $2500 \sim 3000\text{m}$ 范围内时, 气温差平均值为 0.706°C ; 海拔大于 5000m 时, 气温差平均值为 0.637°C 。

2.6 空间化误差和地形分区的关系

以四川省地貌区划和气候区划为参照, 将全省分为4个区^[26], 见图5。采用131个站点, 5种方法分别对各区进行插值后拼接, 与直接对全省131个站点插值的结果进行对比, 仍然以13个校验站点进行分析, 其中IDWE表现仍为最优(图6), 其MAE、RSME分别为 0.44°C 、 0.62°C , 按不同地貌区划进行分区插值, 各种方法的精度明显优于全域(图7), 且MAE与RSME值的差距缩小。

对比分区与不分区IDWE模型下的插值结果图, 不分区拟合结果年均温的值域为 $-9.64 \sim 21.76^\circ\text{C}$,

表6 站点海拔与站点DEM值参与插值的各海拔高度其温差 $^\circ\text{C}$

海拔 (m)	IDWE	OKE	MRM
0 ~ 500	0.06	0.02	0.07
500 ~ 1 000	0.15	0.31	0.16
1 000 ~ 1 500	0.14	0.24	0.16
1 500 ~ 3 000	0.59	0.58	0.67
1 500 ~ 2 000	0.32	0.35	0.38
2 000 ~ 2 500	0.69	0.67	0.76
2 500 ~ 3 000	0.68	0.66	0.78
3 000 ~ 4 000	0.33	0.32	0.37
4 000 ~ 5 000	0.48	0.54	0.52
> 5 000	0.61	0.67	0.63

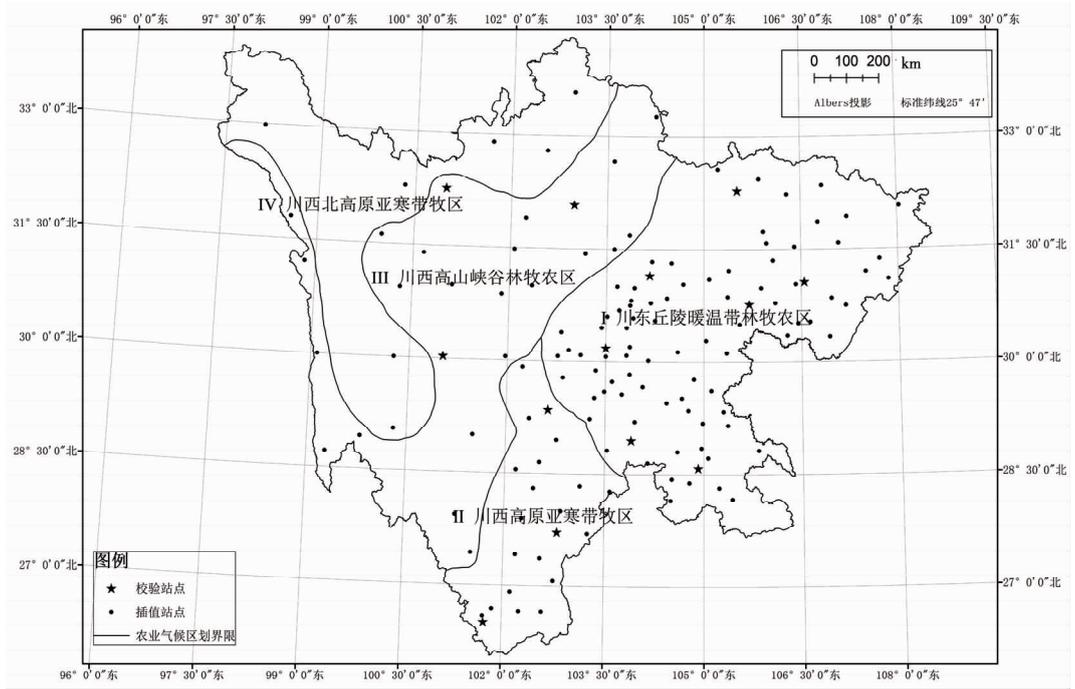


图5 四川省农业气候区划

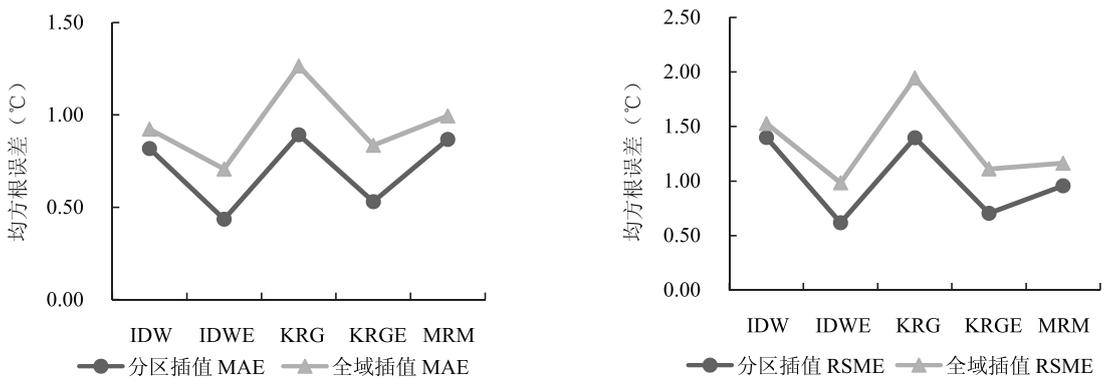


图6 采用5种方法分区与不分区插值精度比较

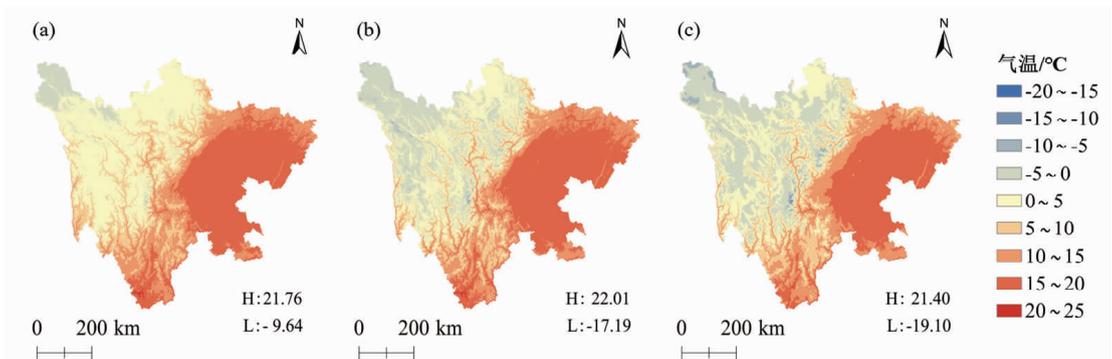


图7 IDWE模型下不分区与分区插值结果图 (a. 不分区插值; b. 分区插值; c. 已有的四川省多年平均气温分布数据)

分区拟合结果年均温的值域为 $-17.19 \sim 22.01^{\circ}\text{C}$ ，结合已有的四川省多年平均气温统计数据，可以看出分区拟合后气温的值域跨度更符合真实气温分布状况。

3 结论与讨论

3.1 空间化模型精度

空间化模型的选择对空间化精度有较大影响。通过精度分析，以空间自相关为主导的传统空间化模型，如IDW和OK，虽然理论和方法都比较成熟，但对真实气温环境的模拟能力较低，其相对误差较大。而IDWE、OKE和MRM等方法较全面的考虑了海拔、经纬度等因素的影响，对于海拔变化较大地区的气温有更好的表现。

3.2 数据源密度对空间化误差的影响

研究表明，数据源密度（站点数量）对空间化结果有一定影响，但不显著。随着站点数量的逐渐增加，所有空间化方法的插值精度均有所提高。加入海拔因子的IDWE、OKE、MRM等模型除整体精度高于不考虑海拔因子的方法外，随着站点密度的增大，空间化误差总体减小，精度逐步提高，但3种数据源密度插值精度之间的差异较小，表明即使减少气象站点数据，加入海拔因子的插值模型也能较好地模拟气温。其原因应该是空间自相关模型是直接对站点观测值进行插值，没有考虑其他影响因素，因此站点数量越多即样本数越多，插值精度自然越高。而考虑了海拔、经纬度等因素的异相关模型，方法核心是建立气温因子与其他环境因子的关系，样本数大小反倒对插值精度影响不大。由于一般科研人员能共享的地面气象观测站点数较少，这一验证结果也对我我国地面气象站点观测结果插值精度提供了有利的佐证。

3.3 栅格尺度对空间化误差的影响

栅格大小对空间化精度有一定的影响。一般空间化模型的栅格大小有经纬度和地理空间距离两种，两种方法的空间化尺度一般从 100km (1°) 到 30m ($1''$) 不等。该次研究随着空间化尺度从 100m 到 10km 的逐级变化，对于只考虑空间自相关性的模型，如IDW和OK，因未考虑海拔因素，随栅格尺度的变化，误差变化很小。但是，考虑了海拔影响的模型，如IDWE、OKE和MRM，随着空间化栅格单元的增大，表现出较显著的误差增大、空间化精度减小的趋势。究其原因，是因为考虑海拔因素的情况下，栅格尺度越大，栅格之间的方差缩小，栅格面趋于平滑连续，模拟真实情况的能力下降，又因垂直递减率在 $0 \sim 1$ 之间，在垂直递减率的作用下这种拟合能力下降的效应就被放大。

3.4 海拔数据源对空间化误差的影响

采用气象局提供的海拔数据和提取DEM的海拔数据进行插值，不同范围内均存在误差，尤其是四川的西北部、中北部以及东南部的气温差异较大。海拔在 $2\ 000 \sim 3\ 000\text{m}$ 和大于 $5\ 000\text{m}$ 时，气温差异达到 0.6°C 以上。因此，在地形复杂的地区，对气温插值时要适当考虑两种不同海拔数据源带来的差异。但从四川省插值结果栅格统计面积来看，误差绝对值大于 1°C 的区域占全省面积仅为 1% 左右，在相对宏观的数据分析时可以不考虑两种海拔数据源对空间化误差的影响。

3.5 分区拟合对空间化误差的影响

在地貌类型多样，气候条件复杂的地区，采用分区拟合能降低空间化误差，类似的研究思路在其他社会经济数据空间化时多有体现^[27]。区域地貌、气候的差异导致地区与地区间存在明显气压差异的，其垂直递减率亦不相同，这是导致分区拟合优于全域拟合的重要原因。由于历史原因，川西高原和四川盆周山区属于气象观测资料稀疏区，虽然分区拟合能提高空间化精度，但对川西地区提升有限。未来需要加密自动观测站网，在地质灾害易发区、输油管线和输电通道等重要基础设施布局区、高原牧区完善气象灾害监测能力，提升中小尺度天气变化监测能力，发挥分区拟合等不同空间化方法的优势提升插值精度。

3.6 减小空间化误差的途径

根据上述分析，为减少空间化误差，空间化过程中宜采用空间异相关模型，如考虑了海拔因子影响的空间自相关模型，或者同时考虑了海拔、经纬度、距海岸线距离等因素的多元回归模型并用残差校正；栅

格单元应符合气象观测数据所能代表的范围,一般宜采用500m(15")至2000m(1')的栅格单元;气象站点应具有一定的代表性,覆盖各种地貌类型,空间分布上相对均匀,在气象观测资料稀疏区加密自动观测站点。由于观测资料本身受到时空代表性有限、分辨率不足、时空覆盖范围有限及不同观测资料之间的系统偏差等问题的限制,以观测资料为基础进行多元资料的融合分析及利用数值模式进行资料的同化分析与再分析,是获得质量可靠、空间覆盖完整、分辨率高的气象数据产品的普遍做法^[28]。对于四川省累年年平均气温、累年月平均气温等数据空间化产品生产加工,宜采用IDWE、OKE、MRM和Anusplin等空间异相关模型,栅格单元的大小控制在2km以内,可以保障产品的平均绝对误差低于1℃。

参考文献

- [1] 廖顺宝,张赛. 多年平均气温数据空间化误差的尺度效应. 地球信息科学学报, 2014, 16(1): 8-14.
- [2] 张明洁,赵艳霞. 日光温室气候适宜性研究——以北方地区为例. 中国农业资源与区划, 2012, 33(05): 40-48.
- [3] 于贵瑞,何洪林,刘新安,等. 中国陆地生态信息空间化技术研究(I)——气象/气候信息的空间化技术途径. 自然资源学报, 2004, 19(4): 537-544.
- [4] 何洪林,于贵瑞,刘新安,等. 中国陆地生态信息空间化技术研究(II)——太阳辐射要素. 自然资源学报, 2004, 19(5): 679-687.
- [5] 刘新安,于贵瑞,范辽生,等. 中国陆地生态信息空间化技术研究(III)——温度、降水等气候要素. 自然资源学报, 2004, 19(6): 818-825.
- [6] Eleanor R. Cross, M. S., B. S. Rick Perrine, P. D. Charles Sheffield, et al. Predicting Areas Endemic for Schistosomiasis Using Weather Variables and a Landsat Data Base. Military Medicine, 1984, 149(10): 542-544.
- [7] Marquinez, J., J. Lastra, P. Garcia. Estimation models for precipitation in mountainous regions: the use of GIS and multivariate analysis. Journal of Hydrology, 2003, 270(1-2): 1-11.
- [8] Daly, C., G. Taylor, W. Gibson. The Prism Approach to Mapping Precipitation and Temperature. in Proceedings of the 10th AMS Conference on Applied Climatology. 1997. Amer. Meteorological Soc., Reno, NV.
- [9] Hutchinson, M. F. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 1995, 9(4): 385-403.
- [10] J. Jeffrey, S., J. O. Carter, K. B. Moodie, et al. Using spatial interpolation to construct a comprehensive archive of Australian climate data. Environmental Modelling & Software, 2001, 16(4): 309-330.
- [11] Thornton, P. E., S. W. Running, M. A. White. Generating surfaces of daily meteorological variables over large regions of complex terrain. Journal of Hydrology, 1997, 190(3-4): 214-251.
- [12] 傅抱璞. 关于山地气候资料的延长和推算问题. 气象学报, 1982(4): 453-463.
- [13] 傅抱璞. 地形和海拔高度对降水的影响. 地理学报, 1992(4): 302-314.
- [14] 沈国权. 考虑宏观地形的小网格温度场分析方法及应用. 气象, 1984(6): 22-27.
- [15] 陈晓峰,刘纪远,张增祥,等. 利用GIS方法建立山区温度分布模型. 中国图象图形学报, 1998(3): 66-70.
- [16] 许民,王雁,周兆叶,等. 长江流域逐月气温空间插值方法的探讨. 长江流域资源与环境, 2012, 21(3): 327-334.
- [17] 蔡福,于贵瑞,祝青林,等. 气象要素空间化方法精度的比较研究——以平均气温为例. 资源科学, 2005(5): 173-179.
- [18] 刘球球. 气象要素插值的空间化精度提高方法研究. 气象科学, 2017, 37(2): 139-143.
- [19] 廖玉芳,汪扩军,赵福华,等. 湖南省现代农业气候区划. 长沙: 湖南大学出版社, 2010.
- [20] 冯晓云,王建源. 基于GIS的山东农业气候资源及区划研究. 中国农业资源与区划, 2005(2): 60-62.
- [21] 游松财,李军. 海拔误差影响气温空间插值误差的研究. 自然资源学报, 2005, 20(1): 140-144.
- [22] 谢云峰,张树文. 基于数字高程模型的复杂地形下的黑龙江平均气温空间插值. 中国农业气象, 2007, 28(2): 205-211.
- [23] 郑小波,罗宇翔,于飞,等. 西南复杂山地农业气候要素空间插值方法比较. 中国农业气象, 2008, 29(4): 458-462.
- [24] 翟丹平,白红英,冯海鹏,等. 基于气象数据和遥感影像的太白山气温直减率. 山地学报, 2016, 34(04): 496-503.
- [25] 江蕾. 基于CIMISS四川地区气温观测数据的格点化应用研究. 第34届中国气象学会年会 S20 气象数据: 深度应用和标准化论文集. 2017.
- [26] 甘书龙,四川省农业资源与区划. 成都: 四川省社会科学院出版社, 1986.
- [27] 廖顺宝,姬广兴,侯鹏敏,等. 面向粮食产量空间化的多元回归分析模型中的两个关键问题探讨. 自然资源学报, 2015, 30(11): 1922-1932.
- [28] 黄嘉佑,李庆祥,气象数据统计分析方法. 北京: 气象出版社, 2014.

STUDY ON SPATIAL INTERPOLATION METHOD AND INFLUENCING FACTORS OF ANNUAL MEAN AIR TEMPERATURE DATA IN SICHUAN PROVINCE *

He Peng^{1, 2}, Zhang Yuan^{2, 3}, Gao Wenbo^{1, 2}, Jian Dongnan², Lin Zhengyu^{4**}

(1. Institute of Agricultural Information and Rural economy, Sichuan Academy of Agricultural Sciences, Chengdu, Sichuan 610066, China;

2. Big data Center, Sichuan Academy of Agricultural Sciences, Chengdu, Sichuan 610066, China;

3. School of Tourism and Urban and Rural Planning, Chengdu University of Technology, Chengdu, Sichuan 610059, China;

4. College of Resources, Sichuan Agricultural University, Chengdu, Sichuan 611130, China)

Abstract This research aims to explore annual mean temperature data spatial interpolation method that is suitable for the topographical features of Sichuan province. Considering the influence of altitude, data source density and spatialized grid size, five methods were selected to interpolate the annual average temperature observation data of 144 meteorological observation stations in Sichuan province from 1981 to 2010, and the results were evaluated by two methods: reserved inspection site and cross-validation. The interpolation precision of spatial model was IDWE > OKE > MRM > OK > IDW; the interpolation precision of medium-density weather station was the highest when altitude factor was not considered, and the interpolation precision of high-density station was the highest when altitude factor was taken into account. With the increasing of the cell size, the error of the interpolation method without considering the altitude factor did not change much, and the error of the interpolation method considering the altitude factor increases gradually, and the error increased rapidly when the spatial cell size was larger than 4 000m. The interpolation results were affected by different altitude data sources. The temperature difference of the three interpolation methods was between -0.98 and 3.05℃, but the absolute temperature difference greater than 1℃ only accounts for 0.76% ~ 1.52% of the total area of the province. The main conclusions are showed as follows. (1) The choice of spatial model method has a great influence on the interpolation precision of air temperature, and the accuracy of the model considering the environmental factors such as altitude is higher than that of the simple spatial autocorrelation model; (2) With the decrease of the number of data sources (meteorological stations), the spatialization error increases, but it is not obvious; (3) With the grid size increasing gradually, the spatialization error increases gradually, and the suitable scale is 500 ~ 2000m; (4) The elevation error of meteorological stations will affect the spatial accuracy of air temperature, but it is not significant from the area of the whole province; (5) Complex terrain will affect the effect of spatialization of air temperature, and the fitting accuracy of zoning is better than that of non-partitioned fitting.

Keywords annual mean temperature; spatial error; elevation; influencing factors; Sichuan province