

# 安徽省气象数据空间插值方法比较与分布特征

赵冰雪, 王雷, 程东亚

(池州学院 资源环境学院, 安徽 池州 247000)

**摘要:**以安徽省境内 80 个气象站点 2011—2014 年的降水量和 2015 年的逐日气温数据为基础,结合该区 30 m 的 DEM 数据,选取反距离权重加权法(IDW)、径向基函数法(RBF)、普通克里金法(OK)和协同克里金法(CK)对降水和气温数据分别进行了空间插值,并对模拟结果进行了交叉验证。结果表明:考虑了纬度和高程的协同克里金插值效果最好,降水和气温的插值误差排序均为:CK<OK<RBF<IDW,并对插值结果进行分析,发现研究区降水和气温空间分布均呈现由南向北逐渐递减的特点,并与经纬度、高程进行相关系数和回归分析,表明气象要素纬度地带性显著,降水量主要由纬度和高程决定,气温则主要受纬度的影响,月均气温与经度和纬度有一定的季节性规律。

**关键词:**GIS; 降水; 气温; 空间插值

中图分类号:P468

文献标识码:A

文章编号:1005-3409(2017)03-0141-05

## Comparison of Spatial Interpolation Method for Meteorological Data and Distribution Characteristic in Anhui Province

ZHAO Bingxue, WANG Lei, CHENG Dongya

(College of Resources and Environment, Chizhou University, Chizhou, Anhui 247000, China)

**Abstract:** By using the precipitation data from 2011 to 2014 and the daily temperature data in 2015 in 80 meteorological stations in Anhui Province, combining with the DEM data of 30 m, we selected inverse distance weighting (IDW), radial basis function (RBF), ordinary Kriging (OK) and collaborative Kriging (CK) to interpolate the precipitation and temperature, and compared with simulation results using cross validation method. The results show that the CK considering latitude and elevation has the most effective, error estimates of precipitation and temperature are: CK<OK<RBF<IDW, by analyzing the precipitation results, the precipitation and temperature show a trend of gradually increasing from the south to the north, and present a clear zonal latitude, the precipitation is mainly decided by latitude and elevation, the temperature is mainly affected by latitude, monthly average temperature has certain seasonal patterns with longitude and latitude.

**Keywords:** GIS; precipitation; temperature; spatial interpolation

气象要素是指一定地点和特定时刻的天气状况,是地学研究和气候模型的重要指标。降水和气温作为农作物生长发育的重要指标,很大程度上决定了农作物的分布类型和生长状况,只有了解降水的空间分布情况,才能合理利用水资源,防治自然灾害,为农业生产服务。但在实际工作中,由于人力、物力和财力的原因,气象站点的布设不可能十分密集和均匀,尤其是在地形复杂的偏远山区,因此,利用有限的气象站点进行插值,进而得到全区的气象数据是一种常用

的方法。孟庆香<sup>[1]</sup>、李飞<sup>[2]</sup>等以气象要素为基础,对几种常规插值方法进行比较;蔡迪花<sup>[3]</sup>、刘智勇<sup>[4]</sup>、许民<sup>[5]</sup>等基于 DEM 和经纬度等地理要素,建立多元回归分析插值法,提高了插值精度;徐超等<sup>[6]</sup>以山东省为研究区,通过增加模拟站点提高气象要素的插值效果;王红霞等<sup>[7]</sup>将 IDW,OK,CK 插值与“多元回归+残差分析法(AMMRR)”进行对比,结果表明 AMMRR 对年积温和年降水的插值精度最高;周婷婷等<sup>[8]</sup>以福州市为研究区,利用 MODIS 影像数据反演

地表温度,在参数使用正确情况下比插值方法反演精度高,适用于气象数据不易获得或站点较少的情况。

安徽省地形复杂多样,南北跨度较大,季节性差异明显。因此,文章以安徽省为研究区,通过收集该省近几年的气象数据,利用交叉验证法从气温和降水两方面对几种插值方法的插值精度进行比较,并分析气象要素与数字高程模型(DEM)以及经纬度之间的相关关系,建立多元回归方程,研究该区气象要素的空间分布特征,为区域农业生产及气候要素的评价提供一定的参考。

## 1 研究区概况与数据来源

### 1.1 研究区概况

安徽省地处中国华东地区,地理位置位于东经 $114^{\circ}54'$ — $119^{\circ}37'$ ,北纬 $29^{\circ}41'$ — $34^{\circ}38'$ ,土地总面积 $13.94$ 万 $\text{km}^2$ ,2015年总人口 $6\,979$ 万,GDP达到 $22\,005.6$ 亿元。气候属暖温带与亚热带的过渡地区,境内淮河是我国重要的南北地理分界线,淮河以北属暖温带半湿润季风气候,淮河以南属亚热带湿润季风气候。地形由淮北平原、江淮丘陵、皖南山区组成,省内最高峰为黄山莲花峰海拔 $1\,873$  m,南北高程变化较大。受季风和复杂地形的影响,安徽省降水和气温存在较为复杂的空间分布特征。

### 1.2 数据来源与处理

国家气象站点的布设综合考虑了自然条件与行政体系,根据安徽省主要由平原和丘陵组成的地形条件,该省的气象站点布设基本按行政区进行,即各县级行政区基本设置一个气象观测站,全省共布设 $80$ 个国家气象观测站,包括国家基本气象站和国家一般气象站,由于各站点均为国家气象观测站,所以具有一定的代表性。

本研究所用数据包括 $80$ 个气象站点的降水量、气温、经纬度、高程。其中,降水数据为 $2011$ — $2014$ 年的年平均降水量,气温数据为 $2015$ 年的逐日气温,来源于安徽省气象局和安徽省统计年鉴,另外还包括 $30$  m的DEM,坐标为Albers等积投影,中央经线为 $E110^{\circ}$ ,标准纬线为 $N25^{\circ}$ 和 $N47^{\circ}$ (图1)。

数据来源的准确性对插值结果的影响显著,因此,要首先对原始数据进行可靠性验证。由于各气象站点每日都有实测气温,每月都有实测降水量,因此,通过对 $2011$ — $2014$ 年各站点的逐月降水量和 $2015$ 年的逐日气温进行统计,计算出各站点的年均降水量和年均气温。运用ArcGIS 10.2软件的地统计模块进行数据探索

性分析,对数据进行对数变换,结果显示数据质量良好,整体呈正态分布,且南北方向趋势明显。

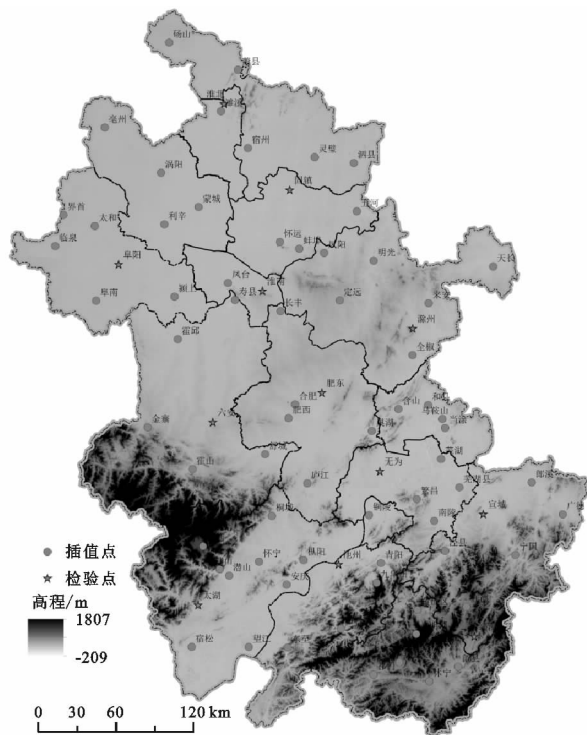


图1 安徽省气象站点分布

## 2 空间插值方法与精度验证

目前常用的空间插值方法主要包括反距离权重插值、样条函数插值、克里金插值、多元线性回归法等。这些插值方法按照数学原理总体上可以归纳为确定性插值和地统计插值,前者包括反距离权重插值、径向基插值、全局多项式插值和局部多项式插值,后者主要指克里金插值和协同克里金插值;另外,根据创建的表面是否经过所有的采样点,插值方法又可以分为精确性插值和非精确性插值。在对插值结果精度进行验证时,标准平均值最接近于 $0$ ,均方根最小,平均值误差最接近均方根误差,平均标准误差最接近于 $1$ ,模型的拟合情况越好<sup>[9]</sup>。论文基于以上两种插值分类方法,考虑本研究所用数据主要为降水量和气温,数据自身具有一定的空间自相关性,确定选取反距离权重插值、径向基函数插值、克里金插值和协同克里金插值几种方法进行插值误差的比较。

### 2.1 反距离权重法

反距离权重加权属于精确性插值,所谓精确插值就是指表面必须经过每一个已知样点。它是基于相似的原理,离试验点的距离越近,权值越大,比较适用于山区或者降水站点不是很密集的地区,缺点是没有考虑地形因素的影响。

$$Z(x_o)=\sum_{i=0}^n\frac{Z(x_i)}{(d_{io})^p}/\sum_{i=0}^n\frac{1}{(d_{io})^p}$$

(1)

式中: $Z(x_o)$ 为 $x_o$ 点的估计值; $Z(x_i)$ 为 $x_i$ 点的真实值; $n$ 为用于插值的气象站点数; $d_{io}$ 为预测点 $x_o$ 到已知站点 $x_i$ 的距离; $p$ 为距离的幂。

2.2 径向基函数插值法

径向基函数包括五种不同的基本函数:平面样条函数、张力样条函数、规则样条函数、高次曲面函数和反高次曲面样条函数。该方法适用于对大量点数据进行插值计算,同时要求获得平滑表面的情况。优点是方法相对比较稳健,缺点是当采样点数据具有很大的不确定性时,不合适使用此种方法。以样条函数为例,计算公式如下<sup>[10]</sup>:

$$Z(x_o)=\sum_{i=0}^nA_id_{io}^2\lg d_{io}+ax_o+by_o+c$$

(2)

式中: $Z(x_o)$ 为 $x_o$ 点的估计值; $n$ 为用于插值的气象站点数; $d_{io}$ 为待插值点与临近点之间的距离; $ax_o+by_o+c$ 为气温的局部趋势函数; $A_i,a,b,c$ 为方程系数。

2.3 普通克里金插值法

又称地统计法,它以空间自相关为基础,利用原始数据和半方差函数的结构性,对区域化变量的未知采样点进行无偏估值的插值方法。

$$Z(x_o)=\sum_{i=1}^n\lambda Z(X_i)$$

(3)

式中: $Z(x_o)$ 为 $x_o$ 点的估计值; $Z(x_i)$ 为 $x_i$ 点的真实值; $n$ 为用于插值的气象站点数; $\lambda$ 为克里金权重系数。

2.4 协同克里金法(CoKriging)

适用于采样点的多个属性之间存在空间分布,且某些属性不易获得,而另一些属性则易于获取时的情况。研究表明,高程与年平均降水量的相关性达到0.175<sup>[11]</sup>,所以采用高程作为影响降水量空间插值的第二影响因素是可取的,除了高程之外,还可以考虑经纬度、坡度、坡向、距离海岸的距离等因素,这有利于提高协同克里金插值算法的精度<sup>[3-5]</sup>。

$$Z(x_o)=\sum_{i=1}^n\lambda_iZ(x_i)+\lambda;[y(x_o)-m_y+m_z]$$

(4)

式中: $Z(x_o)$ 为 $x_o$ 点的估计值; $Z(x_i)$ 为 $x_i$ 点的真实值; $n$ 为用于插值的气象站点数; $y(x_o)$ 是 $x_o$ 点的高

程; $m_y$ 和 $m_z$ 为海拔高程和降水量的全局平均值; $\lambda_i$ 和 $\lambda$ 是协同克里金的权重系数。

2.5 插值精度验证

为了对比以上几种插值方法的插值效果,选取精度最高的方法进行气象要素空间分布特征的分析,本研究采用直接检验和交叉验证的方法进行插值精度的比较。基本思路是通过选取一部分样本点进行插值运算,其余的样本点作为检验点,然后分别从插值表面中提取检验点的预测值,并将预测值与真实值进行比较,最后选用平均绝对误差(Mean Absolute Error, MAE)、平均相对误差(Mean Relative Error, MRE)和均方根误差(Root Mean Square Error, RMSE)几种指标进行交叉验证<sup>[12]</sup>。插值点和检验点的个数将直接影响插值效果,插值点过少不能全面概况全省的气象分布情况,插值点过多、检验点过少又可能会引起交叉验证结果不准确,因此,文章在参考李飞<sup>[2]</sup>、林金煌<sup>[13]</sup>等研究的基础上,最终选取85%的站点作为插值点,其余15%站点作为检验点,且检验点分布整体较为均匀,平原、丘陵和山区地形均有布设。交叉验证选择的几种指标表达式为:

$$MRE=\frac{1}{n}\sum_{i=1}^n\left|\frac{Z_i-Z(X)}{Z_i}\right|$$

(5)

$$MAE=\frac{1}{n}\sum_{i=1}^n|Z_i-Z(X)|$$

(6)

$$RMSE=\sqrt{\frac{\sum_{i=1}^n[Z_i-Z(X)]^2}{n}}$$

(7)

式中: $n$ 为检验点个数; $Z_i$ 为第 $i$ 个站点的实测值; $Z(X)$ 为估计值。MRE, MAE, RMSE 值越小,表示表面插值误差越小、精度越高,模拟效果也就越好。

3 结果与分析

3.1 插值结果与分析

运用IDW, RBF, OK 和 CK 四种插值方法分别对多年平均降水量和2015年平均气温进行插值,并根据交叉验证公式进行计算得到不同插值方法的效果(表1)。

表 1 不同插值方法交叉验证结果

验证指标	降水量				气温			
	IDW	CK	OK	RBF	IDW	CK	OK	RBF
MAE	50.670	38.120	44.310	47.820	0.4015	0.3600	0.3924	0.3932
MRE	0.049	0.038	0.043	0.045	0.0243	0.0220	0.0238	0.0239
RMSE	62.850	43.800	50.880	58.960	0.5034	0.4709	0.4695	0.4933

由表1可以看出,无论是降水量插值还是气温插值,在4种插值方法中,考虑高程和纬度的协同克里金

插值的误差均最小,模拟效果最好。年均降水量的MAE, MRE 和 RMSE 误差排序均为CK<OK<RBF<

IDW, 气温的 MAE 和 MRE 误差排序为  $CK < OK < RBF < IDW$ , RMSE 误差排序为  $OK < CK < RBF < IDW$ 。这是由于气象数据本身具有较强的空间自相关性, 且研究区地势南高北低, 对降水影响较大, 纬向差异性对气温影响较为显著。因此在选择插值方法时, 应根据研究区实际情况尽可能考虑高程、经纬度等因素。

### 3.2 降水空间分布特征分析

通过对 80 个气象站点的降水量数据分别采用不同的插值方法生成降水表面栅格图(图 2)。从图 2 中可以看出: 安徽省降水量分布整体呈现明显的自南

向北递减的规律, 其中, 皖南山区的黄山市降水量居于全省最高值, 而皖北平原的淮北市、亳州市由于整体地势平坦, 且处于内陆, 受季风影响较弱, 降水量在全省处于较低值, 全省年降水量分布较高的区域基本与黄山、九华山、天柱山、大别山等几大山系的空间位置趋于一致。此外, 霍邱县、天长市的降水量分布不符合纬向差异性, 这主要受霍邱县地势较高, 而天长市距离海岸较近的影响。另外, 阜南、淮南、合肥、蚌埠、滁州市的降水量均在 1 000 mm 左右, 与地理上划分南方和北方的秦岭—淮河线基本相符。

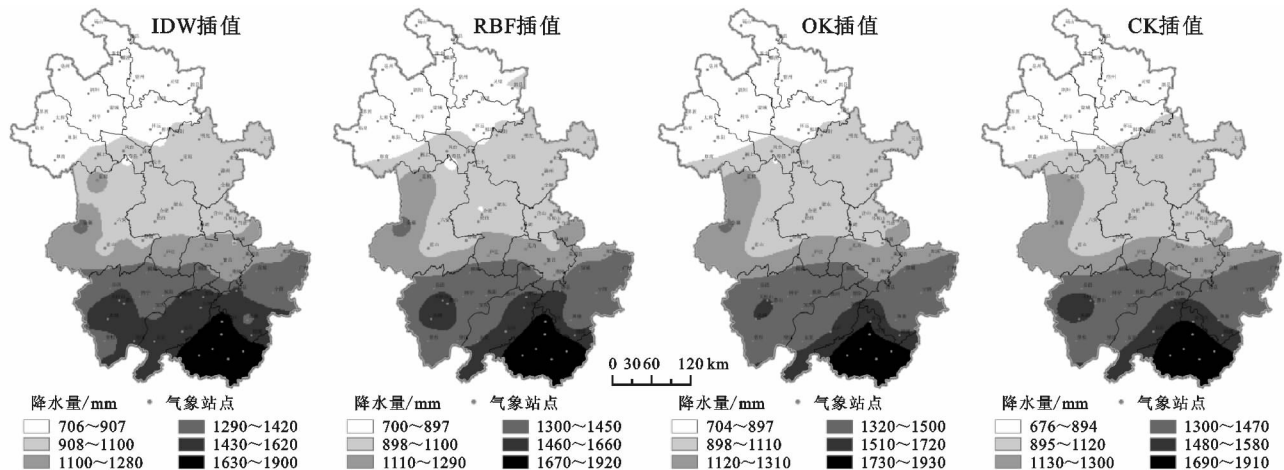


图 2 降水量插值效果

### 3.3 气温空间分布特征分析

本研究通过对安徽省 80 个气象站 2015 年的逐日气温进行统计, 得到该年的逐月和年平均气温, 对年平均气温进行插值分析得到气温模拟栅格图(图 3), 并结合每个气象站的经纬度和高程值, 计算每个月的平均气温与经纬度、高程之间的相关性, 建立气温与地理空间特征的多元线性回归方程(表 2)。结果表明: 对于年均气温而言, 安徽省气温呈明显的南高北低现象, 南北跨度 570 km, 气温相差 3.06℃。其中, 以休宁、祁门、东至和宿松县的气温最高, 以砀

山、萧县和灵璧县的气温最低; 对于月均气温而言, 气温与经纬度、地形有一定的季节性规律。从表 2 可以看出, 除 7 月份之外, 其他月份的平均气温与纬度均呈负相关, 尤其是 1 月、2 月、11 月、12 月, 相关性均在 0.80 以上。其次, 气温与经度也有一定的相关性, 除 6 月和 7 月以外, 其他月份气温与经度均成正相关, 基本符合水平地带性规律。这是因为安徽省夏季南方多雨水, 降水后气温也会有所下降, 而北方属于暖温带半湿润气候, 7 月多为高温炎热天气, 气温较高。

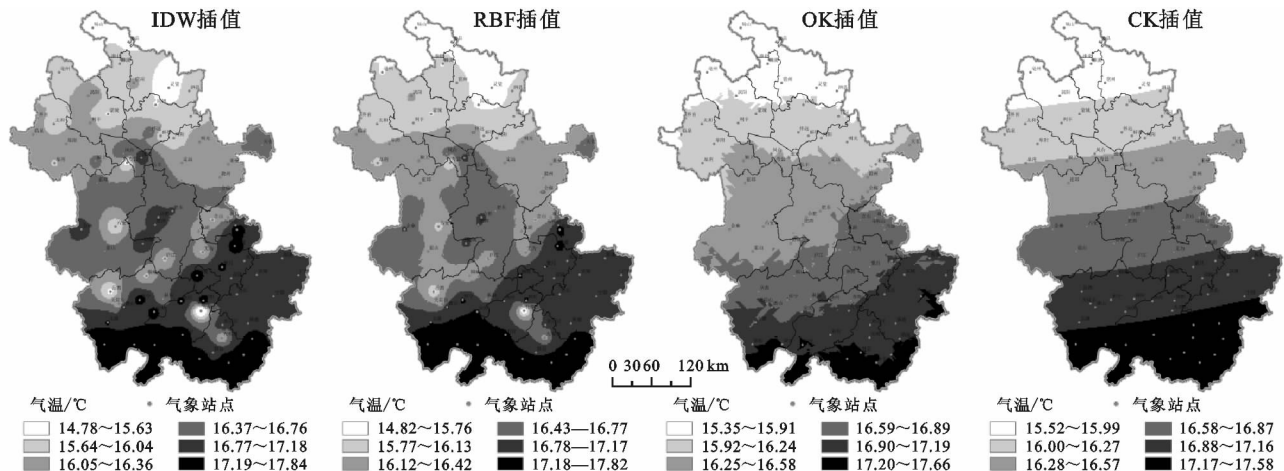


图 3 气温插值效果

表 2 月平均气温与经、纬度和高程间的相关系数与回归方程

月份	纬度	经度	高程	$R^2$	回归方程
1	-0.84	0.18	0.02	0.807	$T_1=49.781739-0.183357 X-0.720210 Y-0.002395 H$
2	-0.89	0.26	0.09	0.851	$T_2=49.971815-0.142017 X-0.842837 Y-0.002263 H$
3	-0.76	0.20	0.01	0.663	$T_3=41.747793-0.110595 X-0.551620 Y-0.001942 H$
4	-0.83	0.30	0.08	0.745	$T_4=43.516685-0.047692 X-0.676589 Y-0.001894 H$
5	-0.36	0.03	-0.35	0.406	$T_5=53.685467-0.153748 X-0.416542 Y-0.004090 H$
6	-0.10	-0.01	-0.33	0.169	$T_6=42.023714-0.086463 X-0.204365 Y-0.003386 H$
7	0.36	-0.02	-0.69	0.499	$T_7=15.393175+0.069660 X+0.103939 Y-0.004132 H$
8	-0.07	0.00	-0.44	0.249	$T_8=42.927489-0.071623 X-0.222033 Y-0.004678 H$
9	-0.31	0.32	-0.38	0.391	$T_9=12.513239+0.176536 X-0.300516 Y-0.004271 H$
10	-0.51	0.22	-0.24	0.458	$T_{10}=31.195408-0.008879 X-0.364845 Y-0.002788 H$
11	-0.85	0.62	0.11	0.846	$T_{11}=-17.512407+0.499675 X-0.927018 Y-0.002295 H$
12	-0.86	0.48	0.04	0.819	$T_{12}=12.470216+0.158891 X-0.787062 Y-0.002637 H$

注:  $T_i$  为第  $i$  个月的回归气温;  $X$  为经度;  $Y$  为纬度;  $H$  为高程;  $R^2$  为多元相关系数的平方值。

4 结 论

(1) 根据研究区地理位置和气象数据特征,选取 IDW,RBF,Kriging 和 CoKriging 四种插值方法进行插值计算,并通过交叉运算比较降水和气温模拟效果。结果表明,降水和气温的插值误差均为:CK<OK<RBF<IDW,即考虑经纬度的协同克里金插值效果最好。

(2) 安徽省降水量分布整体呈自南向北递减的规律,黄山市降水量最多,淮北市、亳州市降水量最少,高程和纬度对降水量影响作用明显。

(3) 安徽省年均气温南高北低的趋势明显,休宁、祁门、东至和宿松县的气温最高,砀山、萧县和灵璧县的气温最低,月均气温与经纬度、地形有一定的季节性规律,气温主要受纬度和经度的影响。

考文献:

[1] 孟庆香,刘国斌,杨勤科. 基于 GIS 的黄土高原气象要素空间插值方法[J]. 水土保持研究,2010,17(1):10-17.

[2] 李飞,孙勇,郑佳佳. 安徽省降水量空间插值研究[J]. 水土保持研究,2010,17(5):183-186.

[3] 蔡迪花,郭锐,李崇伟. 基于 DEM 的气温插值方法研究[J]. 干旱气象,2009,27(1):10-17,28.

[4] 刘智勇,张鑫,方睿红. 基于 DEM 的榆林市降水空间插值方法分析[J]. 西北农林科技大学学报:自然科学版,

2010,38(7):227-234.

[5] 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[J]. 长江流域资源与环境,2012,21(3):327-334.

[6] 徐超,吴大千,张治国. 山东省多年气象要素空间插值方法比较研究[J]. 山东大学学报:理学版,2008,43(3):1-5.

[7] 王红霞,柳小妮,李纯斌,等. 甘肃省近 42 年降水量变化时空分布格局分析[J]. 中国农业气象,2013,34(4):384-389.

[8] 周婷婷,陈文惠. 基于 MODIS 数据和气象观测数据的气温空间插值方法比较[J]. 地理科学进展,2011,30(9):1143-1151.

[9] 汤国安,杨昕. ArcGIS 地理信息系统空间分析试验教程 [M]. 2 版. 北京:科学出版社,2012.

[10] 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.

[11] Hevesi J A, Istok J D, Flint A L. Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis[J]. Journal of Applied Meteorology, 1992,31(7):661-676.

[12] 吴昌广,林德生,周志翔,等. 三峡库区降水量的空间插值方法及时空分布[J]. 长江流域资源与环境,2010,19(7):752-758.

[13] 林金煌,林广发. 福建省降水空间插值方法比较与区域分布特征研究[J]. 海南师范大学学报:自然科学版,2015,28(1):61-65.