

基于 GIS 的气象要素空间插值方法研究

马轩龙¹, 李春娥², 陈全功¹

(1. 兰州大学草地农业科技学院 农业部草地农业生态系统学重点实验室, 甘肃 兰州 730020;

2. 山西省农业遥感中心, 山西 太原 030002)

摘要:基于地理信息系统软件 ArcMap 的统计分析模块,对我国及周边地区 2 114 个气象站点 1961—1990 年的年均温度、年降水量以及年积温等数据,分别使用反距离权重法、样条函数法和普通克里格法,选取不同的气象站点进行了空间插值,并利用交叉检验方法对插值精度进行了评估,结果表明:对于同一种插值方法,参与插值的气象站点数目不同,插值结果也不同。对 3 种气象要素的插值结果进行验证发现,普通克里格法均具有最好的插值精度。对年均温度和年降水量来说,样条函数法的插值精度优于反距离权重法,而对年积温,反距离权重法的插值精度优于样条函数法。通过对原始数据进行一定的处理,可以有效提高最终插值结果的精度。研究表明,我国水热的空间分布呈现明显的东西、南北分界,与胡焕庸线大致相一致,此线以东,水热条件充分,此线以西水热条件较差。

关键词:气象要素;空间插值;反距离权重法;样条函数法;普通克里格法

中图分类号: TP79

文献标识码: A

文章编号: 1001-0629(2008)11-0013-07

*¹ 气象要素是草原综合顺序分类法的重要指标^[1,2],对生态系统中物种的组成、生长、演替更新及干物质的积累有很大的影响,也与生物多样性和土地利用潜力有密切的联系^[3],具有直接生态环境指示的作用。然而,气象要素很难由卫星遥感直接获取,需要利用站点数据与插值方法在 GIS 环境下生成^[2]。

空间插值的实质是通过已知样点的数据来估算未知点的数据。常用于气象要素的空间插值方法按照是否考虑海拔高度等相关因素,可以分为以下 2 类,第一类是不考虑这些因素的方法:有泰森多边形法、反距离权重法、样条函数法、趋势面分析和普通克里格法等。第二类是考虑相关因素的插值方法:多元回归技术、协克里格插值法、Lapse 比率和梯度距离反比法等^[3-7]。

模拟我国农牧交错带分布以及制作草原综合顺序类型图的过程,需对我国及周边地区的年均温度、年积温、年降水量等气象要素进行内插计算,研究使用反距离权重法、样条函数法和普通克里格法对我国及周边地区 2 114 个气象站点 1961—1990 年的年均温度、年积温和年降水量进行空间插值以得到全国气象要素的栅格空间数据库。

1 数据来源

气温和降水数据来自我国及周边 2 300 个气象站点,去掉异常值和缺值,实际参与空间插值的气象站点为 2 114 个,1961—1990 年各气象站点 30 年的均值。使用 ArcMap 软件将空间数据转换成 Albers 等积投影,投影参数为:第 1 条纬线, N 25°;第 2 条纬线, N 47°;中央经线, E 105°。图 1 为我国及周边气象站点的分布情况。

2 研究方法与验证

2.1 反距离权重法(IDW) 反距离权重法是对采样点进行线性的加权来决定输出的栅格值,加权与距离成反比,输入点离输出栅格越远,对输出栅格的影响越小。使用公式(1)表示^[8]:

$$Z = \frac{\sum_{i=1}^n \frac{1}{(d_i)^p} Z(x_i)}{\sum_{i=1}^n \frac{1}{(d_i)^p}} \quad (1)$$

其中, Z 为待估计的气温栅格值, $Z(x_i)$ 为第 i ($i = 1, 2, \dots, n$) 个气象站点的数据, n 为用于气

收稿日期:2007-12-03

基金项目:“863”计划数字农业技术课题(2006AA10Z241)

作者简介:马轩龙(1983-),男,甘肃陇南人,在读硕士生,

主要从事草业遥感和“3S”技术应用研究。

E-mail: mxlsw@yahoo.com.cn

通讯作者:陈全功 E-mail: chenqg@lzu.edu.cn

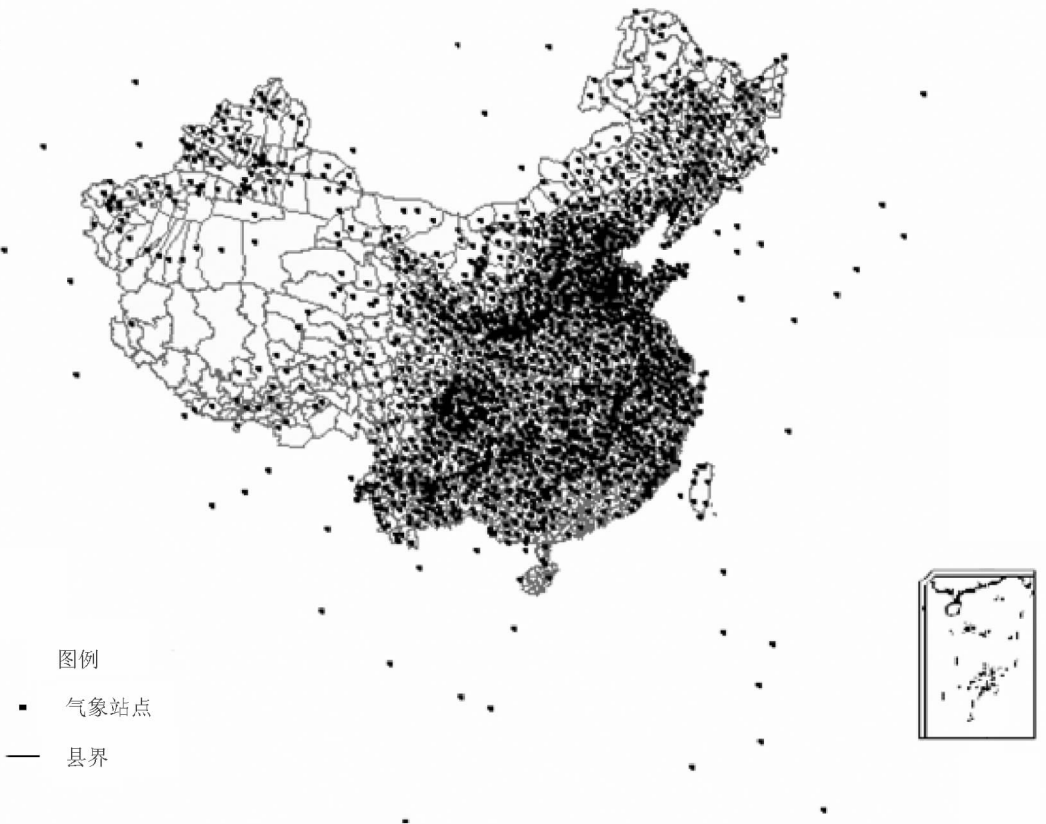


图1 我国及周边地区气象站点分布图

象数据插值的站点数目, d_i 为插值点到第 i 个气象站点的距离, p 为距离的幂。

2.2 样条函数法 (Spline) 样条函数法是使用一种数学函数, 对一些限定的点值, 通过控制估计方差, 利用一些特征节点, 用多项式拟合的方法来产生平滑的插值曲线。这种方法适用于逐渐变化的表面, 如温度、高程、地下水位高度或污染程度等。使用公式(2)表示:

$$Z = \sum_{i=1}^n A_i d_i^2 \ln d_i + a + bx + cy \quad (2)$$

其中, Z 为待估值的栅格值, d_i 为插值点到第 i 个气象站点的距离, $a + bx + cy$ 为气温的局部趋势函数, x, y 为插值点的地理坐标, $\sum_{i=1}^n A_i d_i^2 \ln d_i$ 为一个基础函数, 通过它可以获得最小化表面的曲率, A_i, a, b 和 c 为方程系数, n 为用于插值的气

象站点的数目。

样条函数法又分为张力样条函数法和规则样条函数法。由于规则样条函数法容易产生极值现象^[9], 本文选用张力样条函数法进行插值, 并在插值时对参数进行了优化。

2.3 普通克里格法 (OK) 普通克里格法是利用区域化变量的原始数据和变异函数的结构特点, 对未采样点的区域化变量的取值进行线性无偏最优估计的一种方法。使用公式表示为^[10]:

$$Z = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3)$$

其中, Z 为待估值的气象数据栅格值, λ_i 为赋予站点气象数据的一组权重系数, n 为用于气象数据插值的站点数目, $Z(x_i)$ 为站点的气象数据值。

为满足无偏性和最优性 2 个条件,通过建立如下克里格方程组来确定权重系数^[11]。

$$\begin{cases} \sum_{j=1}^n \lambda_j C(v_i, v_j) - \mu = C(v_i, V) \\ \sum_i \lambda_i = 1 \end{cases} \quad (4)$$

其中, $C(v_i, v_j)$ 为气象站点之间的协方差函数, $C(v_i, V)$ 为气象站点与插值点之间的协方差函数, μ 为拉格朗日乘数。

2.4 插值气象站点年均温度为负值时的处理方法

对于一些气象站点,年平均温度有可能出现负值。而插值的方法是将每个地区所有邻近站点的值都参与进插值过程。所以对年平均温度为负值的气象站点,有必要通过一定的数学处理,使之统一成正值,然后进行插值。确保插值结果的精度与可信度。经过对我国及周边地区气象站点原始数据的分析,发现年平均温度最低为 $-7.4\text{ }^{\circ}\text{C}$,所以对所有参与气象插值的气象站点年平均温度加上 8 以后,再进行插值。插值出的结果再减去 8。希望通过这样的处理,消除负值站点对插值整体精度的影响,处理方法使用数学公式(5)表示:

$$\begin{aligned} Z_i &= A_i + 8 \\ P &= Y - 8 \end{aligned} \quad (5)$$

其中, i 为参与插值的气象站点数目。 Z_i 为处理后的气象站点年均温度值, A_i 为气象站点原始年均温度值, P 为最终生成的栅格数据库每一点的年均温度值, Y 为经过插值后栅格数据库每一点的年均温度值。

2.5 不同气象站点数目的选择 对反距离权重法,普通克里格法和对照 3 种插值方法,各选择 3、5、10、15、20、25、30 个站点数据,分别对年均温、年积温和年降水量进行空间插值比较分析,从中选择最适合的站点数目。

2.6 检验标准 采用交叉验证法来验证插值的效果^[12]。即首先假定每一站点的气象要素值未知,都用周围站点的值来估算,然后计算所有站点实际观测值与估算值的误差,以此来评判插值方法的优劣。本文采用平均误差与均方根误差来

作为评估几种插值方法的插值效果的标准。选择平均误差和均方根误差相对最小的插值方法来建立全国气象数据的栅格空间数据库。平均误差反映总体估计误差的大小,均方根误差反映利用样点数据的估值灵敏度和极值效应。均方根误差表达式为:

$$y = \sqrt{\frac{\sum_{i=1}^n (Z_{a,i} - Z_{e,i})^2}{n}} \quad (9)$$

式中: y 为均方根误差, $Z_{a,i}$ 为第 i 个站点的实际观测值, $Z_{e,i}$ 为估计值, n 为用于检测的站点数目。

3 结果分析

3.1 插值气象站点年均温度为负值的处理结果

通过上述方法,将我国及周边地区气象站点年均温度统一为正值后再进行插值,与使用原始数据插值的结果进行对比发现,在不同参与插值气象站点数目下,均可以在一定程度上提高插值结果的精度,如表 1 所示。

3.2 不同插值站点数的验证结果

由表 2 可以看到反距离权重法对年均温度适宜空间插值的气象站点数为 20 个,平均误差为 $0.1287\text{ }^{\circ}\text{C}$; 年降水量适宜空间插值的气象站点数为 3 个,平均误差为 -5.667 mm ; 年积温适宜空间插值的气象站点数为 10 个,平均误差为 $-1.258\text{ }^{\circ}\text{C}$ 。样条函数法对年均温度适宜空间插值的气象站点数为 10 个,平均误差为 $-0.01466\text{ }^{\circ}\text{C}$; 年降水量适宜空间插值的气象站点数为 3 个,平均误差为 -3.399 mm ; 年积温适宜空间插值的气象站点数为 5 个,平均误差为 $-2.816\text{ }^{\circ}\text{C}$ 。普通克里格法对年均温度适宜空间插值的气象站点数为 30 个,平均误差为 $-0.01032\text{ }^{\circ}\text{C}$; 对年降水适宜空间插值的气象站点数为 30 个,平均误差为 -0.1860 mm ; 对年积温适宜空间插值的气象站点数为 20 个,平均误差为 $-0.1383\text{ }^{\circ}\text{C}$ 。从上述结果可知,不同插值方法对于适宜气象站点数目的要求是不同的,插值的站点数目不同,插值结果的精度不同。空间插值时,适宜数量的站点插值结果最优,并非是越多越好,此结果与李军龙等的研究相一致^[4]。

3.3 误差分析 由表 2 可知,对于年均温来说,ME 的排序为反距离权重法>样条函数法>普通克里格法,其值分别为 $-0.01287\text{ }^{\circ}\text{C}$, $-0.01466\text{ }^{\circ}\text{C}$, $-0.01032\text{ }^{\circ}\text{C}$;RMSE 的排序为反距离权重法>样条函数法>普通克里格法,其值分别为 $1.995\text{ }^{\circ}\text{C}$ > $1.959\text{ }^{\circ}\text{C}$ > $1.049\text{ }^{\circ}\text{C}$ 。普通克里格法具有最小的 ME 和 RMSE。对于年降水来说,ME 的排序为反距离权重法>样条函数法>普通克里格法,其值分别为 -5.667 , -3.399 , -0.1860 mm ,RMSE 的排序为,反距离权重法>样条函数法>普通克里格法,其值分别为 4.134 , 3.341 , 2.459 mm 。普通克里格法具有最小的 ME 和 RMSE。对于年积温来说,样条函数

法>反距离权重法>普通克里格法,其值分别为 $-2.816\text{ }^{\circ}\text{C}$, $-1.258\text{ }^{\circ}\text{C}$, $-0.1383\text{ }^{\circ}\text{C}$;RMSE 的排序为样条函数法>反距离权重法>普通克里格法,其值分别为 $6.614\text{ }^{\circ}\text{C}$ > $6.069\text{ }^{\circ}\text{C}$ > $4.464\text{ }^{\circ}\text{C}$ 。普通克里格法具有最小的 ME 和 RMSE。对 3 种气象要素的插值结果进行验证发现,普通克里格法均具有最好的插值精度。对年均温度和年降水量来说,样条函数法的插值精度优于反距离权重法,而对于年积温,反距离权重法的插值精度优于样条函数法。

3.4 空间插值结果 将插值结果转换为 $1\text{ km}\times 1\text{ km}$ 的我国年均温度、年降水量和年积温的栅格空间数据库,如图 2—图 4。

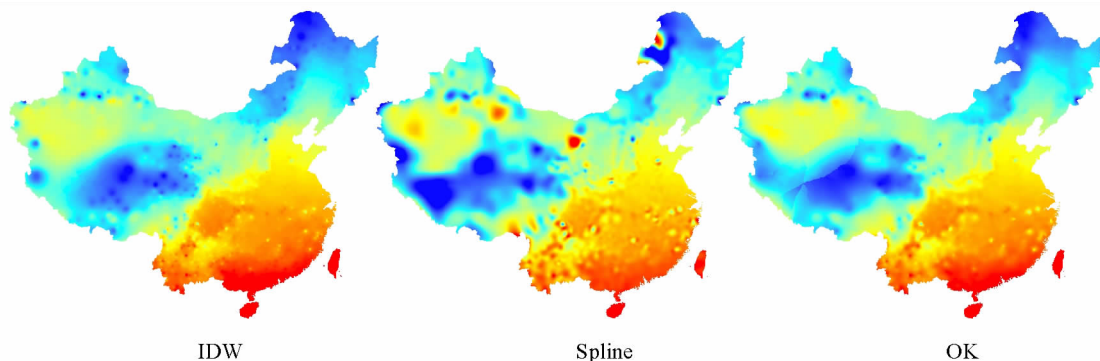


图 2 我国年均温度空间分布图

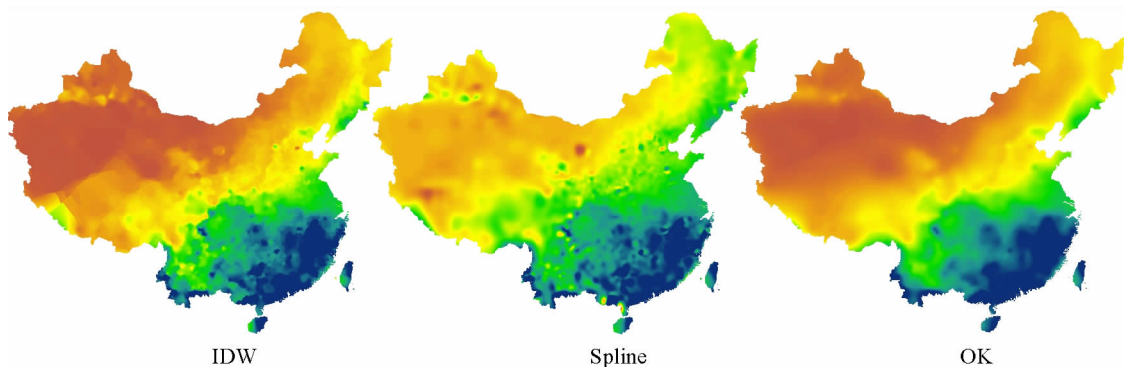


图 3 我国年降水量空间分布图

从图 2 可以看出,3 种插值方法均能较好地反映出年均温度的纬度地带性变化。在西部山区和河谷地区,普通克里格法和样条函数法较反距离权重法能更好地反映出温度随地形变化的趋势。普通克里格法和样条函数法插值出来的图较

平滑,在反映温度的变化趋势上,样条函数法的能力接近于普通克里格法,而反距离权重法生成的图上存在很多比较小的“牛眼”。在青藏高原地区,由于气象站点较少,3 种方法都很难反映出温度随地形和海拔的变化。并且无论采用哪一种插

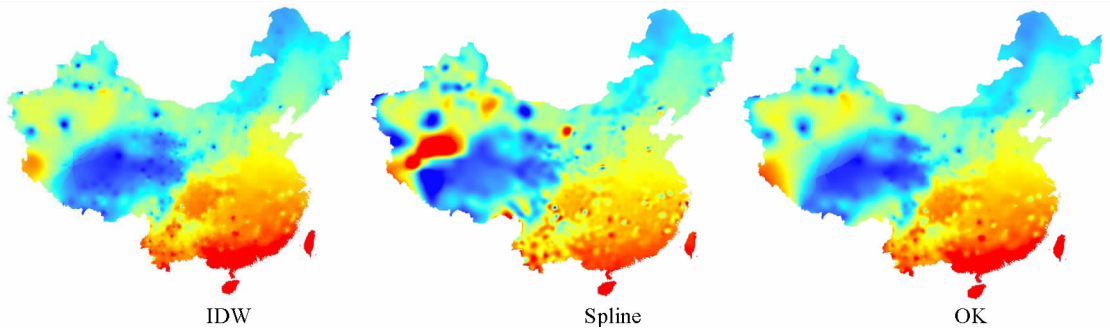


图4 我国年积温空间分布图

值方法,西部的温度差异都大于东部,这与西部气象站点数量少有关。

从图3可以看到,我国降水空间分布的趋势大致呈从东南向西北逐渐减少的趋势,东南沿海的年降水量达到了1939mm,而西北年降水量最低的地区只有24mm,东—西、南—北的空间差异非常显著。样条函数法和反距离权重法生成的图非常相似,在局部样条函数法较反距离权重法的层次更加清楚。相比较来说,普通克里格法生成的年降水量分布图更加接近于真实。

年积温是决定植物生长期从而决定了生态类型的重要因子^[13]。从图4可以看到我国年积温和年降水量的分布趋势非常一致,说明水热条件的分布大致是相同的。反距离权重法生成的图上具有许多非常明显的极值分布,样条函数法生成的图与反距离权重法比较相似,在西北地区,3种方法的插值结果都较差,没有反映出年积温随地形起伏的变化。

4 讨论和结论

1) 纵观我国年均温度、年降水量和年积温的空间分布图,可以发现一个明显的趋势,即水热条件的东南和西北分界非常明显,大致与胡焕庸线相一致,即东南水热条件较为充分,而西北水热条件较差。同时这条线也是我国很多地理、生态格局的分界线,譬如我国水灾呈现为东西分异线,与胡线相一致,此线以西水灾发生频率低,此线以东水灾发生频繁^[14]。根据殷培红等人的研究,我国的常年缺粮区主要分布集中在胡焕庸线以西地区,潜在缺粮区主要分布在此线以东地区^[15]。王铮等人的研究指出,我国存在一条平均走向为腾

冲—黑河方向的生态环境过渡带,东西两侧环境分异非常明显^[16]。根据张剑等人对中国农牧交错带的模拟,可以发现我国的农区基本上分布在这条“水热线”的东南,而我国的牧区则基本上处于这条线的西北,而农牧交错带正是位于这条线上。所以研究清楚这条线上的水热分布,对于我国生态建设以及国家宏观政策的制定,无疑具有非常重要的意义。

2) GIS软件对于空间数据强大的处理能力为地理学、生态学和社会学的研究提供了获取高精度、大范围数据的可能性,但是也应该看到空间插值的局限。从本质上来说,空间内插方法是研究区域变量空间分布的基本方法,各种方法都有其特定假设、适用范围、算法和优缺点^[17],目前常用的办法是对多种方法进行比较研究,然后根据插值精度选择其中精度最高的一种。对于不同的空间变量,在不同的地域和不同的时空尺度内所谓的“最优”内插法是相对的^[18]。选择合适的插值方法仍然要结合数据本身的特点和空间特性决定,并不是越复杂的方法就能产生越好的效果。

3) 通过本次研究发现,影响插值结果的因素很多,参数的选择、气象站点的分布和数目等都会影响到最后的插值结果。通过对原始数据进行一定的处理,可以在一定程度上提高最终插值结果的精度。对于年均温度和年降水量的模拟上,普通克里格法的效果相对最好,其次是样条函数法,反距离权重法最差。对于年积温的模拟上,普通克里格法效果最好,其次是反距离权重法,样条函数法最差。

参考文献

[1] 任继周,胡自治,牟新待,等. 草原的综合顺序分类

- 法及其草原发生学意义[J]. 中国草地学报, 1980, (1):12-24.
- [2] 崔庆虎, 蒋志刚, 刘季科, 等. 青藏高原草地退化原因述评[J]. 草业科学, 2007, 24(5):20-26.
- [3] 姚玉璧, 邓振镛, 尹东, 等. 黄河首曲流域草地生态与自然退化成因及对策研究[J]. 草业科学, 2007, 24(12):87-93.
- [4] 李军龙, 张剑, 张丛, 等. 气象要素空间插值方法的比较分析[J]. 草业科学, 2006, 23(8):6-11.
- [5] 李新, 程国栋, 卢玲. 空间内插方法比较[J]. 地球科学进展, 2000, 15(3):260-265.
- [6] Ashraf M, Loftis J C, Hubbard K G. Application of geostatics to evaluate partial station networks[J]. *Agric. For. Meteorol.*, 1997, 85:225-271.
- [7] Gottschalk L, Batchvarova B, Gryning S B, *et al.* Scale aggregation-comparison of flux estimates from XOPEX [J]. *Agric. For. Meteorol.*, 1999, 98:103-119.
- [8] Nalder I A, Wein R W. Spatial interpolation of climate normals: test of a new method in the Canadian boreal forest[J]. *Agric. For. Meteorol.*, 1998, 92:211-255.
- [9] Price D T, Mekenney D W, Nalder I A. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data[J]. *Agric. For. Meteorol.*, 2000, 101:81-94.
- [10] Jill M, Kevin J. Using ArcMap Spatial Analyst [M]. USA: Redlands, 2001, 136-148.
- [11] 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.
- [12] Kevin J, Jay M H, Konstantin K, *et al.* Using ArcMap Geostatistical Analyst[M]. USA: Redlands, 2001, 116-162.
- [13] Holdaway M R. Spatial modeling and interpolation of monthly temperature using Kriging[J]. *Clim. Res.*, 1996, 24:1835-1845.
- [14] 王铮, 张丕远, 刘啸雷, 等. 中国生态环境过渡的一个重要地带[J]. 生态学报, 1995, 15(3):319-326.
- [15] 王静爱, 王瑛, 黄晓霞, 等. 18 世纪中叶以来不同时段的中国水灾格局[J]. 自然灾害学报, 2001, 10(1):1-7.
- [16] 殷培红, 方修琦, 马玉玲, 等. 21 世纪初我国粮食供需的新空间格局[J]. 自然资源学报, 2006, 21(4):625-631.
- [17] 赵慧颖, 王广生, 魏学占. GIS 支持下的天然牧草产草量区域预报模型研究[J]. 草业学报, 2007, 16(4):100-106.
- [18] 方书敏, 钱正堂, 李远平. 甘肃省降水的空间内插方法比较[J]. 干旱区资源与环境, 2005, 19(3):47-50.
- [19] Haining R. *Spatial Data Analysis in the Social and Environmental Sciences* [M]. Cambridge: Cambridge University Press, 1990, 291-312.

Study on the method of GIS based spatial interpolation of climate factors in China

MA Xuan-long¹, LI Chun-e², CHEN Quan-gong¹

(1. College of Pastoral Agriculture Science and Technology, Lanzhou University, Key Laboratory of Grassland Agro-ecosystem, Ministry of Agriculture, Lanzhou 730020, China;

2. Center for Agriculture Remote Sensing of Shanxi Province, Taiyuan 030002, China)

Abstract: Annual average temperature data, annual rainfall and annual accumulated temperature data from 2114 meteorological stations in China and surrounding countries from 1961 to 1990 were interpolated by using inverse distance weighing (IDW), ordinary kriging (OK) and spline with ArcMap. Cross-validation was applied to evaluate the three interpolation methods. The result indicated that the precision of interpolation results were varied significantly along with the number of selected meteorological stations for interpolating. For the three climate factors, ordinary Kriging shown the best precision. For annual average temperature and annual rainfall, spline shown better result than IDW, but for annual accumulated temperature, IDW was better than spline. The research indicated that the spatial distribution of moisture and temperature in China shown an obvious west-east, south-north line as same with the HU population line. In the west of this line, the moisture and temperature conditions were poor and in the east of this line, the conditions were good.

Key words: climate factor; spatial interpolation; inverse distance weighing; spline; ordinary kriging