

基于 GIS 和地统计学的作物需水量等值线图

王景雷, 孙景生, 张寄阳, 肖俊夫

(水利部农田灌溉研究所, 新乡 453003)

摘 要: 利用“全国灌溉试验资料数据库”中作物需水量试验数据, 在 ERSI 的 ARCGIS 和微软电子表格 EXCEL 支持下, 将 GIS 的空间数据管理功能和地统计学的空间分析功能有机结合, 用多种插值方法进行山东省冬小麦需水量等值线图的绘制, 结果表明: 地统计学方法优于常规的反距离加权方法、多项式方法和径向基函数法; 地统计方法中的普通克里金 (Ordinary Kriging) 球状模型的预测误差的均值、预测误差的均方根、平均预测标准差、平均标准差、标准均方根预测误差等 5 个指标均优于普通克里金指数模型、简单克里金 (Simple Kriging) 球状模型和简单克里金指数模型, 具有较高的插值精度。

关键词: GIS; 地统计学; 作物需水量; 等值线图

中图分类号: S274.1

文献标识码: A

文章编号: 1002-6819(2004)05-0051-04

0 引言

作物需水量是水土资源平衡计算、灌溉工程设计与管理中不可缺少的基本数据, 同时也是制定科学合理的灌溉制度的基础。但是作物需水量受气象、土壤及作物等因素的影响, 在空间上往往表现出很大差异^[1]。通过试验观测只能得到个别点上的需水量值, 而定点观测到的数据大多不能直接用于其它点上, 更不能代替某一较大面积上的平均值^[2]。但是人们又不能无限制地增设试验站点来满足这种要求, 唯一可行的途径就是合理地布置观测站点, 然后由这些站点资料提供的信息使用估值方法来确定众多的非测点或某个区域的需水量值^[2]。

地统计学方法是以反映变量空间结构特征的结构函数 (如半变异函数) 为基础, 以取得估计方差最小为目标, 在无偏性约束条件下求优的一种估计方法^[3,4]。与经典的插值方法不同, 地统计学考虑样点的方向、位置和彼此间的距离, 直接测定和分析空间依赖性, 用于研究有一定随机性和有一定结构性的各种变量的空间分布规律。近年来, 地统计学不仅广泛用于土壤、水土保持等项研究中, 为研究土壤水、盐以及泥沙的空间迁移、扩散和分布关系等提供了强有力的工具^[5-8], 而且在作物需水量的估值和空间结构分析方面也得到了初步应用。对作物需水量这种具有空间属性的变量进行分析和估算, 往往要面对行政区域图、地形图等空间图形数据, 以及降雨量、温度、湿度、地下水位等属性数据, 用人工方法对这些繁杂的数据作统计、分析的工作量较大且有一定的难度, 成图时也难以把这些属性数据都表示在相应的地图上, 另外地统计学通常要求均匀取样^[5], 这给大区域范围的地统计学研究带来一定的困难, 这也是地统计学在大尺度上应用较少的主要原因。

GIS 可以把空间和属性信息有机地结合起来, 通过

地理数据能够比较容易地确定样本点之间的距离, 进而根据同性数据可以计算出变量之间的差异, 从而得到地统计学所需要的步长和半方差的函数关系, 使分析大尺度的作物需水量的变异规律变得较为方便^[9-12]。事实上, GIS 虽然具有较强的空间数据管理功能, 但缺乏对一些问题的空间分析能力; 地统计学则具有较强的空间分析功能, 但其空间数据管理功能较弱; 二者的结合为地统计学在大尺度作物需水量估值中的广泛应用提供了技术支持和条件。

1 主要研究方法

1.1 数据基础

本项研究充分利用国内已有的研究成果, 基础属性数据主要来自“全国主要作物需水量等值线图研究”和“全国灌溉试验资料整编”两个课题的研究成果, 并利用“八五”和“九五”期间国内取得的有关研究成果 (如南水北调局资助的黄河流域主要作物需水量研究等) 对上述两项成果的数据进行充实和订正。空间图形数据主要参照了 1:400 万的全国行政区域图、水系图和等高线图等 (中国资源与环境数据光盘, 中国科学院地理研究所资料与环境信息系统国家重点实验室 1996 年 6 月出版)。

1.2 软件工具

本项研究采用美国环境系统研究所 (ERSI) 的 GIS 桌面平台系统 ARCGIS 8.3 以及目前常用的微软电子表格工具 EXCEL 2000。ARCMAP 主要用于空间数据显示、编辑、查询、检索、统计报表生成、空间分析和高级制图等, ArcCatalog 主要用于空间数据的浏览, Geodatabase 结构定义、数据导入、导出、网络模型生成、对象关系和规则的定义、元数据的定义和编辑修改等, ArcToolbox 主要用于空间数据格式转换、叠加处理、缓冲区生成和坐标转换等。电子表格工具 EXCEL 2000 主要用于原始数据的输入和数据的简单计算等。

1.3 数据处理转换

上述研究成果多以微软电子表格 xls 文件形式保存, 由于目前 ARCGIS 尚不能很好地支持 xls 文件格式

收稿日期: 2003-10-27 修订日期: 2004-05-20

基金项目: 国家“863”计划 (2002AA2Z4031) 及水利部创新基金项目 (SCX2002-22) 联合资助

作者简介: 王景雷, 助理研究员, 主要从事节水灌溉及地下水环境研究。河南省新乡市 水利部农田灌溉研究所, 453003

(必须加载 EXCEL-TOOLS 工具才能实现), 因此首先把含有各灌溉试验站点编号、经纬度和不同水文年的作物需水量的 xls 文件存为 dbf 或 txt 文件格式, 然后通过 ARCTOOLBOX 中的转换工具将此文件生成全国或各省的灌溉试验站点图。为了与其它图件(如全国或省的行政区域图)进行叠加, 必须进行坐标变换。

1.4 研究方法

研究方法依据文献[3, 5]。
作物需水量具有空间分布特征, 既有随机性又有结构性, 取样后不可能再次取到同样的样品, 样品间具有空间相关性, 所以将依空间位置 x 而变化的作物需水量 $z(x)$, 解释为随机变量 $z(x)$ 的一次具体实现。设位于 x_0 处的作物需水量估计值为 z_0^* , 它是周围若干站点实测需水量 $z(x_i)$, ($i = 1, 2, \dots, n$) 的线性组合, 即

$$z_0^* = \sum_{i=1}^n \lambda_i z(x_i) \tag{1}$$

式中 λ ——权重系数; z_0^* —— x_0 处的作物需水量估计值; i ——试验站点编号。

只要 $\sum_{i=1}^n \lambda_i = 1$, 就能满足无偏条件, 估值问题就转化为在约束条件下求极值问题。

目标函数 $\min(\sigma_k^2) = \min\{E[z_0 - z_0^*]^2\}$
约束条件 $\sum_{i=1}^n \lambda_i = 1 \tag{2}$

式中 σ_k^2 ——估计方差; E ——数学期望。
根据空间变异分析理论, 可将估计方差 σ_k^2 写成半变异函数 $\gamma(h)$ 的函数

$$\sigma_k^2 = 2 \sum_{i=1}^n \lambda_i \cdot \gamma(x_i - x_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \cdot \lambda_j \cdot \gamma(x_i - x_j) \tag{3}$$

式中 $\gamma(x_i - x_0)$, $\gamma(x_i - x_j)$ ——分别表示 x_i 与 x_0 , x_i 与 x_j 之间的半变异函数值。

为了求估计方差的极小值, 引入拉格朗日函数 S

$$S = 2 \sum_{i=1}^n \lambda_i \cdot \gamma(x_i - x_0) - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \cdot \lambda_j \cdot \gamma(x_i - x_j) - 2u(\sum_{i=1}^n \lambda_i - 1) \tag{4}$$

要使估计方差最小, 必须满足

$$\frac{\partial S}{\partial \lambda_i} = 0 \quad i = 1, 2, \dots, n \tag{5}$$

从(5)式可得如下的方程组

$$\begin{cases} \sum_{i=1}^n \lambda_i \cdot \gamma(x_i - x_0) + u = \gamma(x_j - x_0) & j = 1, 2, \dots, n \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \tag{6}$$

也即

$$\begin{bmatrix} \gamma_{11} & \dots & \gamma_{1n} & 1 \\ \dots & \dots & \dots & \dots \\ \gamma_{n1} & \dots & \gamma_{nn} & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \\ u \end{bmatrix} = \begin{bmatrix} \lambda_{01} \\ \dots \\ \lambda_{0n} \\ 1 \end{bmatrix}$$

式中 $\gamma_{ij} = \gamma(x_i - x_j)$, 表示 x_i 和 x_j 之间的变异函数值; u ——拉格朗日值。

解上述方程组即可得到所有的权重 $\lambda_1, \dots, \lambda_n$ 和拉格朗日值 u 。利用计算所得到的权重拉格朗日值, 不仅可通过(1)式求得估计值 z_0^* , 还可利用(7)式计算 Kriging 估计方差。

$$\sigma_k^2 = \sum_{i=1}^n \lambda_i \gamma(x_0 - x_i) + u \tag{7}$$

上述介绍的只是常用的普通克里金 (Ordinary Kriging) 方法, 经过几十年的进一步发展, 还形成了指示克里金 (Indicator Kriging)、概率克里金 (Probability Kriging)、泛克里金 (Universal Kriging) 等, 基本原理同普通克里金方法, 但使用条件有所不同。

2 等值线的生成和结果分析

虽然全国 30 多个省、自治区、直辖市对于不同的作物所选用的插值方法不同, 但等值线的生成过程及步骤大体相同, 下面以山东省冬小麦需水量等值线图的绘制过程为例简述其过程。

2.1 数据预处理

克里金方法要求样本数据服从正态分布, 因此首先需要对样本数据进行预处理和相关检验。平稳性检验选用相关分析方法来检验作物需水量值在空间分布的平稳性。若某一作物的需水量与空间坐标 (x, y) 显著相关, 则认为该作物需水量值存在漂移现象, 即为非平稳的; 否则即为平稳。平稳性检验通过后还必须进行分布类型检验, 此处采用偏度峰度检验法。

偏度 $R = m_3/S^3 \tag{8}$

峰值 $P = m_4/S^4 - 3 \tag{9}$

式中 标准方差 $S = \sqrt{\frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2}$;

$m_3 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^3$ (样本 3 阶中心矩);

$m_4 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^4$ (样本 4 阶中心矩)。

如果 $|R| < 2\sqrt{6/n}$ 且 $|P| < 2\sqrt{24/n}$, 则作物需水量值样本可认为是正态分布; 当为非正态分布时, 可将其取对数后再计算偏度和峰值; 若满足上述条件, 则为对数正态分布, 否则为其它分布。

山东省冬小麦需水量的样本数据共有 30 个, 基本的统计分析结果见表 1。

表 1 冬小麦需水量统计分析结果

Table 1 Statistical results of the winter wheat water requirements

样本数目	最大值/mm	最小值/mm	平均值/mm	中值/mm	标准方差	偏度	峰值	$2 \times \sqrt{6/n}$	$2 \times \sqrt{24/n}$
30	576.32	414.05	475.72	485.11	43.125	0.22256	-0.6105	80.49	1.788

从表 1 可看出, 该样本数据基本满足正态分布, 可以直接运用克里金方法进行插值分析。

2 2 不同插值方法比较

本项研究采用简单克里金、普通克里金两种克里金方法进行插值, 每种方法分别采用球状模型和指数模型进行样本内插, 为了与传统插值方法进行比较, 还同时运用了反距离加权插值、多项式插值和径向基函数插值方法。

2 2 1 常规方法与地统计学方法(克里金)的比较

将以上插值方法所得拟合值与相应测定值进行比较, 计算误差均值(ME)和误差均方根(RMS)。结果见表 2。

表 2 不同插值方法检验比较结果

Table 2 Results comparison of different interpolation methods

插值方法		参数模型	误差均值 (ME)	误差均方 根(RMS)
常规 方法	反距离加权插值	指数为 7.3957	- 2.744	21.7
	多项式插值	指数为 2	- 0.9149	29.78
	径向基函数插值	完全规则样条函数参数为 0.000042905	- 1.49	21.24
克里金 方法	简单克里金插值	球状模型	- 0.2022	15.82
		指数模型	- 0.7478	19.76
	普通克里金插值	球状模型	- 0.7117	15.53
		指数模型	- 0.965	18.15

一般来说, 各种插值方法的ME和RMS总体最小者, 具有较好的插值效果。从表 2 可知, 地统计学方法优于传统的确定性内插法。

2 2 2 不同克里金插值方法的比较

不同克里金插值方法的模型特征参数见表 3。

表 3 不同方法的理论模型及参数

Table 3 Theoretical models and parameters of different methods

方法		块金值	基台值	最大变程方向	最大变程/km	最小变程/km	相异性比
普通	球状	0	2243.8	73.5	527.52	201.84	2.6136
克里金	指数	0	2239.5	73.0	527.52	248.37	2.1239
简单	球状	0	2226.6	73.8	487.86	207.63	2.3497
克里金	指数	0	2490.1	67.2	487.86	121.4	4.0186

从表 3 可看出, 半变异函数的“块金效应”均很小, 块金值均为 0, 说明它对于分析冬小麦需水量的空间变异几乎不起作用。从表中还可看出, 各种方法的最小变程都在 200 km 左右, 出现这种较大范围的空间结构性, 主要归因于气象因子(如湿度、风力、温度等)在区域上的连续性。此外也与该省地势相对较平坦有关。不论哪种方法, 样本数据都呈现出明显的各向异性, 从大方向来说, 在东偏北-西偏南方向变程较大, 说明这个方向的变异较小, 而在其相垂直的方向即北偏西-南偏东方向的变异较小, 初步说明山东冬小麦需水量受纬度影响较大。

在地统计学方法中通常采用下面 5 个指标来评价其预测精度, 即: 预测误差的均值(Mean); 预测误差的

均方根(Root-Mean-Square)、平均预测标准差(Average Standard Error)、平均标准差(Mean Standardized)、标准均方根预测误差(Root-Mean-Square Standardized)。其中前 4 个指标越小越好, 第 5 个指标越大越好。

表 4 不同克里金方法的检验比较结果

Table 4 Comparison between the Ordinary Kriging and the Simple Kriging

项 目	普通克里金		简单克里金	
	球状	指数	球状	指数
误差均值	- 0.7117	- 0.965	- 0.2022	- 0.7478
误差均方根	15.53	18.15	15.82	19.76
平均预测标准差	24.62	30.8	24.72	37.31
平均标准差	- 0.001979	- 0.008848	0.01507	- 0.02383
标准均方根预测误差	0.5619	0.5261	0.5723	0.4706

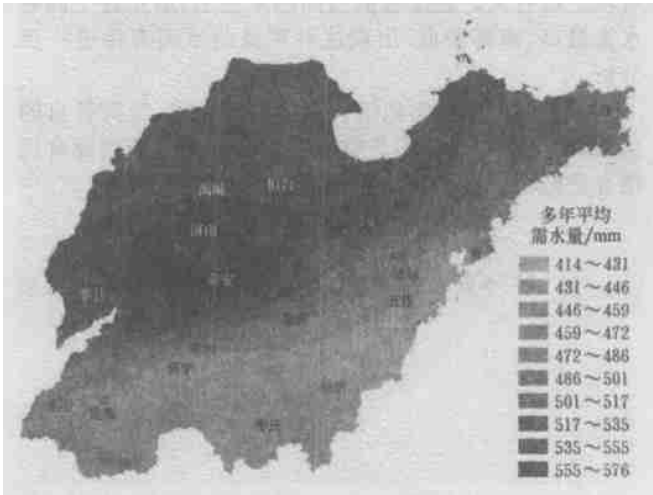


图 1 普通克里金法山东冬小麦多年平均需水量等值线图(球状模型)

Fig. 1 Yearly annual mean water requirements isoline of winter wheat with Ordinary Kriging in Shandong(spherical model)

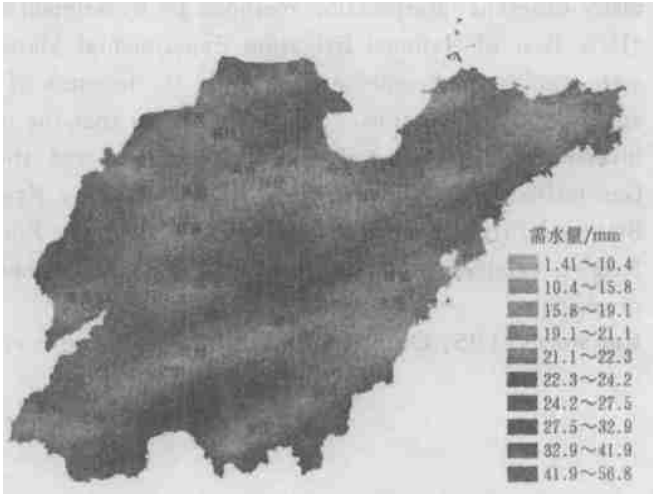


图 2 普通克里金法山东冬小麦多年平均需水量预测标准误差等值线图(球状模型)

Fig. 2 Prediction standard error isoline of yearly annual mean water requirements of the winter wheat with Ordinary Kriging in Shandong(spherical model)

从表 4 可看出, 两种克里金方法都是球状模型优于指数模型。但普通克里金的估值精度高于简单克里金方法。普通克里金球状模型的插值结果见图 1, 预测标准误差等值线图见图 2。

3 结 论

通过对山东省冬小麦需水量的数据分析, 冬小麦需水量基本符合正态分布, 克里金方法估值精度高于传统的插值方法, 球状普通克里金方法较简单克里金方法具有较高的插值精度; 研究结果表明: 冬小麦需水量具有明显的各向异性, 在东偏北-西偏南方向变程较大, 而在其相垂直的方向即北偏西-南偏东方向的变异较小, 空间自相关距离在 201.84~527.52 km, 出现这种较大范围的空间结构性, 主要归因于气象因子(如湿度、风力、温度等)在区域上的连续性, 此外可能与该省地势相对较平坦有关。克里金插值图结果显示: 山东省中部需水量最高, 南部最低, 出现这种现象的原因有待进一步分析。

GIS 和地统计学相结合能够取长补短, 发挥各自的优势, 研究结果也证明将空间数据管理和分析功能有机结合进行大尺度的作物需水量的估值是可行的。

[参 考 文 献]

[1] 袁 新, 李恩羊 参照作物需水量的空间变异性[J]. 水利

学报, 1990, (2): 33- 37.

- [2] 李恩羊 地质统计学在作物需水量研究中的应用[J]. 武汉水利电力学院学报, 1991, (1): 1- 9.
- [3] 李恩羊, 袁 新 作物需水量的最优估计[J]. 水利学报, 1989, (10): 45- 49.
- [4] 崔远来, 茆 智, 李远华 作物水分敏感指标空间变异规律及其等值线图研究[J]. 中国农村水利水电, 1999, (11): 16 - 17.
- [5] 赵 斌, 蔡庆华 地统计学分析方法在水生态系统研究中的应用[J]. 水生生物学报, 2000, (5): 514- 520.
- [6] 自由路, 金继运, 杨俐萍, 等 基于 GIS 的土壤养分分区管理模型研究[J]. 中国农业科学, 2001, (1): 46- 50.
- [7] 温广玉, 侯锡铭, 陈华豪 用地统计学方法内插气象台站资料预报林火发生[J]. 东北林业大学学报, 2002, (4): 19- 21.
- [8] 王晓春, 韩士杰, 邹春静, 等 长白山岳桦种群格局的地统计学分析[J]. 应用生态学报, 2002, (7): 781- 784.
- [9] 史 舟, 王人潮, 吴宏海 基于 GIS 的山区年均温分布模拟与制图[J]. 山地研究, 1997, (15): 264- 268.
- [10] 韦红波, 任红玉, 杨勤科 中国多年平均输沙模数的研究[J]. 泥沙研究, 2003, (1): 39- 45.
- [11] 李正泉, 于贵瑞, 刘新安 东北地区降水与湿度气候资料的栅格化技术[J]. 资源科学, 2003, (1): 73- 77.
- [12] 牛振国, 李保国, 张凤荣, 等 参考作物蒸散量的分布式模型[J]. 水科学进展, 2002, (3): 303- 307.

Crop water requirement isoline based on GIS and geostatistics

Wang Jinglei, Sun Jingsheng, Zhang Jiyang, Xiao Junfu

(Fam land Irrigation Research Institute, Ministry of Water Resources, Xinxiang 453003, China)

Abstract The water requirement isogram of the winter wheat in Shandong province could be mapped through many different interpolation methods by experimental data of the crop water requirement, which comes from "Data Base of National Irrigation Experimental Materials". The EXCELL of Microsoft and ARCGIS of ERIS were employed to combine organically the function of the space data management in GIS with that of the space analysis in Geostatistics. The result shows that the method of the Geostatistics is better than the conventional inverse distance weighting, the polynomial and the radial basis functions. That the five indexes in the Geostatistics, namely, the mean of the Ordinary Kriging spherical model, the Root-Mean-Square the Average Standard Error, the Mean Standardized, and the Root-Mean-Square Standardized are superior to the Ordinary Kriging exponential model, the simple Kriging spherical model and the Kriging exponential model with high precision.

Key words: GIS; Geostatistics; crop water requirement; isoline