

文章编号: 1672—7940(2007)01—0052—06

Surfer 8.0 等值线绘制中的十二种插值方法

陈欢欢¹, 李星², 丁文秀¹

(1. 中国地质大学资源学院, 武汉 430074;

2. 中国地质大学数理学院, 武汉 430074)

摘要: 美国 GOLDEN 软件公司的 Surfer 软件是一套在 Windows 操作环境下运行的二维和三维图形绘制软件, 可以利用它绘制等值线图、3D 立体图、阴影地貌图、矢量图等。目前, 该软件的最高版本是 Surfer 8.0, 其中提供了包括克里格法、距离倒数加权法、多元回归等十二种插值方法, 可以满足不同应用的插值需求。本文主要介绍了该软件中的十二种插值方法, 并以月球表面铀的属性数据为例进行了简要分析。

关键词: Surfer 软件; 等值线; 插值

中图分类号: P631

文献标识码: A

收稿日期: 2006—11—30

Twelve Kinds of Gridding Methods of Surfer 8.0 in Isoline Drawing

Chen Huanhuan¹, Li Xing², Ding Wenxiu¹

(1. The Faculty of Earth Resources, China University of Geosciences, Wuhan 430074, China;

2. Maths and Physics Department, China University of Geosciences, Wuhan 430074, China)

Abstract: The Surfer software made by American GOLDEN software company is a two & three-dimensional graphic drawing software, which can be run in windows operating environment. It can be used for drawing isoline map, 3D stereogram, shadow geomorphological map, vectorgraph and so on. The Surfer 8.0 is the highest edition at present. It provides 12 kinds of interpolation methods, including Kriging, inverse distance to a power, polynomial regression, which can fulfill interpolation requirement of different application direction. This article introduces 12 kinds of interpolation methods in Surfer 8.0, and using the attribute data of U which is from Moon surface to make a simple analysis.

Key words: surfer; isoline; gridding

1 引言

在地质工作中,总是可以得到大量研究对象的各方面特性或特征数据,在对其进行分析时,往

往借助于其等值线来对其进行一些必要的分析。要将大量数据进行处理并以等值线的形式呈现,这就需要一套能方便绘制等值线的绘图工具。在众多的商业化绘图软件中,美国 GOLDEN 软件公司的 Surfer 软件,以其方便、直观、快捷、安装

作者简介: 陈欢欢(1982—),女,中国地质大学(武汉)资源学院硕士研究生,主要从事地学信息研究。E-mail: cleverhuan@163.com

李星(1962—),男,博士,教授,主要从事数学地质、盆地古地热场及有机质演化动态模拟等教学及研究工作。

E-mail: lixing@cug.edu.cn

简单、对系统要求低等优点得到广大用户的青睐,成为普及度最高的绘图软件之一。其最高版本 Surfer8.0 提供了十二种插值方法,用户可以根据不同的需要选择不同方法来进行插值,来对其进行分析,以达到自己想要的效果。要科学地选择插值方法和灵活地进行参数设置,必须要熟悉各种插值方法的基本理论知识,下面将介绍 Surfer8.0 中的十二种插值方法及其应用实例。

2 Surfer8.0 中十二种插值介绍

2.1 反距离加权插值法 (Inverse Distance to a Power)

反距离加权插值法是 20 世纪 60 年代末提出的计算区域平均降水量的一种方法^[1]。它实际上是一种加权移动平均方法,设平面上分布一系列离散点 $P(x, y, z)$, 已知其位置坐标 $P(x_i, y_i)$ 和属性值 $z_i (i=1, 2, \dots, n)$, 根据周围离散点的属性值,通过距离加权插值求 P 点属性值。其插值原理是待插值点邻域内已知散乱点属性值的加权平均,权的大小与待插值的邻域内散乱点之间的距离有关,是距离 $k (0 \leq k \leq 2)$ (k 一般取 2) 次方的倒数。即:

$$z_p = \frac{\sum_{i=1}^n \frac{z_i}{[d_i(x, y)]^k}}{\sum_{i=1}^n \frac{1}{[d_i(x, y)]^k}}$$

其中, $d_i(x, y) = \sqrt{(x-x_i)^2 + (y-y_i)^2}$, 表示由离散点 (x_i, y_i) 至 $P(x, y)$ 点的距离。这种方法的优点是可以通过权重调整空间插值等值线的结构。

2.2 克里金插值法 (Kriging)

克里金插值法,又称克立格法是法国 G·马特隆教授以南非矿山地质工程师 D·G·克立格的名字命名的一种方法^[2]。它是以区域化变量理论为基础,以变差函数为主要工具,在保证估计值满足无偏性条件和最小方差条件的前提下求得估计值。设区域化变量 $f(x)$ 满足二阶平稳假设或本征假设,则待插点 P 的估计值为 $f_p = \sum_{i=1}^n w_i f_i$, 其中 f_i 是 n 个已知点的函数值, w_i 是 n 个已知点的全系数。由无偏的条件,有: $\sum_{i=1}^n w_i = 1$ 。再根据估计的方差最小的条件, $\sum_{i=1}^n w_i \gamma(x_j - x_i) + \mu = \gamma(x_p - x_i) (i=1, 2, \dots, n)$ 。其中, μ 为拉格朗日算子, $\gamma(x_j - x_i)$ 为已知点间的变差函数值 γ

$(x_p - x_i)$ 为已知点与待插点间的变差函数值。求出待插点 P 的估计值。简单地说,克里金插值法就是一种特定的滑动加权平均法。

2.3 最小曲率法 (Minimum Curvature)

最小曲率法广泛应用于地球科学,是构造出具有最小曲率的曲面,使其穿过空间场的每一点,并尽可能使曲面变得光滑^[3]。使用最小曲率法时要涉及到两个参数:最大偏差参数 (Maximum Residuals) 和最大循环次数 (Maximum Iteration parameter) 参数来控制最小曲率的收敛标准,而且最小曲率法要求至少有四个点。最小曲率法试图在尽可能严格地尊重数据的同时,生成尽可能圆滑的曲面。最小曲率法主要考虑曲面的光滑性,因此插值的成果容易失真,往往超出了最大值和最小值的范畴,由此绘出的等值线与实际相差较大。实际应用中此法只能作为平滑估值,绘出的降水量等值线主要用于定性研究降水的空间分布及走向。

2.4 改进谢别德法 (Modified Shepard's Method)

使用反距离加权插值法,当增加、删除或改变一个点时,需要重新计算权函数 $w_i(x, y)$, 为了克服反距离加权插值法的这一缺陷,改进谢别德法同样使用距离倒数加权的最小二乘方的方法,主要有以下两个方面的改进^[3,4]:

1) 通过修改反距离加权插值法的权函数 $w_i(x, y) = 1/[d_i(x, y)]^k$, 使其只能在局部范围内起作用,以改变反距离加权插值法的全局插值性质。

2) 同时引用节点函数 $Q_i(x, y)$ (插值于 (x_i, y_i) 点的二次多项式) 来代替离散点 (x_i, y_i) 的属性值 z_i 。 $Q_i(x, y)$ 在点 (x_i, y_i) 附近与函数属性值 $z(x, y)$ 具有局部近似的性质,因此,如果认为距离 (x_i, y_i) 较远的点对 $Q_i(x, y)$ 影响不大,则可以认为在 (x_i, y_i) 点附近, $Q_i(x, y)$ 就可以近似地表示函数属性值 $z(x, y)$ 了。

2.5 自然邻点插值法 (Natural Neighbor)

自然邻点插值是基于 Voronoi 结构的一类插值方法。自然邻点 Laplace 插值由 Belikov 等以及 Sugihara 和 Hiyoshi 等先后从不同角度提出的。对于二维空间,自然邻点插值型函数(基函数)的形式为^[5]:

$$\begin{aligned} \varphi_i(x) &= \alpha_i(x) / \sum_{j=1}^n \alpha_j(x) \\ \alpha_j(x) &= s_j(x) / h_j(x) \end{aligned}$$

其中 $s_i(x)$ 是与节点 i 关联的 Voronoi 边的长度, $h_i(x)$ 是插值点 x 到节点 i 的 Voronoi 边的垂直距离, n 为插值点的自然邻点个数。

自然邻点 Laplace 插值型函数满足在插值节点等于 1、单位分解性和线性完备性等插值型函数的基本性质。

2.6 最近邻点插值法 (Nearest Neighbor)

最近邻点插值法又称泰森多边形方法, 是荷兰气象学家 A. H. Thiessen 提出的一种方法, 最初用于根据离散分布的气象站的降水来计算平均降水, 而现在的 GIS 和地理分析中也经常用其进行快速的赋值^[3,6]。实际上, 最近邻点插值的一个隐含的假设条件是任一网格点 $P(x, y)$ 的属性值都使用距它最近的位置点的属性值, 用每一个网格节点的最邻点值作为待的节点值, 具体算法如下:

假若有一块总面积为 s 的区域, 其中共有 N 个网格点 $n(n=1, 2, \dots, N)$ 。各个网格点的属性值分别为 $q_i(i=1, 2, \dots, N)$ 。分别以各网格点之间连线的垂直平分线, 把流域划分为若干个多边形, 然后以各个多边形的面积 $s_i(i=1, 2, \dots, N)$ 为权系数, 计算各网格点的加权平均值, 并将其作为区域的平均属性值:

$$Q = \frac{\sum_{i=1}^N s_i \cdot q_i}{\sum_{i=1}^N s_i}$$

2.7 多元回归法 (Polynomial Regression)

多元回归是用来确定数据的大规模的趋势和图案, 它只是根据空间的采样数据, 拟合一个数学曲面, 用该数学曲面来反映空间分布的变化情况, 实际上是一个趋势面分析作图程序。

设随机变量 y 随 n 个自变量 x_1, x_2, \dots, x_n 变化, 并有 m 组观测数据 $x_{1i}, x_{2i}, \dots, x_{mi}, y_i, (i=1, 2, \dots, m)$, 确定随机变量 y 与各个自变量 x_1, x_2, \dots, x_n 是否存在相关关系, 若存在则给出合适的线性组合或可以化成线性组合的关系式, 这就是多元回归分析要解决的主要问题之一。假设 $y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$, 可用最小二乘法来确定 $b_0, b_1, b_2, \dots, b_n$ 的值^[7]。

对所给出的一组数据拟合一个线性模型之后, 为了对拟合的合适程度作出评价, 需要进行方差分析。利用统计量 F 值对整个回归过程进行显著性检验:

$$F = \frac{U/n}{Q/(m-n-1)}$$

$$= \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2 / n}{\sum_{i=1}^m (\hat{y}_i - \bar{y})^2 / (m-n-1)}$$

式中 U 是回归平方和, Q 是残差平方和, \hat{y}_i 是与观测值相对应的计算值; \bar{y} 为均值。

给定检验水平 α , 当 F 值大于临界值 F_α 时, 则 x_1, x_2, \dots, x_n 对 y 有显著的影响, 并且这种影响是线性的, 亦即回归方程有实际意义; 否则, 回归方程无意义。

2.8 径向基函数法 (Radial Basis Function)

径向基函数, 又称距离基函数, 它是多个数据插值方法的组合, 其基函数是由单个变量的函数构成的。一个点 (x, y) 的这种基函数的形式往往是 $h_i(x, y) = h(d_i)$, 这里的 d_i 表示由点 (x, y) 到第 i 个数据点的距离。所有径向基函数插值法都是准确的插值器, 它们都能尽量适应的数据。若要生成一个更圆滑的曲面, 对所有这些方法都可以引入一个圆滑系数^[3]。

函数类型: 最基本的函数类似于克里金中的方差图。当对于一个网格点插值时, 这些函数为数据点规定了一套最佳权重。

基函数类型有:

倒转复二次函数 (Inverse Multiquadric):

$$B(h) = 1/\sqrt{h^2 + R^2};$$

复对数 (Multilog):

$$B(h) = \log(h^2 + R^2);$$

复二次函数 (Multiquadratic):

$$B(h) = \sqrt{(h^2 + R^2)};$$

自然三次样条函数 (Natural Cubic Spline):

$$B(h) = (h^2 + R^2)^{3/2};$$

薄板样条法函数 (Thin Plate Spline):

$$B(h) = (h^2 + R^2) \log(h^2 + R^2).$$

式中 h 为表示由点 (x, y) 到第 i 个数据点的距离; R 参数是平滑因子。

其中的复二次函数方法在水文测量、大地测量、地质及采矿、地球物理等领域都得到了广泛应用, 效果良好, 在数据点数量不太大的情况下计算也不太复杂。薄板样条法实质是使插值函数所代表的弹性薄板受限于插值点, 并且具有最小的弯曲能量。

2.9 线性插值三角网法 (Triangulation with Linear Interpolation)

线性插值三角网法是使用最佳的 Delaunay

三角形, 将连接数据点间的连线形成三角形。原始数据点连结的规则是这样的: 所有三角形的边都不能与另外的三角形相交, 其结果构成了一张由三角形拼接起来的覆盖网格范围的网。每一个三角形定义了一个覆盖该三角形内网格节点的面^[3]。三角形的倾斜和标高由定义这个三角形的三个原始数据点确定。给定三角形内的全部节点都要受到该三角形的表面的限制。因为各个三角形都是用原始数据点来定义的, 这样就把三角形和数据紧密联系起来。线性插值三角网法将在网格范围内均匀分配数据, 地图上稀疏的区域将会形成截然不同的三角面。

2.10 移动平均法 (Moving Average)

移动平均法是一种简单平滑预测方法。它的基本思想是^[8]: 用大于或等于取样间隔为半径的搜索圆在插值区域内连续搜索移动, 以落在搜索圆内所有样点的均值作为待插值点(圆心)取值, 所得插值曲面即为所求。

2.11 数据度量法 (Data Metrics)

数据度量用来提供有关的数据信息。根据度量所得的数据资料, 可以再次利用一个网格数据网格的其他方法。它其实不是一种插值方法, 它是一种数据的度量方法, 通过这种方法可以找到

比较合适的插值方法。

2.12 局部多项式法 (Local Polynomial)

多项式插值也是常用的方法之一。但是, 在进行多项式插值时, 要找到一个合理的函数并不是那么容易的, 而且当多项式的阶数太大, 其波动也很大。鉴于此, 采用局部多项式法, 即对插值对象给定搜索领域内所有点插值出适当特定阶数的多项式, 局部多项式插值产生的曲面根多依赖于局部的变异。在 Surfer8.0 中, 这些多项式的形式有以下三种:

$$F(X,Y) = a + bX + cY$$

$$F(X,Y) = a + bX + cY + dXY + eX^2 + fY^2$$

$$F(X,Y) = a + bX + cY + dXY + eX^2 + fY^2 + gX^2Y + hXY^2 + iX^3 + jY^3$$

3 应用实例

笔者选用月球表面铀含量的一组离散数据, 分别使用 Surfer8.0 中的十二种不同的插值方法对其数据网格化, 得到规则的网格文件, 然后分别绘制等值线图, 如图 1~图 12 所示。图中所示的横坐标代表月球的经度, 纵坐标代表纬度, 单位为度; 插值的数据为铀的含量, 单位为 ppm。

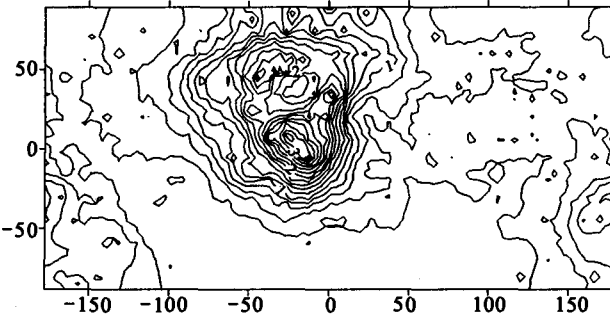


图 1 反距离加权插值法
Fig. 1 Inverse distance to a power

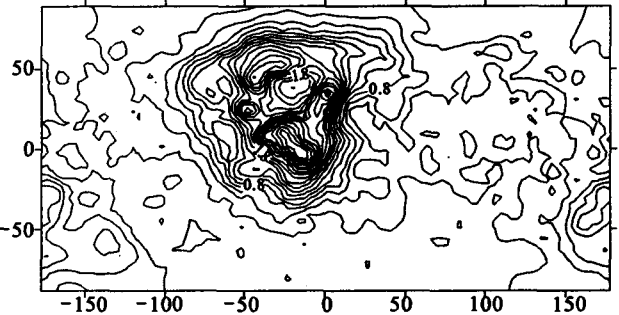


图 2 克里金插值法
Fig. 2 Kriging

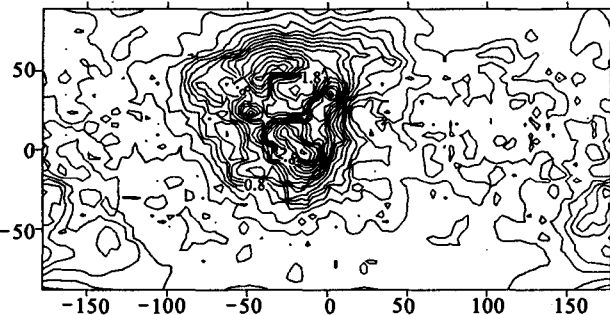


图 3 最小曲率法
Fig. 3 Minimum curvature

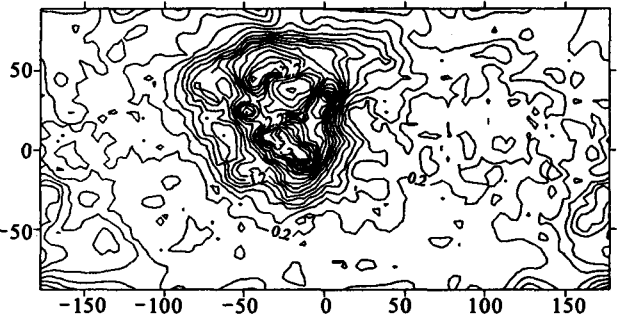


图 4 改进谢别德法
Fig. 4 Modified Shepard's method

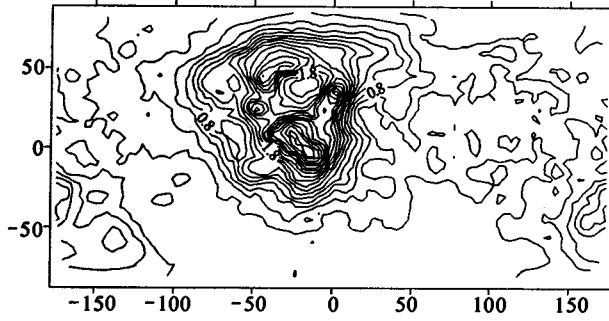


图 5 自然邻点插值法
Fig. 5 Natural neighbor

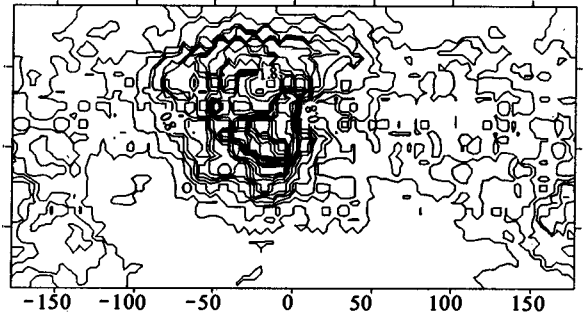


图 6 最近邻点插值法
Fig. 6 Nearest neighbor

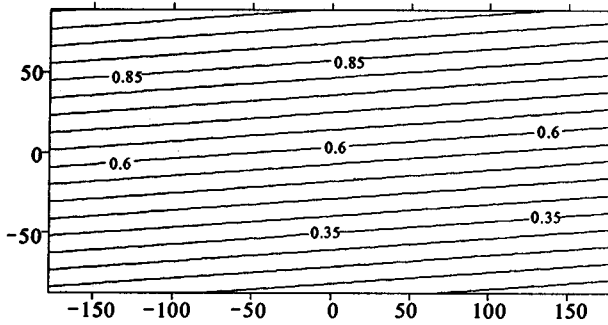


图 7 多元回归法
Fig. 7 Polynomial regression

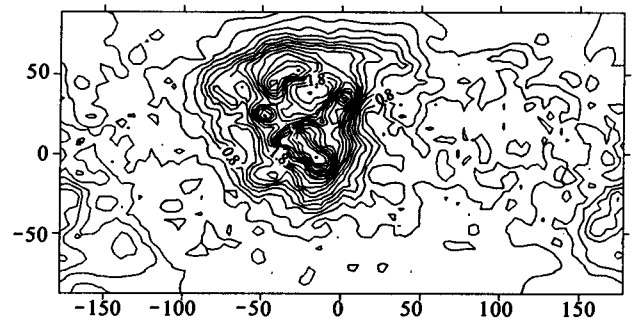


图 8 径向基函数插值法
Fig. 8 Radial basis function

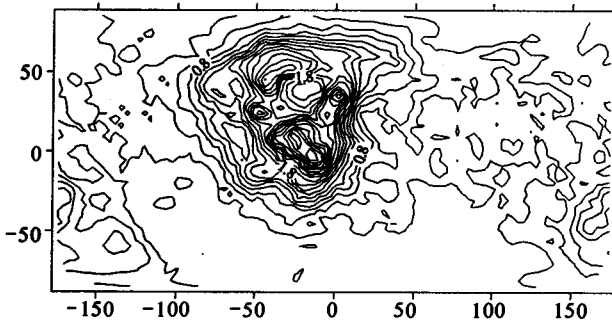


图 9 线性插值三角网法
Fig. 9 Triangulation with linear interpolation

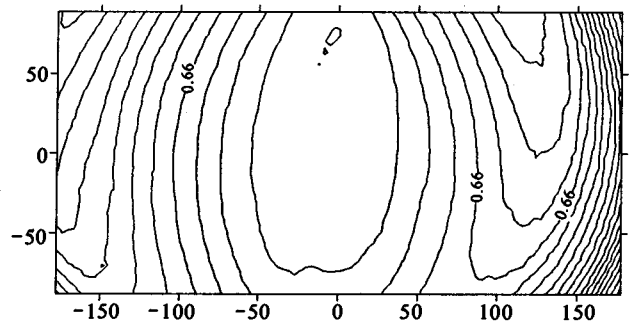


图 10 移动平均方法
Fig. 10 Moving average

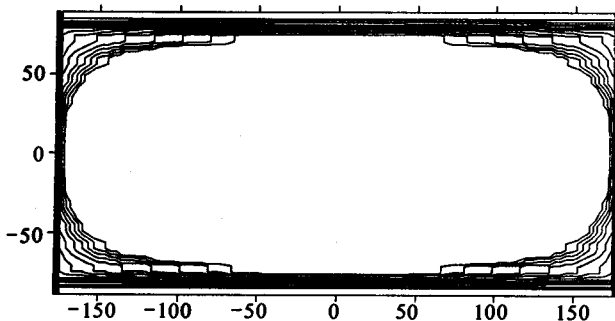


图 11 数据度量方法
Fig. 11 Data metrics

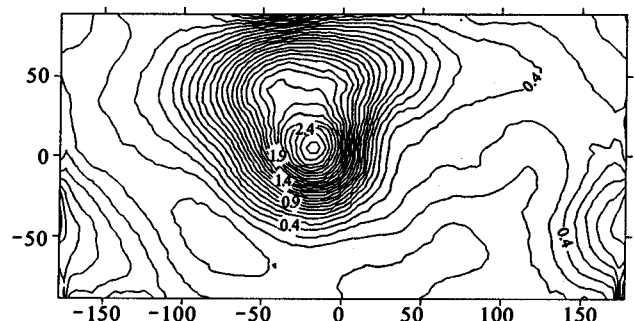


图 12 局部多项式法
Fig. 12 Local polynomial

4 结 语

Surfer8.0中所采用的十二种离散数据的插值方法,基本上包括了目前所有的插值方法。通过对Surfer8.0中十二种离散数据的插值方法的基本原理的了解,可以根据各类数据不同的特点,以及要进行的各种不同的分析,科学地选择正确的插值方法,从而进行参数设置生成网格文件,然后绘制正确的、有意义的等值线图。比如,进行面降水量分析,当站网足够密,站网密度水平较低时,可以选用泰森多边形插值,即最近邻点插值法对其进行插值;对于山区或者降水站点不是很密集的地区,则用反距离加权法有助于提高所预测数据的精度;而考虑到空间连续性变化的属性非常不规则时,克里金插值则解决了这一问题^[9]。

参考文献:

[1]孟庆香,刘国彬,杨勤科.黄土高原降水量的空间插值

方法研究[J].西北农林科技大学学报(自然科学版),2006(3):83~88.

[2]王仁铎,胡光道.线性地质统计学[M].北京:地质出版社,1988.4~10.

[3]白世彪,陈晔,王建.等值线绘图软件 Surfer 7.0 中九种插值法介绍[J].物探化探计算技术,2002(5):157~162.

[4]唐泽圣.三维数据场可视化[M].北京:清华大学出版社,1999.

[5]王兆清,冯伟.高度不规则网格多边形单元的有理函数插值格式[J].固体力学学报,2005(6):199~202.

[6]冯仲科,郭清文,朱萍.Voronoi图-泰森多边形法在三角规则树中的应用[J].林业资源管理,2006(6):44~47.

[7]卫海桥,舒歌群,梁兴雨,等.燃烧噪声一级影响模型的多元回归分析[J].内燃机工程,2006(8):66~69.

[8]胡慧萍.土壤特性的空间差异及其空间插值方法研究[J].湘潭师范学院学报(自然科学版),2001(9):99~101.

[9]何红艳,郭志华,肖文发.降水空间插值技术的研究进展[J].生态学杂志,2005,24(10):1187~1191.