

基于多重网格求解的土壤属性 高精度曲面建模

史文娇, 杜正平, 宋印军, 岳天祥

(中国科学院地理科学与资源研究所, 北京 100101)

摘要: 高精度曲面建模 (HASM) 是近几年发展起来的可用于地理信息系统和生态建模的一种较高精度的曲面建模方法。本研究选择南方红壤丘陵区江西省吉安市辖区、吉安县和泰和县为研究区, 采集 150 个表层土壤 (0~20cm) 样品, 分别随机选取 60、90 和 120 个点作为模拟数据集, 90、60 和 30 个点作为验证数据集, 基于多重网格 (MG) 作为土壤属性高精度曲面建模的求解方法 (HASM-MG), 比较 HASM-MG、Kriging、IDW 和 Spline 四种方法的插值精度, 检验 HASM-MG 在土壤 pH 的插值效果, 并分析土壤 pH 在典型南方红壤丘陵区的空间分异规律。结果表明: HASM-MG 的模拟结果与测定值最为接近; 与其他三种方法相比, 其模拟结果在三组数据集中均具有最小的平均绝对误差、均方根误差和与测定值之间最高的相关系数; 研究区水稻土分布较多的中部和东南部地区的土壤 pH 明显高于红壤分布较多的地区。多重网格求解的高精度曲面建模 (HASM-MG) 在提高了求解速度的前提下, 保证了土壤属性的插值精度, 将土壤属性的高精度曲面建模向实用性方面向前推进了一步。

关键词: 土壤 pH; HASM-MG; 插值精度; 南方红壤丘陵区

文章编号: 1000-0585(2011)05-0861-10

1 引言

土壤的空间变异将显著影响农业和环境管理^[1], 研究土壤属性的空间分异特征, 既可在宏观和整体上把握分布规律, 又可对感兴趣区进行微观和局部的深入分析, 因此, 数字化土壤制图对于农业和环境管理至关重要^[2]。精准农业的实施和土壤环境的管理迫切需要对土壤属性的空间异质性进行准确模拟, 而科学、有效的区域环境政策的制定必然依赖于大量区域范围数据的获取和高精度的空间插值, 这些在很大程度上提高了对土壤属性插值精度的要求。因此, 研究高精度的土壤属性曲面建模方法十分必要^[1, 2]。

传统的土壤属性插值方法主要有 Kriging^[3~6]、IDW^[7]和 Spline^[8]。Kriging 在理论上提供了最好的线性无偏估计、准确的数据空间结构以及估计误差分布^[5, 6], 但它基于二阶平稳假设, 并采用单一的半方差函数描述土壤属性的空间变异, 不符合实际情况。IDW 插值假设每个输入点对局部的影响随距离增大而逐渐消失^[7], 无需对数据另作其他假设^[6]。虽操作简单, 但不具备 Kriging 方法在统计上的一些优点。Spline 的原理是将零散

收稿日期: 2010-04-29; **修订日期:** 2010-09-06

基金项目: 自然科学基金青年科学基金项目 (41001057); 国家杰出青年科学基金项目 (40825003); 中国科学院知识创新工程重要方向项目 (KZCX2-YW-429); 资源与环境信息系统国家重点实验室自主研究项目 (081105)

作者简介: 史文娇 (1982-), 女, 辽宁葫芦岛人, 博士, 助研, 研究方向为土壤属性空间模拟、全球变化与区域生态。E-mail: shiwj@lreis.ac.cn

的线或面的多项式较光滑地连接在一起^[8], 仅适用于变化较小的曲面^[9]。

高精度曲面建模 (High accuracy surface modeling, HASM) 是近几年来发展起来的用于地理信息系统和生态建模的一种曲面建模方法, 在数值实验、DEM 和气候、生态系统变化趋势的模拟上, 精度均明显高于传统方法^[10]。Shi 等^[11]首次将其应用在土壤属性插值中, 证明了其在土壤属性插值中的可行性。但是, HASM 在土壤属性插值过程中求解偏微分方程组时存在高运算量和高存储量两大难题, 与其他方法相比具有较长的计算时间, 这将制约它在实践中的广泛应用。本文应用多重网格法 (Multi-grid, MG) 求解 HASM 的微分方程组 (HASM-MG)^[3~4, 12], 选择南方红壤丘陵区江西省吉安市辖区、吉安县和泰和县为研究区, 应用 HASM-MG 对其土壤 pH 的空间分异规律进行分析, 并将 HASM-MG 的插值精度与 Kriging、IDW 和 Spline 方法进行对比, 旨在缩短计算时间的前提下, 验证 HASM-MG 在土壤属性中的插值精度, 将高精度曲面建模在土壤属性插值方面的应用推向实用化。

2 研究区概况

研究区位于 $26^{\circ}26'55''\sim 27^{\circ}37'35''\text{N}$ 、 $114^{\circ}17'49''\sim 115^{\circ}30'13''\text{E}$, 包括江西省中部的吉安市辖区、吉安县和泰和县 (图 1), 总面积 6156.92km^2 。该区为典型的南方红壤丘陵区, 地势起伏较大, 土壤类型丰富, 土地利用类型多样。气候属于典型的中亚热带季风气候, 吉安市辖区、吉安县和泰和县的年均降水量分别为 1458mm 、 1438mm 和 1381mm , 年均温度分别为 18.1°C 、 18.4°C 和 19.0°C 。研究区的高程从四周 (最高处 1204.5m) 到中心 (最低处 42.0m) 逐渐变小。根据 1995 年国家土壤调查办公室编制的 $1:100$ 万土壤类型图 (图 2), 研究区土壤分为红壤、水稻土、紫色土、新积土和石灰 (岩) 土。通过对研究区 2005 年 $30\times 30\text{m}^2$ 分辨率的 TM 影像进行几何纠正、光学纠正和目视解译, 把研究区分为林地、耕地、草地、建筑用地、水域和未利用土地 6 种土地利用类型 (图 3)。

3 材料与方法

3.1 样品采集与分析

根据研究区的主要土壤类型和土地利用类型, 兼顾植被、地形和交通条件等因素, 在布置数目与密度满足高精度曲面建模、地统计学和 GIS 空间分析要求的基础上, 样点在研究区内的分布尽量均匀。2007 年 10 月, 在研究区内采集了 $0\sim 20\text{cm}$ 土壤样品 150 个 (图 1), 同时记录采样点的经纬度、高程、土壤类型、土地利用类型、植被、开垦利用历

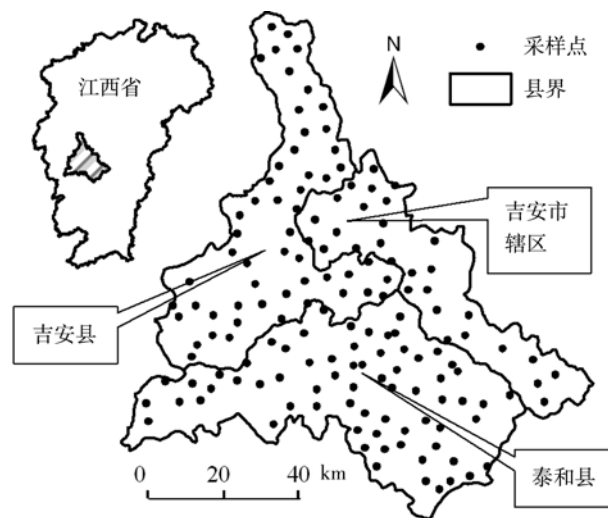


图 1 研究区位置和土壤样点分布图

Fig. 1 Location of the study area and distribution of the soil samples

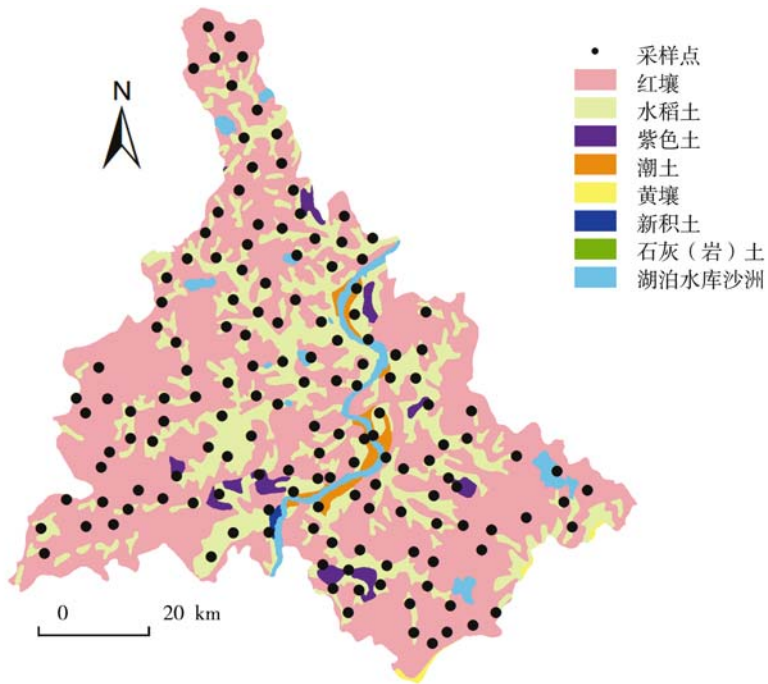


图 2 研究区土壤类型分布图

Fig. 2 The distribution of soil types of the study area

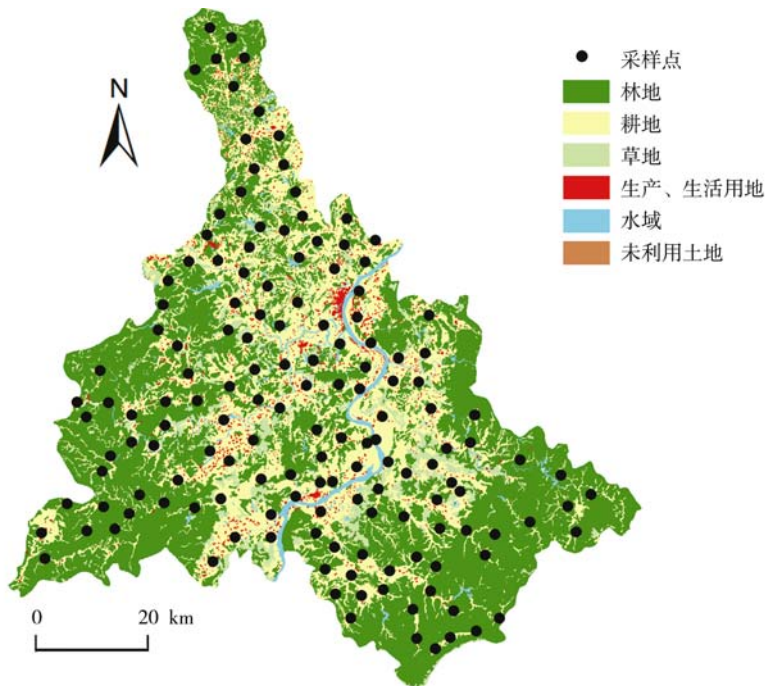


图 3 研究区土地利用类型分布图

Fig. 3 The distribution of land use types of the study area

史、耕作制度、施肥水平、周围可能的土壤污染源状况及其他景观信息。将土样风干后过 2mm 筛，土壤 pH 的测定采用 1 : 2.5 的水土比，pH 计测定。

3.2 HASM-MG

假设 $\{\bar{u}_{i,j}\}$ 是采样点 $\{(x_i, y_j)\}$ 在曲面 u 上的采样值， $u_{i,j}^n$ (其中 $n \geq 0, 0 \leq i \leq I+1, 0 \leq j \leq J+1$) 是中心点为 (x_i, y_j) 的栅格值的第 n 次迭代值，其中 $u_{i,j}^0 = \tilde{u}_{i,j}$ ， $\{\tilde{u}_{i,j}\}$ 是基于采样点 $\{\bar{u}_{i,j}\}$ 的插值结果。根据数值计算^[13]，得出 HASM 的第 $n+1$ 次迭代的有限差分基本表达式为^[3, 4, 10, 14~17]：

$$\frac{u_{i+1,j}^{n+1} - 2u_{i,j}^{n+1} + u_{i-1,j}^{n+1}}{h^2} = (\Gamma_{11}^n)_{i,j}^n \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2h} + (\Gamma_{11}^e)_{i,j}^n \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2h} + \frac{L_{i,j}^n}{\sqrt{E_{i,j}^n + G_{i,j}^n} - 1} \tag{1}$$

$$\frac{u_{i,j+1}^{n+1} - 2u_{i,j}^{n+1} + u_{i,j-1}^{n+1}}{h^2} = (\Gamma_{22}^n)_{i,j}^n \frac{u_{i+1,j}^n - u_{i-1,j}^n}{2h} + (\Gamma_{22}^e)_{i,j}^n \frac{u_{i,j+1}^n - u_{i,j-1}^n}{2h} + \frac{N_{i,j}^n}{\sqrt{E_{i,j}^n + G_{i,j}^n} - 1} \tag{2}$$

其中 $n \geq 0; 0 < i < I+1; 0 < j < J+1; E_{i,j}^n, F_{i,j}^n, G_{i,j}^n, L_{i,j}^n, N_{i,j}^n, (\Gamma_{11}^n)_{i,j}^n, (\Gamma_{22}^n)_{i,j}^n, (\Gamma_{11}^e)_{i,j}^n, (\Gamma_{22}^e)_{i,j}^n$ 等变量的定义见文献^[10]； $u_{0,j}^{n+1} = u_{0,j}^0 (0 \leq j \leq J+1)$ ； $u_{i,0}^{n+1} = u_{i,0}^0 (0 \leq i \leq I+1)$ ； $u_{i+1,j}^{n+1} = u_{i+1,j}^0 (0 < j < J+1)$ ； $u_{i,j+1}^{n+1} = u_{i,j+1}^0 (0 < i < I+1)$ ； $u_{i,j}^{n+1}, u_{i,0}^{n+1}, u_{i+1,j}^{n+1}$ 和 $u_{i,j+1}^{n+1}$ 是 HASM 的边界条件。

(1) 式和 (2) 式的矩阵表达形式为：

$$A_1 U^{n+1} = b_1^n \tag{3}$$

$$A_2 U^{n+1} = b_2^n \tag{4}$$

其中 $U^{n+1} = (u_{i,1}^{n+1}, \dots, u_{i,J}^{n+1}, u_{i+1,1}^{n+1}, \dots, u_{i+1,J}^{n+1}, \dots, u_{I-1,1}^{n+1}, \dots, u_{I-1,J}^{n+1}, u_{I,1}^{n+1}, \dots, u_{I,J}^{n+1})^T$ ； A_1 和 b_1^n 分别是 (1) 式的系数矩阵和右端常数项矩阵； A_2 和 b_2^n 分别是式 (2) 式的系数矩阵和右端常数项矩阵。

假设 $Z = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, q^n = \begin{bmatrix} b_1^n \\ b_2^n \end{bmatrix}$ ，为了保证土壤采样点的真实值与采样点的估计值相等或相近，HASM 的表达式可转化为等式约束的最小二乘问题^[10, 14]：

$$\begin{cases} \min \| ZU^{n+1} - q^n \|_2 \\ \text{s. t. } CU^{n+1} = d \end{cases} \tag{5}$$

其中， C 和 d 分别为采样点系数矩阵和采样点的值，表达式分别为： $C(k, (i-1) \cdot J + j) = 1, d(k) = \bar{u}_{i,j}$ ，即第 k 个采样点 (x_i, y_j) 的值为 $\bar{u}_{i,j}$ 。

为了求解最小二乘问题的方程组 (5)，引入一个正的足够大的参数 $\lambda^{[18]}$ ， λ 是赋予采样点的权重，决定于采样点对模拟曲面的贡献^[13]。对于足够大的 λ ，(5) 式可被转化为无约束的最小二乘问题^[18]，

$$\min_f \left\| \begin{bmatrix} Z \\ \lambda C \end{bmatrix} U^{n+1} - \begin{bmatrix} q^n \\ \lambda d \end{bmatrix} \right\|_2 \tag{6}$$

也就是求解

$$\begin{bmatrix} Z^T & \lambda C^T \end{bmatrix} \begin{bmatrix} Z \\ \lambda C \end{bmatrix} U^{n+1} = \begin{bmatrix} Z^T & \lambda C^T \end{bmatrix} \begin{bmatrix} q^n \\ \lambda d \end{bmatrix} \tag{7}$$

设 $A_h = \begin{bmatrix} Z^T & \lambda C^T \\ \lambda C \end{bmatrix}, b_h^n = \begin{bmatrix} Z^T & \lambda C^T \\ \lambda C \end{bmatrix} \begin{bmatrix} q^n \\ \lambda d \end{bmatrix}$ ， h 为模拟步长，则问题转换为求解

线性代数方程组，

$$A_h U^{n+1} = b_h^n \tag{8}$$

以往的 HASM 模型^[10]需要存储最细网格层系数矩阵 A_h ，需占用大量内存。本文采用多重网格法求解 (8) 式^[4]，针对系数矩阵非零元素的分布规律，只需存储不同网格层 (A_h 、 A_{2h} ...) 的对角线元素，从而大大减少了存储量。其解算过程的一般步骤如下：

(1) 建立一系列粗网格层，其像元大小分别为 $2h \times 2h$ ， $4h \times 4h$ ，...，重复 $h \times h$ 网格层上建立方程组 (8) 的过程，得到 A_{2h} ， A_{4h} ，...；

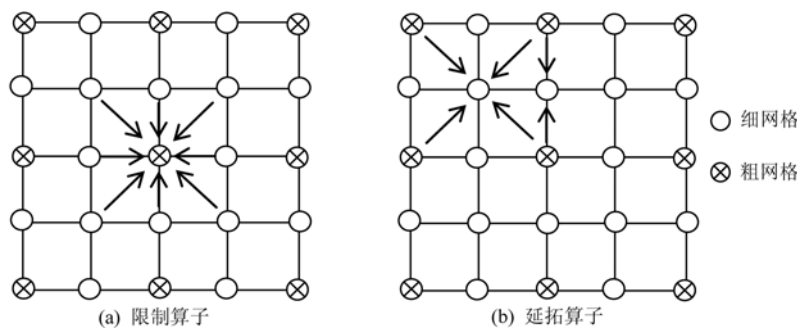
(2) 在最细网格层上方程为 $A_h U^{n+1} = b_h^n$ ，以 $U_h^0 = U^n$ 为迭代初值，使用 Gauss-Seidel 迭代法迭代 v_1 步，得到近似解 U_h^1 ；

(3) 设 $r_h^0 = b_h^n - A_h U_h^1$ ，计算 r_h^0 ；

(4) 限制算子将 r_h^0 的信息传递到下一层网格 (图 4a)；令 $b_{2h} = I_{2h}^h r_h^0$ 为下一网格层线性方程组的右端项向量，其中 I_{2h}^h 表示限制算子；

(5) 重复第二步到第四步直到已经达到最粗网格层。最粗网格层方程为 $A_H U^{n+1} = b_H^n$ ，其中 H 为最粗网格层步长，直接精确求解此方程，得到 U_H^1 ；

(6) 延拓算子将 U_H^1 的信息传到上一层网格 (图 4b)，上一网格层方程为 $O_X = P_H^X U_H^1$ ，其中 X 为上一网格层步长， P_H^X 表示延拓算子。令 $U_X^0 = U_X^1 + O_X$ ，其中 U_X^1 是向下传递过程中计算得到的近似解；



(a) 限制算子：细网格信息向粗网格传递；(b) 延拓算子：粗网格信息向细网格传递

图 4 多重网格信息传递

Fig. 4 Information fusion on multigrids

(a) Information on the fine grid \bigcirc is transferred to the coarse grid \otimes , and (b) information on the coarse grid \otimes is transferred to the fine grid \bigcirc

(7) 以 U_X^0 为初值，对方程 $A_X U^{n+1}$ 进行 v_2 步 Gauss-Seidel 迭代，得到近似解 U_X^2 ；

(8) 重复第六步到第七步，直到返回到最细网格层，得到最细网格层上的近似解 U_h^2 ；

(9) 以 $U_h^0 = U_h^2$ 为初值，重复第二步到八步，不断更新 U_h^i ，直至得到满意精度。

以往的 HASM 模型需要存储最细网格层系数矩阵 A_h ，因而占用大量内存；使用多重网格法，针对系数矩阵非零元素的分布规律，只需要存储 $A_h, A_{2h}, A_{4h}, \dots$ 的对角线元素，从而大大减少了存储量。用多重网格法求解 HASM 模型，代数方程组求解过程的计算时间随着栅格总数呈线性增长^[19]，解决了以往 HASM 模型运算量随栅格总数几何增长的问题，大幅度提高了模型的求解速度，解决了 HASM 高运算量的瓶颈^[4]。

3.3 模型精度的验证

通常用来比较插值方法之间效果的验证方法有交叉验证和独立数据集验证。交叉验证包括以下几个步骤^[20]：（1）轮流消除某个点；（2）用每次未消除的那些点估计已消除的那个点；（3）比较模拟值与测定值。这是比较模拟值与测定值的较为快速和廉价的方法^[18]。但是，在许多情况下并不能描述空间模拟误差^[21, 22]。以 Kriging 方法为例，在交叉验证过程中，他每次都应用相同的变异函数，但对于真正的交叉验证而言，在每次去除观测点后都应重新计算变异函数^[21]。独立数据集检验可以避免交叉验证的缺点^[21, 23~24]，是一个直接估计空间不确定性的更独立的方法，验证点随机分布在研究区内^[21, 23]。所以，可用独立数据集验证法来评价 HASM-MG、Kriging、IDW 和 Spline 这四种方法的插值效果。

选择平均误差（ME）、平均绝对误差（MAE）、均方根误差（RMSE）和验证点的测定值与模拟值的相关系数（ r ）4 个指标来评价插值方法的精度。ME、MAE 和 RMSE 的计算公式如下^[11]：

$$ME = \frac{1}{n} \sum_{i=1}^n \{z^*(x_i) - z(x_i)\} \tag{9}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n \{|z^*(x_i) - z(x_i)|\} \tag{10}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \{z^*(x_i) - z(x_i)\}^2} \tag{11}$$

其中 n 为验证点个数； $z^*(x_i)$ 为验证点的模拟值； $z(x_i)$ 为验证点的测定值。ME 是预测无偏的量度，越接近 0 越说明方法是无偏的；MAE 和 RMSE 是预测精度的量度，越小则说明插值方法越精确。

4 不同方法模拟值与测定值的比较

将四种方法对 150 个土壤样点的 pH 进行插值后的模拟值与测定值相比较（表 1），

表 1 土壤 pH 的测定值与四种模拟方法模拟值的比较

Tab. 1 Summary statistics for both measured data and simulated results of soil pH

	平均值	最小值	最大值	0~25%	25%~50%	50%~75%	75%~100%
测定值	5.07	4.50	6.18	4.50~4.80	4.80~4.99	4.99~5.25	5.25~6.18
HASM-MG	5.06	4.49	6.18	4.49~4.84	4.84~5.01	5.01~5.24	5.24~6.18
Kriging	5.06	4.61	5.79	4.61~4.94	4.94~5.07	5.07~5.21	5.21~5.79
IDW	5.05	4.50	6.18	4.50~4.85	4.85~5.01	5.01~5.23	5.23~6.18
Spline	5.05	4.20	6.67	4.20~4.78	4.78~4.97	4.97~5.26	5.26~6.67

HASM-MG 和 Kriging 的模拟结果的平均值均为 5.06，在四种方法中，二者与测定值的平均值 5.07 最为接近。由于 Kriging 方法的平滑作用，模拟的最小值 4.61 和最大值 5.79 分别高于和低于测定值的最小值 4.50 和最大值 6.18。相反，Spline 有着较为明显的震荡，模拟结果的最小值和最大值分别是 4.20 和 6.67。然而，HASM-MG 和 IDW 模拟的最小值分别为 4.49 和 4.50，最大值均为 6.18，与测定值的最小值和最大值最为接近。从模拟值和测定值分别升序排列的 0~25%、25%~50%、50%~75% 和 75%~100% 范围的统计来看，HASM-MG 也与测定值最为接近。

5 精度验证与对比

随机产生三组不同的数据集用来模拟和验证，验证点数据集分别为采样点总数（150 个）的 20%（30 个）、40%（60 个）和 60%（90 个），分别用 $V_{20\%}$ 、 $V_{40\%}$ 和 $V_{60\%}$ 表示；模拟数据集分别为采样点总数（150 个）的 80%（120 个）、60%（90 个）和 40%（60 个），分别用 $S_{80\%}$ 、 $S_{60\%}$ 和 $S_{40\%}$ 表示。用 ME、MAE、RMSE 以及模拟值与观测值的相关系数 r 这 4 个参数比较不同方法的预测效果（表 2）。四种方法的 ME 都接近于 0，说明他们都是无偏的。三组数据集中 HASM-MG 的 MAE 和 RMSE 均为最小， r 均为最大，说明 HASM-MG 是四种方法中精度最高的方法。HASM-MG 在对土壤 pH 曲面进行模拟时，首先根据采样数据计算其数值模拟方程组的系数矩阵和常数项矩阵，采用最小二乘法和多次迭代步骤使插值的精度得到了保证；另外，HASM-MG 的迭代消除了传统方法模拟土壤 pH 时常见的平滑现象（Kriging）和震荡现象（Spline）^[11]；最后，用 MG 法求解 HASM 模型，求解时间与待模拟格点总数呈线性关系^[19]，解决了以往 HASM 模型运算量随栅格总数几何增长的问题^[3]，在大幅度提高了模拟速度的前提下，同样保证了模拟精度，将高精度曲面建模向土壤属性插值的实用性向前推进了一大步。

6 土壤 pH 的空间分布

土壤 pH 是土壤的基本化学性质之一，是盐基状况的综合反映，它不仅影响土壤微生物的生长、植物所必需的营养元素的转化及有效性，同时还会影响土壤重金属元素的含量和存在形态，进而影响到土壤生态环境和食品安全。

从图 5 可以看出，研究区的中部和东南部土壤 pH 与其他地区相比相对较高，这在很大程度上与土壤类型的分布有关（图 2）。水稻土的 pH 明显比红壤的 pH 高，而研究区中部和东南部大部分为水稻土，所以中部和东南部地区的土壤 pH 明显高于其他地区。《江

表 2 HASM-MG 与三种传统方法的 ME, MAE, RMSE 和 r 比较

Tab. 2 MEs, MAEs, RMSEs and correlation coefficients (r) of HASM-MG and the classical methods

		HASM-MG	Kriging	IDW	Spline
$V_{20\%}$	ME	-0.02	0.06	0.03	0.01
	MAE	0.14	0.19	0.16	0.22
	RMSE	0.20	0.25	0.21	0.29
	r	0.82**	0.73**	0.81**	0.66**
$V_{40\%}$	ME	0.01	0.05	0.03	-0.03
	MAE	0.19	0.24	0.21	0.23
	RMSE	0.27	0.31	0.29	0.32
	r	0.64**	0.48**	0.55**	0.58**
$V_{60\%}$	ME	-0.04	0.01	0.02	-0.05
	MAE	0.20	0.26	0.22	0.30
	RMSE	0.28	0.34	0.30	0.40
	r	0.67**	0.46**	0.62**	0.59**

注：** 表示 0.01 水平的极显著相关。

西土壤》将土壤 pH 的分级如下^[25]：1 级 8.5~9.0（强碱性）、2 级 7.5~8.5（微碱性）、3 级 6.5~7.5（中性）、4 级 5.5~6.5（微酸性）、5 级 4.5~5.5（酸性）、6 级 ≤ 4.5 （强酸性）。吉安地区第二次土壤普查（20 世纪 80 年代）的土壤 pH 的分级情况为：1 级占土壤总面积的 0.28%、2 级占 0.92%、3 级占 2.10%、4 级占 17.35%、5 级占 68.58%、6 级占 10.77%（表 3）。可见，在 20 世纪 80 年代时，吉安地区的土壤 pH 绝大部分在微酸性（5.5~6.5）、酸性（4.5~5.5）和强酸性（ ≤ 4.5 ）范围内。2007 年研究区土壤 pH 则主要分布在微酸性和酸性范围内（图 5），其中微酸性（5.5~6.5）的土壤占总面积的 32.07%，酸性（4.5~5.5）的占总面积的 67.93%。与吉安地区 20 世纪 80 年代的状况相比，研究区土壤 pH 在酸性（4.5~5.5）范围的分布面积十分相近，而 80 年代的强酸性土壤（10.77%）、中性土壤（2.10%）、微碱性土壤（0.92%）和强碱性土壤（0.28%）在 2007 年的研究区范围内已基本消失，转化为酸性土壤或微酸性土壤。说明与 30 年前相比，由于长期的人为活动如耕作、施肥等的影响，研究区强酸和强碱的土壤已明显向他们的中间状态即酸性或微酸性发展。

表 3 吉安地区和研究区土壤 pH 分布面积统计表^[25]

Tab. 3 The statistics of distribution areas of soil pH in Ji'an and the study area

土壤分类	总面积	1 级	2 级	3 级	4 级	5 级	6 级
自然土壤（万亩）	2847.21	9.85	25.58	50.41	384.98	1996.84	379.55
耕地土壤（万亩）	746.73	0.15	7.46	25.13	238.65	467.82	7.52
面积汇总（万亩）	3593.94	10	33.04	75.54	623.63	2464.66	387.07
80 年代吉安地区（%）		0.28	0.92	2.10	17.35	68.58	10.77
2007 年研究区（%）		0	0	0	32.07	67.93	0

7 结论

本文应用多重网格求解高精度曲面建模（HASM-MG），刻画了南方典型红壤丘陵区土壤 pH 的空间分布规律，结果表明：研究区水稻土分布较多的中部和东南部地区的土壤 pH 明显高于红壤分布较多的地区。将 HASM-MG 的插值效果与其他三种传统方法（Kriging、IDW 和 Spline）对比表明：HASM-MG 模拟的土壤 pH 结果与测定值最为接近；从 MAE、RMSE 和 r 来看，模拟效果均优于 Kriging、IDW 和 Spline 三种传统方法。

HASM-MG 在土壤属性插值中表现出来的高精度和实用性的

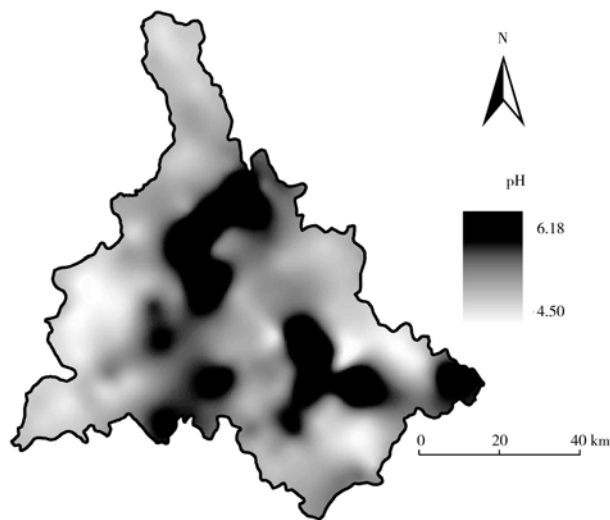


图 5 HASM-MG 对研究区土壤 pH 空间分布的模拟结果

Fig. 5 The spatial pattern of soil pH in the study area interpolated by HASM-MG

理论根源在于: (1) HASM 在对土壤 pH 曲面进行模拟时, 首先根据采样数据计算其数值模拟方程组的系数矩阵和常数项矩阵, 用最小二乘法和多次迭代步骤保证其精度; (2) HASM 迭代足够多的次数, 不仅可以获得比较精确的第一基本量和第二基本量的系数, 从而获得较高精度的模拟曲面, 还可解决传统方法模拟土壤 pH 时常见的平滑 (Kriging) 和震荡问题 (Spline); (3) 用多重网格法求解 HASM 模型, 求解的计算时间随着栅格总数的增加表现为线性增长, 解决了以往 HASM 模型运算量随栅格总数几何增长的问题, 在大幅度提高了模拟速度的前提下, 同样保证了模拟精度, 将高精度曲面建模向土壤属性插值的实用性向前推进了一大步。

土壤属性的空间异质性较高, 如果高精度曲面建模结合与土壤属性关系密切的地学信息进行模拟, 其模拟精度有可能有大幅度提高^[26]。因此, 将地学信息作为辅助信息与 HASM 相结合对土壤属性进行插值, 将是土壤属性高精度曲面建模领域下阶段的重点研究方向之一。

参考文献:

- [1] Cook S, Jarvis A, Gonzalez J. A new global demand for digital soil information. In: Hartemink A. Digital Soil Mapping with Limited Data. Berlin; Springer, 2008. 31~41.
- [2] Lagacherie P, McBratney A. Spatial soil information systems and spatial soil inference systems: Perspectives for digital soil mapping. *Developments in Soil Science*, 2006, 31: 3~22.
- [3] Yue T X, Du Z P, Song Y J. Ecological models: Spatial models and geographic information systems. In: Jorgensen SE, Fath B. *Encyclopedia of Ecology*. England; Elsevier Limited, 2008. 3315~3325.
- [4] Yue T X, Song Y J. The YUE-HASM method. In: Li D, Ge Y, Foody G M. *Proceeding of the 8th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*. Liverpool; World Academic Press, 2008. 148~153.
- [5] Kravchenko A, Bullock D G. A comparative study of interpolation methods for mapping soil properties. *Agronomy Journal*, 1999, 91(3): 393~400.
- [6] Jongman R H G, Ter Braak C J F, Van Tongeren O F R. *Data Analysis in Community and Landscape Ecology*. Cambridge; Cambridge University Press, 1995.
- [7] Panagopoulos T, Jesus J, Antunes M D C, *et al.* Analysis of spatial interpolation for optimising management of a salinized field cultivated with lettuce. *European Journal of Agronomy*, 2006, 24(1): 1~10.
- [8] Webster R, Oliver M A. *Geostatistics for Environmental Scientists*. Chichester; Wiley, 2001.
- [9] Robinson T P, Metternicht G. Testing the performance of spatial interpolation techniques for mapping soil properties. *Computers and Electronics in Agriculture*, 2006, 50(2): 97~108.
- [10] Yue T X, Du Z P, Song D J, *et al.* A new method of surface modeling and its application to DEM construction. *Geomorphology*, 2007, 91(1-2): 161~172.
- [11] Shi W J, Liu J Y, Du Z P, *et al.* Surface modelling of soil pH. *Geoderma*, 2009, 150(1-2): 113~119.
- [12] 范泽孟, 岳天祥, 宋印军. 基于 YUE-HASM 方法的气温与降水时空变化趋势. *地理研究*, 2009, 28(3): 643~652.
- [13] Quarteroni A, Sacco R, Saleri F. *Numerical Mathematics*. New York; Springer, 2000.
- [14] Yue T X, Song D J, Du Z P, *et al.* High-accuracy surface modelling and its application to DEM generation. *International Journal of Remote Sensing*, 2010, 31(8): 2205~2226.
- [15] 岳天祥, 杜正平, 宋敦江. 高精度曲面建模: HASM4. *中国图象图形学报*, 2007, 12(2): 343~348.
- [16] 岳天祥, 杜正平, 刘纪远. 高精度曲面建模与误差分析. *自然科学进展*, 2004, 14(3): 300~306.
- [17] 岳天祥, 杜正平. 高精度曲面建模最佳表达形式的数值实验分析. *地球信息科学*, 2006, 8(3): 83~87.
- [18] Van Loan C F. On the method of weighting for equality-constrained least-squares problems. *SIAM Journal on Numerical Analysis*, 1985, 22(5): 851~864.
- [19] 刘超群. *多重网络法及其在计算流体力学中的应用*. 北京: 清华大学出版社, 1995.

- [20] Voltz M, Webster R. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. *European Journal of Soil Science*, 1990,41(3): 473~490.
- [21] Isaaks E H, Srivastava R M. *An Introduction to Applied Geostatistics*. New York: Oxford University Press, 1989.
- [22] Mueller T G, Pierce F J, Schabenberger O, *et al.* Map quality for site-specific fertility management. *Soil Science Society of America Journal*, 2001,65(5): 1547~1558.
- [23] Laslett G M, McBratney A B, Pahl P J, *et al.* Comparison of several spatial prediction methods for soil pH. *Journal of Soil Science*, 1987,38(2): 325~341.
- [24] Triantafilis J, Odeh I O A, McBratney A B. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. *Soil Science Society of America Journal*, 2001,65(3): 869~878.
- [25] 江西省土壤普查办公室. 江西土壤. 北京: 中国农业科技出版社, 1991.
- [26] Shi W J, Liu J Y, Du Z P, *et al.* Surface modelling of soil properties based on land use information. *Geoderma*, 2011,162(3-4):347~357.

High accuracy surface modeling of soil properties based on multi-grid

SHI Wen-jiao, DU Zheng-ping, SONG Yin-jun, YUE Tian-xiang

(Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China)

Abstract: Kriging, inverse distance weighting (IDW) and spline are three classical methods for interpolating soil properties. High accuracy surface modeling (HASM) has been developed in recent years, which is successfully used in geographic information system and ecological modeling. In order to assess its feasibility of soil property interpolation, HASM based on multi-grid (HASM-MG) was used to interpolate soil pH for a red soil hilly region of Jiangxi Province, China. A total of 150 samples of topsoil (0—20 cm) were collected in Ji'an municipal district, Ji'an County and Taihe County in October, 2007. Three different data sets were randomly created for simulation and validation to comparatively simulate performance of HASM, kriging, IDW and spline. A total of 30, 60 and 90 samples were removed as validation points, thus 120, 90 and 60 data points were selected as simulation points, respectively. The results show that HASM simulation is the closest to the measured data among the four methods. It has much smaller mean absolute error, root mean square error and much higher correlation coefficient of simulated value and measured value than the three classical methods. Soil pH values in the central and southeastern parts of the study area dominated by paddy soils are much higher than those in the red soil area. Therefore, HASM-MG can be considered as an accurate method for interpolating soil properties on the premise of improving the calculation speed, and it advances the practicality of HASM applied in soil property interpolation.

Key words: soil pH; high accuracy surface modeling based on multi-grid; interpolation accuracy; the red soil hilly region in South China