

# 中国表层土壤有机质空间分布模拟分析方法研究

李启权<sup>1,2</sup>, 岳天祥<sup>1\*</sup>, 范泽孟<sup>1</sup>, 杜正平<sup>1</sup>

(1. 中国科学院 地理科学与资源研究所, 北京 100101; 2. 中国科学院 研究生院, 北京 100049)

**摘要:** 气候变化背景下,对土壤有机碳的研究是目前大尺度上土壤性质研究的热点。基于第二次全国土壤普查 5 374 个典型土壤剖面数据,分析表层土壤有机质(20 cm)与环境因素的相关关系,利用多元回归模型和 HASM 模型结合的方法模拟中国国家尺度上表层土壤有机质含量的空间分布格局,探讨该方法的模拟误差,为国家尺度上有机碳的估算提供方法参考。研究结果表明,对 350 个检验点模拟结果的平均绝对误差和平均相对误差为 15.61 g·kg<sup>-1</sup>和 56.59%,与普通克里格法相比分别降低了 1.61 g·kg<sup>-1</sup>和 20.84%;对样点分布较少以及无样点的西北地区 and 台湾省的模拟结果也更符合实际情况。建模样点减少一半的情况下,模拟结果的平均绝对误差和平均相对误差仅分别增加了 0.14 g·kg<sup>-1</sup>和 1.07%。因此,论文方法可作为模拟国家尺度上有机质空间分布相对有效的方法,同时如何使模型解释更多的土壤有机质空间变异将是进一步提高模拟精度的关键。

**关键词:** HASM 方法; 土壤有机质; 国家尺度; 样点密度; 空间模拟

**中图分类号:** S153.6<sup>+</sup>21; TP319 **文献标志码:** A **文章编号:** 1000-3037(2010)08-1385-15

土壤性状在不同尺度上的空间变异是近年来土壤空间研究的热点之一。国家尺度上,高精度的土壤信息可为生态地理区划、准确的土壤碳储量估算以及土壤环境、农业宏观管理决策等提供可靠的数据基础。然而,由于研究范围太大,国家尺度上土壤性状变异研究不仅受到数据源获取的限制;同时,由于各种成土因素、土壤形成过程以及土壤利用与管理等在时空上的复杂性,土壤性质表现出极强的变异性,这为国家尺度上定量化、高精度土壤空间信息的获取和处理提出了挑战。

根据研究过程中基于的理论基础,土壤性质空间分布模拟的方法可分为空间自相关模型和空间异相关模型。空间自相关模型基于 Tobler 定理,利用要素自身特征和地理位置来预测未采样区域的值,如反比距离方法<sup>[1]</sup>、趋势面分析法<sup>[2]</sup>、普通克里格(Ordinary Kriging, OK)法<sup>[3-5]</sup>等。空间异相关模型基于经典的土壤发生学理论,利用土壤性质与其影响因素之间的关系来推测该土壤性质的空间分布,如多元线性回归方法<sup>[6]</sup>、回归树方法<sup>[7]</sup>、协同克里格方法<sup>[8]</sup>、模糊推理方法<sup>[9-10]</sup>等。空间自相关模型往往需要大量的样本数据或满足特定的假设条件。对空间异相关模型而言,由于土壤与各种影响因素之间是一种极其复杂的非线性关系<sup>[11-12]</sup>,如何利用模型表达这种关系仍需要进一步研究。

目前中国国家尺度上土壤性质空间分布的研究较少。张慧智等<sup>[13]</sup>和王绍强等<sup>[14]</sup>分别

收稿日期: 2010-02-28; 修订日期: 2010-04-30。

基金项目: 国家杰出青年科学基金(40825003); 国家科技支撑计划课题(2006BAC08B04); 国家科技支撑计划专题(2006BAC08B0505); 中国科学院知识创新工程重要方向项目(kzcx2-yw-429); 国家 973 项目(2009CB421105)。

第一作者简介: 李启权(1980-),男,四川泸县人,博士生,主要从事资源环境与系统模拟。E-mail: liqq@reis.ac.cn

\*通信作者简介: 岳天祥(1963-),男,甘肃庆阳人,教授,博士生导师,主要从事资源环境与系统模拟。E-mail: yue@reis.ac.cn

对全国土壤温度和土层厚度进行了研究。而其它国家尺度上土壤性质的研究则主要集中在土壤有机碳上<sup>[15-20]</sup>。在对全国土壤有机碳的研究中,大多采用分类统计的方法,这使得研究结果最大相差达到3倍多。如何获得国家尺度上准确的土壤有机碳信息还需要深入研究。

基于曲面论的 HASM方法 (High Accuracy Surface Model, HASM)是近几年发展起来的一种新的可用于地理要素空间模拟的方法,在数值模拟和建立 DEM 等方面比经典插值方法精度高很多<sup>[21-23]</sup>;在土壤性质空间分布模拟中也取得了较好的效果<sup>[24]</sup>。本文以第二次全国土壤普查中记录的典型土壤剖面为数据基础,以表层土壤有机质 (Soil Organic Matter, SOM)为模拟对象,应用空间异相关模型中应用最多的多元线性回归方法与 HASM方法相结合,探讨中国国家尺度上表层土壤有机质的空间分布格局,并分析样点数对模拟结果的影响,为国家尺度上土壤性质的研究提供参考。

## 1 数据与方法

### 1.1 土壤有机质数据

我国第二次全国土壤普查期间 (1979—1994年)采集分析了大量的土壤剖面,其中各土壤类型的典型土壤剖面资料记录在《中国土种志》及各省级土种志中,本研究收集整理了这些典型土壤剖面数据,去除《中国土种志》和各省级土种志中重复记录的剖面,建立典型土壤剖面属性数据库。表层土壤性质空间变化特征明显,对环境条件变化的响应也更为直接;基于样点和模型模拟的国家尺度土壤有机质空间分布格局研究也未见报道;同时,土壤有机碳关系全球气候变化,是国内外研究的热点。因此,本文以表层土壤有机质作为模拟对象。

### 1.2 土壤剖面准确定位

由于典型土壤剖面数据对于剖面位置的描述没有坐标信息而只以地名来记录,因此,在利用该数据进行模拟前,本研究先依据典型土壤剖面数据记录的地名以及环境信息对这些典型土壤剖面进行逐一定位,使其尽可能地接近原采样点位置,并符合该典型剖面所处的环境条件,包括成土母质或母岩、土壤类型、植被、地形 (坡度、坡位、海拔等)。具体定位过程如下:首先按剖面数据中记录的地名在全国居民点数据 (矢量格式)中找到相应的乡或村 [图 1(a)~(c)];再以 1:400万植被图 (来源于中国科学院资源环境数据中心)、1:100万土壤图 (由中国科学院南京土壤研究所史学正等制作<sup>[25]</sup>,下载于 [www.isric.nl](http://www.isric.nl))确定该区域土壤类型、植被的分布;最后结合 90 m分辨率的 DEM (来源于 <http://csi.cgiar.org>)和剖面记录的地形信息 (坡度、坡位、高程)确定该剖面的空间位置,获得其相对准确的坐标位置 [图 1(d)]。定位时使用 1954Beijing坐标系统。

通过以上方法对收集到的典型土壤剖面进行逐一定位,去掉无有机质数据记录的剖面点,共计获得 5374个具有相对准确空间坐标信息的包含土壤有机质含量记录的典型土壤剖面 (图 2)。

### 1.3 影响因素的选择与处理

较大尺度上影响土壤有机质的因素很多。国家尺度上影响土壤性质的因素包括 Jenny 方程<sup>[26]</sup>中的所有要素 (母质、气候、地形、植被及人类活动等)。根据相关研究<sup>[11,26-27]</sup>、数据的可获取情况、典型土壤剖面记录的详细程度,本研究选择了气候、地形、植被、人口和全国土壤类型方面的相关因子作为影响因素。其中,气象数据为 1971—2000年全国 671个气象站的多年平均数据,研究过程中运用普通克里格法得到全国多年平均的气温 (Temperature,  $T$ )、降水 (Precipitation,  $P$ )、相对湿度 (Relative humidity,  $RH$ )、日照时数 (Duration of sunshine,  $DS$ )、

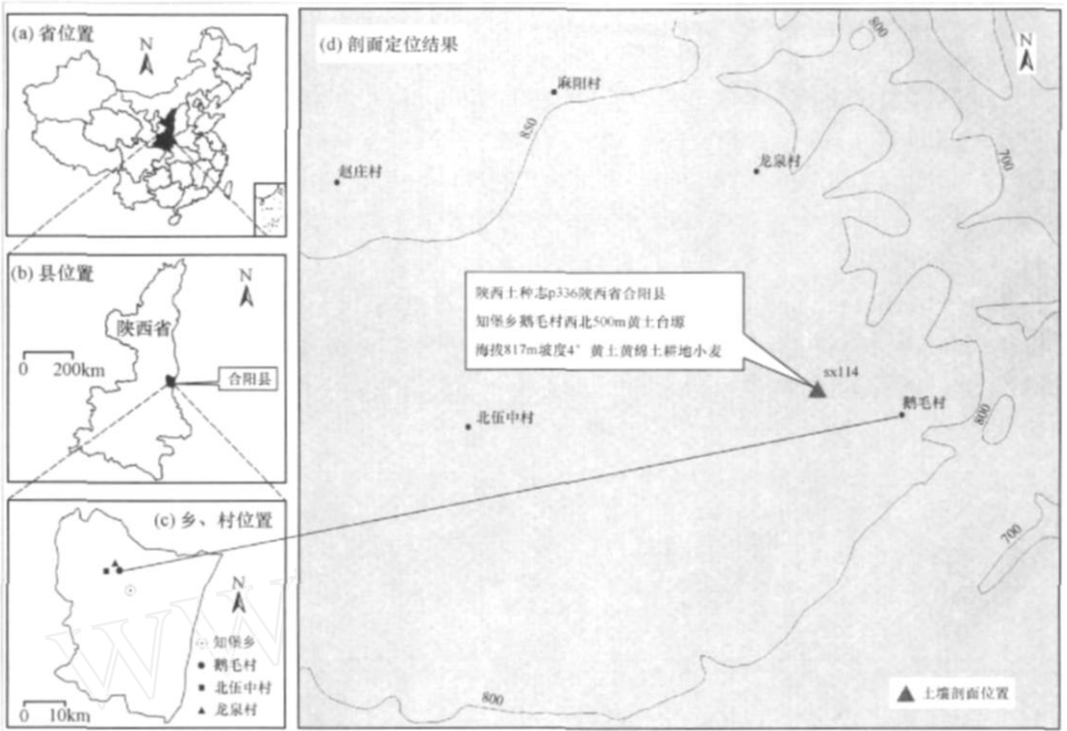


图1 典型土壤剖面定位过程示意图

Fig. 1 An example for process of spatial positioning of a typical soil profile

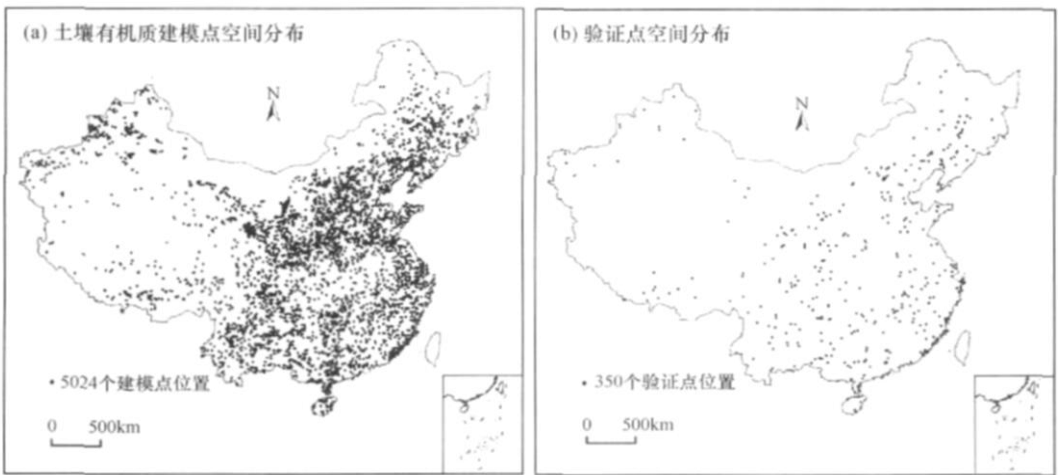


图2 全国5374个典型土壤剖面位置图(台湾省无数据)

Fig. 2 Map of the 5374 typical soil profiles in China(no data from Taiwan Province)

太阳辐射 (Solar radiation,  $SR$ ) 1 km分辨率空间分布 [图 3(a) ~ (e)]。基于全国 1 km DEM 数据,在 ArcGIS9.2中生成坡度 (Slope,  $S$ )、汇流面积 (Catchment area,  $CA$ )、地形指数 (Topographic index,  $TI$ ,  $TI = \ln(CA/S)$ ),其中汇流面积数值较大,对其作自然对数转换 [图 3(f) ~ (i)]。考虑到第二次土壤普查的时间,归一化植被指数 ( $NDVI$ )选择了能获取到的、分辨率较高的、与

土壤普查时间最接近的 1992 - 04至 1993 - 03共 36旬 1 km AVHRR 数据 (来源于 <http://edc2.usgs.gov/1KM/1kmhomepage.php>),将 36旬数据作平均获得全国 1 km 分辨率年平均 *NDVI*[图 3(j)],其取值为经线性拉伸至 [0, 200]区间的结果。人类对土地的利用在一定程度上与人口密度 (Density of population, *DP*)有关。人口密度在一定程度上能反映人类对土壤的影响。受地形、历史等原因影响,人口分布在空间上具有相对的稳定性。本研究采用与采样时间较接近的 1995年的 1 km人口密度分布数据 (来源于中国科学院资源环境数据中心)作为影响因素之一 [图 3(k)],以此代表人类活动的强度。成土母质虽然是土壤形成的主要物质来源,但在气候等因素的长期作用下,大多数土壤性质与母质有较大差异。而相同土壤类型与环境因素的关系却具有更大的相似性;同时考虑到典型剖面数目,本文将 12个土纲作为模型模拟时考虑的因子之一。土地利用方式也是影响土壤性质的重要因素,但受原始数据记录等方面的限制,研究中暂时未能考虑。

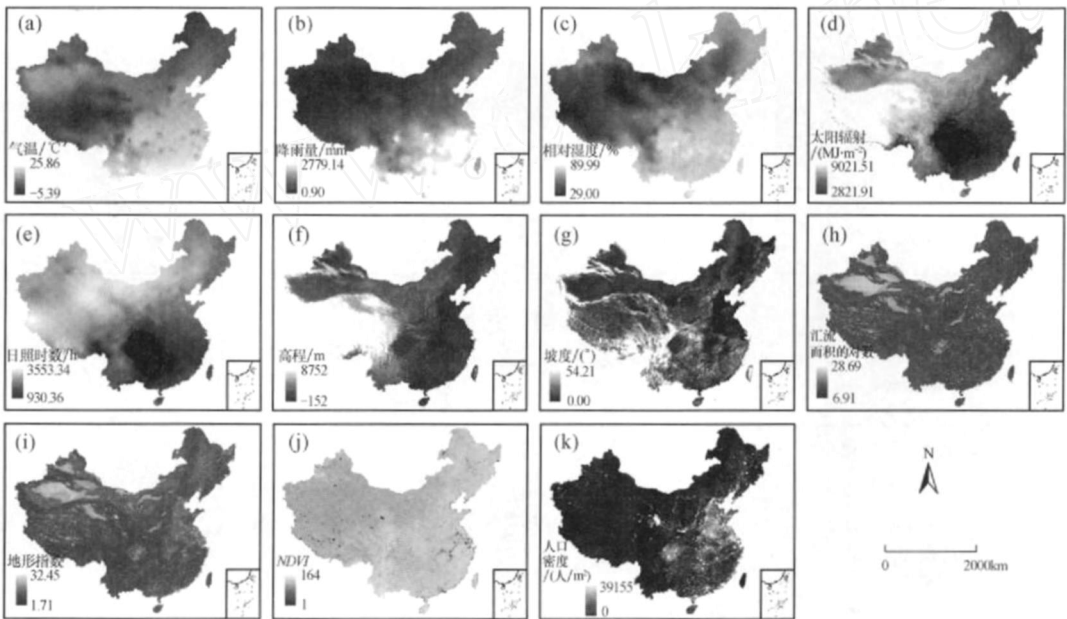


图 3 11 个环境影响因子空间分布

Fig. 3 Spatial distribution of 11 environmental factors

### 1.4 研究方法

#### 1.4.1 多元线性回归方法

回归分析能够反映地理要素之间具体的数量关系。在多要素的地理系统中,一个地理要素往往受到多个因素的影响。多元回归分析方法可建立表达某个地理要素与其影响因素间具体数量关系的数学模型。本文中采用多元线性回归来揭示各种环境因子综合作用与土壤有机质的具体数量关系。

#### 1.4.2 HASM 方法

在多元回归模拟的基础上,本文进一步用 HASM 方法来逼近真实的土壤有机质空间分布格局。HASM 方法基于曲面理论发展而来,现已有一套基于 MATLAB 的较成熟的程序,其方法推导过程可参阅相关文献<sup>[25-29]</sup>。HASM 方法最终可表达为等式约束的最小二乘问题:

$$\begin{cases} \min ZU^{n+1} - q^n \\ \text{s. t. } CU^{n+1} = d \end{cases} \quad (1)$$

式中:  $C$ 和  $d$ 分别为采样点系数矩阵和采样点的值,  $Z$ 、 $U^{n+1}$ 、 $q^n$ 为初始曲面上的参数。该式可理解为:将建模点的值看作土壤性质空间分布曲面上的真值,并以此真值作为目标,通过不断迭代,使得初始曲面不断逼近真实的曲面。

#### 1.4.3 普通克里格法

普通克里格法是一种最基本的克里格方法,它假设数据的变化呈正态分布,认为区域化变量的期望值是未知的,未采样点可以根据邻近区域中观测点数据进行加权平均计算获得。这种方法是目前土壤性质空间预测中实际应用最多的一种方法。

#### 1.4.4 模拟过程

在 ArcGIS9.2的地统计模块中将 5 374个典型剖面随机划分为两部分:5 024个建模点和 350个检验点(图 2);对建模点土壤有机质含量数据进行自然对数转换;利用建模点进行半方差分析并在此基础上进行普通克里格插值;在 SPSS14.0中按不同土纲分析土壤有机质与各影响因素的相关关系,建立不同土纲内表土有机质与其有显著影响的因素间的多元回归模型;利用 12个线性回归方程和全国各土纲空间分布数据,在 ArcGIS9.2中编程实现基于多元回归方法的土壤有机质空间分布模拟;以步骤 5的结果作初值,以建模点为约束条件,运用 HASM方法实现对土壤有机质空间分布真实曲面的进一步逼近。具体做法是:在 ArcGIS中将多元回归分析的结果转换为文本文件,然后在 MATLAB中运行 HASM方法的相关程序,得到的结果在 ArcGIS中转换为 Grid栅格格式。

将 5 024个建模点作为数据集 A,利用 ArcGIS9.2地统计模块中的 Create Subsets从数据集 A中随机取出 50%的样点,即 2 512个样点作为数据集 B,以探讨分析样点数减少一半的情况下各种方法模拟精度的变化。

### 1.5 验证方法

平均绝对误差反映模拟值与真值的平均偏离量;而平均相对误差可反映模拟值与真值的平均偏离程度,能更好地反映模型优劣。本研究以平均绝对误差和平均相对误差对 350个检验点的模拟值和实际观测值进行分析,检验各种方法模拟结果的精度。平均绝对误差(MAE)和平均相对误差(MRE)的表达式分别为:

$$MAE = \frac{1}{n} \sum_{i=1}^n | \hat{Z}_i - Z_i | \quad (2)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{| \hat{Z}_i - Z_i |}{Z_i} \times 100\% \quad (3)$$

式中:  $\hat{Z}_i$ 为模拟值,  $Z_i$ 为实际观测值,  $n$ 为检验样点数。显然,  $MAE$ 和  $MRE$ 越小,则误差越小,模拟精度越高。

## 2 结果与分析

### 2.1 描述性统计分析

基于对 5 024个样点的描述性统计分析表明,第二次土壤普查所得的中国表层(20 cm)土壤有机质含量的平均值为  $30.22 \text{ g} \cdot \text{kg}^{-1}$ (表 1)。其中最大值为  $678.90 \text{ g} \cdot \text{kg}^{-1}$ ,出现在四川省若尔盖草原的泥炭土中,隶属水成土土纲;最小值为  $0.20 \text{ g} \cdot \text{kg}^{-1}$ ,出现在新疆伽师县的龟裂土上,隶属初育土土纲。依据变异系数(Coefficient of Variation,  $C_v$ )的大小粗略估

计土壤有机质的变异程度:  $C_v < 10\%$  为弱变异性;  $C_v$  介于  $10\% \sim 100\%$  为中等变异性;  $C_v > 100\%$  属于强变异性<sup>[14]</sup>。表 1 中的统计结果表明, 中国表层土壤有机质的变异系数达到 134.58%, 表现出极强的空间变异性。偏度系数和峰度系数计算结果表明, 原始数据呈现明显的偏态分布。将原始数据作自然对数转换后, 数值分布接近正态分布 (表 1)。

表 1 土壤有机质统计特征 (单位:  $\text{g} \cdot \text{kg}^{-1}$ )

Table 1 Descriptive statistic of topsoil SOM (unit:  $\text{g} \cdot \text{kg}^{-1}$ )

数据集	样点数	数据预处理	最小值	最大值	平均值	标准差	变异系数 / %	偏度系数	峰度系数
A	5 024	原始数据	0.20	678.90	30.22	40.67	134.58	6.00	56.05
		对数变换后	-1.61	6.52	2.98	0.88	29.57	0.12	0.94
B	2 512	原始数据	0.30	678.90	30.52	42.70	139.90	6.44	63.56
		对数变换后	-1.20	6.52	2.98	0.89	29.82	0.16	0.85

注: A 为包含 5 024 个建模点的数据集, B 为包含 2 512 个建模点的数据集。下同。

样点数减少一半后, 表层土壤有机质的平均含量略有变化, 为  $30.52 \text{ g} \cdot \text{kg}^{-1}$ 。变异系数增加了 5.32%, 达到 139.90%, 仍属于极强的变异性。对数据作自然对数转换后, 数据分布接近标准的正态分布。

不同土纲的土壤有机质含量明显不同 (表 2)。水成土土纲有机质含量普遍较高, 平均值高达  $108.70 \text{ g} \cdot \text{kg}^{-1}$ , 居各土纲首位; 其次是高山土, 平均含量为  $63.89 \text{ g} \cdot \text{kg}^{-1}$ ; 有机质含量最低的是主要分布于西北的漠土和干旱土, 平均值分别为  $9.08 \text{ g} \cdot \text{kg}^{-1}$  和  $11.51 \text{ g} \cdot \text{kg}^{-1}$ 。各土纲土壤有机质变异程度不同, 其中, 漠土、干旱土、人为土、铁铝土属于中等程度的变异, 盐碱土、钙层土接近强变异性, 其它 6 个土纲为强变异性。各土纲中, 除漠土接近正态分布外, 其余 11 个土纲均呈偏态数据分布。

表 2 不同土纲土壤有机质含量的统计特征 (单位:  $\text{g} \cdot \text{kg}^{-1}$ )

Table 2 Descriptive statistic of topsoil SOM in each soil order (unit:  $\text{g} \cdot \text{kg}^{-1}$ )

土纲	样本数	最小值	最大值	平均值	标准差	变异系数 / %	偏度系数	峰度系数
淋溶土	457	4.50	528.30	50.70	59.30	116.97	3.48	17.84
半淋溶土	507	2.40	227.60	29.19	32.22	110.40	2.90	10.70
钙层土	363	2.30	212.80	25.82	25.12	97.29	3.21	14.66
干旱土	94	1.50	38.00	11.51	6.86	59.60	1.56	3.13
漠土	68	1.30	26.20	9.08	5.32	58.58	0.93	0.74
初育土	720	0.20	269.10	21.78	25.20	115.67	3.79	21.71
半水成土	836	0.82	282.10	22.52	30.07	133.53	4.09	21.30
水成土	109	0.80	678.90	108.70	148.25	136.39	1.88	3.03
盐碱土	159	1.10	79.70	13.63	11.26	82.61	2.22	7.37
人为土	1046	1.97	305.20	27.59	16.69	60.49	5.64	76.92
高山土	132	2.00	343.30	63.89	65.61	102.69	1.85	3.70
铁铝土	533	0.72	144.30	31.83	21.53	67.64	1.72	3.99

## 2.2 半方差分析

半方差分析是进行克里格插值的前提。分析结果表明 (表 3, 图 4), 两个样点密度下, 表土有机质均符合理论半方差模型中的球状模型, 模型决定系数分别为 0.931 和 0.945。模型中参数表明, 块金值与基台值之比分别为 0.560 和 0.467, 即表土有机质具有中等程度的

空间自相关性<sup>[14]</sup>。样点密度变化对半方差拟合结果影响不大,仅对基台值有一定的影响。

表 3 不同样点密度下土壤有机质的理论半方差模型及参数

Table 3 Parameters of theoretical semivariogram model of SOM under different sampling densities

数据集	模型	块金值 $C_0$	基台值 $C + C_0$	$C_0 / C_0 + C$	变程 /km	决定系数 $R^2$
A	球状模型	700.00	1 250.00	0.560	500.00	0.931
B	球状模型	700.00	1 500.00	0.467	500.00	0.945

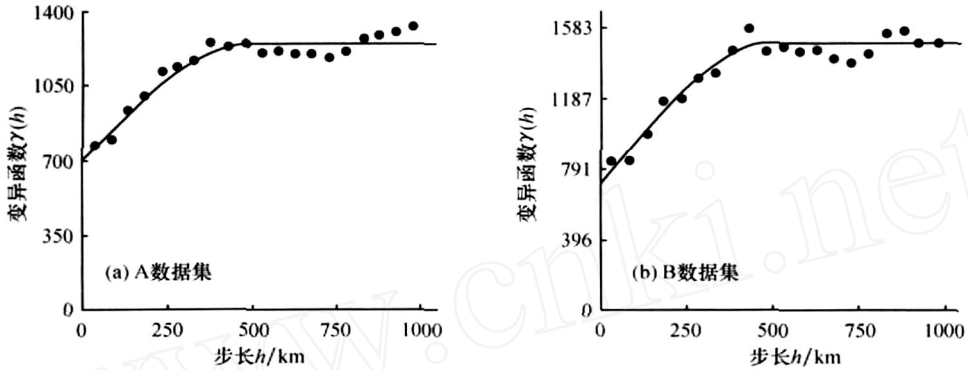


图 4 不同样点密度下土壤有机质的半方差

Fig. 4 Semivariogram map of SOM under different sampling densities

### 2.3 相关分析与多元回归模拟

国家尺度上,不同的土纲往往分布于一定的区域,如,铁铝土分布于我国降水量较大的南方,高山土分布于我国海拔较高的青藏高原,而漠土则分布于我国降水量较小的西北地区,等。不同区域内影响土壤过程的控制因子明显不同<sup>[26]</sup>。按土纲进行相关分析可以更容易找出特定条件下影响土壤性质的主要因素。对数据集 A 中各土纲土壤有机质与环境因素的相关分析结果表明(表 4),绝大多数土纲内都能找到与该土纲表土有机质显著相关( $p < 0.05$ )或极显著相关( $p < 0.01$ )的环境因素;不同的土纲内影响表土有机质含量的环境因素个数以及影响程度明显不同。这与相应土纲的成土环境要求及其分布面积等因素有关。如水成土土纲,该土纲主要为沼泽土和泥炭土,这两类土壤有机质含量较高;低气温高湿度有利于有机质的保存而不利其分解。因此水成土土纲表土有机质含量与气温和日照时数均呈现极显著的负相关关系,相关系数分别为 - 0.463 和 - 0.281;与高程和相对湿度呈极显著的正相关关系,相关系数分别为 0.381 和 0.396,其中与高程相关主要是因为海拔与气温联系在一起,一般地,气温随海拔升高而降低。

总的来看,影响因素对表土有机质含量总的作用趋势是(表 4):地形因子中高程和坡度以正相关为主,汇流面积和地形指数以负相关为主;气候因素中,温度因子主要起到负的作用,而湿度因子起到正的作用;植被以正相关为主;而人类活动强度不利于有机质的积累。总趋势外也有例外的情况,这仍与土纲所处的成土环境有关。如高山土主要分布于海拔较高的地方,高山土有机质与气温的关系不像多数土纲一样呈负相关,而是呈显著的正相关( $p < 0.05$ ),这可能是由于过低的温度不利于地表动植物的活动。样点数量减少一半后,各土纲表土有机质含量与影响因子间的相关系数有所改变,但趋势并没有大的变化(表 5)。

表 4 数据集 A 中各土壤有机质与影响因素间相关分析

Table 4 Correlation analysis between SOM and factors in each soil order in database A

土壤	淋溶土	半淋溶土	钙层土	干旱土	漠土	初育土	半水成土	水成土	盐碱土	人为土	高山土	铁铝土
样本个数	457	507	363	94	68	720	836	109	159	1 046	132	533
高程 DEM	0.586**	0.566**	0.278**	-0.047	-0.008	0.049	0.353**	0.381**	-0.032	-0.059	-0.267**	0.341**
坡度 S	0.411**	0.429**	0.242**	0.158	-0.097	0.353**	0.505**	0.325**	0.019	0.209**	0.12	0.427**
汇流面积 CA	-0.111*	0.006	0.001	-0.088	0.037	-0.106**	-0.100**	0.03	0.051	-0.099**	-0.232**	-0.116**
地形指数 TI	-0.296**	-0.175**	-0.026	-0.142	0.046	-0.248**	-0.278**	-0.149	0.043	-0.194**	-0.242**	-0.289**
气温 T	-0.467**	-0.519**	-0.608**	-0.230*	-0.240*	0.199**	-0.325**	-0.463**	0.054	0.231**	0.170*	-0.307**
降水 P	0.033	-0.089*	-0.325**	-0.132	-0.005	0.405**	0.161**	0.091	0.027	0.386**	0.305**	-0.001
相对湿度 RH	-0.075	-0.121*	0.280**	0.249*	0.243*	0.434**	0.105**	0.396**	0.009	0.424**	0.611**	0.065
日照时数 DS	-0.046	0.005	0.071	-0.066	-0.271*	-0.417**	-0.137**	-0.281**	-0.015	-0.415**	-0.632**	-0.120**
太阳辐射 SR	0.293**	0.364**	0.089	-0.250*	-0.126	-0.307**	0.044	-0.005	-0.021	-0.334**	-0.584**	-0.003
NDVI	0.082	-0.028	0.193**	0.228*	0.258*	0.199**	-0.007	0.112	0.076	0.081**	0.414**	0.201**
人口密度 DP	-0.182**	-0.116**	0.057	-0.122	0.217	0.02	0.003	-0.067	0.004	-0.071**	-0.116	-0.159**

注: \* 和 \*\* 分别表示在  $p < 0.05$  和  $p < 0.01$  水平下显著。下同。

表 5 数据集 B 中各土壤有机质与影响因素间相关分析

Table 5 Correlation analysis between SOM and factors in each soil order in database B

土壤	淋溶土	半淋溶土	钙层土	干旱土	漠土	初育土	半水成土	水成土	盐碱土	人为土	高山土	铁铝土
样本个数	229	257	186	47	36	377	396	54	80	510	73	267
高程 DEM	0.487**	0.607**	0.293**	-0.167	-0.052	0.037	0.421**	0.341*	-0.089	-0.023	-0.192	0.385**
坡度 S	0.378**	0.428**	0.296**	0.013	0.293	0.278**	0.526**	0.398**	0.133	0.204**	0.262*	0.375**
汇流面积 CA	-0.115	0.014	-0.054	0.041	-0.175	-0.056	-0.035	0.14	-0.004	-0.111*	-0.255*	-0.08
地形指数 TI	-0.263**	-0.173**	-0.092	0.025	-0.216	-0.165**	-0.220**	-0.058	-0.027	-0.201**	-0.310**	-0.233**
气温 T	-0.427**	-0.589**	-0.615**	0.016	-0.249	0.191**	-0.405**	-0.533**	0.015	0.200**	0.024	-0.274**
降水 P	0.025	-0.177**	-0.292**	-0.13	-0.082	0.396**	0.052	0.208	-0.014	0.375**	0.218	-0.015
相对湿度 RH	-0.024	-0.178**	0.304**	0.300*	0.410*	0.432**	0.04	0.442**	0.082	0.417**	0.577**	0.044
日照时数 DS	-0.073	0.016	0.046	-0.129	-0.353*	-0.402**	-0.08	-0.333*	0.016	-0.407**	-0.646**	-0.057
太阳辐射 SR	0.212**	0.433**	0.091	-0.309*	-0.229	-0.313**	0.103*	0.007	-0.02	-0.327**	-0.551**	0.066
NDVI	0.122	0.026	0.171*	0.441**	0.239	0.223**	-0.038	-0.023	-0.005	0.047	0.333**	0.146*
人口密度 DP	-0.330**	-0.1	0.046	-0.279	0.149	0.024	-0.024	-0.125	0.026	-0.052	-0.004	-0.199**



利用相关分析结果,选择与各土纲表土有机质显著相关以上或接近显著相关的环境因素,建立不同土纲内表土有机质含量与环境因素间的多元线性回归模型,其中土壤有机质采用自然对数数据变换后的数值。全国 12 个土纲表土有机质与环境因素间的线性回归模型及相应的模型检验结果如表 6、表 7 所示。表 6 是总样点数为 5 024 个样点时各土纲与环境因素间的线性回归方程及方程的决定系数和显著水平;表 7 是总样点数为 2 512 个样点时各土纲与环境因素间的线性回归方程及方程的决定系数和显著水平。

土壤性质空间分布模拟研究中,多元线性回归模型一般可以解释土壤性质变异的 20%~88%<sup>[28]</sup>,本研究结果处于这个范围内。从表 6 中可以看到,除盐碱土土纲外,其余 11 个土纲所建立的多元回归模型均达到显著或极显著水平。模型的决定系数多在 0.2~0.5 之间,平均值为 0.366,最大达到了 0.535。当样点数减少一半以后,各土纲模型的决定系数除钙层土和盐碱土外均有一定程度的下降,平均值为 0.347,减小了 0.019。从模拟的结果来看,模型所解释的变异仍然较少,得到的有机质含量空间分布曲面与真实的分布曲面仍将较大的差距。因此,在多元回归分析的基础上,进一步利用 HASM 方法来模拟表土有机质含量的空间分布格局。

表 6 数据集 A 中各土纲多元回归模型

Table 6 Regression models and their parameters for each soil order in database A

土纲	多元回归方程	决定系数 $R^2$	$P$
淋溶土	$\ln(SOM) = 5.603\ 981 + 0.000\ 494\ DEM - 0.021\ 216\ S + 0.072\ 376\ CA - 0.130\ 889\ TI - 0.077\ 648\ T - 0.000\ 304\ SR - 0.000\ 006\ DP$	0.512	0.000
半淋溶土	$\ln(SOM) = 0.640\ 842 + 0.000\ 232\ DEM + 0.019\ 249\ S - 0.005\ 320\ TI - 0.075\ 373\ T + 0.000\ 194\ P + 0.025\ 280\ RH + 0.000\ 196\ SR + 0.000\ 037\ DP$	0.426	0.000
钙层土	$\ln(SOM) = 1.261\ 645 - 0.000\ 046\ DEM + 0.066\ 465\ S - 0.176\ 252\ T - 0.001\ 411\ P + 0.052\ 320\ RH$	0.535	0.000
干旱土	$\ln(SOM) = 5.842\ 880 - 0.138\ 641\ T + 0.000\ 777\ RH - 0.000\ 562\ SR + 0.005\ 018\ NDVI$	0.252	0.000
漠土	$\ln(SOM) = 1.640\ 955 - 0.088\ 651\ T + 0.011\ 407\ RH - 0.000\ 630\ DS + 0.019\ 126\ NDVI + 0.000\ 127\ DP$	0.248	0.003
初育土	$\ln(SOM) = 0.098\ 677 - 0.001\ 088\ S + 0.125\ 925\ CA - 0.163\ 178\ TI - 0.082\ 045\ T + 0.000\ 803\ P + 0.029\ 189\ RH - 0.000\ 419\ DS + 0.000\ 243\ SR + 0.004\ 799\ NDVI$	0.365	0.000
半水成土	$\ln(SOM) = 4.269\ 508 + 0.000\ 053\ DEM + 0.080\ 643\ S - 0.010\ 018\ CA + 0.002\ 986\ TI - 0.117\ 738\ T + 0.000\ 664\ P + 0.006\ 726\ RH - 0.000\ 504\ DS$	0.440	0.000
水成土	$\ln(SOM) = 0.000\ 377\ DEM - 6.147\ 265 - 0.012\ 427\ S - 0.101\ 128\ T + 0.127\ 630\ RH + 0.000\ 885\ DS$	0.520	0.000
盐碱土	$\ln(SOM) = 0.970\ 231 + 0.000\ 017\ DEM + 0.013\ 222\ S + 0.022\ 611\ CA - 0.017\ 800\ TI + 0.024\ 574\ T + 0.000\ 160\ P + 0.000\ 483\ DS - 0.000\ 12\ SR + 0.002\ 551\ NDVI$	0.021	0.954
人为土	$\ln(SOM) = 2.176\ 376 + 0.000\ 132\ DEM + 0.008\ 773\ S + 0.000\ 946\ CA - 0.000\ 150\ TI - 0.045\ 887\ T + 0.000\ 477\ P + 0.011\ 931\ RH - 0.000\ 330\ DS + 0.000\ 123\ SR + 0.002\ 035\ NDVI - 0.000\ 016\ DP$	0.271	0.000
高山土	$\ln(SOM) = 4.119\ 678 - 0.000\ 102\ DEM - 0.064\ 619\ CA + 0.052\ 758\ TI + 0.028\ 626\ T - 0.000\ 287\ P + 0.021\ 087\ RH - 0.001\ 268\ DS + 0.000\ 082\ SR + 0.018\ 435\ NDVI$	0.482	0.000
铁铝土	$\ln(SOM) = 1.237\ 276 + 0.000\ 273\ DEM + 0.015\ 388\ S + 0.087\ 280\ CA - 0.122\ 353\ TI + 0.005\ 753\ T - 0.000\ 558\ DS + 0.021\ 021\ NDVI + 0.000\ 014\ DP$	0.322	0.000

表 7 数据集 B 中各土纲多元回归模型

Table 7 Regression models and their parameters for each soil order in database B

土纲	多元回归方程	决定系数 $R^2$	$p$
淋溶土	$\ln(SOM) = 6.134420 + 0.000486DEM - 0.008475S - 0.046453TI - 0.076987T - 0.000391SR - 0.000218DP$	0.437	0.000
半淋溶土	$\ln(SOM) = 1.670036 + 0.000284DEM + 0.022386S - 0.000699TI - 0.074768T - 0.000243P + 0.018825RH + 0.000105SR$	0.478	0.000
钙层土	$\ln(SOM) = -0.004580 + 0.000005DEM + 0.063754S - 0.176976T - 0.001350P + 0.051326RH + 0.009894NDVI$	0.551	0.000
干旱土	$\ln(SOM) = -4.194696 + 0.010399RH + 0.000030SR + 0.047904NDVI$	0.204	0.019
漠土	$\ln(SOM) = 1.564144 + 0.169223S + 0.021658RH - 0.000294DS$	0.231	0.037
初育土	$\ln(SOM) = -1.372440 + 0.044817S - 0.009797TI - 0.098512T + 0.000836P + 0.040263RH - 0.000454DS + 0.000350SR + 0.007343NDVI$	0.344	0.000
半水成土	$\ln(SOM) = 6.769252 + 0.000289DEM + 0.104942S - 0.003617TI - 0.077343T - 0.000670SR$	0.415	0.000
水成土	$\ln(SOM) = -8.543350 + 0.000211DEM + 0.062686S - 0.112544T + 0.140256RH + 0.001609DS$	0.535	0.000
盐碱土	$\ln(SOM) = 2.076483 - 0.000095DEM + 0.111342S + 0.008176TI + 0.000758RH$	0.030	0.678
人为土	$\ln(SOM) = 2.640654 + 0.014003S + 0.004664CA - 0.005102TI - 0.055008T + 0.000450P + 0.009276RH - 0.000536DS + 0.000255SR$	0.264	0.000
高山土	$\ln(SOM) = 6.539317 + 0.027414S - 0.070325CA + 0.088174TI + 0.006733RH - 0.001370DS - 0.000159SR + 0.012862NDVI$	0.455	0.000
铁铝土	$\ln(SOM) = 2.031376 + 0.000224DEM + 0.032316S - 0.015169TI - 0.017127T + 0.010832NDVI - 0.000059DP$	0.225	0.000

## 2.4 模拟结果分析

分别对数据集 A 和 B 的数据进行半方差分析,并在此基础上进行普通克里格插值。以多元回归分析结果为初值进行 HA SM 模型模拟,得到 2 种方法、2 种样点密度条件下中国表层土壤有机质含量的空间分布。

### 2.4.1 模拟图比较

图 5 是 2 种方法、2 种样点密度条件下中国表层土壤有机质含量空间分布的模拟结果。可以看出,两种方法模拟得到的中国表土有机质含量空间分布比较一致。其中,高值区分布于 3 个区域:分布面积最大的高值出现在西藏东部、青海省东南、甘肃南部和四川省西部地区。另一块分布面积相对较小的高值区则分布于东北地区内蒙古自治区的额尔古纳市、陈巴尔虎旗、牙克石市等以及黑龙江省北部的漠河县、塔河县、呼玛县和东部地区。除上述两个高值区外,在新疆的和静县、新源县、尼勒克县和特克斯县一带也出现一个高值区域。其中,极高值出现在四川省若尔盖县境内的若尔盖草原上,与描述性统计分析的结果一致。而低值区出现在东北地区的中部、华北平原、四川盆地丘陵区、黄土高原以及西北地区的大部分区域,其中极低值分布于西北的荒漠区。显然,模拟得到的中国表土有机质空间分布趋势基本符合实际情况。

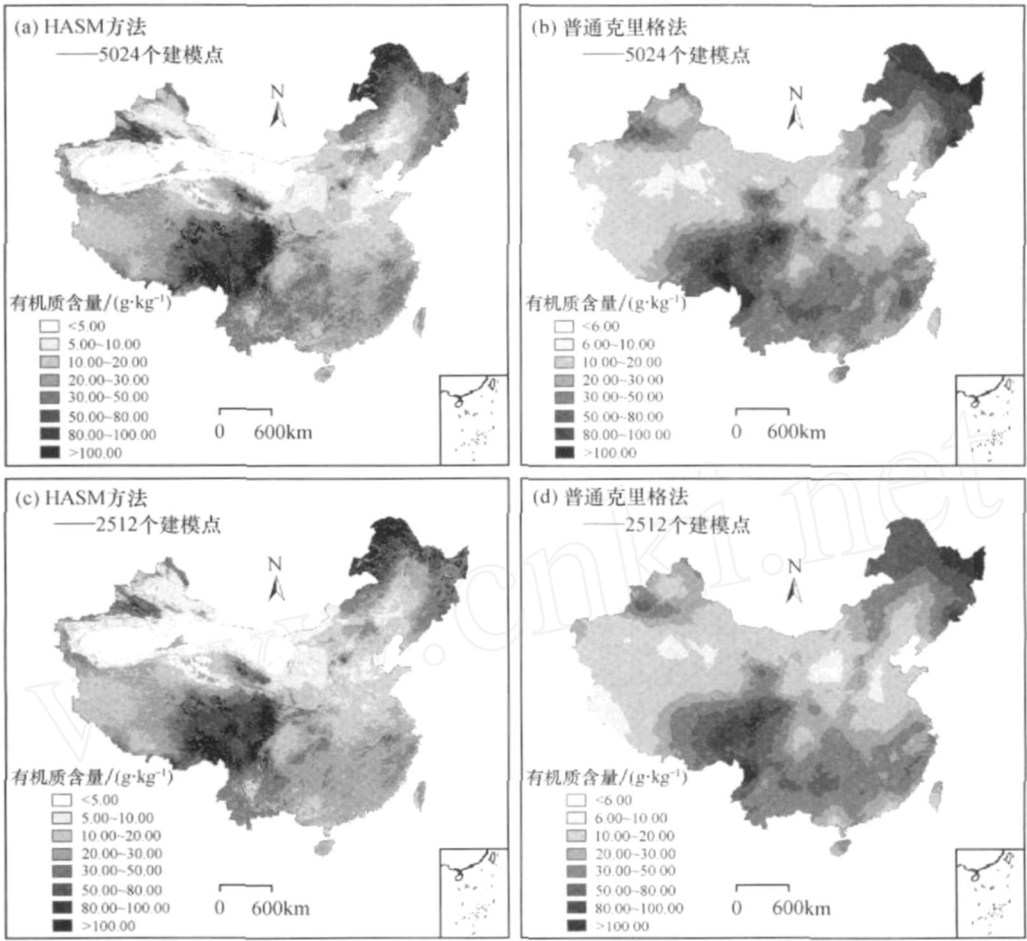


图 5 不同方法模拟的中国表土有机质含量空间分布格局

Fig. 5 Spatial distribution of topsoil SOM in China based on different methods

进一步分析两种方法得到的有机质含量空间分布图可以看出,两种模拟结果的最大差异出现在典型剖面数较少的西北地区和无样点分布数据的台湾省。多元回归与 HASM 模型结合方法得到的结果中,荒漠地区的值小于  $5.00 \text{ g} \cdot \text{kg}^{-1}$ ,明显低于其它地区。而普通克里格法对新疆南部荒漠区至西藏西部这一范围的模拟结果一致,大多处于  $5.00 \sim 20.00 \text{ g} \cdot \text{kg}^{-1}$  的范围内,这个结果显然不符合实际情况。台湾省中部山区植被较好,有机质含量较高,多元回归与 HASM 模型结合方法得到的结果显然更符合实际情况。究其原因,多元回归与 HASM 模型结合的方法充分考虑了环境因素的作用;而普通克里格法仅依据邻近采样点的信息来推测未采样点的值,这在样点密度较高时,可以得到较准确的结果。而西北地区样点密度极小并且分布极不均匀,绝大多数样点都分布于环境条件较好的地段,荒漠中极少有样点分布,这必然使得基于邻近点信息模拟的普通克里格法的结果偏高。对台湾省而言,普通克里格法对其有机质空间分布的模拟是利用较近的福建等省的样点以及这些样点距离插值点的远近来进行的,因而得到的结果是距离大陆近的区域模拟值较高。

将本研究中多元回归与 HASM 模型结合方法得到的有机质含量(质量分数)空间分布与解宪丽等<sup>[19]</sup>和于东升等<sup>[20]</sup>基于统计方法得出的全国土壤有机碳密度空间分布格局进行

比较,可以看出,高值区和低值区分布的位置与范围比较接近。

样点数减少一半的情况下,普通克里格法模拟结果变化较明显,三块高值区的分布面积明显减少。如东北的高值区只剩下东北的东部;四川西北部与西藏处的高值区分布面积明显减少;西北地区的低值区面积也有明显的减小[图5(b)、(d)]。相对而言,样点数变化对多元回归与HASM模型结合方法模拟结果的影响较小,仅在贵州、广西一带模拟值有一个较明显的降低,而其它区域的变化并不明显。

#### 2.4.2 模拟误差分析

对350个检验样点的模拟结果表明(表8),普通克里格法模拟结果的平均绝对误差为 $17.22 \text{ g} \cdot \text{kg}^{-1}$ ,多元回归与HASM模型结合的模拟结果为 $15.61 \text{ g} \cdot \text{kg}^{-1}$ ,比普通克里格法小 $1.61 \text{ g} \cdot \text{kg}^{-1}$ 。而多元回归与HASM模型模拟结果的平均相对误差MRE为56.59%,比普通克里格法减少了20.84%。对350个验证点模拟结果的相对误差进行分段统计表明,HASM方法预测结果的相对误差小于50%的样点数占全部样点的64.57%,而普通克里格法为52.28%;模拟结果较差的样点中( $MRE > 100\%$ ),HASM方法有35个点,而普通克里格法为78个,多出1倍多。

表8 不同方法模拟结果的误差分析

Table 8 Error analysis for spatial simulation of topsoil SOM based on different methods

数据集	研究方法	MAE/ ( $\text{g} \cdot \text{kg}^{-1}$ )	MRE/%	相对误差分段统计/点							
				<10%	10% ~ 30%	30% ~ 40%	40% ~ 50%	50% ~ 60%	60% ~ 80%	80% ~ 100%	>100%
A	HASM	15.61	56.59	50	105	41	30	28	29	32	35
	OK	17.22	77.43	42	76	34	31	32	37	20	78
B	HASM	15.75	57.76	49	96	40	36	36	30	27	36
	OK	17.88	81.25	40	68	32	38	34	39	22	77

注:MAE为平均绝对误差,MRE为平均相对误差。

样点数减少一半后,各种方法的预测精度有所降低,但降低程度并不大。HASM方法预测结果的平均绝对误差和平均相对误差分别增加了 $0.14 \text{ g} \cdot \text{kg}^{-1}$ 和1.07%;普通克里格法分别增加了 $0.66 \text{ g} \cdot \text{kg}^{-1}$ 和3.82%。这表明,样点数从2512个到5024个的成倍增加对预测精度的影响极小,更未出现成倍的变化;但样点数变化对具有不同理论基础的预测方法有不同程度的影响。比较样点数减少后两种方法的预测误差变化程度可以发现,样点数变化对基于要素空间自相关的普通克里格法的影响强于多元回归与HASM模型结合的方法。

### 3 结论与讨论

本研究主要结论如下:多元回归方法与HASM模型相结合,充分利用了两类模型的优点,在对中国国家尺度表层土壤有机质空间分布的模拟预测中取得了明显优于普通克里格法的结果;样点数的成倍变化未能引起模拟精度的大幅度改变,在现有样点数的基础上再增加采样有可能对模拟精度的提高作用不大。

以往中国国家尺度上土壤有机碳的研究多是采用分类统计的方法,同时也未做精度评价,无法与本研究进行对比。然而就本研究模拟结果来看,模拟结果误差仍然较大,有待于进一步提高。受数据的限制,研究过程中某些重要的因素没有利用。如人类的耕种活动使耕地土壤有机碳与温度、降水的相关性远低于非耕地土壤<sup>[29]</sup>。进一步提高模拟精度需要考

虑更多的因素,从而使模型能够解释更多的土壤有机质空间变异。

### 参考文献 (References):

- [1] Brodsky L, Vanek V, Bazakova M, *et al* The differences in the interpolation methods for mapping spatial variability of soil property [J]. *Rostlinna-Vyroba*, 2001, 47(12): 529-535.
- [2] 李雅琦, 田均良, 刘普灵. 黄土高原土壤元素含量地域分异规律 [J]. 西北农业学报, 2000, 9(3): 63-66. [LI Ya-qi, TAN Jun-liang, LU Pu-ling A study on laws of regional variance of soil element content in Loess Plateau through trend surface analysis method *Acta Agriculturae Boreali-occidentalis Sinica*, 2000, 9(3): 63-66.]
- [3] 霍霄妮, 李红, 张微微, 等. 北京耕作土壤重金属多尺度空间结构 [J]. 农业工程学报, 2009, 25(3): 223-229. [HUO Xiao-ni, LI Hong, ZHANG Wei-wei, *et al* Multi-scale spatial structure of heavy metals in Beijing cultivated soils *Transactions of the CSAE*, 2009, 25(3): 223-229.]
- [4] 赵安玖, 胡庭兴, 陈小红. 森林类型对土壤表层有机碳空间异质性的影响 [J]. 自然资源学报, 2009, 24(10): 1748-1756. [ZHAO An-jiu, HU Ting-xing, CHEN Xiao-hong Influence of subtropical forest ecosystems types on spatial heterogeneity of surface soil organic carbon *Journal of Natural Resources*, 2009, 24(10): 1748-1756.]
- [5] Kerry R, Oliver M A. Forest soil acidification assessment using principal component analysis and geostatistics [J]. *Geoderma*, 2007, 140: 374-382.
- [6] Meersmans J, Ridder F D, Canters F, *et al* A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium) [J]. *Geoderma*, 2008, 143: 1-13.
- [7] McKenzie N J, Ryan P J. Spatial prediction of soil properties using environmental correlation [J]. *Geoderma*, 1999, 89: 67-94.
- [8] Odeh IO A, McBratney A B, Chittleborough D J. Further result on prediction of soil properties from terrain attributes: Heterotrophic cokriging and regression kriging [J]. *Geoderma*, 1994, 67: 215-226.
- [9] Zhu A X, Hudson B, Burt J, *et al* Soil mapping using GIS, expert knowledge, and fuzzy logic [J]. *Soil Sci Soc Am. J.*, 2001, 65: 1463-1472.
- [10] 檀满枝, 陈杰. 模糊逻辑在土壤连续分类和制图表达中的应用及展望 [J]. 土壤学报, 2009, 46(1): 136-143. [TAN Man-zhi, CHEN Jie Application of fuzzy logic to soil continuous classification and cartographic representation of soils and its prospect *Acta Pedologica Sinica*, 2009, 46(1): 136-143.]
- [11] McBratney A B, Mendonca Santos M L, Minasny B. On digital soil mapping [J]. *Geoderma*, 2003, 117: 3-52.
- [12] Zhu A X, Band L, Vertessy R, *et al* Derivation of soil properties using a Soil Land Inference Model (SoLIM) [J]. *Soil Sci Soc Am. J.*, 1997, 61(2): 523-533.
- [13] 张慧智, 史学正, 于东升, 等. 中国土壤温度的空间预测研究 [J]. 土壤学报, 2009, 46(1): 1-8. [ZHANG Hui-zhi, SHI Xue-zheng, YU Dong-sheng, *et al* Spatial prediction of soil temperature in China *Acta Pedologica Sinica*, 2009, 46(1): 1-8.]
- [14] 王绍强, 朱松丽, 周成虎. 中国土壤土层厚度的空间变异性特征 [J]. 地理研究, 2001, 20(2): 161-168. [WANG Shao-qiang, ZHU Song-li, ZHOU Cheng-hu Characteristics of spatial variability of soil thickness in China *Geographical Research*, 2001, 20(2): 161-168.]
- [15] 潘根兴. 中国土壤有机碳和无机碳库量的研究 [J]. 科技通报, 1999, 15(5): 330-332. [PAN Gen-xing Study on carbon reservoir in soils of China *Bulletin of Science and Technology*, 1999, 15(5): 330-332.]
- [16] 王绍强, 周成虎, 李克让, 等. 中国土壤有机碳库及空间分布特征分析 [J]. 地理学报, 2000, 55(5): 533-543. [WANG Shao-qiang, ZHOU Cheng-hu, LI Ke-rang, *et al* Analysis on spatial distribution characteristics of soil organic carbon reservoir in China *Acta Geographica Sinica*, 2000, 55(5): 533-543.]
- [17] Ni J. Carbon storage in terrestrial ecosystems of China: Estimates at different spatial resolutions and their responses to climate change [J]. *Climate Change*, 2001, 49: 339-358.
- [18] 李克让, 王绍强, 曹明奎. 中国植被和土壤碳贮量 [J]. 中国科学 D 辑, 2003, 33(1): 72-78. [LI Ke-rang, WANG Shao-qiang, CAO Min-kui Vegetation and soil carbon storage in China *Science in China: Series D*, 2003, 33(1): 72-80.]
- [19] 解宪丽, 孙波, 周慧珍, 等. 中国土壤有机碳密度和储量的估算与空间分布分析 [J]. 土壤学报, 2004, 41(1):

- 35-43. [XIE Xian-li, SUN Bo, ZHOU Hui-zhen, *et al* Organic carbon density and storage in soils of China and spatial analysis *Acta Pedologica Sinica*, 2004, 41(1): 35-43. ]
- [20] 于东升, 史学正, 孙维侠, 等. 基于 1:100 万土壤数据库的中国土壤有机碳密度及储量研究 [J]. 应用生态学报, 2005, 16(12): 2279-2283. [YU Dong-sheng, SHI Xue-zheng, SUN Wei-xia, *et al* Estimation of China soil organic carbon storage and density based on 1:1000000 soil database. *Chinese Journal of Applied Ecology*, 2005, 16(12): 2279-2283. ]
- [21] Yue T X, Chen S P, Xu B, *et al* A curve-theorem based approach for change detection and its application to Yellow River delta [J]. *International Journal of Remote Sensing*, 2002, 23(11): 2283-2292.
- [22] Yue T X, Du Z P, Song D J, *et al* A new method of surface modeling and its application to DEM construction [J]. *Geomorphology*, 2007, 91(1): 161-172.
- [23] 岳天祥, 杜正平, 宋敦江. 高精度曲面建模: HASM4 [J]. 中国图象图形学报, 2007, 12(2): 343-348. [YUE Tian-xiang, DU Zheng-ping, SONG Dun-jiang High accuracy surface modelling: HASM4. *Journal of Image and Graphics*, 2007, 12(2): 343-348.
- [24] Shi W J, Liu J Y, Du Z P, *et al* Surface modelling of soil pH [J]. *Geodema*, 2009, 150: 113-119.
- [25] Shi X Z, Yu D S, Wamer E D, *et al* Soil database of 1:1,000,000 digital soil survey and reference system of the Chinese genetic soil classification system [J]. *Soil Survey Horizons*, 2004, 45: 129-136.
- [26] Jenny H. Factors of Soil Formation, A System of Quantitative Pedology [M]. McGraw-Hill, New York, 1941.
- [27] Herrick J E, Whitford W G Assessing the quality of rangeland soils: Challenges and opportunities [J]. *Journal of Soil and Water Conservation*, 1995, 50: 237-242.
- [28] Thompson J A, Eugenia M, Pena Y, *et al* Soil-landscape modeling across a physiographic region: Topographic patterns and model transportability [J]. *Geodema*, 2006, 133: 57-70.
- [29] 周涛, 史培军, 王绍强. 气候变化及人类活动对中国土壤有机碳储量的影响 [J]. 地理学报, 2003, 58(5): 727-734. [ZHOU Tao, SHI Pei-jun, WANG Shao-qiang Impacts of climate change and human activities on soil carbon storage in China *Acta Geographica Sinica*, 2003, 58(5): 727-734. ]

## Study on Method for Spatial Simulation of Topsoil SOM at National Scale in China

LI Qi-quan<sup>1, 2</sup>, YUE Tian-xiang<sup>1</sup>, FAN Ze-meng<sup>1</sup>, DU Zheng-ping<sup>1</sup>

(1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;

2. Graduate University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** Given the importance of soil organic carbon (SOC) as a pool in the global carbon cycle and an indicator for soil quality, there exists a need to simulate this soil property at large scale (regional or national). However, few researches focus on simulating spatial distribution of SOC at national scale in China by using model and samples. In this paper, based on 5374 typical soil profiles collected during the second national soil survey period (1979 - 1994), correlation between topsoil organic matter content (20 cm) and 11 environmental factors were analyzed, spatial distribution of topsoil organic matter (SOM) at national scale in China was simulated with the combination of multiple regression model and HASM model, and prediction error of this method was discussed, in order to provide a new method for spatial simulation of soil organic carbon. Results indicated that, mean absolute error and mean relative error of the predicted value for 350 validation points were  $15.61 \text{ g} \cdot \text{kg}^{-1}$  and 56.59%; compared with ordinary Kriging method, the two errors were reduced by  $1.61 \text{ g} \cdot \text{kg}^{-1}$  and 20.84% respectively. Besides, simulation result for Northwest China and Taiwan Province, where the density of sample points was much smaller and even no samples distributed, was much closer to the actual situation. When the samples were cut by half, the two errors were only increased by  $0.14 \text{ g} \cdot \text{kg}^{-1}$  and 1.07% respectively. Consequently, the method in this paper can be used as a relatively effective method for simulating spatial distribution of SOM at national scale, and attaining higher levels of precision largely depend on making the model explain much more spatial variability of SOM as well as increasing the number of sampling sites used to establish the model.

**Key words:** HASM method; soil organic matter; national scale; sampling density; spatial simulation