

doi:10.6041/j.issn.1000-1298.2015.01.020

基于环境因子和 R-STPS 的林地土壤有机质预测模型*

刘二永¹ 刘健^{2,3} 余坤勇^{2,3} 何平^{2,3} 赵振贺³

(1. 福建农林大学资源与环境学院, 福州 350002; 2. 福建农林大学 3S 技术应用研究所, 福州 350002;

3. 福建农林大学林学院, 福州 350002)

摘要: 研究了基于环境因子和混合插值的林地土壤有机质预测模型。首先应用数字地形与遥感影像分析技术获取地形因子与遥感指数, 然后分析土壤有机质与环境因子的相关性, 最后用环境因子对土壤有机质进行空间预测。针对回归克里格法(RK)需要计算半变异函数的缺陷, 提出了一种空间插值方法, 即回归-光滑薄板样条插值法(R-STPS)。将这2种插值方法用于顺昌县土壤有机质的空间预测。结果表明, RK与R-STPS的预测精度、计算效率、预测的研究区土壤有机质空间分布的总体趋势相近。R-STPS无需计算半变异函数, 使用方便, 因此更有优势。

关键词: 林地 土壤有机质 预测模型 环境因子 混合插值 回归克里格法

中图分类号: S158.2 **文献标识码:** A **文章编号:** 1000-1298(2015)01-0133-05

Spatial Prediction of Forest Soil Organic Matter Based on Environmental Factors and R-STPS Interpolation Methods

Liu Eryong¹ Liu Jian^{2,3} Yu Kunyong^{2,3} He Ping^{2,3} Zhao Zhenhe³

(1. College of Resource and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China

2. Institute of Geomatics Application, Fujian Agriculture and Forestry University, Fuzhou 350002, China

3. College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract: The spatial prediction model of forest soil organic matter was studied based on environmental factors and mixed interpolation methods. Firstly, digital terrain and remote sensing image analysis technologies were applied to get topographic factors and index of remote sensing. Then, the correlation of soil organic matter and environmental factors was analyzed. In the end, soil organic matter was predicted spatially according to the environmental factors. Aiming at the flaw of regression Kriging (RK) which needs to compute semi-variogram, a spatial interpolation method named regression-smoothing thin plate spline (R-STPS) was presented. This two interpolation methods were applied to predict soil organic matter of Shunchang county spatially. The results showed that the prediction accuracy and computation efficiency of RK and R-STPS were almost consistent. The overall trend of spatial prediction distribution of soil organic matter of study area was similar. However, R-STPS was not needed for calculation of semi-variogram and easy to use. Therefore, R-STPS has more advantages.

Key words: Forest land Soil organic matter Prediction model Environmental factors Mixed interpolation Regression Kriging

引言

快速准确地获取土壤性质的空间分布是精准农

业和精准林业的基础。因此,对土壤属性空间异质性的研究方法——空间插值技术研究得到了广泛关注和重视。土壤有机质是土壤最具代表性的属性之

收稿日期: 2014-05-21 修回日期: 2014-07-08

* 国家自然科学基金资助项目(40971043)、国家林业局引进国际先进林业科学技术资助项目(2013-4-70)和福建农林大学青年教师科研基金资助项目(2013XJJ15)

作者简介: 刘二永, 讲师, 主要从事土壤养分遥感反演以及 DEM 精度分析研究, E-mail: cumtley@hotmail.com

通讯作者: 刘健, 教授, 博士生导师, 主要从事森林经营管理与 3S 技术应用研究, E-mail: fjlujian@126.com

一,是评价土壤肥力和土壤质量的一项重要指标,在环境保护、农业可持续发展等方面起着极其重要的作用。由于成土母质、地形、植被以及人类耕作等的影响,土壤有机质在空间分布上往往呈现较大的空间变异性。本文以土壤有机质为例展开研究^[1]。

研究表明,量化的环境因子和土壤属性之间存在很好的相关性,可以用来预测土壤属性。随着3S技术的发展,环境相关法日益受到重视,越来越多的研究倾向于利用更为详尽的辅助变量来指导土壤制图和相关属性空间分布。近年来,结合克里格法和辅助数据进行土壤属性空间分布预测的混合插值技术得到了快速发展,结合回归分析和普通克里格法进行预测的回归克里格法(Regression Kriging, RK)是其中最有效的方法之一^[2]。RK的基本原理是:首先以环境因子作为自变量,通过线性拟合后采用普通克里格法来预测残差,最后将拟合趋势和残差求和可得RK的土壤属性预测值。然而,RK中的普通克里格法在插值之前需要先计算半变异函数,并且半变异函数的拟合曲线的选择会受主观因素影响较大,进而会影响预测精度。光滑薄板样条插值(Smooth thin plate spline, STPS)已被证实是一种有效且精确的空间插值方法^[3],并且不用计算半变异函数。本文针对RK需要计算半变异函数这个缺陷,提出一种混合插值方法回归-光滑薄板样条插值法(Regression-smooth thin plate spline, R-STPS)。结合3S技术,利用数字地形和遥感影像分析技术,提取相关地形因子与遥感指数,分析土壤有机质与环境因子之间的关系,并利用这些环境因子分别采用RK和R-STPS对福建省南平市顺昌县土壤有机质空间分布进行预测,从预测精度、计算效率以及空间分布趋势预测效果3个方面进行对比分析^[1]。

1 材料与方法

1.1 试验区概况

选取福建省南平市顺昌县作为试验区。顺昌县位于福建省中部偏北,介于117°30'E~118°14'E、26°39'N~27°29'N之间,总面积1992 km²。境内以低山、高丘的地貌类型为主,出露地层主要是前震旦系和震旦系变质岩。该区属中亚热带海洋性季风气候,年平均气温19℃,土壤以山地红壤为主,腐殖质层1~5 cm,平均土层厚度1 m以上。气候土壤资源得天独厚,是杉木生长繁衍最佳地区,森林覆盖率82.8%,绿化程度95.1%,超发达国家水平,素有“林海粮仓果乡”之美誉。

1.2 数据采集

土壤样品采集于2010年9月。此次采集了顺

昌境内的80个国家森林资源连续清查样地(一类样地),样地间距4 km×6 km;30个福建省森林资源规划设计调查样地(二类样地),样地间距1 km×2 km;21个经营小班。沿样地对角线或小班中心位置重复取3个土样,取样深度为0~30 cm,并记录GPS坐标,共采集样品131个,采样点的分布如图1所示。土壤样品用铝盒带回实验室,风干、研磨、过筛后,用重铬酸钾氧化外加加热法测定有机质含量^[4]。

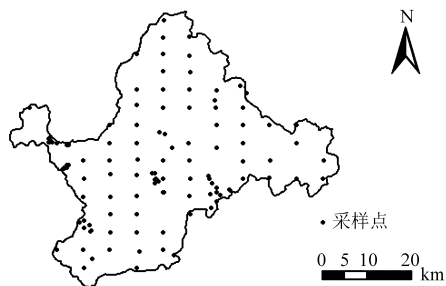


图1 采样点分布

Fig. 1 Distribution map of sample points

1.3 地形因子提取

利用县域1:50 000地形图所生成的DEM(栅格为25 m×25 m),通过数字地形分析技术获取相关地形因子。地形因子包括相对高程(H_r)、坡度(β)、垂直曲率(C_v)、水平曲率(C_h)、地形起伏度(Q_{FD})、坡度变率(S_{OS})、地面粗糙度(M)、地形湿度指数(ψ)、河流动能指数(Ω)和坡度坡长因子(L_s)。其中地形湿度指数、河流动能指数和坡度坡长因子对地表径流、植物的生长和分布等水文、生态过程产生影响,进而影响土壤属性空间分布,各地形因子的计算方法参照文献[5]。

1.4 遥感指数提取

本文用到的遥感影像为顺昌县2008年9月上旬的ALOS影像,空间分辨率为2.5 m,包含4个波段,其波长范围依次为0.42~0.50 μm (蓝光),0.52~0.60 μm (绿光),0.61~0.69 μm (红光),0.76~0.89 μm (近红外)。本文保持野外土壤样本采样时间与遥感影像获取时间在月份上的一致性,即季节一致性,从而最大程度地降低由时间一致性所带来的误差^[6]。归一化植被指数(NDVI)是最常用的反映植被覆盖的植被指数,本文以ALOS影像为数据源提取NDVI,用于土壤有机质的预测。

1.5 研究方法

1.5.1 回归克里格法原理

假设土壤变量的观察值可表示为: $z(s_1)$, $z(s_2)$, ..., $z(s_n)$,这里 $s_i(x_i, y_i)$ 是空间位置, n 是观测点的个数。对于回归克里格法,未知点 $s_0(x, y)$ 的土壤特征可以通过估计趋势和残差求得^[7],即

$$\hat{z}(s_0) = \hat{m}(s_0) + \hat{e}(s_0) \quad (1)$$

趋势 \hat{m} 通常采用线性回归分析来拟合, 残差 \hat{e} 采用普通克里格法 (OK) 插值求得, 则式 (1) 可表示为

$$\hat{z}(s_0) = \sum_{k=0}^p \hat{\beta}_k q_k(s_0) + \sum_{i=1}^n w_i(s_i) e(s_i) \quad (2)$$

其中

$$q_0(s_0) = 1$$

式中 $\hat{\beta}_k$ ——估计趋势的模型系数

$q_k(s_0)$ ——点 s_0 处的第 k 个外部解释变量
(本文中对应的是环境因子)

$w_i(s_i)$ ——由残差的变异函数决定的权重

$e(s_i)$ ——回归残差

1.5.2 回归-光滑薄板样条插值法原理

RK 中, 残差采用的是普通克里格法插值求得, 而普通克里格法需要首先求出半变异函数然后才能进行插值。光滑薄板样条插值 (STPS) 可以避免计算半变异函数。本文提出的回归-光滑薄板样条插值法 (R-STPS) 是首先以环境因子作为自变量进行线性拟合, 然后采用 STPS 插值求残差, 最后将拟合趋势和残差求和即可得最后的土壤属性预测值。

STPS 的基本原理是拟合的曲面应尽可能通过或接近采样点的值并且尽可能光滑, 计算式为^[3]

$$y = \min \sum_{i=1}^n (z_i - f_i)^2 + \lambda J_d^l(f) \quad (3)$$

式中 z_i, f_i ——采样点处的量测值、拟合值

λ ——控制拟合程度和光滑程度的光滑因子

本文中的光滑因子 λ 的最优值是通过广义交叉验证 (GCV) 方法求得, GCV 方法采用依次移去一个采样点, 用剩余采样点在一定的光滑因子下进行曲面拟合得到该点的估计值, 再计算观测值与估计值的方差, 最小方差对应的 λ 即为最优的光滑因子^[8]; $J_d^l(f)$ 是根据 d 维空间中 l 阶偏导数构造的函数 f 的粗糙度, 有

$$J_d^2(f) = \iint \left[\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right] dx dy \quad (4)$$

式 (3) 中, 对函数 f 的估计采用线性组合的形式^[9], 即

$$\hat{f}(s_0) = \sum_{j=1}^3 a_j \phi_j(s_0) + \sum_{i=1}^n b_i \psi(h_i) \quad (5)$$

其中 $\phi_1 = 1$ $\phi_2 = x$ $\phi_3 = y$ $\psi(h) = h^2 \ln h$

$$h_i = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

式中 h_i —— s_0 和 s_i 之间的距离

a_j, b_i ——系数, 可通过求解 n 阶的线性方程组求得

1.5.3 数据分析及验证

先剔除采样点中的异常数据, 共 5 个, 从剩下的

126 个样点数据中随机抽取 110 个样点用来建立预测模型, 用剩余的 16 个点作为验证数据集。通过比较验证数据点位置上土壤有机质的实际观测值和估计值可以评价预测精度, 本文采用验证数据集的平均误差 (ME) 和均方根误差 (RMSE) 来评价预测的准确性^[2]。

2 结果和分析

2.1 描述性统计分析

对剔除异常值后的土壤有机质数据进行统计分析, 结果如表 1 所示。从表 1 可以看出, 变异系数为 43%, 所以顺昌县土壤有机质属于中等变异。经偏度和峰度检验, 样本数据基本服从正态分布。

表 1 顺昌县土壤有机质统计特征值
Tab.1 Statistic characteristic values of soil organic matter in Shunchang county

参数	数值	参数	数值
最大值/(g·kg ⁻¹)	95.24	变异系数/%	43
最小值/(g·kg ⁻¹)	9.67	峰度	-0.656
均值/(g·kg ⁻¹)	44.38	偏度	0.4

2.2 土壤有机质与环境因子相关性分析

表 2 为研究区土壤有机质与环境因子间的相关性。从表 2 可以看出, 有机质与坡度 (β)、垂直曲率 (C_v)、地形起伏度 (Q_{FD})、坡度变率 (S_{OS})、地面粗糙度 (M) 和植被指数 (NDVI) 具有较显著的相关关系。

表 2 土壤有机质和环境因子间的相关性
Tab.2 Correlations between environmental attributes and soil organic matter

环境因子	相关系数	环境因子	相关系数
H_r	-0.029	M	0.256
β	0.244	ψ	-0.068
C_v	-0.126	Ω	-0.057
C_h	0.022	L_s	-0.023
Q_{FD}	0.307	NDVI	0.120
S_{OS}	0.136		

2.3 混合插值预测过程

表 3 为应用逐步回归分析拟合土壤有机质的多元线性回归方程。从表 3 可以看出, 坡度 (β) 和地形起伏度 (Q_{FD}) 是预测土壤有机质的最优因子。从多元线性回归模型可以看出, 拟合方程的决定系数不高, 因此需要对残差进一步分析。

首先采用普通克里格法 (OK) 对残差进行插值。图 2 为土壤有机质的多元线性回归残差的半方差模型。从图 2 中可以看出, 土壤有机质的半方差函数表现出孔穴效应, 孔穴效应的产生主要由于区域化变量周期性变化引起的, 说明研究区土壤有机质存

在周期性的变化^[10]。采用图2中的半方差模型对残差进行普通克里格法插值,然后将回归分析预测结果和普通克里格法对残差的预测结果求和,就是RK的预测结果。

表3 土壤有机质多元回归方程逐步拟合过程

Tab.3 Stepwise process of multiple regression equation of soil organic matter

环境因子	回归系数	环境因子	回归系数
H_r		ψ	
β	-1.013 6	Ω	
C_v		L_s	
C_h		NDVI	
Q_{FD}	1.202 7	常数项	32.689 7
S_{OS}		R^2	0.13
M			

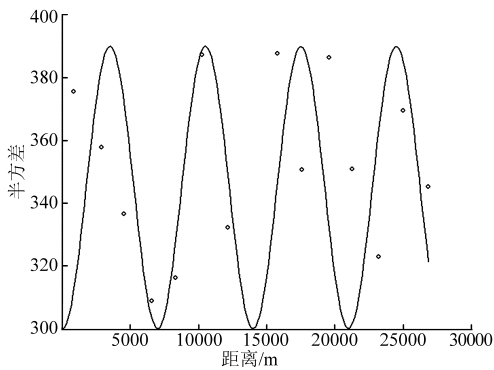


图2 土壤有机质的半方差函数

Fig.2 Semivariogram of soil organic matter

进一步采用STPS方法对残差进行插值,然后将回归分析预测结果和STPS方法对残差的预测结果求和,就是R-STPS的预测结果。

2.4 精度分析

表4为采用RK、R-STPS 2种插值方法对土壤有机质的预测精度。可以看出,RK与R-STPS的预

测精度基本相当,ME相差0.094, RMSE相差0.059。因此,2种插值方法对研究区土壤有机质的整体特征的描述能力基本相当。然而,R-STPS方法无需计算半变异函数,使用方便。另外,R-STPS仅需少量采样点即可计算,而RK需较多采样点才能计算半变异函数,因此R-STPS更有优势。

表4 2种插值方法的土壤有机质预测精度

Tab.4 Prediction accuracy of soil organic matter of two interpolation methods

指标	RK	R-STPS
ME	-3.729	-3.823
RMSE	18.743	18.684

2.5 土壤有机质空间分布

图3为采用以上2种插值方法预测的以50 m分辨率的栅格格式显示并输出的土壤有机质空间分布图。本文中插值过程采用的是R软件,然后将插值结果文本文件导入ArcGIS 10.0进行制图。计算机配置为: Intel (R) Core (TM) i5-2400 CPU @ 3.10 GHz 3 GB内存。RK方法和R-STPS的计算效率基本相当,分别为28.06 s和30.27 s。从图3中可以看出,2种插值方法预测的研究区土壤有机质空间分布的总体趋势相似,不同的插值方法预测结果的高值区和低值区分布位置大致相同,土壤有机质含量呈西高东低的趋势,具体表现为西部最高,南北部和中部次之,东部最低^[4]。因此R-STPS对土壤有机质的预测比RK更有优势。

3 结论

(1) 本文选取福建省南平市顺昌县作为研究区,利用环境因子与混合插值方法对该县的土壤有机质空间分布进行了预测。针对RK需要计算半变

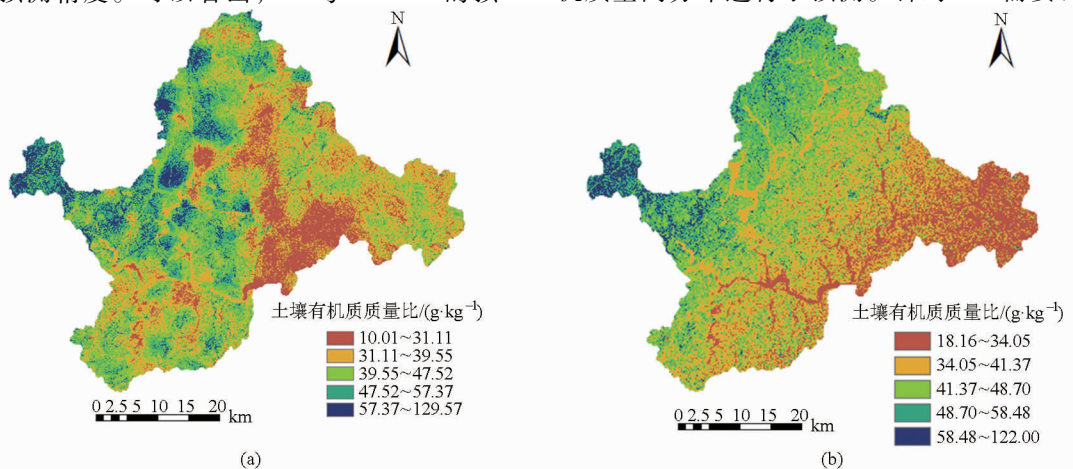


图3 2种方法预测的土壤有机质空间分布

Fig.3 Distribution maps of soil organic matter of two interpolation methods

(a) RK (b) R-STPS

异函数的缺陷,提出了 R-STPS 方法。将这 2 种方法用于土壤有机质预测。

(2)结果表明,RK 和 R-STPS 对土壤有机质的整体特征的描述能力、计算效率以及对土壤有机质空间分布的总体趋势的描述基本相近。但 R-STPS 无需计算半变异函数,仅需少量采样点即可计算,因此,R-STPS 对土壤有机质的预测比 RK 更有优势。

(3)本文中仍有一些问题值得进一步关注:选取的环境因子较多,难免有相关性,通过主成分分析消除数据冗余,进一步提升预测精度。从表 1 中可

以看出,样本仅仅近似服从正态分布,即使采用变换方法也很难将样本转换成正态分布,因此需要采用稳健估计方法来进行预测以提升预测精度。本研究是采用线性模型来描述土壤有机质与环境因子的关系,模型过于简单,很难描述因素之间的真实关系。因此,需要建立具有非线性映射能力的模型,比如神经网络模型。可进一步将研究区按照不同的土地利用类型、土壤类型、植被类型进行分区,分区进行插值以提高土壤有机质的预测精度。

参 考 文 献

- 李启权,王昌权,岳天祥,等. 基于 RBF 神经网络的土壤有机质空间变异研究方法[J]. 农业工程学报,2010, 26(1): 87-93.
Li Qiquan, Wang Changquan, Yue Tianxiang, et al. Method for spatial variety of soil organic matter based on radial basis function network[J]. Transactions of the CSAE, 2010, 26(1): 87-93. (in Chinese)
- 连刚,郭旭东,傅伯杰,等. 基于环境相关法和地统计学的土壤属性空间分布预测[J]. 农业工程学报,2009, 25(7):237-242.
Lian Gang, Guo Xudong, Fu Bojie, et al. Prediction of the spatial distribution of soil properties based on environmental correlation and geostatistics[J]. Transactions of the CSAE, 2009, 25(7):237-242. (in Chinese)
- Chen C F, Li Y Y. A robust method of thin plate spline and its application to DEM construction[J]. Computer & Geosciences, 2012, 48: 9-16.
- 杨子清,陈平留,刘健,等. 基于 Kriging 法的森林土壤养分空间插值[J]. 福建农林大学学报:自然科学版,2012,41(3): 296-300.
Yang Ziqing, Chen Pingliu, Liu Jian, et al. Spatial interpolation of forest soil nutrients based on Kriging method[J]. Journal of Fujian Agriculture and Forestry University: Nature Sciences, 2012,41(3):296-300. (in Chinese)
- 张素梅,王宗明,张柏,等. 利用地形和遥感数据预测土壤养分空间分布[J]. 农业工程学报,2010,26(5): 188-194.
Zhang Sumei, Wang Zongming, Zhang Bai, et al. Prediction of spatial distribution of soil nutrients using terrain attributes and remote sensing data[J]. Transactions of the CSAE,2010,26(5): 188-194. (in Chinese)
- 游浩辰. 林地土壤有机碳遥感反演及空间分异研究[D]. 福州:福建农林大学,2012.
You Haochen. Research on the remote sensing inversion and spatial distribution of soil organic carbon in forest[D]. Fuzhou: Fujian Agriculture and Forestry University, 2012. (in Chinese)
- Hengl T, Heuvelink G B M, Stein A. A generic framework for spatial prediction of soil variables based on regression-Kriging[J]. Geoderma,2004,120: 75-93.
- 刘志红, Li L T, McVicar T R, 等. 专用气候数据空间插值软件 ANUSPLIN 及其应用[J]. 气象,2008,34(2): 92-100.
Liu Zhihong, Li Lingtao, McVicar T R, et al. Introduction of the professional interpolation software for meteorology data: ANUSPLIN[J]. Meteorological Monthly, 2008, 34(2): 92-100. (in Chinese)
- Boer E P J, Beurs K M D, Hartkamp A D. Kriging and thin plate splines for mapping climate variables[J]. International Journal of Applied Earth Observation and Geoinformation,2001,3(2): 146-154.
- 王亮,朝伦巴根,金菊良,等. 基于孔穴效应的时空变差函数在水文数据插补延长中的应用[J]. 水文,2008,28(6):15-20.
Wang Liang, Chaolunbagen, Jin Juliang, et al. Application in the hydrological data interpolation and extension based on spatial and temporal variation function of cavity effect[J]. Journal of China Hydrology, 2008, 28(6):15-20. (in Chinese)
- Florinsky I V, Eilers R G, Manning G R, et al. Prediction of soil properties by digital terrain modeling[J]. Environmental Modelling & Software, 2002, 17(3): 295-311.
- Sumfleth K, Duttman R. Prediction of soil property distribution in paddy soil landscapes using terrain data and satellite information as indicators[J]. Ecological Indicators, 2008, 8(5): 485-501.
- 刘双印,徐龙琴,李振波,等. 基于 PCA-MCAFA-LSSVM 的养殖水质 pH 值预测模型[J]. 农业机械学报,2014,45(5): 240-246.
Liu Shuangyin, Xu Longqin, Li Zhenbo, et al. Forecasting model for pH value of aquaculture water quality based on PCA-MCAFA-LSSVM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(5): 240-246. (in Chinese)
- 陈冬花,邹陈,王苏颖,等. 基于 DEM 的伊犁河谷气温空间插值研究[J]. 光谱学与光谱分析,2011,31(7):1925-1929.
Chen Donghua, Zou Chen, Wang Suying, et al. Study on spatial interpolation of the average temperature in the Yili river valley based on DEM[J]. Spectroscopy and Spectral Analysis, 2011, 31(7):1925-1929. (in Chinese)
- Thompson J A, Bell, J C, Butler C A. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling[J]. Geoderma, 2001, 100(1-2): 67-89.