任向宁,董玉祥.基于地理探测器的区域土壤耕层有机碳含量多元复合模型构建——以珠三角核心区为例[J].热带地理, 2018, 38(4):546-556.

REN Xiangning and DONG Yuxiang. 2018. Construction of Multivariate Composite Calculation Model of Soil Organic Carbon Content in Plough Horizon Based on Geodetector[J]. *Tropical Geography*, 38 (4): 546-556.

基于地理探测器的区域土壤耕层有机碳含量 多元复合模型构建——以珠三角核心区为例

任向宁^{1,2},董玉祥^{1,3}

(1. 中山大学地理科学与规划学院//广东省城市化与地理环境空间模拟重点实验室,广州 510275;
2. 华南农业大学资源环境学院,广州 510642; 3. 中山大学 新华学院,广州 510520)

摘 要: 区域土壤耕层有机碳含量的精确测算,对于研究全球碳循环至关重要,但其影响因素多,空间变异性强,现有插值测算方法精度偏低。从提高测算精度出发,基于地理探测器技术,改进传统的协同克里格插值方法,构建多元复合模型进行区域土壤耕层有机碳含量的测算,并以珠三角核心区为例进行实际验证。结果表明:1)珠三角核心区土壤耕层有机碳含量空间变异与地形、水文、土壤和农田利用方式等有关,不同因素的贡献力存在较大差异,各因素贡献力(q统计量)在0.076~0.201之间,其中土壤理化性状与农田利用方式的贡献力大于地形、水文,区别不同因素的贡献力进行土壤耕层有机碳含量的精确测算应是客观要求。2)依照地理探测器分析结果,基于不同因素贡献力构建土壤耕层有机碳含量多元复合模型,插值测算精度明显提升,在珠三角核心区普通克里格、地理加权回归克里格和 BP 神经网络克里格的土壤耕层有机碳含量插值结果精度较多元复合模型下降 16.62%、23.86%和 37.33%。由结果发现,多元复合模型突破了现有算法中辅助因素数量的局限,有效抑制了区域 SOCC测算过程中的影响因素耦合噪声,降低了测算结果的不确定性。 关键词: 土壤耕层;有机碳含量;地理探测器;多元复合模型;插值噪声;珠三角核心区 中图分类号:S153 文献标志码;A 文字编号:1001-5221(2018)04-0546-11

DOI: 10.13284/j.cnki.rddl.003063

土壤耕层有机碳是全球碳循环中的重要组成 部分,是可在较短时间内进行适度调节的碳库之一, 具有提升区域固碳潜力的重要作用(赵永存等, 2018)。受限于采样数量和土壤空间异质性的影响, 区域土壤耕层有机碳含量(Soil organic carbon content, SOCC)测算方法对于精确测算区域土壤 耕层有机碳密度和碳库储量显得尤为重要。目前, 常用区域SOCC测算方法有土壤类型法(解宪丽 等,2003)、多元线性回归法(赵永存等,2005) 和空间插值方法等,其中空间插值方法主要有土壤 类型平均法、普通克里格插值(Ordinary Kriging, OK)、协同克里格法(Cooperative Kriging, COK)、 地理加权回归克里格(Geographically weighted regression-Kriging, GWRK)和人工神经网络克里 格(Neural network model-Kriging, NNK)等。土 壤类型平均法是以区域SOCC插值为基础(Batjes, 1996; 栾文楼 等, 2011),采用样点平均值代替同 一土壤类型空白点。普通克里格插值是以SOCC空 间自相关性为基础,利用原始数据和半方差函数的 结构性,对区域化变量的未知采样点进行线性最优 无偏估计的一种插值方法,其实质是实行局部估计 的加权平均值(汪媛媛 等, 2011)。协同克里格法 是OK的扩展形式,能够综合考虑SOCC空间变异的 结构性因子与随机性因子,把区域化变量的最佳估 值方法从单一属性发展到2个或2个以上的协同属 性,能够将主变量的自相关性和辅变量的交互相关 性结合起来用于无偏最优估值中,从而提高主要预 测值的精度(文雯 等, 2013)。GWRK是将回归克 里格法中的全局拟合转换为地理加权回归的局部拟 合,然后将插值点的局部拟合残差进行OK插值,再

作者简介:任向宁(1978-),男,河北正定人,博士生,主要研究方向为土地资源开发利用与保护,(E-mail)xnren@scau.edu.cn。

收稿日期: 2018-05-30; 修回日期: 2018-07-16

与GWR法拟合的趋势相加的测算方法(杨顺华 等, 2015)。NNK是基于SOCC空间分布与影响因素之间 非线性关系的假设(McBratney et al, 2003), 通过 训练构建传递函数预测区域SOCC, 预测结果残差 再通过OK法进行空间内插,将预测结果与残差插值 结果进行叠加的测算方法(李启权 等, 2013)。

实际应用表明:上述区域SOCC测算方法仍存 在一些问题,如土壤类型法适用于空间尺度较大且 土壤类型比较单一的区域,但对土壤类型的多样性 反映不足(周涛等,2003)。在大空间大尺度下, OK插值结果具有较好的内在空间关联性和精确性 (Sherpa et al., 2016; Gouri et al., 2018), 但在情景 复杂、局部变异较强的地区,其空间特征表达能力 和预测精度较差(汪媛媛等,2011;赵广帅等, 2014)。一般情况下,辅助气象、地形、土地利用方 式等各类影响因素的COK较OK具有更高的插值精 度 (Phachomphon et al., 2010; 顾成军, 2014; Ye H C et al., 2017), 且能够有效地减少回归残差值(王 库, 2013; 杨顺华 等, 2015), 应用尺度更为广泛, 空间变异的细节表达也较为细致(文雯 等, 2013; Zhang Z Q et al., 2015; Fitzpatrick et al., 2016; Ghorbani et al., 2018), 但超过4个协同因素后,因 素越多耦合作用产生的插值噪声越大, 对未知数据

的预测能力越差(郭治兴等,2017)。 GWRK和NNK对空间尺度和采样密度的 依赖性较强,空间尺度越小、采样密度越 大,测算精度越高,但在低采样密度下, GWRK与NNK的测算误差较大,表现出比 OK更差的性能 (Ye H C et al., 2017; Emamgholizadeh et al., 2017)。因此, 在 大尺度、采样密度较低的复杂情景下,为 保证区域SOCC测算精度,增加辅助影响 因素成为唯一有效手段。但增加辅助因 素,又导致各辅助因素之间相互耦合产生 的插值噪声增加与保证测算精度的高要 求形成了难以调和的矛盾。为解决这一矛 盾,有关学者做出了非插值算法的探索, 避免辅助因素增加后,插值测算模型复杂 化和结果偏离度的增加,如构建土壤碳与 单个土地利用类型的映射关系(Wu O et al., 2017), 构建基于因子自组织映射聚 类的土壤有机质与相关因子之间的非线 性关系(Huang Y J et al., 2017)等。同 时,引入SOCC影响因素的贡献力,降低

插值噪声的强度,如郭治兴等(2017)在对丹麦日 德兰半岛SOCC插值过程中,通过Cubist模型引入了 相对坡位、高程等4个变量的贡献率。总体来说,目 前解决区域SOCC插值噪声问题的研究刚刚起步, 成果相对较少。

547

为进一步抑制插值噪声对区域SOCC测算结果的影响,根据非插值算法的思路,尝试以珠三角核心区为研究区,引入地理探测器建立SOCC影响因素贡献力大小序列,在分层协同克里格的基础上构建多元复合模型(Multielement Composite Model,MCM)进行区域SOCC测算,通过验证集进行验证,并与OK、GWRK及NNK中的BP神经网络克里格(BP neural network model-Kriging, BPK)进行精度比较,遴选最优区域SOCC测算方法,以期为区域土壤有机碳研究提供更为科学的空间基础数据。

1 研究方法、数据与研究区域

1.1 研究方法

1.1.1 土壤采样及处理方法 2015年,根据珠三角 核心区土壤类型图(研究区第二次土壤普查县级成 果)、土地利用类型图(县级土地利用变更调查成 果)、地形图(国家基础数据中心ASTER GDEM数 据)等资料,采用叠置法划分18 113个土地利用类



Fig.1 Distribution of soil samples in plough layer

型-土壤类型集合体单元。根据统计学样本设置原则,在区内均匀设置耕层土壤野外采样点942个,约占单元总数的5.20%(图1)。采用棋盘式采样法采 集距田面0~20 cm深度的土壤样品,用于有机碳含量、pH值测定,其中坑塘水面主要采集塘基农田土 壤样品。土壤样本风干后挑去根系和>2 mm的砾 石,利用FW-100粉碎机粉碎后过100目筛,采用重 铬酸钾外加热氧化法测定土壤有机碳含量,比重计 法测定土壤质地,电位法测定土壤pH值(鲍土旦, 2013)。在耕层土壤样本采样过程中,利用洛阳铲采 集距田面0~100 cm深度的土壤剖面,判断样本土体 构型,并记录不同样点的地下水位。

1.1.2 地理探测器 地理探测器(王劲峰 等,2017) 是研究某现象空间分布不一,即空间变异性的统计 学理论方法。该模型假设自变量对因变量产生了影 响,在空间上的分布就会相应产生一定的相关性; 主要包括分异及因子探测、交互作用探测、风险区 探测和生态探测4个探测器,其中分异及因子探测器 即是探测地理属性Y的空间变异性和探测因子 X_1 、 X_2 ······对Y空间变异的解释力(即贡献力),用q值 度量。地理探测器现已被广泛应用于地理空间相似 **性与分异性研究中**,如地理空间影响因素识别 (Wang J F et al., 2010; Du Z, 2016; 李雨 等, 2017)、 区域分异与空间优化(杨忍 等, 2016; 董玉祥 等, 2017)、生态环境与公共健康(Shen J et al., 2015; 周磊 等, 2016)等。区域SOCC具有较强的空间变 异性,且影响因素众多,通过地理探测器可以有效 测度各影响因素的贡献力。同时,区域SOCC影响 因素中包含有海拔、坡度等连续性数值变量和表土 质地、农田利用方式等类型变量,而地理探测器擅 长分析类型变量,且连续型变量可以通过分类分级 转换为类型变量后进行地理探测。地理探测器可用 于测度区域SOCC空间变异影响因素贡献力,较主成 分分析法、回归法等更为直观和快捷。

根据珠三角核心区自然与农田利用特征,采用 文献法遴选海拔高度、地形坡度、坡向、表土质地、 剖面构型、地下水位、土壤pH和农田利用方式等8 个因素作为珠三角核心区SOCC空间变异的协助变 量(Anderson et al., 1985; Bayer et al., 2006; LI X J et al., 2011)。

1.1.3 多元复合模型 自然土壤不同方向上的变异 程度与水热、土壤、地形等变化的方向性有关(席 小康 等, 2017), SOCC空间特征较为平稳。而农 田尤其是经济先发地区农田的利用方式复杂,土壤 理化性状空间差异大,影响因素众多,区域SOCC 空间异质性更强(吴子豪等,2018),常规改进算 法的思路难以满足经济先发地区农田SOCC测算精 度的要求。因此根据非算法的思路,首次提出MCM, 基于COK构建SOCC样本与单因素之间的非线性关 系,进行区域SOCC插值,获取该因素层的区域 SOCC插值结果;再采用地理探测器探测不同因素 对区域SOCC空间变异的贡献力,并将不同因素贡 献力转换为模型权重,构建区域SOCC测算模型; 采用加权求和法,综合各层单因素COK插值结果, 测算出区内各空间单元的SOCC。其公式表达为:

$$Z_{0,ck}(x_0) = \sum_{i=1}^{n} \lambda_i Z_i(x_0)$$
 (1)

式中: $Z_{0,ck}(x_0)$ 为 x_0 处最终的有机碳含量估算值; $Z_i(x_0)$ 为 x_0 处第i个协同变量(影响因素)克里格 插值的有机碳含量估算值; λ_i 为第i个协同变量的权 重系数。

1.1.4 常用区域插值测算方法 常用的区域插值测 算方法中,OK是利用区域化变量的原始数据和变异 函数的结构特点,对未采样点的区域化变量的取值 进行线性无偏最优估计的一种方法(汤国安 等, 2015)。GWRK是对GWR法的延伸与扩展,即对局 部模型GWR拟合后得到的残差进行OK法插值,然 后与GWR法拟合的趋势相加(杨顺华 等, 2015)。 BPK则是采用Matlab软件运行BP神经网络模型,以 空间样本值作为BP神经元,同时输入不同影响因 子,通过迭代训练,自我完成描述耕地土壤复杂系 统中多元影响因素与SOCC的非线性映射关系。设 置最大训练次数为10000,学习率为0.01,限时训练 迭代过程为20,训练要求精度为0.000 000 001。对 神经网络预测结果残差进行半方差分析后,采用OK 完成对神经网络预测结果残差的空间内插。再将预 测结果与残差插值结果进行叠加,得到最终的区域 SOCC预测结果 (李启权 等, 2013)。

1.2 研究区域

珠三角核心区地处中国大陆南部最大的热带三 角洲——珠江三角洲平原的核心区域,区域范围为 21°48′-23°56′N、112°23′-114°37′E,总面积为2.31 万km²;包括广州、深圳2个副省级城市,珠海、佛 山、东莞、中山4个地级市,以及江门鹤山、蓬江、 江海、新会与肇庆四会市5个市(县、区)(图2)。

珠三角核心区绝大部分在北回归线以南,属亚 热带海洋季风气候,热量充足,年均日照时数为1 751.9 h,年均气温22.5℃,年均降水量1 817.6 mm;



图 2 珠三角核心区范围图 Fig.2 The core area map of Pearl River Delta

其四周多为山地、丘陵和岗地,中间为河口三角洲 沉积平原,整体呈南向"铲"状地形。区内农田海 拔高度在-52~801 m之间,平均为17.32 m;地形坡 度在0~26.65°之间,平均为1.89°;各坡向比例为 5.91%~21.94%,其中东南向和南向分别占19.46%和 21.94%,北向和西北向分别占5.91%和6.74%(图 3-a)。区内水网密集,水域面积占区域总面积的 18.69%,各行政村水网密度在0~0.87之间,平均为 0.15(图3-b)。根据第二次土壤普查资料,珠三角 核心区农田土壤主要为赤红壤、水稻土、潮土等8 个土类,赤红壤、潜育水稻土、淹育水稻土等15个 亚类,潮泥田、河泥田、洪积泥田、麻沙质田、砂 页岩赤红砂地等70多个土种(亚种、变种)。区内农 田土壤面积最大的是水稻土和赤红壤,分别占总面 积的47.55%和42.44%。2015年区域耕地面积为22.58 万hm²,占研究区总面积的9.76%,其中水田13.97 万hm²,占比61.87%;水浇地6.38万hm²,占28.25%; 旱地2.23万hm²,占9.88%。区域农田利用方式复杂 多样,利用方式主要为蔬菜、水果种植和水产养殖。

珠三角核心区地形、水文条件复杂,人口集中, 社会经济发展速度较快,农田利用方式多样,对区 域SOCC空间插值测算提出了更高要求。在该区域划 分300 m×300 m栅格单元257 390个,分别采用OK、 GWRK、BPK和MCM四种方法测算区域SOCC。

1.3 数据来源

2015年珠三角核心区耕地、水域等土地利用/ 覆被数据及行政区划界线,主要根据研究区年度土 地利用变更调查数据(耕地),并采用地理空间数据 云下载的高分一号遥感影像解译进行检验、修正。 研究区DEM数据采用ASTER GDEM数据(国家基础 数据中心),主要用于海拔高度、地形坡度和坡向等 信息提取,空间分辨率30 m。2015年采集耕层土壤 样本942个,其中土壤类型为水稻土的占79.62%, 赤红壤的占15.39%;利用类型为水田的占40.13%, 水浇地的占31.21%,旱地的占11.36%(表1)。采用 ArcGIS10.2完成土壤样本实测定位点经纬度数据的



图 3 珠三角核心区地形图(a)和水系图(b) Fig.3 Topographic(a) and water system distribution(b) of the core area of the Pearl River Delta

Tab.1 The description table of soil samples in plough layer								
样本属性					样本情况			
土壤类型	土类	赤红壤	红壤	黄壤	滨海盐土	水稻土	潮土	石灰土
	数量/个	145	1	1	2	750	41	2
利用类型	地类	水田	水浇地	旱地	园地	林地	坑塘水面	—
	数量/个	378	294	107	96	18	49	_
表土质地	表土质地	黏土	重壤	中壤	轻壤	砂壤	砂土	—
	数量/个	80	166	195	169	175	157	_
剖面构型 —	剖面构型	通体黏	壤/黏/黏	壤/黏/壤	通体壤	壤/砂/壤	壤/砂/砂	黏心心
	数量/个	69	128	48	418	111	87	2
	剖面构型	黏/砂/黏	砂/黏/砂	砂/黏/黏	通体砂	_	_	_
	数量/个	14	5	14	46	_	_	_
pH 值	pH 值	(3,4]	(4,5]	(5,6]	(6,7]	(7,8]	(8,9]	_
	数量/个	13	141	410	258	101	19	—
地下水位	埋深/cm	(15,20]	(20,40]	(40,60]	(60,100]	>100	_	_
	数量/个	35	44	339	470	54	_	_

表 1 珠三角核心区耕层土壤样本情况说明表

注: 园地、林地、坑塘水面皆为可调整地类,即由耕地转换成的其他利用方式,但仍具耕地生产功能。

精确投影转换,构建用于地统计分析的土壤样点空间数据库。采用Arcgis的"Subset Features(Geostatisical Analyst)"工具从数据库中随机选择70%的SOCC样本(不含最高值和最低值)作为空间插值集,剩余的30%(空间不连续)作为验证集,验证区域SOCC测算精度。研究区SOCC插值集643个样本,验证集275个样本,并通过留出法交叉验证(周志华,2016)。

2 结果与分析

2.1 地理探测器探测结果

珠三角核心区SOCC影响因素众多,经地理探测器探测,海拔高度、地形坡度、坡向、表土质地、 剖面构型、地下水位、土壤pH和农田利用方式8个 影响因素之间为增强后非线性弱相关关系,满足地 理探测器对因素共线性的要求,各影响因素对研究 区 SOCC 空间 变 异 的 贡 献 力 (q statistics)在 0.076~0.201之间 (p=0.000)。各因素贡献力都在2倍 标准差 (0.059, 0.207) 内,其影响真实存在 (表2)。

表 2 珠三角核心区 SOCC 空间变异影响因素 贡献力探测结果

Tab.2 The result of influencing factors' contribution to spatial variability of SOCC is detected

探测指标	q statistics	<i>p</i> 值
高程	0.157	0
坡度	0.093	0
坡向	0.141	0
表土质地	0.196	0
剖面构型	0.201	0
地下水位	0.076	0
土壤 pH	0.162	0
农田利用方式	0.159	0

2.2 不同方法的测算结果

2.2.1 普通克里格测算结果 采用OK测算的珠三 角核心区SOCC平均值为11.63 g/kg,值域范围为 3.48~20.12 g/kg(表3),其中珠海市西北部和江门 市的新会区SOCC较高,广州市与东莞市接壤地带 SOCC较低(图4-a)。验证集与插值集OK测算结果 的残差范围在11.07 g/kg以内(见表3),主要集中在 -2.00~+2.00 g/kg之间(图5-a)。

表 3 插值集测算结果及其与验证集的残差分析表

Tab.3 Calculation results based on interpolation set and its residual analysis table with validation set

测答面	H	区域 SOCC/(g·kg ⁻¹)				
侧异坝口		OK	GWRK	BPK	MCM	
托估住托店	平均值	11.63	11.57	17.94	10.95	
佃 但 果 佃 但 涮 笛 姑 里	最大值	20.12	17.15	17.94	19.1	
初开印水	最小值	3.48	6.36	5.78	4.95	
	平均值	1.33	-0.15	-0.7	-0.19	
验证集残差	最大值	5.17	7.41	7.51	6.85	
	最小值	-5.9	-7.34	-9.54	-5.46	

2.2.2 地理加权克里格测算结果 采用GWRK测算的珠三角核心区SOCC平均值为11.57 g/kg,值域范围为6.36~17.15 g/kg(见表3),其中四会市西南部SOCC较高,珠海市西北部和江门市新会区次之,广州市增城区东南部较低(图4-b)。验证集与GWRK测算结果的残差范围在14.75 g/kg以内(见表3),主要集中在-4.50~+4.00 g/kg之间(图5-b)。

2.2.3 BP神经网络克里格测算结果 采用BPK测算的珠三角核心区SOCC平均值为17.94 g/kg,值域范围为5.78~17.94 g/kg(见表3),其中珠海市东南沿海和中山市南部SOCC较高,广州市增城区中南部较低(图4-c)。验证集与BPK测算结果的残差范围

551

在17.05 g/kg以内(见表3),主要集中在-4.00~+4.00 g/kg之间(图5-c)。

2.2.4 基于地理探测器的多元复合模型测算结果 结合珠三角核心区SOCC影响因素遴选结果和地理 探测器影响因素贡献力探测结果,构建研究区 SOCC的MCM如下:

 $Z_{0,ck}(x_0) = 0.132 \times Z_1(x_0) + 0.078 \times Z_2(x_0) + 0.119$ $\times Z_3(x_0) + 0.165 \times Z_4(x_0) + 0.170 \times Z_5(x_0) + 0.064 \times Z_6(x_0) + 0.137 \times Z_7(x_0) + 0.134 \times Z_8(x_0)$ (2)

式中: $Z_{0,ck}(x_0)$ 为 x_0 处最终的SOCC估算值; $Z_1(x_0)$ 、 $Z_2(x_0)、Z_3(x_0)、Z_4(x_0)、Z_5(x_0)、Z_6(x_0)、Z_7(x_0)$ 和 $Z_8(x_0)$ 为 x_0 处分别协同高程、坡度、坡向、表土质地、剖 面构型、地下水位、土壤pH、农田利用方式后单因 素COK的SOCC估算值。

根据珠三角核心区MCM测算的区域SOCC平均值为10.95 g/kg,值域范围为4.95~19.10 g/kg(见表3),其中珠海市西北部、江门市的新会区及中山市东南部SOCC较高,广州市西南部与四会市西南



图 4 珠三角核心区 SOCC 插值测算结果空间分布图(a. OK 插值结果; b. GWRK 插值结果; c. BPK 插值结果; d. MCM 测算结果) Fig.4 Spatial distribution of SOCC calculation results in the core area of Pearl River Delta



Fig.5 The interpolation(a) and frequency histogram of the verifying set(b) of OK

部次之,广州市增城区中南部,佛山市、东莞市和 江门市的北部SOCC最低(图4-d)。验证集与MCM 测算结果的残差范围在12.31 g/kg内,主要集中在 -2.00~+2.00 g/kg之间(图5-d)。

2.3 不同方法比较分析

根据OK、GWRK、BPK和MCM四种方法测算的珠三角核心区SOCC具有基本一致的空间分布特征,即珠海市西北部和江门市的新会区的SOCC相对较高,广州市增城区东南部较低(图4-a、图5-a、图6-a、图7-a)。但在局部地区4种方法的测算结果仍存在较大差异,以中山市南部为例,GWRK和BPK测算结果显著高于OK和MCM。该区域主要为沿海沉积平原上突起的丘陵岗地,地形、土壤、农田利用方式复杂,在周边平原环境和自身多影响因素的综合噪声影响下,GWRK和BPK测算结果与验证集之间的残差平均为2.73;MCM的空间插值和测算逻辑则更为清晰,其测算结果与验证集之间的残差平均为1.47,较OK更小,有效避免了在该区域出现低值高估状况。

由于研究区样本较多, 难以比对验证集样本预 测值与实测值的差距。为进一步比较4种方法在区域 剖面线上的SOCC测算结果精度, 在研究区设置能 够穿越地貌、土壤类型、主要行政辖区和耕地斑块 最多的两条典型样带对不同方法测算结果的精度进 行比较观察, 即"东北一西南"样带和"西北一东 南"样带, 宽度为5 km(图6)。OK、MCM、GWRK 和BPK四种方法的插值测算结果空间变化趋势基本 一致, 但GWRK和BPK的插值结果空间变化更为平 缓, 插值结果明显收敛在一定范围内, 而MCM的测 算结果的空间变化最为显著, 细节表达较好(图7)。 以"东北一西南"样带第40 km和"西北一东南" 样带第90 km附近为例, MCM显著高于OK、GWRK 和BPK三种传统常用方法的SOCC测算结果。以上 两处的农田表土质地主要为轻壤和砂壤, 剖面构型



函化一东南(b)样带位置图 Fig.6 The position map of SOCC typical sample band in the core area of Pearl River Delta

主要为通体壤或壤/壤/砂,农田利用方式大部分为 菜地,MCM通过权重设置,充分体现了以上3个影 响因素的正向影响作用,有效避免了高值低估。

珠三角核心区地形水文等地理环境复杂,城郊 型现代农业特征突出,农田利用方式多样,人工扰 动频繁,农田土壤空间变异性较强。OK、GWRK 和BPK基于克里格方法生成的空间数据趋向于将数 据进行均一化,空间异质性难以得到合理、有效的 表达(郭治兴等,2017),其中BPK受插值噪声的 影响,精度进一步降低。MCM则充分考虑了SOCC 空间分布的地形、土壤理化性状、农田利用方式等 结构性与随机性因素的影响,突破了一般算法中对 辅助因素的数量限制,SOCC空间变异局部特征的



图 7 东北一西南(a)和西北一东南(b)样带 SOCC 空间插值测算结果对比图 Fig.7 Comparison of SOCC calculated results in the sample

表达效果优于OK、GWRK和BPK。同时,通过地理 探测器探测出不同影响因素的贡献力,在测算模型 中建立了因素贡献力序列,避免了因素间的相互干 扰,抑制了插值噪声。MCM测算的区域SOCC结果 与验证集之间的平均误差、平均绝对误差和均方根 误差均小于OK、GWRK和BPK,而相关性(皮而森 R)、一致性(Kappa)高于OK、GWRK和BPK。将 OK、GWRK和BPK三种方法插值结果的相关性、一 致性与MCM测算结果进行比较,其差值除以MCM 的皮而森R和Kappa值,测算出3种方法与MCM的相 关性差距和一致性差距,进一步简单算术平均相关 性差距和一致性差距测算出综合相对精度差距分 别下降16.62%、23.86%和37.33%(表4)。通过交叉

表 4	不同方法区域	SOCC	测算精度对比表
-----	--------	------	---------

Tab.4 Comparison of SOCC calculation precision in different methods

测答面目	区域 SOCC 测算方法					
侧异坝日	OK	GWRK	BPK	MCM		
平均误差(ME)	-0.006	0.078	-0.424	-0.003		
平均绝对误差(MAE)	2.073	2.499	2.618	1.419		
均方根误差(RMSE)	2.618	3.138	3.29	1.905		
皮而森 R	0.685	0.643	0.523	0.806		
与 MCM 相关性差距/%	-15.01	-20.22	-35.11	—		
卡帕 (Kappa)	0.583	0.517	0.431	0.713		
与 MCM 一致性差距/%	-18.23	-27.49	-39.55	—		
综合相对精度差距/%	-16.62	-23.86	-37.33	_		

注:综合相对精度差距:OK、GWRK和 BPK 三种传统常用方法与 MCM 相关性和一致性差距的简单算术平均。 验证,MCM能够显著提高珠三角核心区SOCC测算结果的相关性、一致性,可以作为一项适用于复杂地理环境和农田利用情景的高精度SOCC预测技术,进一步扩展研究区域、深入探讨MCM的适用性。

3 结论

农田耕层土壤发生、发育过程中不断受到区域 地形、水文及农田利用方式等因素影响,使得区域 SOCC产生较强的时空变化,尤其是在珠三角核心 区,改革开放以来农田利用方式和利用强度剧烈变 化进一步增强了区域SOCC的空间变异性和不确定 性。通过交叉验证, MCM根据区域农田自然经济特 征,充分考虑地形、水文、土壤、农田利用方式等 影响因素,将清晰的分层单因素COK插值结果与基 于因素贡献力序列融入统一的模型, 其测算结果相 比OK、GWRK和BPK三种传统常用方法,误差最小, 精度提升显著。MCM突破了现有算法中辅助因素数 量的局限,能够在SOCC空间插值过程中协同更多易 获取的辅助变量,同时通过在模型中设置权重体现 不同因素对空间变异的贡献力,兼顾了区域SOCC 测算过程中的秩序与协同,降低了不确定性,对 SOCC的空间差异刻画更为细致,有效揭示了SOCC 空间变异规律。综合模型复杂度、预测精度以及空 间表达力, MCM有效体现了农田人工土壤SOCC空 间变异的多因性及不同因素的贡献力差异,可以作

为复杂自然经济环境下区域SOCC预测的优先模型 进一步深入研究。

参考文献 (References):

- ANDERSON D W and COLEMAN D C. 1985. The dynamics of organic matter in grassland soils. *Journal of Soil and Water Conservation*, 40: 211-216.
- 鲍士旦. 2013. 土壤农化分析. 3版. 北京:中国农业出版社. [BAO Shidan. 2013. Soil Agro-chemistrical Analysis. 3rd edition. *Beijing: China Agriculture Press.*]
- BAYER C, LOVATO T, DIECKOW J, ZANATTA J A and MIELNICZUK J. 2006. A method for estimating coefficients of soil organic matter dynamics based on long-term experiments. *Soil&Tillage Research*, 91: 217–226.
- BATJES N H. 1996. Total carbon and nitrogen in the soils of the world. *European Journal of Soil Science*. 47: 151-163.
- 董玉祥,徐茜,杨忍,徐成东,王钰莹. 2017. 基于地理探测器的中国 陆地热带北界探讨. 地理学报,72(1):135-147. [DONG Yuxiang, XU Qian, YANG Ren, XU Chengdong and WANG Yuying. 2017. Delineation of the northern border of the tropical zone of China's mainland using Geodetector. *Acta Geographica Sinica*,72(1): 135-147.]
- DU Z, XU X, ZHANG H, WU Z and LIU Y. 2016. Geographical detector-based identification of the impact of major determinants on aeolian desertification risk. *PLoS One*, 11 (3): e0151331.
- EMAMGHOLIZADEH S, SHAHSAVANI S and ESLAMI M A. 2017. Comparison of artificial neural networks, geographically weighted regression and Cokriging methods for predicting the spatial distribution of soil macronutrients(N, P, and K). *Chinese Geographical Science*, 27 (5): 747-759.
- FITZPATRICK B R, LAMB D W and MENGERSEN K. 2016. Ultrahigh dimensional variable selection for interpolation of point referenced spatial Data: A digital soil mapping case study. *PLoS One*, 11 (9): 1-19.
- GHORBANI A, MOGHADDAM S M, MAJD K H and DADGAR N. 2018. Spatial variation analysis of soil properties using spatial statistics: a case study in the region of Sabalan mountain, Iran. *Ecomont-Journal on Protected Mountain Areas Research*, 10 (1): 70-80.
- 顾成军. 2014. 克里格插值在区域土壤有机碳空间预测中的应用. 中国 土壤与肥料, 2014(3): 93-97. [GU Chengjun. 2014. Application of Kriging method in spatial prediction of regional soil organic carbon. *Soil and Fertilizer Sciences in China*, 2014(3): 93-97.]
- 郭治兴,袁宇志,郭颖,孙慧,柴敏,陈泽鹏, Mogens H GREVE. 2017. 基 于地形因子的土壤有机碳最优估算模型. 土壤学报,54(2):
 331-343. [GUO Zhixing, YUAN Yuzhi, GUO Yin, SUN Hui, CHAI Min, CHEN Zepeng and Mogens H GREVE. 2017. Optimal estimation model of soil organic carbon based on the terrain factor. *Acta Pedologica Sinica*, 54(2): 331-343.]

- GOURI S B, PRAVAT K S and RAMKRISHNA M. 2018. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). Journal of the Saudi Society of Agricultural Sciences, 17 (2): 114-126.
- HUANG Y J, YE H C, ZHANG L P, ZHANG S W, SHEN C Y, LI Z and HUANG Y F. 2017. Prediction of soil organic matter using ordinary Kriging combined with the clustering of self-organizing map: A case study in Pinggu district, Beijing, China. *Soil Science*, 182 (2): 52-62.
- 李启权, 王昌全, 张文江, 余勇, 李冰, 杨娟, 白根川, 蔡艳. 2013. 基 于神经网络模型和地统计学方法的土壤养分空间分布预测. 应用生 态学报, 24(2): 459-466. [LI Qiquan, WANG Changquan, ZHANG Wenjiang, YU Yong, LI Bing, YANG Juan, BAI Genchuan and CAI Yan. 2013. Prediction of soil nutrients spatial distribution based on neural network model combined with goestatistics. *Chinese Journal of Applied Ecology*, 24 (2): 459-466.]
- 栾文楼,宋泽峰,李随民,崔邢涛. 2011. 河北平原土壤有机碳含量的变化. 地质学报,85(9):1528-1535. [LUAN Wenlou, SONG Zefeng, LI Suimin and CUI Xingtao. 2011. Changes of soil organic carbon content in Hebei plain. Acta Geologica Sinica, 85(9): 1528-1535.]
- LI X J, ZHANG X Z and ZHANG Y J. 2011. Modeling the effects of climate change and elevated CO₂ on soil organic carbon in an alpine steppe. *Journal of Resources and Ecology*, 2 (2): 168-174.
- 李雨,韩平,任东,罗娜,王纪华. 2017. 基于地理探测器的农田土壤 重金属影响因子分析. 中国农业科学, 50(21): 4138-4148. [LI Yu, HAN Ping, REN Dong, LUO Na and WANG Jihua. 2017. Influence Factor Analysis of Farmland Soil Heavy Metal Based on the Geographical Detector. *Scientia Agricultura Sinica*, 50 (21): 4138-4148.]
- MCBRATNEY A B, SANTOS M L M and MINASNY B. 2003. On digital soil mapping. *Geoderma*, 117 (1/2): 3-52.
- PHACHOMPHON K, DLAMINI P and CHAPLOT V. 2010. Estimating carbon stocks at a regional level using soil information and easily accessible auxiliary variables. *Geoderma*, 155 (3/4): 372-380.
- SHEN J, ZHANG N, GEXI G D R, HE B, LIU C Y, LI Y, ZHANG HY, CHEN X Y and LIN H. 2015. Construction of a GeogDetector-based model system to indicate the potential occurrence of grasshoppers in Inner Mongolia steppe habitats. *Bulletin of Entomological Research*, 105 (3): 335-346.
- SHERPA S R, WOLFE D W and VAN ES H M. 2016. Sampling and data analysis optimization for estimating soil organic carbon stocks in agroecosystems. *Soil Science Society of America Journal*, 80 (5): 1377-1392.
- 汤国安,杨昕. 2015. ArcGIS地理信息系统空间分析实验教程. 2版. 北京:科学出版社. [TANG Guoan and YANG Xin. 2015. Spatial Analysis Experimental Course of Geography Information System based on Arc GIS. 2nd edition. Beijing: Science Press.]

王劲峰,徐成东. 2017. 地理探测器: 原理与展望. 地理学报, 72 (1):

116-134. [WANG Jinfeng and XU Chengdong. 2017. Geodetector:Principle and prospective. *Acta Geographica Sinica*, 72(1): 116-134.]

- WANG J F, LI X H, CHRISTAKOS G, LIAO Y L, ZHANG T, GU X and ZHENG X Y. 2010. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. International Journal of Geographical Information Science, 24 (1): 107-127.
- 王库. 2013. 基于地理权重回归模型的土壤有机质空间预测. 土壤通报, 44(1): 21-28. [WANG Ku. 2013. Spatial estimation of soil organic matter by using geographically weighted regression model. *Chinese Journal of Soil Science*, 44(1): 21-28.]
- 汪媛媛,杨忠芳,余涛,文字博,夏学齐,白荣杰. 2011. 土壤碳储量 计算中不同插值方法对比研究:以吉林省大安市为例. 中国岩溶, 30(4):479-486. [WANG Yuanyuan, YANG Zhongfang, YU Tao, WEN Yubo, XIA Xueqi and BAI Rongjie. 2011. Contrastive studies on different interpolation methods in soil carbon storage calculation in Da'an city, Jilin Province. *Carsologica Sinica*, 30(4):479-486.]
- 文雯,周宝同,汪亚峰,黄勇. 2013. 基于辅助环境变量的土壤有机碳 空间插值:以黄土丘陵区小流域为例. 生态学报,33(19):6389-6397. [WEN Wen, ZHOU Baotong, WANG Yafeng and HUANG Yong. 2013. Soil organic carbon interpolation based on auxiliary environmental covariates: a case study at small watershed scale in Loess Hilly region. *Acta Ecologica Sinica*, 33(19): 6389-6397.]
- WU Q, LI Q L, GAO J B, LIN Q Y, XU Q F, GROFFMAN P M and YU S. 2017. Non-algorithmically integrating land use type with spatial interpolation of surface soil nutrients in an urbanizing watershed. *Pedosphere*, 27 (1): 147-154.
- 吴子豪,刘艳芳,陈奕云,郭龙,姜庆虎,王少辰. 2018. 综合土地利 用及空间异质性的土壤有机碳空间插值模型. 应用生态学报, 29 (1):238-246. [WU Zihao, LIU Yanfang, CHEN Yiyun, GUO Long, JIANG Qinghu and WANG Shaochen. 2018. Spatial interpolation model of soil organic carbon density considering land-use and spatial heterogeneity. *Chinese Journal of Applied Ecology*, 29(1):238-246.]
- 席小康,朱仲元,郝祥云. 2017. 锡林河流域土壤有机碳空间变异分析. 水 土保持研究, 24(6): 97-104. [XI Xiaokang, ZHU Zhongyuan and HAO Xiangyun. 2017. Spatial variability of soil organic carbon in Xilin River Basin. *Research of Soil and Water Conserbation*, 24(6): 97-104.]
- 解宪丽,孙波,周慧珍,李忠佩,李安波. 2004. 中国土壤有机碳密度
 和储量的估算与空间分布分析. 土壤学报,41(1):35-43. [XIE
 Xianli, SUN Bo, ZHOU Huizhen, LI Zhongpei and LI Anbo. 2004.
 Acta Pedologica Sinica, 41(1):35-43.]
- 杨忍,刘彦随,龙花楼,王洋,张怡筠. 2016. 中国村庄空间分布特征 及空间优化重组解析. 地理科学,36(2):170-179. [YANG Ren, LIU Yansui, LONG Hualou, WANG Yang and ZHANG Yiyun. 2016. Spatial distribution characteristics and optimized reconstructing analysis of rural settlement in China. *Scientia Geographica Sinica*, 36(2): 170-179.]

- YE H C, HUANG W J, HUANG S Y, HUANG Y F, ZHANG S W, DONG Y Y and CHEN P F. 2017. Effects of different sampling densities on geographically weighted regression Kriging for predicting soil organic carbon]. Spatial Statistics, 20: 76-91.
- 杨顺华,张海涛,郭龙,任艳. 2015. 基于回归和地理加权回归Kriging 的土壤有机质空间插值. 应用生态学报,26(6):1649-1656. [YANG Shunhua, ZHANG Haitao, GUO Long and REN Yan. 2015. Spatial interpolation of soil organic matter using regression Kriging and geographically weighted regression Kriging. *Chinese Journal of Applied Ecology*, 26 (6): 1649-1656.]
- ZHANG Z Q, YU D S, SHI X Z, WANG N and ZHANG G X. 2015. Priority selection rating of sampling density and interpolation method for detecting the spatial variability of soil organic carbon in China. *Environmental Earth Sciences*, 73 (5): 2287-2297.
- 赵广帅,李发东,李运生,李静,欧阳竹. 2014. GIS空间插值模拟法 与土壤类型法估算比较土壤碳储. 农业工程学报, 30 (20): 155-162. [ZHAO Guangshuai, LI Fadong, LI Yunsheng, LI Jing and OUYANG Zhu. 2014. GIS spatial interpolation compared with soil type method for estimating soil carbon storage. *Transactions of the Chinese Society of Agricultural Engineering*, 30 (20): 155-162.]
- 赵永存,史学正,于东升,赵彦锋,孙维侠,王洪杰. 2005. 不同方法 预测河北省土壤有机碳密度空间分布特征的研究. 土壤学报,42 (3): 379-385. [ZHAO Yongcun, SHI Xuezheng, YU Dongsheng, ZHAO Yanfeng, SUN Weixia and WANG Hongjie. 2005. Different methods for prediction of spatial patterns of soil organic carbon density in Hebei province, China. *Acta Pedologica Sinica*, 42(3): 379-385.]
- 赵永存, 徐胜祥, 王美艳, 史学正. 2018. 中国农田土壤固碳潜力与速 率:认识、挑战与研究建议. 中国科学院院刊, 33(2): 191-197. [ZHAO Yongcun, XU Shengxiang, WANG Meiyan and SHI Xuezheng. 2018. Carbon sequestration potential in Chinese cropland soils: Review, challenge, and research suggestions. *Bulletin of Chinese Academy of Sciences*, 33(2): 191-197.]
- 周磊,武建军,贾瑞静,梁念,张凤英,倪永,刘明. 2016. 京津冀PM₂₅
 时空分布特征及其污染风险因素.环境科学研究, 29 (4):
 483-493. [ZHOU Lei, WU Jianjun, JIA Ruijing, LIANG Nian, ZHANG
 Fengying, NI Yong and LIU Ming. 2016. Investigation of temporal-spatial characteristics and underlying risk factors of PM₂₅
 pollution in Beijing-Tianjin- Hebei Area. *Research of Environmental Sciences*, 29 (4): 483-493.]
- 周涛,史培军,王绍强. 2003. 气候变化及人类活动对中国土壤有机碳 储量的影响. 地理学报,58(5):727-734. [ZHOU Tao, SHI Peijun and WANG Shaoqiang. 2003. Impacts of climate change and human activities on soil carbon storage in China. *Acta Geographica Sinica*, 58(5):727-734.]
- 周志华. 2016. 机器学习. 1版. 北京:清华大学出版社. [ZHOU Zhihua. 2016. Machine Learning. 1st edition. Beijing: Tsinghua University Press.]

Construction of Multivariate Composite Calculation Model of Soil Organic Carbon Content in Plough Horizon Based on Geodetector

REN Xiangning^{1,3}, DONG Yuxiang^{1,2}

 Guangdong Provincial Key Laboratory of Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China; 2. Xinhua College of Sun Yat-sen University, Guangzhou 510520, China; 3. College of Natural Resources and Environment, South China Agricultural University, Guangzhou 510642, China)

Abstract: The accurate calculation of organic carbon content in regional soil plough is very important for the study of global carbon cycle, but its influence factors are many, the spatial variability is stronger, and the accuracy of the existing interpolation calculation method is low. The Geodetector provides a new method of spatial differentiation and factor detection, which can effectively measure the contribution of the spatial differentiation of soil organic carbon content in plough layer. Combining the traditional Cooperative-Kriging interpolation method with the Geodetector, according to the detection results of the contribution of the geographical detector to the influencing factors, this paper constructs a multivariate composite model based on hierarchical Cooperative-Kriging to calculating regional soil organic carbon content in plough layer. The core area of the Pearl River Delta is used as the study area. Through the setting of soil plough organic carbon content interpolation set and verification set, Ordinary Kriging, Geographically weighted regression-Kriging, BP neural network model-Kriging and Multielement composite model are used to cross validation. The results showed that: 1) the spatial variation of soil organic carbon content in the core area of the Pearl River Delta is related to topography, hydrology, soil and farmland utilization, and the contribution of different factors is different, the contribution of various factors (q Statistics) is from 0.076 to 0.201, among which the contribution of soil physical and chemical properties and farmland utilization methods is greater than that of topography, hydrology. The objective difference of contribution of different factors has an important impact on the accurate estimation of soil organic carbon content. 2) On the basis of the detection of the geographic detector, the factor contribution sequence is introduced into the Multielement composite model, which inhibits the interpolation noise generated by the interference among the factors, and effectively avoids the low value overestimation or the high value underestimation. Through cross validation, the mean error (ME), mean absolute error (MAE) and root mean square error (RMSE) between the results of soil organic carbon content calculated by Multielement composite model and the validation set are less, the correlation (Pearson R) and consistency (Kappa) are higher, and the accuracy of Ordinary Kriging, Geographically weighted regression-Kriging, BP neural network model-Kriging is 16.62%, 23.86% and 37.33%, respectively, lower than that of Multielement composite model. 3) Multielement composite model fully considers the effects of structural and random factors such as topography, soil physical and chemical properties and farmland utilization on the spatial differentiation of soil organic carbon content in plough layer. It breaks through the limitation of the number of auxiliary factors in the existing algorithms, and can cooperate more auxiliary variables in the calculation of soil organic carbon content in plough layer. At the same time, by setting weights in the model, the contribution of different factors to spatial variation is reflected, the order and coordination in the process of regional soil organic carbon content are taken into account, the uncertainty is reduced, the spatial difference is depicted more meticulously and effectively reveals the spatial variability of soil organic carbon content in plough layer. The construction of Multielement composite model has made a positive exploration for further improving the prediction accuracy of soil organic carbon content in plough layer, which provides a new research idea for the study of soil organic carbon in plough layer with strong variation characteristics under complex geographical environment.

Key words: soil in plough layer; soil organic carbon content; Geodetector; Multielement composite model; the core area of the Pearl River Delta