

江叶枫, 郭 熙, 叶英聪, 等. 省域尺度土壤有机质空间分布的神经网络法预测[J]. 江苏农业学报, 2017, 33(4): 828-835.
doi:10.3969/j.issn.1000-4440.2017.04.016

省域尺度土壤有机质空间分布的神经网络法预测

江叶枫^{1,2}, 郭 熙^{1,2}, 叶英聪¹, 孙 凯¹, 饶 磊¹

(1. 江西农业大学国土资源与环境学院/江西省鄱阳湖流域农业资源与生态重点实验室, 江西 南昌 330045; 2. 南方粮油作物协同创新中心, 湖南 长沙 410000)

摘要: 土壤有机质空间分布预测方法研究对指导省域尺度下土壤有机质空间插值模型选取和精度优化具有重要意义。以江西省为例, 利用 BP 神经网络模型与普通克里金结合的方法 (BPNN-OK)、RBF 神经网络模型与普通克里金结合的方法 (RBFNN-OK) 以及普通克里金法 (OK) 3 种方法, 预测省域尺度下耕地表层 (0~20 cm) 土壤有机质的空间分布。16 109 个土壤样点随机分成 12 887 个建模样点, 3 222 个测试样点。结果表明: 在省域尺度下, BPNN-OK 法、RBFNN-OK 法较 OK 法在土壤有机质空间预测精度上有较大提升, 三者的预测精度为 BPNN-OK > RBFNN-OK > OK。BPNN-OK 法对土壤有机质预测结果的均方根误差、平均绝对误差、平均相对误差较 OK 法分别降低 28.66%、30.71%、34.76%, RBFNN-OK 法较 OK 法分别降低 27.76%、29.74%、33.71%。在省域尺度下, 神经网络模型与普通克里金结合的方法能很好地捕捉土壤有机质的复杂空间变异关系。研究结果可指导江西省土壤有机质空间插值模型选取。

关键词: 土壤有机质预测; BP 神经网络模型; RBF 神经网络模型; 普通克里金法; 省域尺度; 空间插值; 精度

中图分类号: S153.6 文献标识码: A 文章编号: 1000-4440(2017)04-0828-08

Prediction of spatial distribution of soil organic matter at provincial scale with neural network

JIANG Ye-feng^{1,2}, GUO Xi^{1,2}, YE Ying-cong¹, SUN Kai¹, RAO Lei¹

(1. College of Land Resource and Environment, Jiangxi Agricultural University/Key laboratory of Poyang Lake Watershed Agricultural Resources and Ecology, Nanchang 330045, China; 2. Southern Regional Collaborative Innovation Center for Grain and Oil Crops in China, Changsha 410000, China)

Abstract: Soil organic matter (SOM) is one of the most important indicators of soil quality. Accurate spatial information about SOM at provincial scale is critical for farmland use and soil environmental protection. In this study, BP neural network combined with ordinary Kriging (BPNN-OK), RBF neural network combined with ordinary Kriging (RBFNN-OK), and ordinary Kriging (OK) were adopted to predict the spatial distribution of SOM in Jiangxi province. It has great significance for the model selection of spatial interpolation and precision optimization of SOM. 16 109 soil samples were collected and randomly divided into two groups, as modeling samples (12 887) and validation samples (3 222). Three

methods generated similar SOM maps. The prediction precision followed the order of BPNN-OK > RBFNN-OK > OK. The root mean square errors (RMSE), the mean absolute errors (MAE) and the mean relative errors (MRE) of BPNN-OK were 28.66%, 30.71% and 34.76%, smaller than those of OK, and the RMSE, MAE and MRE of RBFNN-OK were 27.76%, 29.74% and 33.71%, smaller than those of OK. It is indicated that the neural network

收稿日期: 2016-12-30

基金项目: 国家自然科学基金项目 (41361049); 江西省自然科学基金项目 (20122BAB204012); 江西省赣鄱英才“555”领军人才项目 (201295)

作者简介: 江叶枫 (1994-), 男, 江西余干人, 硕士研究生, 主要从事土壤环境与系统模拟研究。(E-mail) jiangyf0308@163.com

通讯作者: 郭 熙, (E-mail) xig435@163.com

model combined with ordinary Kriging could capture the complex spatial variation of SOM in provincial scale. This method provides a useful tool for the accurate prediction of soil properties in Jiangxi province.

Key words: soil organic matter; back propagation neural network model; radial basis function neural network model; ordinary Kriging; provincial scale; spatial interpolation; precision

土壤养分受生物^[1]、气候^[2]、母质^[3]和地形^[4]等因素影响,导致其在空间分布上呈现非均匀性。快速准确地获取土壤养分空间分布是有效管理土壤、科学规划土地利用、精准农业施肥的基础^[5]。土壤有机质(Soil organic matter, SOM)是土壤养分的重要组成部分,其含量的高低将直接影响其空间上生态系统的生产力^[6-7]。目前基于实地采样获取的 SOM 含量信息远远不能满足精准农业和环境保护的实际需求。由于成本限制,大规模采样并不现实,且仍然无法得到研究区域空间上 SOM 含量。因此,预测 SOM 的空间分布一直是土壤学研究的热点问题之一^[8-14]。

目前,针对 SOM 含量空间分布的预测方法有很多,其中普通克里金法(Ordinary kriging, OK)是应用最普遍也最广泛的方法^[15]。但在许多情况下因不满足普通克里金法的 3 个重要前提条件而导致精度偏低甚至不能应用,而且普通克里金法容易造成平滑效应^[16],尤其是难以表达土壤属性含量突变的区域,导致突变区域信息丢失严重^[17]。为解决这一问题,近年来国内外众多学者将人工神经网络引入 SOM 的空间分布预测中,人工神经网络模型能够较准确地揭示自变量与 SOM 间的复杂非线性关系,并取得了显著成就^[18-20]。目前应用于 SOM 空间插值的神经网络主要有 BP 神经网络(Back propagation neural network, BPNN)和 RBF 神经网络(Radial basis function neural network, RBFNN),但大多数集中在县域尺度,有关省域尺度下的研究相对薄弱。由于县域尺度下地形因子变化较明显而省域尺度下地形因子较为模糊,导致仅基于地形因子的神经网络模型精度不高。近年来,越来越多的学者将邻近信息纳入土壤属性的空间预测中来。徐剑波等^[21]利用欧式距离来寻找离建模点最近的 5 个土壤属性值,人工神经网络模型在不同样本下的误差较普通克里金法都有明显的降低。李启权等^[22]根据空间距离衰减规律,选取离插值点最近的 3 个土壤属性值,其预测误差较普通克里金法低。通过把邻近信息纳入 SOM 的空间分布预测,能够更好地揭示土壤

属性空间分布的局部变化特征,能在提高精度的同时更加真实地反映土壤属性空间变异情况。

本研究以江西省耕地表层 SOM 含量为研究对象,基于耕地地力评价的数据,以地理坐标(x, y)、高程、坡度以及邻近信息(P_1, P_2, P_3, P_4)作为 2 种神经网络的输入变量,利用 BP 神经网络模型与普通克里金结合法(BPNN-OK)、RBF 神经网络模型与普通克里金法(RBFNN-OK)来预测江西省耕地表层 SOM 的空间分布,并与普通克里金法进行对比,以期在省域尺度下土壤性质的空间插值模型选取和精度优化提供参考。

1 材料和方法

1.1 研究区域概况

江西省($24^{\circ}29'14''-30^{\circ}04'41''N, 113^{\circ}34'36''-118^{\circ}28'58''E$)总面积为 $1.69 \times 10^5 \text{ km}^2$,地处北回归线附近,全省气候温暖湿润,四季比较分明。年平均气温 $20 \sim 27^{\circ}\text{C}$,雨量充沛,年均降水量 $1\,341 \sim 1\,940 \text{ mm}$;全年无霜期约 $240 \sim 307 \text{ d}$,为亚热带湿润气候。土壤类型以红壤分布最广,其次为黄壤、黄棕壤、紫色土、石灰(岩)土和草甸土。耕作土壤以水稻土居多,约占全省耕地总面积的 85%。

1.2 土壤采样与数据处理

土壤样品采集于 2012–2013 年,按照全国耕地地力调查与质量评价技术规程,同时考虑地形部位的基础上,采用多点混合的方法采集 $0 \sim 20 \text{ cm}$ 的耕地表层土壤样品 16 582 个(图 1)。每个采样点均以 GPS 记录其海拔和坐标,每个样点采集样品 1 000 g,土壤样品经过自然风干后,在实验室磨碎过筛,用重铬酸钾($\text{K}_2\text{Cr}_2\text{O}_7$)油浴加热测定 SOM 含量。

样品采集与农化分析过程中受采样及指标测试误差的影响,导致 SOM 含量的测量结果存在异常值。本研究采用拉依达准则法对采样点数据的 SOM 含量值进行异常值处理(3 倍的标准差),后续的研究均采用剔除异常值后的数据。对处理后的数据导入 ArcGIS10.2 中运用地统计模块生成样本数据子集,剔除后总样点 16 109 个。随机均匀选取 80% 样点进入

训练子集(训练子集12 887个)用于插值;剩下 20%样点进入测试子集(测试子集3 222个,不参与插值过

程,只在后续精度分析中出现)用于验证插值精度。

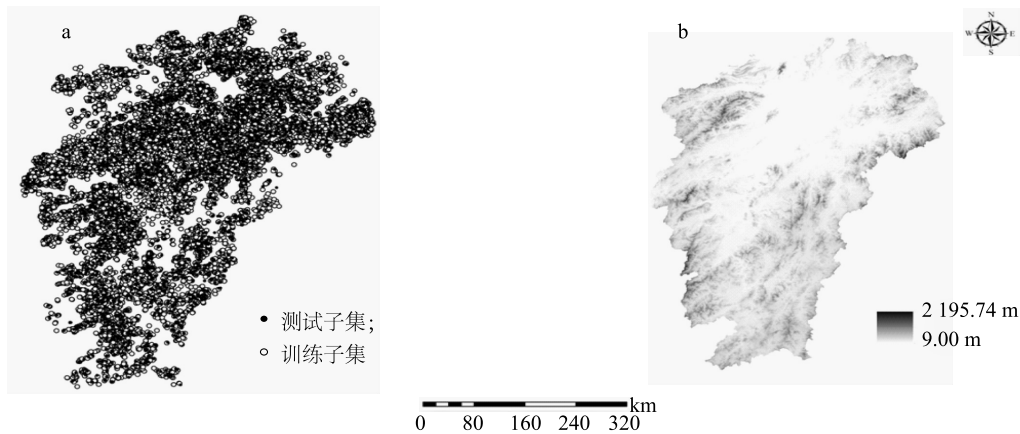


图1 江西省耕地表层土壤样点分布图(a)及DEM数据(b)

Fig.1 Distribution of farmland surface soil sampling sites(a) and DEM (b) of Jiangxi province

1.3 数据获取

1.3.1 邻近样点信息的获取 利用四方位搜索法(Four-direction search method)获取插值点附近 SOM 含量值。四方位搜索法基于 IDW 算法思想^[23]和空间自相关理论^[24],根据采样点之间的距离越近影响越大,距离越远影响越小的思想,以采样点为原点,变程 a 为半径,在 4 个象限内选择 4 个邻近点 SOM 含量值(P_1 、 P_2 、 P_3 、 P_4)。四方位搜索法步骤:

- 1) 在 MATLAB 中新建一个 mat 文件,导入所有采样点地理坐标(x, y)和 SOM 含量值(z);
- 2) 从文件中随机抽取 1 个点,赋值给 M_i ;
- 3) 在 mat 文件剩余点中随机抽取 1 个点,赋值给 N_i ;
- 4) 计算 M, N 2 点之间的欧氏距离,设为 d ,令 $Temp1 = Temp2 = Temp3 = Temp4 = a$,若 $d > a$,则回到步骤 3);
- 5) 若 $N_{i(x)} > M_{i(x)}$, $N_{i(y)} \geq M_{i(y)}$ 且 $d \leq Temp1$,则 $Temp1 = d, P_1 = N_{i(z)}$;
- 6) 若 $N_{i(x)} \leq M_{i(x)}$, $N_{i(y)} > M_{i(y)}$ 且 $d \leq Temp2$,则 $Temp2 = d, P_2 = N_{i(z)}$;
- 7) 若 $N_{i(x)} < M_{i(x)}$, $N_{i(y)} \leq M_{i(y)}$ 且 $d \leq Temp3$,则 $Temp3 = d, P_3 = N_{i(z)}$;
- 8) 若 $N_{i(x)} \geq M_{i(x)}$, $N_{i(y)} < M_{i(y)}$ 且 $d \leq Temp4$,则 $Temp4 = d, P_4 = N_{i(z)}$;
- 9) 重复步骤 2) ~ 步骤 8),直至遍历 mat 文件中

所有点,将 P_1 、 P_2 、 P_3 、 P_4 存放在文本中。

1.3.2 地形数据的获取 高程、坡度是能够较好反映土壤发生和发育的地形因子,因此常被用来预测土壤养分的空间分布^[25-26]。高程数据和坡度数据来源于地理空间数据云(<http://www.gscloud.cn/>)下载的 30 m 分辨率的江西省 DEM 数据和坡度栅格数据,借助 ArcGIS10.2 软件处理提取。

1.4 研究方法

人工神经网络(Artificial neural network, ANN)诞生于 20 世纪 40 年代,通常简称为神经网络,是由大量的人工神经元相互连接进行计算,根据外界的信息改变自身的结构,主要通过调整神经元之间的权值来对输入的数据进行建模,最终具备解决实际问题的能力。BP 神经网络和径向基神经网络是 2 种具有代表性的神经网络。

1.4.1 BP 神经网络与普通克里金法结合 BP 神经网络由输入层、隐含层、输出层构成,按误差信号反向传播、工作信号正向传播对网络权值进行修正,建立一种输入-输出映射关系。BPNN-OK:以地理坐标(x, y)、高程和坡度以及邻近样点信息(P_1 、 P_2 、 P_3 、 P_4)作为输入变量,SOM 含量值作为输出变量。通过 BP 神经网络建立输入与输出的最佳非线性映射关系,得到代表确定性部分的趋势项和代表随机性部分的残差项,然后把趋势项的预测结果导入 ArcGIS10.2 中转成栅格文件,同时利用普通克里金法对

残差项进行插值,最后将两者的结果进行叠加。其过程可表示为:

$$F_1(x) = G_1(x) + M_1(x) \quad (1)$$

式中: $F_1(x)$ 为BPNN-OK在 x 处的插值结果, $G_1(x)$ 为BPNN在 x 处的趋势项, $M_1(x)$ 为BPNN残差项在 x 处的普通克里金插值结果。在matlab中调用函数为newcf建立一个3层网络,由于BPNN对输入样本进行了归一化处理,所以隐含层和输出层分别采用tansig和purelin函数,训练函数使用trainlm函数,网络的拓扑结构为8-20-1(8表示网络的输入节点数,20表示网络的隐含层节点数,1表示输出的SOM含量)。步数为50、学习率为0.05,修改权值采用LMS(Least mean square algorithm)算法。

1.4.2 RBF神经网络与普通克里金法结合 RBF神经网络具有最佳逼近性能和全局逼近性质,是目前应用预测研究的热点研究之一^[21-22]。RBFNN-OK:以地理坐标(x, y)、高程和坡度以及邻近样点信息(P_1, P_2, P_3, P_4)作为输入变量,SOM含量作为输出变量。通过RBF神经网络建立输入与输出的最佳非线性映射关系,得到代表确定性部分的趋势项和代表随机性部分的残差项,然后把趋势项的预测结果导入ArcGIS10.2中转为栅格文件,同时利用普通克里金法对残差项进行插值,最后将两者的结果进行叠加。其过程可表示为:

$$F_2(x) = G_2(x) + M_2(x) \quad (2)$$

式中: $F_2(x)$ 为RBFNN-OK在 x 处的插值结果, $G_2(x)$ 为RBFNN在 x 处的趋势项, $M_2(x)$ 为RBFNN残差项在 x 处的普通克里金法插值结果。以误差最小为约束条件,调用MATLAB中的newrb函数对网络进行训练,设置误差容为 1×10^{-8} ,扩散因子为22,最大神经元个数为101。

1.4.3 普通克里金法 根据前人研究的结果^[10-11],本次研究将普通克里金法作为参照方法,普通克里金法作为传统的地统计学方法被广泛应用于SOM的空间插值中。利用普通克里金对江西省耕层表层

SOM含量的空间分布进行插值,提取样点插值结果并进行分析。

1.5 插值精度分析

通过ArcGIS10.2得到建模子集和测试子集,以均方根误差(RMSE)、平均绝对误差(MAE)、平均相对误差(MRE)对建模点和测试的SOM含量预测值与实测值进行对比分析,得出精度评价结果,其公式分别为:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n [SOM_{ob(i)} - SOM_{pv(i)}]^2}{n}} \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |SOM_{ob(i)} - SOM_{pv(i)}| \quad (2)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \frac{|SOM_{ob(i)} - SOM_{pv(i)}|}{Z_i} \times 100\% \quad (3)$$

式中: $SOM_{pv(i)}$ 为样点的预测值; $SOM_{ob(i)}$ 为样点的实测值; n 为样点数,RMSE、MAE和MRE值越小,说明误差越小、模拟精度越高。

本研究中常规性统计分析均在SPSS(version 22)中计算,四方位搜索法、BP神经网络和RBF神经网络均在MATLAB R2014a中通过编程实现,半方差函数分析主要借助于GS+(version 7)软件,空间分析主要在matlab和ArcGIS10.2中进行计算和叠加。

2 结果与分析

2.1 SOM的统计特征

描述性统计分析结果(表1)表明,SOM含量的平均值为30.22 g/kg,值域范围为12.00~50.00 g/kg。从K-S检验可以判断SOM含量服从正态分布。根据变异系数的大小,江西省SOM的变异系数为0.1~1.0,属于中等变异。同时对训练子集和测试子集进行常规统计分析,训练子集和测试子集的变异系数分别为0.29和0.30,为中等变异。

表1 SOM含量的描述性统计

Table 1 Descriptive statistics of soil organic matter (SOM)

项目	样本数	最大值 (g/kg)	最小值 (g/kg)	平均值 (g/kg)	中位数 (g/kg)	标准差	偏度系数	K-S 检验	变异系数
全部样点	16 109	50.00	12.00	30.22	29.90	8.88	0.12	0.56	0.29
训练子集	12 887	50.00	12.00	30.27	30.00	8.84	0.11	0.43	0.29
测试子集	3 222	50.00	12.10	30.04	29.60	9.05	0.14	0.38	0.30

数据的正态性也是 OK 法的前提,只有当数据服从正态分布时,普通克里金法才有效^[15]。图 2 是对 16 109 个 SOM 含量的样点值、RBFNN-OK 预测残

差值和 BPNN-OK 预测残差值做出的直方图和正态分布曲线,表明三者基本符合正态分布,可以进行半方差分析和普通克里金法插值。

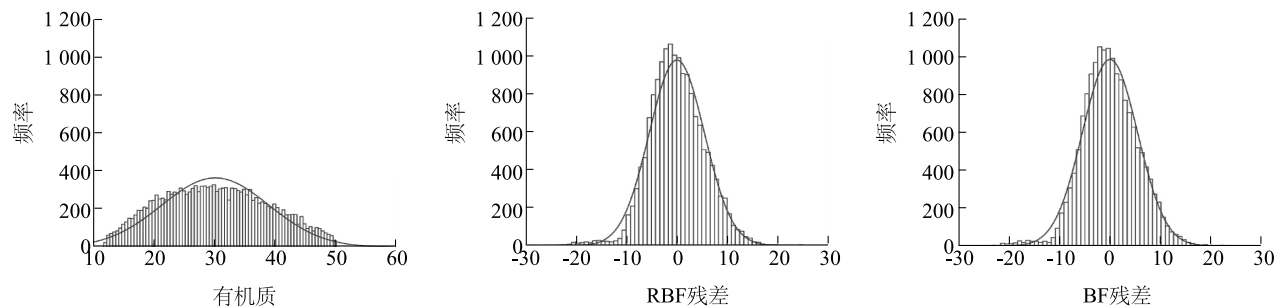


图 2 SOM、RBFNN 残差和 BPNN 残差的分布频率
Fig.2 Histogram of the SOM and the SOM residuals from RBF and BP regression

2.2 半方差函数分析

半方差函数是描述 SOM 含量空间变异性的最有效方法,是 SOM 含量等地理属性进行空间插值的基础^[15]。经过反复试验得到的最优模型均为球状模型。从模型的参数(表 2)来看,SOM 含量的块金

效应值达到 0.80,变程为 18 734.25 m,表明 SOM 含量的空间相关性较弱,空间自相关的范围比较大。2 种神经网络残差的半方差模型参数与原变量变化较小,基本保留了原变量的空间结构特征。

表 2 SOM、RBF 预测残差和 BP 预测残差的半方差函数参数

Table 2 Semi-variogram parameters of log-transformed SOM and predicted residuals by BPNN and RBFNN							
项目	数据项	模型	块金值	基台值	块金效应	变程(m)	决定系数
SOM	SOM 含量值	球状	55.91	69.99	0.80	18 734.25	0.88
	RBF 残差	球状	26.89	32.29	0.83	8 259.94	0.92
	BP 残差	球状	26.35	31.77	0.83	8 088.98	0.93

2.3 预测能力对比分析

用 BPNN-OK、RBFNN-OK 和普通克里金法对 SOM 含量的预测结果与实测值进行相关分析,结果表明 2 种神经网络对 SOM 的预测值与实测值的相关系数要大于普通克里金法(表 3)。其中,BP 神经网络模型对 SOM 含量的训练样点、测试样点和全部样点的预测结果与实测值的相关系数较普

通克里金法分别提高了 0.177、0.270 和 0.190;RBF 神经网络模型与比普通克里金法分别提高了 0.171、0.265 和 0.189。表明在省域尺度下,神经网络模型较普通克里金法能够更加真实地反映 SOM 的空间分布,有效揭示 SOM 含量和因变量之间的复杂非线性关系。

表 3 神经网络模型和普通克里金预测结果与实测值的相关性

土壤养分	训练样点			测试样点			全部		
	OK	RBFNN-OK	BPNN-OK	OK	RBFNN-OK	BPNN-OK	OK	RBFNN-OK	BPNN-OK
SOM 含量	0.616 **	0.787 **	0.793 **	0.524 **	0.789 **	0.794 **	0.598 **	0.787	0.788 **

OK:普通克里金法;RBFNN-OK:RBF 神经网络模型与普通克里金结合法;BPNN-OK:BP 神经网络模型与普通克里金结合法。* 表示相关差异显著($P<0.05$);** 表示相关极显著($P<0.01$)。

2.4 插值精度评价

预测精度分析结果(表4)表明,运用神经网络进行 SOM 空间分布预测的方法(BPNN-OK 和 RBFNN-OK)要明显优于普通克里金法。从训练集预测效果看,与普通克里金法进行比较,BPNN-OK 对 SOM 预测结果的 *RMSE*、*MAE*、*MRE* 与普通克里金法比较分别降低了 22.90%、25.13%、28.56%,RBFNN-OK 分别降低了 22.03%、24.06%、27.61%。从测试集预测效果看,BPNN-OK 对 SOM 预测结果

的 *RMSE*、*MAE*、*MRE* 与普通克里金法比较分别降低了 28.66%、30.71%、34.76%,RBFNN-OK 分别降低了 27.76%、29.74%、33.71%。从全部样点预测效果看,BPNN-OK 对 SOM 预测结果的 *RMSE*、*MAE*、*MRE* 与普通克里金法比较分别降低了 23.53%、25.48%、28.84%,RBFNN-OK 分别降低了 23.25%、25.31%、28.98%。运用神经网络模型进行 SOM 空间分布预测精度得到了较为明显的提高。

表4 SOM 预测方法精度对比

Table 4 Precisions of different approaches for predicting SOM

项目	方法	训练集预测效果			测试集预测效果			全部		
		<i>RMSE</i>	<i>MAE</i>	<i>MRE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MRE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MRE</i>
SOM	OK	6.99	5.61	21.04	7.71	6.22	23.82	7.14	5.73	21.60
	RBFNN-OK	5.45	4.26	15.23	5.57	4.37	15.79	5.48	4.28	15.34
	BPNN-OK	5.39	4.20	15.03	5.50	4.31	15.54	5.46	4.27	15.37

OK、RBFNN-OK、BPNN-OK 见表3注。

图3显示,3种方法对 SOM 空间分布的预测结果非常相似,与普通克里金法相比,RBFNN-OK 和 BPNN-OK 法较普通克里金法在预测范围、预测误差、模型拟合度方面都表现出明显的优势。从宏观角度可知,RBFNN-OK 和 BPNN-OK 预测结果中 SOM 的空间分布与普通克里金法预测结果非常接近,高低趋势分布相当,说明 RBFNN-OK 和 BPNN-OK 预测 SOM 的空间分布可信度很高。从微观角度看,RBFNN-OK 和 BPNN-OK 预测 SOM 的空间分布克服了普通克里金法仅基于邻近相关采样点权重均值而难以捕捉复杂地形下 SOM 空间分布的缺点,同

时普通克里金法预测结果较平滑,难以表达 SOM 的空间变异信息,而神经网络方法预测结果高低值斑块丰富,能在一定程度上体现 SOM 的空间变异特征。在省域尺度上,SOM 的空间自相关性范围大(表2),同时地形因子也不如县域那么明显^[27],普通克里金法难以表达 SOM 含量突变的区域。而神经网络方法根据 SOM 的空间自相关性和影响 SOM 的地形因素把邻近信息(P_1 、 P_2 、 P_3 、 P_4)、高程和坡度同时纳入神经网络模型,能够有效提高 SOM 空间分布预测精度,使 SOM 空间分布模拟更加符合实际情况。

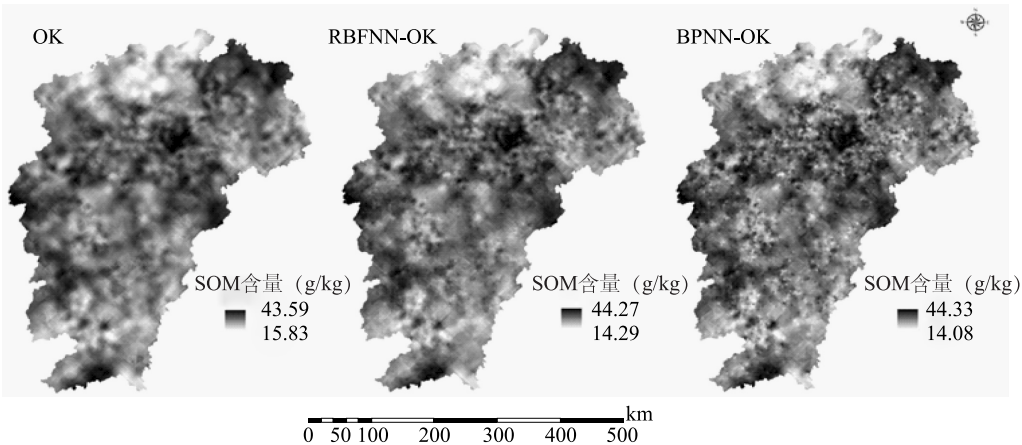


图3 江西省 SOM 空间分布预测结果

Fig.3 Spatial distribution of SOM in Jiangxi province predicted with different approaches

3 讨论

对省域尺度下 SOM 的空间分布进行模拟及准确预测是进行 SOM 数字制图、地表过程模拟及土壤肥力有效调控与提升的基础工作,是开展其他土壤学相关工作的前提之一。地形与邻近信息在 SOM 形成及运移过程中均扮演着重要角色,起到了不可忽视的作用,也是 SOM 分布模拟研究中必须考虑的因素。

由于土壤结构的复杂性、气候带的渐变性以及一些非地带性因素的影响,使得 SOM 在一定范围内存在空间自相关性,这是空间自相关理论的基础^[28]。前人研究结果表明,仅基于地理坐标(x, y)而不考虑地形与邻近信息的神经网络模型空间分布模拟效果不理想,误差较大^[29],而融合地形因子或邻近信息的神经网络模型能够在尊重原始数据的情况下更加真实地反映土壤性质空间变异信息^[20-22],与本研究结果一致。这主要是由于神经网络具有极强的映射能力,能以任意精度逼近任意连续函数,因此对 SOM 的拟合效果、预测精度和异质性的描述均较为理想。在各种预测方法中,普通克里金法是应用最普遍也最具代表性的,虽然能够通过邻近相关采样点土壤属性值权重来预测未知点,同时对预测结果给出误差,但该方法没有考虑地形因子和邻近信息的影响。由于普通克里金法理论假设、前提条件多、半变异函数的拟合曲线受主观因素影响大以及预测精度偏低等导致 2 种方法在土壤性质空间异质性描述时产生较大差异^[5,30]。神经网络模型不仅能够构建 SOM 与地理坐标(x, y)、高程、坡度以及邻近信息(P_1, P_2, P_3, P_4)之间的复杂非线性映射关系,而且在残差存在空间自相关的情况下充分利用神经网络预测残差进行普通克里金插值,考虑了采样点的结构性和随机性的空间模拟,对以后指导 SOM 插值模型选取与精度优化提供了有力的工具。

省域尺度下地形条件复杂, SOM 的空间分布预测方法中必须要考虑地形因子和邻近信息引起的 SOM 空间异质性,否则必然会导致 SOM 过高或过低的预测。本研究引入地形因子与邻近信息的神经网络模型能够较好模拟 SOM 异质性,该方法要优于常用的普通克里金法。省域尺度下耕层 SOM 空间分布格局除了高程以及坡度外,其他地形因子如地形起伏度、地形湿度指数、平面曲率等和人为活动因

素影响也都与 SOM 含量高低有着密切的关系,如果有时间资料的累积,也可将时间因子引入 2 种神经网络,同时本研究选取的 DEM 分辨率较低,获取的环境变量仅为高程和坡度,没有考虑其他环境因素或者其他类型数据。因此在预测方法中,充分考虑各种因素和它们之间的关系以及各因素对 SOM 分布的作用机理,才能进一步提高省域尺度下 SOM 的空间分布预测结果使其更加接近实际的空间分布。

本研究以江西省耕地表层 SOM 含量为研究对象,基于 BP 神经网络模型与普通克里金结合法(BPNN-OK)、RBF 神经网络模型与普通克里金结合法(RBFNN-OK),利用地理坐标(x, y)、高程、坡度以及邻近样点信息(P_1, P_2, P_3, P_4)作为网络的输入变量对 SOM 空间分布进行预测,并与普通克里金法进行对比。对 3 222 个测试样点预测结果的误差分析结果表明:在省域尺度下, BPNN-OK 法对 SOM 预测结果的均方根误差、平均绝对误差、平均相对误差较普通克里金法分别降低了 28.66%、30.71%、34.76%, RBFNN-OK 法分别降低了 27.76%、29.74%、33.71%。可以看出,神经网络模型结合普通克里金法预测精度明显高于普通克里金法,该方法可为省域尺度下土壤性质的空间插值模型选取和精度优化提供参考。

参考文献:

- [1] McBRATNEY A B, PRINGLE M J. Estimating average and proportional variograms of soil properties and their potential use in precision agriculture[J]. Precision Agriculture, 1999, 1(2): 125-152.
- [2] DAI W, HUANG Y. Relation of soil organic matter concentration to climate and altitude in zonal soils of China[J]. Catena, 2006, 65(1): 87-94.
- [3] 李子忠, 龚元石. 农田土壤水分和电导率空间变异性及确定其采样数的方法[J]. 中国农业大学学报, 2000, 5(5): 59-66.
- [4] 连 纲, 郭旭东, 傅伯杰, 等. 黄土高原县域土壤养分空间变异特征及预测——以陕西省横山县为例[J]. 土壤学报, 2008, 45(4): 577-584.
- [5] 刘二永, 刘 健, 余坤勇, 等. 基于环境因子和 R-STPS 的林地土壤有机质预测模型[J]. 农业机械学报, 2015, 46(1): 133-137.
- [6] 胡克林, 余 艳, 张凤荣, 等. 北京郊区土壤有机质含量的时空变异及其影响因素[J]. 中国农业科学, 2006, 39(4): 764-771.
- [7] 韩 冰, 王效科, 欧阳志云, 等. 辽宁省农田土壤碳库分布及变化的模拟分析[J]. 生态学报, 2003, 23(7): 1321-1327.
- [8] 龙 军, 张黎明, 沈金泉, 等. 复杂地貌类型区耕地土壤有机质空间插值方法研究[J]. 土壤学报, 2014, 51(6): 1270-1281.

- [9] 辛展, 娄华君, 李颖, 等. 泊江海子流域土壤有机质、全氮空间分布特征[J]. 江苏农业科学, 2015, 43(3): 323-326.
- [10] 刘东海, 李双来, 陈云峰, 等. 武汉市汉南区土壤有机质和养分时空变异研究[J]. 中国土壤与肥料, 2014(4): 24-28.
- [11] 杨顺华, 张海涛, 郭龙, 等. 基于回归和地理加权回归 Kriging 的土壤有机质空间插值[J]. 应用生态学报, 2015, 26(6): 1649-1656.
- [12] 黄安, 杨联安, 杜挺, 等. 基于多元成土因素的土壤有机质空间分布分析[J]. 干旱区地理, 2015, 38(5): 994-1003.
- [13] 马驰. 基于 Landsat8 遥感图像的长春中北部地区土壤有机质含量反演研究[J]. 江苏农业科学, 2016, 44(12): 415-419.
- [14] 马泉来, 高凤杰, 张志民, 等. 我国东北黑土丘陵区小流域土壤有机质空间分布模拟[J]. 环境科学研究, 2016, 29(3): 382-390.
- [15] 史舟, 李艳. 地统计学在土壤学中的应用[M]. 北京: 中国农业出版社, 2006.
- [16] 沈掌泉, 施洁斌, 王珂, 等. 应用集成 BP 神经网络进行田间土壤空间变异研究[J]. 农业工程学报, 2004, 20(3): 35-39.
- [17] 董敏, 王昌全, 李冰, 等. 基于 GARBF 神经网络的土壤有效锌空间插值方法研究[J]. 土壤学报, 2010, 47(1): 42-50.
- [18] 李启权, 王昌全, 岳天祥, 等. 基于神经网络模型的中国表层土壤有机质空间分布模拟方法[J]. 地球科学进展, 2012, 27(2): 175-184.
- [19] 沈润平, 丁国香, 魏国栓, 等. 基于人工神经网络的土壤有机质含量高光谱反演[J]. 土壤学报, 2009, 46(3): 391-397.
- [20] 李启权, 王昌全, 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.
- [21] 徐剑波, 宋立生, 夏振, 等. 基于 GARBF 神经网络的耕地土壤有效磷空间变异分析[J]. 农业工程学报, 2012, 28(16): 158-165.
- [22] 李启权, 王昌全, 岳天祥, 等. 不同输入方式下 RBF 神经网络对土壤性质空间插值的误差分析[J]. 土壤学报, 2008, 45(2): 360-365.
- [23] 王玉璟. 空间插值算法的研究及其在空气质量监测中的应用[D]. 开封: 河南大学, 2010.
- [24] 朱会义, 刘述林, 贾绍凤. 自然地理要素空间插值的几个问题[J]. 地理研究, 2004, 23(4): 425-432.
- [25] 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.
- [26] 李启权, 王昌全, 张文江, 等. 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[J]. 应用生态学报, 2013, 24(2): 459-466.
- [27] 夏春兰. 多尺度条件下川中丘陵区土壤氮素分布的模拟研究[D]. 雅安: 四川农业大学, 2013.
- [28] 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): 161-169.
- [29] 李启权, 王昌全, 张文江, 等. 丘陵区土壤有机质空间分布预测的神经网络方法[J]. 农业环境科学学报, 2012, 31(12): 2451-2458.
- [30] 沈掌泉, 周斌, 孔繁胜, 等. 应用广义回归神经网络进行土壤空间变异研究[J]. 土壤学报, 2004, 41(3): 471-475.

(责任编辑: 陈海霞)