

旱作区阳离子交换量土壤转换函数的研究

郭孝理^{1,2*}, 曹梦³

¹煤炭科学研究总院有限公司, 北京

²中煤科工集团北京土地整治与生态修复科技研究院有限公司, 北京

³中国建筑一局(集团)有限公司生态园林分公司, 北京

收稿日期: 2023年2月22日; 录用日期: 2023年3月23日; 发布日期: 2023年3月30日

摘要

为减少大范围测定阳离子交换量(cation exchange capacity, CEC)工作量, 探寻一种利用较容易测定的土壤理化性质预测CEC简单、高效的方法。本文基于旱作区152个表层0~10 cm土壤的pH、质地、有机质等数据, 建立BP神经网络、支持向量机、多元线性回归3种CEC土壤转换函数, 并比较分析了各方法的预测精度和敏感性。结果表明, CEC主要受有机质、黏粒的影响且呈明显的正相关关系; 支持向量机预测精度最高, 其决定系数($R^2 = 0.58$)、效率系数($E = 0.57$)均高于BP神经网络和多元线性回归, 均方根误差($RMSE = 5.41$)均低于其他模型, 多元线性回归模型预测精度最低。支持向量机方法在旱作区内能够较好地预测阳离子交换量。

关键词

旱作区, 阳离子交换量, BP神经网络, 支持向量机, 多元线性回归

Study on Pedo-Transfer Functions of Cation Exchange Capacity in Dry Farming Region

Xiaoli Guo^{1,2*}, Meng Cao³

¹Chinese Institute of Coal Science, Beijing

²CCTEG Beijing Research Institute of Land Renovation and Ecological Restoration Technology Co., Ltd., Beijing

³China Construction First Group Company Limited Ecology and Landscape Branch, Beijing

Received: Feb. 22nd, 2023; accepted: Mar. 23rd, 2023; published: Mar. 30th, 2023

Abstract

In order to reduce the workload of large-scale determination of cation exchange capacity (CEC), a

*通讯作者。

simple and efficient method for predicting CEC by using easily measured soil physical and chemical properties was explored. Researcher have used different input soil properties to derive pedo-transfer functions (PTFs) to predict soil CEC. Based on the pH, texture and organic matter (OM) of 152 topsoil (0~10 cm) samples in dry farming region, the CEC soil transfer function was established by back propagation neural network (BP-NN), support vector machine (SVM) and multiple linear regression (MLR), and the prediction accuracy and sensitivity of each method were compared and analyzed. The results showed that CEC was mainly positively correlated with OM and clay. The prediction accuracy of SVM was the highest, with its decision coefficient ($R^2 = 0.58$) and efficiency coefficient ($E = 0.57$) higher than BP-NN and MLR. The root mean square error of SVM (RMSE = 5.41) is lower than other models, and MLR model has the lowest prediction accuracy. SVM method can better predict cation exchange capacity in dry farming.

Keywords

Dry Farming Region, Cation Exchange Capacity, BP-Neural Network, Support Vector Machines, Multiple Linear Regression

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

阳离子交换量(CEC)作为衡量土壤肥力和生产力的重要指标,它能够直接反应土壤的保肥性能和缓冲性能。我国东北和黄淮海旱地作物种植区是重要商品粮基地,从北到南土壤类型复杂,土壤性质差异性大[1],准确获取CEC对制定土壤肥力调控措施有重要意义。然而,传统的CEC测定方法存在着成本高、耗时长、工作繁重等缺点,利用易测定的土壤基本理化指标建立土壤转换函数(PTFs)间接预测CEC是一种高效、快捷的方法[2]。常用的指标为pH、土壤有机质、黏粒含量等[3] [4]。建立PTFs常用方法是人工神经网络、多元线性[5]与非线性回归、分类树回归等方法。人工神经网络可以从有限的数据集中通过相关性学习复杂的函数关系,并根据环境的变化对模型进行调整,构建非线性、多变量动态系统黑箱模型[6]。BP神经网络因其结构简单、操作性强、收敛速度快等优点广泛用于非线性建模、函数逼近、模式识别等方面[7]。Vapnik提出支持向量机(SVM)可以代替人工神经网络,它是非线性分类、回归和时间序列预测的有力工具[8]。

为此,本文以旱作区表层0~10 cm土壤为研究对象,建立了BP神经网络(BP-NN)、支持向量机、多元线性回归(MLR)三种的PTFs,通过比较分析三种模型的精度和敏感性,提出适用于旱作区预测CEC的PTFs函数,以为合理获取旱作区的土壤参数提供依据,也为土壤、作物等模型参数获取提供理论支持。

本文的组织结构如下:第2节介绍研究区域数据来源与处理、研究方法;第3节介绍SVM、BP-NN、MLR三种PTFs的模型构建结果及精度检验与敏感性分析;第4节介绍了输入变量对PTFs预测精度的影响及三种PTFs在旱作区的有效性 with 适用性;在第5节给出了结论与展望。

2. 材料与方法

2.1. 研究区域

旱作区以地形坡度小于 5° 、 1 km^2 网格内旱地占耕地比在40%以上作为划分依据,界定其内涵与范围。

该区域涵盖东北黑龙江、辽宁和吉林三省, 以及黄淮海的河北、北京、山东、河南、安徽五省市, 主要种植玉米、小麦等粮食作物和大豆、甜菜等经济作物, 是中国重要商品粮基地。地理坐标为 $113^{\circ}03'51''\sim 134^{\circ}5'31''\text{E}$, $32^{\circ}21'17''\sim 48^{\circ}57'51''\text{N}$, 研究区属温带季风气候, 年均气温 $2.4^{\circ}\text{C}\sim 16.1^{\circ}\text{C}$, 年均降水量 $435.9\sim 957.4\text{ mm}$, 土壤类型主要包括潮土、褐土、黑钙土、黑土、砂姜黑土、棕壤、暗棕壤。

2.2. 数据来源与处理

研究区采用网格均匀布点和分层抽样相结合方法的布点采样。该方法在考虑空间自相关基础上, 采用分层抽样避免随机采样中样点的空间聚集, 考虑环境相似性, 可使用少量样点代表相似的环境下土壤的信息, 提高采样效率。采样网格大小为 $15\text{ km}\times 15\text{ km}$, 在此基础上根据土壤类型、土壤颗粒组成与耕地质量等级等进行分层抽样。采样时间为 2017 年 4~6 月, 共取 152 个 0~10 cm 表层土壤样品, 每个样点用 GPS 记录其位置, 样点分布如下(见图 1)。每个样点随机设置 3 个重复。样品经自然风干后混合过 2 mm 筛备用。土壤质地采用激光粒度分析仪测定, 土壤颗粒分级标准采用美国制。土壤有机质采用重铬酸钾外加热法测定。土壤 pH 采用电位法测定。土壤 CEC 采用乙酸铵交换法、乙酸钠-火焰光度法测定。

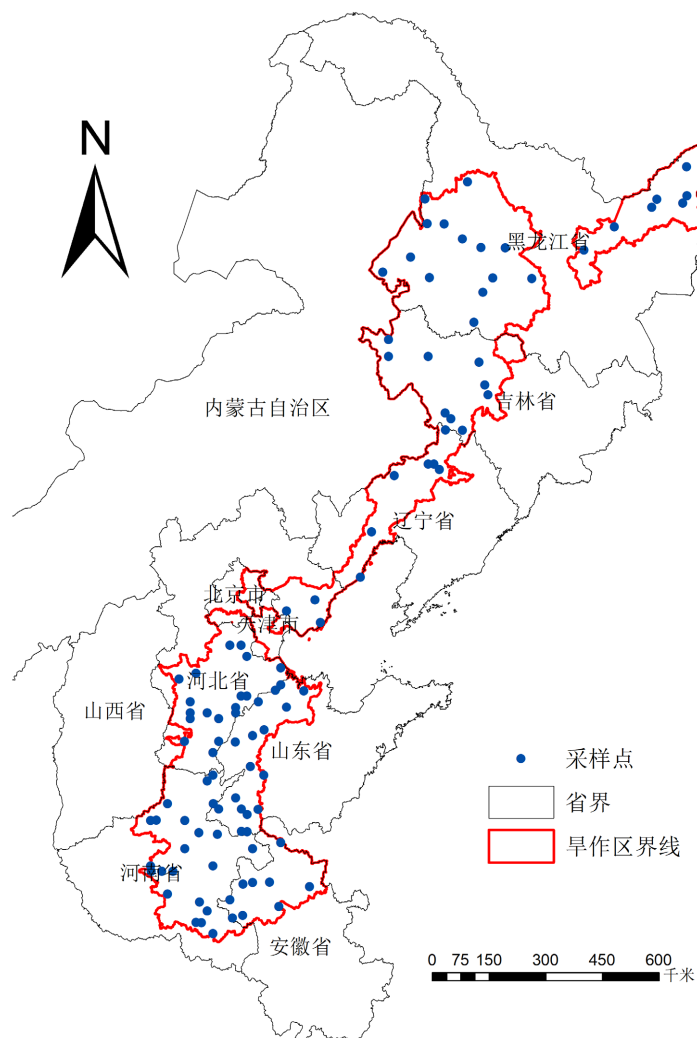


Figure 1. Location and sample distribution of the study area
图 1. 研究区位置及样点分布

样点各数据的描述性统计、相关分析以及多元线性回归采用 SPSS 软件, BP-NN、SVM 建模采用 Matlab 软件。

因各个变量单位不同, 为取消各数据间数量级差别, 避免因输入、输出数据数量级别差别较大而造成预测误差较大[9], 在预测前首先要将数据集通过公式(1)进行归一化处理。

$$X_k = \frac{x_k - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

式中: x_{\min} 为数据序列中最小值; x_{\max} 为数据序列中最大值; X_k 为归一化后的值。

最后将模型输出结果进行反归一化。

2.3. 研究方法

在本研究中选择的自变量为: pH、砂粒、黏粒、有机质含量, 因变量为 CEC, 这与学者们的研究[3][4][9]一致。

2.3.1. 多元线性回归

在 MLR 分析中, 5%显著水平下的预测变量为主要变量[2]。在此回归中将有机质、砂粒、粉粒、黏粒含量和 pH 作为自变量, 并将 CEC 作为因变量利用 SPSS 软件按照公式(2)形式推导回归模型。

$$y = b_1x_1 + b_2x_2 + \dots + b_nx_n + c \quad (2)$$

式中: b_1, b_2, \dots, b_n 为自变量, y 为因变量, c 为常数。

2.3.2. BP 神经网络

神经网络是一种自适应系统, 它从输入和输出数据集学习关系, 并能预测具有与输入集相似特征的数据集[10]。最初是为生物神经系统的性能建模而设计的[11]。BP-NN 是一种多层前馈神经网络, 该网络主要包括三层: 输入层、隐含层、输出层, 每层均由神经元组成, 并通过连接权重和前一层进行连接。网络学习目的是获得恰当的权、阈值, 由正向和反向传播两部分组成[12]。

在 BP-NN 中, 隐含层节点数的选择对 BP 神经网络预测精度有较大的影响: 节点数太少, 网络不能很好地学习, 需要增加训练的次数, 精度也受影响; 节点数太多, 训练时间增加, 网络容易过拟合[13]。通过公式(3)选择最佳隐含层节点数:

$$l < \sqrt{(i+j)} + a \quad (3)$$

式中: a 为 0~10 之间的常数。

隐含层节点数选择首先是参考公式来确定节点数的大概范围, 然后用试凑法确定最佳的节点数。

2.3.3. 支持向量机

支持向量机是建立在统计学习理论和结构风险最小化原则基础上的机器学习方法。它能够根据有限的样本信息, 在模型复杂性和学习能力之间寻求最佳折衷[14]。SVM 提供了一种使用内核函数将超平面表面拟合到训练数据的机制。

对于本研究的回归情况, 首先执行从输入空间到高维特征空间的映射, 然后使用 - 不敏感损失通过特征空间中的超平面执行线性回归。

对于给定的 n 个训练集, 分别为 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 其中 x 是输入向量, y 是对应的输出值, 回归的 SVM 估计量 f 如公式(4)所示:

$$f = \omega \cdot \varphi(x) + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) k(x \cdot x_i) + b \quad (4)$$

式中: w 是权重向量, b 为偏差, “ \cdot ”表示点集, φ 为非线性映射函数, α_i, α_i^* 为拉格朗日乘子, $k(x \cdot x_i)$ 为核函数。

2.3.4. 模型精度检验

本文以平均绝对误差(MAE)、决定系数(R^2)、均方根误差(RMSE)及 Nash-Sutcliff 效率系数(E)四个指标来评估模型预测精度。其中 MAE 越小预测精度越高; R^2 越大预测精度越高, 自变量对因变量的解释程度越高, 自变量引起的变动占总变动的百分比越高; RMSE 可以对模型的拟合优度提供均衡评估, RMSE 越小预测精度越高; E 的范围从 $-\infty$ 到 1, 当 $E = 1$ 时获得的模型是最优拟合。如公式(5)~(8)所示:

$$\text{MAE} = \frac{\sum_{i=1}^n |(CEC_p - CEC_m)|}{n} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (CEC_m - \overline{CEC_p})^2}{\sum_{i=1}^n (CEC_m - \overline{CEC_m})^2} \quad (6)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n ((CEC_m)_i - (CEC_p)_i)^2}{n}} \quad (7)$$

$$E = 1 - \frac{\sum_{i=1}^n [(CEC_m)_i - (CEC_p)_i]^2}{\sum_{i=1}^n [(CEC_m)_i - (\overline{CEC_m})]^2} \quad (8)$$

式中: 下标 “ m ” 和 “ p ” 代表实测值和预测值; i 为第 i 个样本; n 是样本数。

2.3.5. 模型敏感性分析

敏感性分析可以确定参数对模型模拟结果影响的大小。本文用回归分析法对 BP-NN、SVM 模型中输入的土壤参数对 CEC 的影响进行全局敏感性分析, 其中将标准回归系数表示各参数的敏感性。

3. 结果与分析

3.1. 描述性统计分析

样本集随机将 75% 的样本分为训练集, 剩余部分为测试集, 样本集的土壤理化性质变异系数为 19.89%~75.14% (见表 1), 其中有机质变异系数约为 CEC 变异系数的 1.6 倍。根据 Nielsen [15] 和李哈滨 [16] 对变异程度的定义, 变异系数 < 10% 为弱变异性, 变异系数介于 10% 与 100% 之间为中等变异性, 变异系数 > 100% 为强变异性。研究区域内样本集的 pH、砂粒、粉粒、黏粒含量、有机质、CEC 等 6 种土壤理化性质均处于中等强度变异水平。样本集、训练集、测试集中土壤理化性质具有相似的变异系数。

3.2. 相关性分析

对样本集的各变量与 CEC 之间进行相关性分析。在 $p < 0.01$ 情况下, 有机质、黏粒和砂粒含量与 CEC 相关系数分别为: 0.656、0.449、-0.438, 有机质、黏粒含量与 CEC 存在明显的正相关关系, 而砂粒含量与 CEC 之间存在明显的负相关关系(见表 2)。砂粒与粉粒含量之间相关系数为-0.994, 表明它们存在着极为强烈的负相关关系, 这是因为在该研究区域内土壤大多数为粉土、粉壤, 黏粒含量大都在 10% 以下, 而砂粒、粉粒、黏粒含量定和为 1, 因此砂粒和粉粒含量之间会呈现明显的负相关关系。在不显著水平下, pH 与 CEC 的相关系数为-0.110 呈现弱负相关性, 其仅是在统计分析结果。在自然界中, 随着 pH 升高土壤胶体中负电荷数量增加, CEC 也会随之变大。这可能是因为研究区域内土壤性质差异较大导致。

Table 1. Description statistics of soil physico-chemical properties**表 1.** 土壤理化性质描述性统计

	统计量	最小值	最大值	均值	标准差	变异系数(%)
样本集						
pH	152	4.10	9.05	6.97	1.39	19.89
砂粒(%)	152	1.32	83.38	23.86	17.93	75.14
粉粒(%)	152	15.19	90.96	71.45	16.82	23.54
黏粒(%)	152	0.74	12.55	4.69	2.14	45.72
有机质(g·kg ⁻¹)	152	2.88	85.77	22.44	14.20	63.29
CEC (cmol·kg ⁻¹)	152	3.07	53.30	22.54	8.59	38.10
训练集						
pH	116	4.10	8.97	6.89	1.39	20.15
砂粒(%)	116	1.32	77.88	24.25	17.90	73.81
粉粒(%)	116	20.67	90.15	71.16	16.71	23.48
黏粒(%)	116	0.74	12.55	4.59	2.19	47.78
有机质(g·kg ⁻¹)	116	2.88	77.13	22.19	13.26	59.73
CEC (cmol·kg ⁻¹)	116	6.36	53.30	22.20	8.67	39.04
测试集						
pH	36	4.53	9.05	7.23	1.37	18.91
砂粒(%)	36	4.30	83.38	22.61	18.23	80.63
粉粒(%)	36	15.19	90.96	72.38	17.37	24.00
黏粒(%)	36	1.44	11.60	5.01	1.97	39.47
有机质(g·kg ⁻¹)	36	5.32	85.77	23.24	17.09	73.54
CEC (cmol·kg ⁻¹)	36	3.07	41.51	23.66	8.36	35.34

Table 2. Correlation analysis of soil physico-chemical properties**表 2.** 土壤理化性质相关性分析

	pH	粉粒(%)	砂粒(%)	黏粒(%)	有机质(g·kg ⁻¹)	CEC (cmol·kg ⁻¹)
pH	1					
粉粒(%)	-0.158	1				
砂粒(%)	0.177*	-0.994**	1			
黏粒(%)	-0.241**	0.472**	-0.563**	1		
有机质(g·kg ⁻¹)	-0.397**	0.173*	-0.190*	0.237**	1	
CEC (cmol·kg ⁻¹)	-0.110	0.410**	-0.438**	0.449**	0.656**	1

注: *在 0.05 水平上显著相关, **在 0.01 水平上显著相关。

3.3. 土壤转化函数构建

3.3.1. MLR 转换函数

在多元回归分析时需要满足的条件之一是变量总体服从正态分布。因此, 利用 K-S 检验方法来验证数据是否符合正态分布。本研究中将变量进行对数转换后符合正态分布。表 1 中训练集进行多元逐步回归方程如公式(9)所示, 测试集来进行预测 CEC。回归方程的 R^2 为 0.52。

$$\text{CEC} = 0.415 * \text{OM} - 0.113 * \text{sand} + 0.480 * \text{pH} + 0.148 * \text{clay} \quad (9)$$

式中: OM 为有机质含量; sand 为砂粒含量; clay 为黏粒含量。

3.3.2. BP-NN 转换函数

根据公式(3)计算得到隐含层节点数范围为 1~12, 然后通过试凑法得到最佳隐含层神经元数量。误差随隐含层节点数的变化过程, 误差随节点数增加呈现先增加后减少再增加的现象(见图 2)。综合比较隐含层节点数的 RMSE 和 MAE, 当隐含层神经元数为 9 时, BP-NN 的性能最好。因此, BP-NN 的结构为 4-9-1, 即 4 个输入变量、9 个隐含层神经元, 1 个输出变量。模型训练次数为 10000 次, 训练精度为 0.001, 学习速率为 0.01 s, 隐含层、输出层传递函数为 tansig, 训练函数为 traingdx。训练集用来训练 BP-NN, 测试集用来预测 CEC。

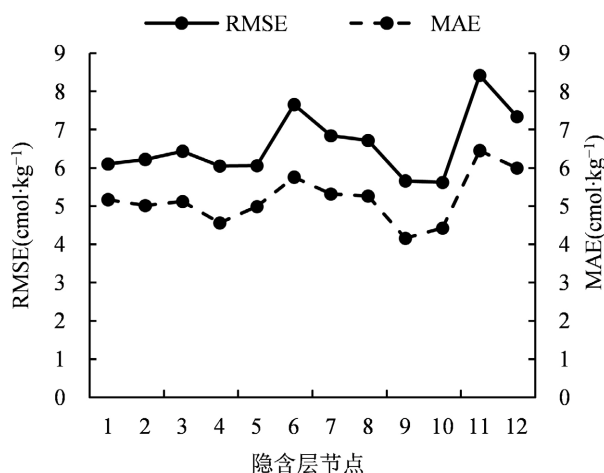


Figure 2. Influence of the number of neurons in the hidden layer on prediction accuracy

图 2. 隐含层神经元节点数对预测精度的影响

3.3.3. SVM 转换函数

在支持向量机中应用最广泛的高斯径向基核(RBF)函数, 具有较宽的收敛域, 且只含 1 个参数, 易于优化[17], 选择 SVM 类型为支持向量回归(SVR), 核函数为径向基核函数, 在 RBF 核函数中需要确定的参数是核函数的参数系数、惩罚因子 c 。本研究通过网格参数寻优来寻找一组最优的 c 、分别为 1.00、0.33, 训练精度设置为 0.01。

3.4. 函数精度分析

MLR、BP-NN 和 SVM 三种模型建立的 PTFs 用于预测测试集土壤样本 CEC。测试集 CEC 的实测值与三种模型所得到的预测值之间的散点图见图 3。BP-NN 的 MAE 最小, 表明 BP-NN 误差整体上小于 SVM 和 MLR (见表 3)。MLR、BP-NN、SVM 三种模型的 R^2 , 其中 SVM 模型 R^2 最大即自变量能解释因变量

程度最高, BP-NN 次之, MLR 最低; SVM 模型 RMSE 最小, BP-NN 次之, MLR 最大; SVM 模型的 Nash-Sutcliffe 效率系数 E 大于其他两种模型且都大于 0 小于 1, 表明三种模型的预测结果都接近于实测值的均值, 其中 SVM 模型的效率系数为 0.57 拟合优度最高。

三种模型预测 CEC 较大值与较小值时有所欠缺(见图 3), 这可能是由于模型中的训练集土壤类型不同、数量不足导致, 当 CEC 处于 15~25 $\text{cmol}\cdot\text{kg}^{-1}$ 范围, SVM 能较好地预测 CEC, BP-NN 次之, MLR 最差。综合考虑以上四个评价参数可知, SVM、BP-NN 两种模型预测精度优于 MLR, 其中 SVM、BP-NN 的四个评价参数都较为接近, 表明 SVM 模型的预测精度略优于 BP-NN。

Table 3. Accuracy comparison of models

表 3. 模型精度比较

模型	MAE	R ²	E	RMSE
MLR	4.70	0.51	0.49	5.91
BP-NN	4.16	0.56	0.53	5.67
SVM	4.45	0.58	0.57	5.41

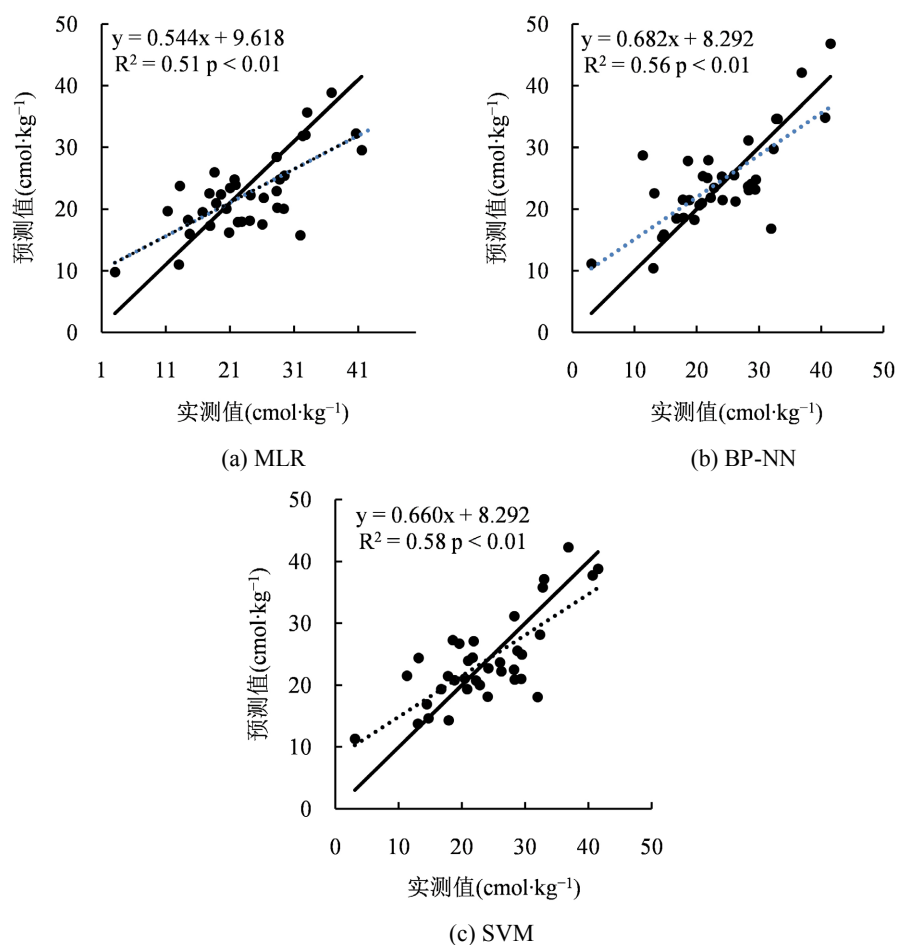


Figure 3. Scatter plots of measured and predicted values by models

图 3. 不同模型实测值与预测值散点图

3.5. 函数敏感性分析

通过执行敏感性分析得到输入变量的相对重要程度。两种模型中输入变量 pH、砂粒、黏粒、有机质对应的敏感性分别为 15%、12%、34%、75%和 24%、17%、26%、87% (见图 4), 敏感性排序从大到小依次为有机质、黏粒、pH、砂粒。有机质在两个模型中都为最重要的变量, 黏粒在 BP-NN 中的重要程度大于 SVM。

4. 讨论

在相关性分析中, CEC 与黏粒、有机质存在明显的正相关关系, 其中有机质与 CEC 之间相关系数最大, 表明 CEC 与有机质关系最为明显, 其次为黏粒含量。这和前人的研究结果并不完全一致, 例如 Liao [3]和 Tang [10]研究表明黏粒含量在预测 CEC 中起最主要的作用, 有机质次之, 而有些学者[18] [19]研究表明 CEC 主要受有机碳影响。这可能是因为在研究区域内有机质在研究区内变化明显, 其变异系数远大于黏粒含量, 且黏粒含量变程较小大都在 10%以下。虽黏粒、有机质在不同研究中所起的主次要作用并不相一致, 但都表明有机质和黏粒对 CEC 有明显的直接作用[20]。

SVM 模型预测精度高于 BP-NN、MLR 模型, 表明 CEC 与土壤理化性质存在非线性关系。这与国内外学者的研究不尽一致。Were [21]和 Liao [3]发现 SVM 模型预测精度优于神经网络和多元回归。而 Jafarzadeh 等[22]比较支持向量机和神经网络两种模型的精度, 研究表明神经网络模型精度要优于支持向量机。这可能是由于样本数量不同, 模型输入变量有所差异导致。此外, 与 BP-NN 相比, SVM 有效地实现对基于小样本的高维非线性系统精确拟合, 同时也避免了 BP-NN 存在的陷入局部最优解的问题, 并且提高了泛化能力。由于土壤的异质性, 想要在大尺度范围内测定土壤 CEC 是很困难的。因此, 对于大尺度上数字土壤图的优化、作物模型等工作中, SVM 可以提供比 MLR 和 BP-NN 更合理的 CEC 参数[3]。

从 4 个输入参数对输出变量 CEC 影响敏感性的统计分析角度看, 影响较大的参数是有机质、黏粒, 表明在预测 CEC 过程中模型的输入变量有机质、黏粒是最重要的两个变量。pH 虽与 CEC 的相关性小于砂粒与 CEC 的相关性, 但 pH 的敏感性要大于砂粒, 这可能是因为砂粒对 CEC 的影响很小。

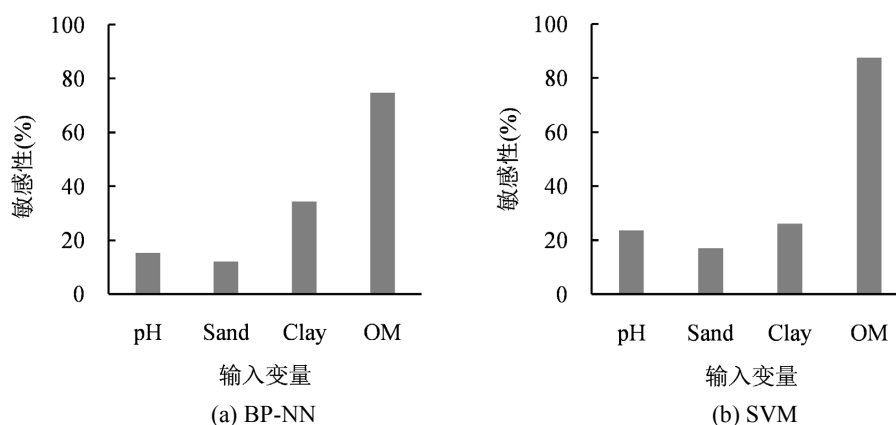


Figure 4. Sensitivity analysis of output variable (CEC) to the input variable of the models
图 4. 输出变量(CEC)对模型输入变量的敏感性分析

5. 结论与展望

5.1. 结论

研究利用 BP-NN、SVM、MLR 三种方法选择 pH、砂粒、粘粒、有机质含量为输入变量来预测旱作

区土壤阳离子交换量。结果表明, 在构建预测 CEC 的 PTFs 时有机质和黏粒是最重要的, 应优先考虑; SVM 和 BP-NN 两种方法在旱作区内预测土壤 CEC 是可行的, 但是 SVM、BP-NN 在预测 CEC 较大值与较小值上不准确; 利用 SVM 建立 PTFs 的精度略优于 BP-NN, 两者预测精度均优于 MLR。基于支持向量机的土壤转换函数在旱作区内能够较好的预测阳离子交换量。

5.2. 展望

在本研究中, 砂粒、黏粒、pH、有机质作为自变量变量, 并没有考虑土壤黏土矿物类型、碳酸钙等含量, 这些都对土壤 CEC 产生影响, 在以后的研究中可进一步考虑它们。此外, 有学者[23] [24]将土壤粒径的分形维数作为输入变量预测土壤 CEC, 以后的研究可将分形维数引入模型。

基金项目

煤科总院技术研发项目(2021-JSYF-007); 中煤科工集团创新创业重点项目(2019-ZD004); 国家重点研发计划项目(2016YFD0300801)。

参考文献

- [1] 孙忠祥, 李勇, 赵云泽, 等. 旱作区土壤有机碳密度空间分布特征与其驱动力分析[J]. 农业机械学报, 2019, 50(1): 255-262.
- [2] Bouma, J. (1989) Using Soil Survey Data for Quantitative Land Evaluation. *Advance Soil Science*, **9**, 177-213. https://doi.org/10.1007/978-1-4612-3532-3_4
- [3] Liao, K., Xu, S., Wu, J., et al. (2014) Using Support Vector Machines to Predict Cation Exchange Capacity of Different Soil Horizons in Qingdao City, China. *Journal of Plant Nutrition and Soil Science*, **177**, 775-782. <https://doi.org/10.1002/jpln.201300176>
- [4] Ghorbani, H., Kashi, H., Hafezi, N., et al. (2015) Estimation of Soil Cation Exchange Capacity Using Multiple Regression, Artificial Neural Networks, and Adaptive Neuro-Fuzzy Inference System Models in Golestan Province, Iran. *Communications in Soil Science and Plant Analysis*, **46**, 763-780. <https://doi.org/10.1080/00103624.2015.1006367>
- [5] Krogh, L., Breuning-Madsen, H. and Greve, M.H. (2000) Cation-Exchange Capacity Pedotransfer Functions for Danish Soils. *Acta Agriculturae Scandinavica, Section B-Soil & Plant Science, Section B-Plant Soil Science*, **50**, 1-12. <https://doi.org/10.1080/090647100750014358>
- [6] Kusuma, H.S., Amenaghawon, A.N., Darmokoesoemo, H., et al. (2021) Evaluation of Extract of Ipomoea Batatas Leaves as a Green Coagulant-Flocculant for Turbid Water Treatment: Parametric Modelling and Optimization Using Response Surface Methodology and Artificial Neural Networks. *Environmental Technology & Innovation*, **24**, e102005. <https://doi.org/10.1016/j.eti.2021.102005>
- [7] 王志强, 李航宇, 刘树阳, 等. 基于 BP 神经网络的咸水层 CO₂ 封存注入能力预测方法[J]. 中南大学学报(自然科学版), 2022, 53(12): 4678-4686.
- [8] Wang, W., Men, C. and Lu, W. (2008) Online Prediction Model Based on Support Vector Machine. *Neurocomputing*, **71**, 550-558. <https://doi.org/10.1016/j.neucom.2007.07.020>
- [9] Khaledian, Y., Brevik, E.C., Pereira, P., et al. (2017) Modeling Soil Cation Exchange Capacity in Multiple Countries. *Catena*, **158**, 194-200. <https://doi.org/10.1016/j.catena.2017.07.002>
- [10] Tang, L., Zeng, G., Nourbakhsh, F., et al. (2009) Artificial Neural Network Approach for Predicting Cation Exchange Capacity in Soil Based on Physico-Chemical Properties. *Environmental Engineering Science*, **26**, 137-146. <https://doi.org/10.1089/ees.2007.0238>
- [11] Shiri, J., Keshavarzi, A., Kisi, O., et al. (2017) Modeling Soil Cation Exchange Capacity Using Soil Parameters: Assessing The Heuristic Models. *Computers and Electronics in Agriculture*, **135**, 242-251. <https://doi.org/10.1016/j.compag.2017.02.016>
- [12] 陈雯柏. 神经网络原理与实践[M]. 西安: 西安电子科技大学出版社, 2016.
- [13] 王小川, 史峰, 郁磊, 等. Mtalab 神经网络 43 个案例[M]. 北京: 北京航空航天大学出版社, 2013.
- [14] 耿睿, 崔德光, 徐冰. 应用支持向量机的空中交通流量组合预测模型[J]. 清华大学学报自然(科学版), 2008, 48(7): 1205-1208.

-
- [15] Nielsen, D.R. and Bouma, J. (1995) Soil Spatial Variability. Pudoc Scientific Publishers, Wageningen, 209-234.
- [16] 李海滨, 王政权, 王庆成. 空间异质性定量研究理论与方法[J]. 应用生态学报, 1998, 9(6): 651-657.
- [17] 张磊. 中国石油安全体系评价研究——基于粗糙集及支持向量机方法[J]. 中国软科学, 2022(11): 13-19.
- [18] Mcbratney, A.B., Minasny, B., Cattle, S.R., *et al.* (2002) From Pedotransfer Functions to Soil Inference Systems. *Geoderma*, **109**, 41-73. [https://doi.org/10.1016/S0016-7061\(02\)00139-8](https://doi.org/10.1016/S0016-7061(02)00139-8)
- [19] Khaledian, Y., Kiani, F., Ebrahimi, S., *et al.* (2016) Assessment and Monitoring of Soil Degradation during Land Use Change Using Multivariate Analysis. *Land Degradation & Development*, **28**, 128-141. <https://doi.org/10.1002/ldr.2541>
- [20] Keshavarzi, A., Sarmadian, F., Shiri, J., *et al.* (2017) Application of ANFIS-Based Subtractive Clustering Algorithm in Soil Cation Exchange Capacity Estimation Using Soil and Remotely Sensed Data. *Measurement*, **95**, 173-180. <https://doi.org/10.1016/j.measurement.2016.10.010>
- [21] Werea, K., Bui, D.T., Dick, B., *et al.* (2015) A Comparative Assessment of Support Vector Regression, Artificial Neural Networks, and Random Forests for Predicting and Mapping Soil Organic Carbon Stocks Across an Afromontane Landscape. *Ecological Indicators*, **52**, 394-403. <https://doi.org/10.1016/j.ecolind.2014.12.028>
- [22] Jafarzadeh, A.A., Pal, M., Servati, M., *et al.* (2016) Comparative Analysis of Support Vector Machine and Artificial Neural Network Models for Soil Cation Exchange Capacity Prediction. *International Journal of Environmental Science and Technology*, **13**, 87-96. <https://doi.org/10.1007/s13762-015-0856-4>
- [23] Hossein, B., Naser, D. and Mohsen, J. (2014) Prediction of CEC Using Fractal Parameters by Artificial Neural Networks. *International Agrophysics*, **28**, 143-152. <https://doi.org/10.2478/intag-2014-0002>
- [24] Khodaverdiloo, H., Momtaz, H. and Liao, K. (2018) Performance of Soil Cation Exchange Capacity Pedotransfer Function as Affected by the Inputs and Database Size. *Clean: Soil Air Water*, **46**, e1700670. <https://doi.org/10.1002/clen.201700670>