宁夏中宁枸杞林地粉壤土阳离子交换量空间分布预测方法对比

包维斌[1,2] 白一茹[1,2] 杨 帆[1,2] 钟艳霞[1,2] 夏子书[1,2] 王幼奇[1,2*]

([1]宁夏大学资源环境学院, 银川 750021; [2]旱区特色资源与环境治理教育部国际合作联合实验室, 银川 750021)

摘 要 土壤阳离子交换量(CEC)是土壤施肥、改良的主要依据和土壤质量的评价指标,研究土壤 CEC 的空间分布及模型预测可为土壤养分监测、管理及精准农业实施提供科学依据。本研究以中宁枸杞林地粉壤土为对象,在自相关、交互相关等分析基础上,采用协同克里格(CoKriging)、普通最小二乘法(OLS)、地理加权回归(GWR)和随机森林(RF)模型对土壤 CEC 进行回归分析,比较了制图效果及模型预测精度。结果表明:中宁枸杞林地粉壤土 CEC 平均值为 13.12 cmol·kg^{-1},属中等肥力;土壤 CEC 的空间分布具有自相关性,并与土壤 pH、有机质、黏粒和电导率在不同滞后距离上存在不同的空间相互关系;RF 模型预测图避免了 CoKriging、OLS 和 GWR 模型预测图中土壤 CEC 图斑边界两侧破碎程度大、突变明显的缺陷,使土壤 CEC 在空间变化上表现为自然、平缓的过渡;RF 模型 RMSE 值分别比 CoKriging、OLS 和 GWR 模型减少 33.82%、20.55%和 19.81%,R^2分别提高 8.84%、51.92%和 7.69%。RF 模型考虑了样点空间位置,明显提高了插值精度且制图效果更加平缓。

关键词 阳离子交换量; 协同克里格; 普通最小二乘法; 地理加权回归; 随机森林

Predicting spatial distribution of cation exchange capacity in silt loam soil under *Lycium barbarum* of Zhongning, Ningxia.

BAO Wei-bin[1,2], BAI Yi-ru[1,2], YANG Fan[1,2], ZHONG Yan-xia [1,2], XIA Zi-shu[1,2], WANG You-qi[1,2*] ([1]*College of Resources and Environment, Ningxia University, Yinchuan* 750021, *China*; [2]*Arid Area Characteristic Resources and Environmental Governance Department of Education International Cooperation Joint Laboratory, Yinchuan* 750021, *China*).

Abstract: Soil cation exchange capacity (CEC) is the basis for soil fertilization and improvement, and thus an evaluation index of soil quality. Studies on the spatial heterogeneity of soil CEC can provide scientific basis for soil nutrient monitoring, management, and precision agriculture implementation. Spatial distribution of CEC in silt loam soil of *Lycium barbarum* in Zhongning was investigated based on the analysis of autocorrelation and interaction correlation. The CoCriging, ordinary least squares (OLS), geographically weighted regression Kriging (GWR) and Random Forest (RF) models were applied to perform regression analysis. After that, spatial interpolation mapping effect and the accuracy of the models were compared. The results showed that the average silt loam soil CEC in *L. barbarum* was 13.12 cmol·kg^{-1}, belonging to moderately fertility. The distribution of soil CEC had spatial autocorrelation. There were different spatial relationships among CEC, pH, soil organic matter, clay and electrical conductivity in different lag distances. The prediction map of RF model avoided the large fragmentation and obvious mutations which occurred on both sides of the CEC map boundary in the prediction maps of CoKriging, OLS and GWR models. Consequently, soil CEC appeared as a natural and gentle transition in the spatial variation map. The RMSE of the RF model was 33.82%, 20.55% and 19.81% low-

国家自然科学基金项目(41867003 和 41761049)、宁夏高等学校科研项目(NGY2017015)、宁夏回族自治区科技重大专项(2018BFG02016)、宁夏自然科学基金项目(2018AAC03027)、宁夏青年科技人才托举工程项目和宁夏环境保护科学技术研究项目(2018-07)资助。
收稿日期: 2018-09-02 接受日期: 2020-01-08
*通讯作者 E-mail: wyq0563@163.com

er than that of CoKriging, OLS and GWR models, and R^2 was increased by 8.84%, 51.92% and 7.69%, respectively. The RF model considered the location of sample space, and obviously improved the accuracy of interpolation, making the mapping more smooth.

Key words: cationic exchange capacity; CoKriging; ordinary least square; geographically weighted regression; Random Forest.

土壤属性由土壤质地、阳离子交换量(CEC)、有机质(SOM)等物理、化学属性构成,是土壤肥力形成的基础和农业生产可持续发展的必要条件(李锦芬等,2017;田圣贤等,2018)。土壤CEC是表征土壤养分及其缓冲性能的重要属性(谢恩泽等,2018),CEC变化直接决定着作物的品质和产量(廖凯华等,2010)。传统的粗放式施肥、灌溉等田间管理措施一方面难以保证作物品质和产量,同时还会导致农业面源污染问题(李锦芬等,2017)。因此,准确和便捷地预测土壤属性空间分布特征,是实施精准施肥和分区管理的前提,对提高作物品质产量、防止农业面源污染有积极作用(田圣贤等,2018)。

由于地理信息技术的迅速发展,越来越多的专家学者开始将地统计、GIS等空间分析方法应用到土壤属性研究中(张继舟等,2015;马坤等,2016;董悦等,2018)。其中,土壤属性之间的空间自相关性研究多集中于普通克里格(OK)和回归克里格(RK)及普通最小二乘法(OLS)等空间分析方法(张继舟等,2015;李锦芬等,2017;董悦等,2018)。虽然上述方法对因变量与解释变量的关系进行了综合考虑,但忽视了空间位置对相关参数产生的影响。易小波等(2017)运用状态空间方程对土壤养分、容重、水分等进行了空间模拟,结果发现,状态空间方程对土壤属性的模拟效果要明显优于OK、RK和OLS等传统方法。虽然其考虑了各变量的空间位置以及不同位置上各变量与其自身和它变量之间的空间依赖关系,但模拟结果仅仅是一维线性的结果,不能呈现二维的土壤属性空间分布规律。地理加权回归模型(GWR)作为一种有效处理回归分析中空间依存和依赖性的建模技术,将数据空间位置嵌入回归参数,给每个变量分配了不同的权重,但容易受带宽和过度拟合的影响(孙钰森等,2019)。而随着计算机和人工智能技术的快速发展,机器学习模型可以通过大量样本数据的反复训练使其目标最优化,具有极强的非线性拟合和约束能力。随机森林(RF)模型作为机器学习模型的一种,被广泛应用于土壤空间属性模拟,RF模型可以有效避免传统方法中存在的过度拟合和对多元线性关系不敏感的缺陷(王卓然等,2016;卢宏亮等,2019),同时能适应土壤属性空间分布上的复杂性和不确定性(Dharumarajan *et al.*, 2017;齐雁冰等,2017)。可以看出,关于土壤属性方面的空间预测研究较多,但针对宁夏枸杞(*Lycium barbarum*)林地粉壤土研究较少,特别是利用协同克里格(CoKriging)、OLS、GWR和RF四种方法对枸杞林地粉壤土CEC的空间分布特征进行预测并分析各种方法的优劣较为缺乏。

中宁县地处西北干旱区被国务院命名的“中国枸杞之乡”,枸杞是该县重要的生态经济型树种,不仅可以防风固沙、减少水土流失,而且有利于增加农民收入、摆脱贫困,是当地政府精准扶贫的重要产业之一。近年来因人类不合理施加有机肥和管理措施不当导致土壤质量变差,出现了枸杞开花结果少、产量低及品质下降等问题(王幼奇等,2017)。因此有必要通过预测中宁枸杞林地粉壤土养分属性为准确地掌握土壤肥力状况,制定合理的田间管理方式提供借鉴。目前已有学者进行了宁夏枸杞地土壤有机碳组分、电导率(EC)的空间异质性研究(李云翔等,2016;王幼奇等,2017),但是关于土壤CEC空间异质性特征研究较少,特别是利用模型预测其空间分布状况的研究较为缺乏。本研究以中宁枸杞林地粉壤土作为研究对象,在分析土壤CEC空间异质性及与相关控制因子间相关性的基础上,采用CoKriging、OLS、GWR和RF模型分别对土壤CEC进行回归分析、制图效果及模型预测精度比较,解释土壤CEC与相关控制因子间空间异质性,同时获取土壤CEC的空间分布预测图,为枸杞地施肥及土壤资源质量评价等方面提供借鉴和参考。

1 研究地区与研究方法

1.1 研究区概况

本研究以宁夏回族自治区中宁县万亩枸杞观光示范园(105°37′46″E—105°38′27″E,37°29′51″N—37°30′09″N)为研究区。海拔1258~1261 m,夏季酷热、冬季寒冷、气候干燥,为典型的温带大陆性气候。

年均温 9.5 ℃,年平均降水 202 mm。土壤质地为沙壤土,种植作物以枸杞为主。

1.2 样品采集与分析

为选取立地条件相近的研究区,排除地形、海拔、植被类型等因素的影响,在"万亩枸杞观光示范园"中以地形平坦、植被单一的枸杞林地为研究对象,按照 100 m×100 m 网格方式设计采样点(图 1),去除表层枯落物后利用土铲采集 0~20 cm 土样,采样点共 119 个,同时用手持 GPS 定位获取每个样点的经纬度坐标信息。土样经自然风干、剔除杂物、研磨后过筛。土壤 SOM 用重铬酸钾外加热法测定,CEC 用乙酸铵法测定,EC 采用电导率仪(EC330)测定(土水比 1∶5),pH 值采用 pH 计(上海雷磁)测定(土水比 1∶5)。CLAY 采用激光粒度仪(MS2000,Malvern Instruments)测定,按美国制标准分级。

1.3 地统计分析

半方差函数,也称为半变异函数,是地统计学中用来描述土壤空间变异性的最重要的函数,其公式表达式为:

$$r(h)=\frac{1}{2N(h)}\sum_{i=1}^{N(h)}[Z(x)-Z(x+h)]^2 \quad (1)$$

式中,$r(h)$为半方差函数,h 为采样点间距,$N(h)$为间距 h 的采样点对数,$Z(x)$为采样点 x 处的实测值,$Z(x+h)$为采样点 $x+h$ 处的实测值

1.4 四种预测模型介绍

1.4.1 协同克里格(CoKriging) CoKriging 是普通克里格的扩展,将单个区域化变量扩展为多个区域化变量,区域化变量的最优估值从单一属性扩展到多个属性的协同区域属性,其主要将主变量的自相关性和辅助变量的交互相关性结合起来,并用于无偏最优估值中,提高对主变量的插值精度,揭示主变量的空间分布规律。计算公式如下:

$$Z^*(B)=\sum_{i=1}^{N}\lambda_{mi}Z_m(X_i)+\sum_{j=1}^{P}\lambda_{nj}X_n(X_j) \quad (2)$$

式中:$Z^*(B)$为待测样点 B 处的预测值;$Z_m(X_i)$和 $Z_n(X_j)$为主变量 Z_m和辅助变量 Z_n的实测值;λ_m 和 λ_n 为 Z_m和 Z_n样点实测值权重,$\sum\lambda_{mi}=1$,$\sum\lambda_{nj}=0$,N 和 P 为 Z_m和 Z_n的样点数目。

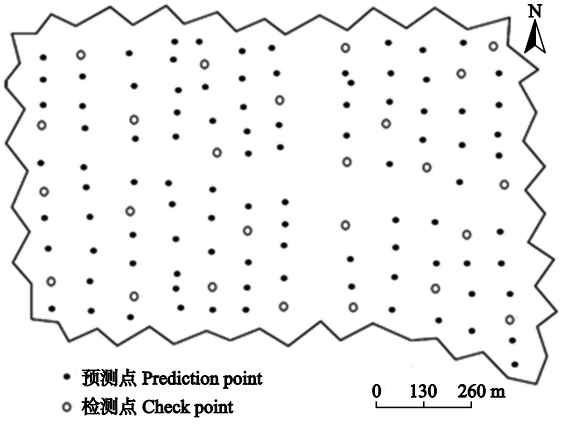


图 1 研究区样点分布示意图

Fig.1 Location of the sampling points

1.4.2 普通最小二乘法(OLS) OLS 是对解释变量与被解释变量相互依存关系进行分析,使所有观察值的残差平方和达到最小。计算公式如下:

$$y_i=\beta_0+\sum^{k}\beta_kX_{ik}+\varepsilon_i \quad (3)$$

式中:y_i 为第 i 点的因变量值;β_0 为模型截距;X_{ik}为第 k 个自变量在第 i 点的值;k 为自变量记数;β_k 为第 k 个自变量的回归系数;ε_1 为残差。

1.4.3 地理加权回归(GWR) 基于地理学第一定律的 GWR 模型扩展了传统 OLS 模型的框架,将数据的空间位置属性嵌入到回归模型参数中,使参数可以进行局部估计,其计算公式如下:

$$y(\mu,v)=\beta_0(\mu,v)+\sum_{i=1}^{P}\beta_1(\mu,v)x_1+\varepsilon(\mu,v) \quad (4)$$

式中:(μ,v)为研究区的空间位置;$\beta_0(\mu,v)$为位置(μ,v)处的模型截距;$\beta_1(\mu,v)$为位置(μ,v)处的自变量 i 的局部回归系数,$\varepsilon(\mu,v)$为位置位置(μ,v)处的误差项。

其中 $B_i(\mu,v)$参数通过加权最小二乘法实现,用矩阵形式表示为:

$$\beta_1(\mu,v)=[X^T-W_i(\mu,v)X]^{-1}X^TW_i(\mu,v)Y \quad (5)$$

式中:$W_i(\mu,v)$为 $m\times m$ 的空间权重矩阵,X 为 $m\times(n+1)$的自变量矩阵。本文采用高斯函数估算 GWR 模型中的参数,计算公式如下:

$$W_{ij}=[1-(d_{ij}/h)^2]^2 \qquad d_{ij}<h$$
$$W_{ij}=0 \qquad d_{ij}\geq h \quad (6)$$

式中:W_{ij}为实测点 j 估计预测 i 点时的权重,d_{ij}为 j 和 i 点间的欧氏距离,h 为带宽,h 由最优 AIC 信息准则值确定,以上 GWR 回归过程均在 GWR 4.0 软件中完成。

1.4.4 随机森林(RF) RF 模型由多棵树随机组成一个森林的机器学习算法,每棵决策树之间不存在联系(卢宏亮等,2019)。基本原理如下:通过

R3.3.2软件中的 Random Forest 软件包进行 RF 模拟,采用 Bootstrap 抽样法构建回归树和组成袋外数据,即从 M 个原始样点中有放回的随机选获得 m 个样点构建回归树,而($M-m$)个样点组成袋外数据,用于检验模型性能;设定决策树(ntree)和节点分裂(mtry)参数,其中 mtry 值按照袋外预测误差最小原则确定,生成多棵树组成的随机森林并以决策树预测值的平均值作为最终结果。本文在实际建模中,经过尝试发现将 ntree 和 mtry 分别设定为 1000 和 4 时,不仅可以产生较小的袋外误差,而且提高了模型的预测精度(王卓然等,2016)。

为验证模型的预测精度,随机选择 94 个样点作为建模样点进行空间插值,25 个样点作为验证样点用于分析插值精度,模型评价采用决定系数(R^2)、均方根误差(RMSE)和(RE)等指标进行评价。

1.5 数据处理

利用 SPSS 19.0 软件对样本数据进行描述性统计、自相关和交互相关分析,在 GWR 4.0 中完成地理加权回归分析,使用 ArcGIS 10.2、Excel 2010 完成地统计学分析及出图。

2 结果与分析

2.1 土壤 CEC 及相关控制因子的经典统计学分析

由表 1 可见,土壤 CEC 均值为 13.12 cmol · kg^{-1},变化范围为 11.02~18.47 cmol · kg^{-1}。根据土壤 CEC>20 cmol · kg^{-1}为保肥能力强、10~20 cmol · kg^{-1}为中等保肥能力、<10 cmol · kg^{-1}为弱保肥能力可知,研究区枸杞土壤保肥性能中等。从 SOM 均值看,也达到全国第二次土壤普查 4 级标准,即土壤肥力属于中等水平(谢恩泽等,2018)。土壤 CEC、SOM、CLAY 的变异系数大于 0.1 小于 1 均属中等变异程度,pH 和 EC 的变异系数小于 0.1 属弱变异程度(李云翔等,2016),说明施肥、耕作等人为活动对其影响较大,导致在空间尺度上存在一定变异性。其中 pH 均值为 8.57,即研究区土壤呈碱性。CEC 和 pH 偏度为负值,呈现左偏的分布特征。其他土壤属性偏度为正值,呈现右偏。经 K-S 非参数检验土壤属性服从正态分布(P>0.05),表明数据满足地统计学分析要求。

2.2 土壤属性的自相关和交互相关分析

土壤属性在空间分布上受多种控制因子影响,控制因子之间也经常存在相互影响,所以会产生多重共线性问题,进而导致统计结果精度降低(李锦芬等,2017)。因此,在进行空间模型预测之前,需确定土壤 CEC 与各控制因子的相关性。

相对于传统的 Pearson 相关而言,自相关、交互相关可以反映不同变量在不同滞后距离下的相关性(易小波等,2017)。研究表明,土壤 CEC 含量与 pH、SOM、CLAY、EC 的联系相当紧密(李锦芬等,2017),因此本文采用自相关和交互相关分析土壤 CEC、pH、SOM、CLAY、EC 之间的空间自相关性,分析结果如图 2 所示。土壤 CEC、pH、SOM、CLAY、EC 在空间上自相关关系不同,且不同的滞后距离下各控制变量的相关性也存在差异。其中,CLAY 在整个滞后距离内均不存在显著自相关性,土壤 CEC、SOM、pH、EC 分别在滞后距离 1、1~3 和 15~16、1~3、1~16 m 上自相关系数大于 95%置信水平下的临界值 0.18,说明具有显著自相关性。交互相关分析结果表明,不同土壤属性在不同滞后距离下具有不同的空间相关性,且不同方向上交互相关关系也存在差异(图 3)。土壤 SOM 和 pH 在滞后距离 0~1 m 时存在交互正相关,滞后距离较小。EC 仅在滞后距离为-4 m 时呈现交互正相关,滞后距离也较小。土壤 pH、SOM、CLAY 和 EC 均在其他滞后距离下不存在交互相关,这可能与研究区采样点距离较大存在一定联系。

2.3 半方差分析及 Moran I 分析

根据土壤CEC半方差函数图可以看出(图4),

表 1 土壤 CEC 及相关控制因子描述统计

Table 1 Descriptive statistics of soil CEC and its control factors

	极大值 Max	极小值 Min	均值 Mean	标准差 SD	变异系数(%) CV	偏度 Skewness	峰度 Kurtosis	K-S 检验 P 值 K-S test P value
CEC (cmol · kg^{-1})	18.47	11.02	13.12	1.59	12.12	-0.22	-0.01	0.07
SOM (g · kg^{-1})	23.13	10.99	17.39	2.57	14.77	0.18	-0.44	0.51
pH	8.94	7.45	8.57	0.24	2.80	-1.54	0.35	0.79
CLAY (%)	26.19	10.82	14.78	2.07	14.01	2.52	11.47	0.09
EC (μS · cm^{-1})	605	120	220	0.25	0.11	1.41	0.28	0.88

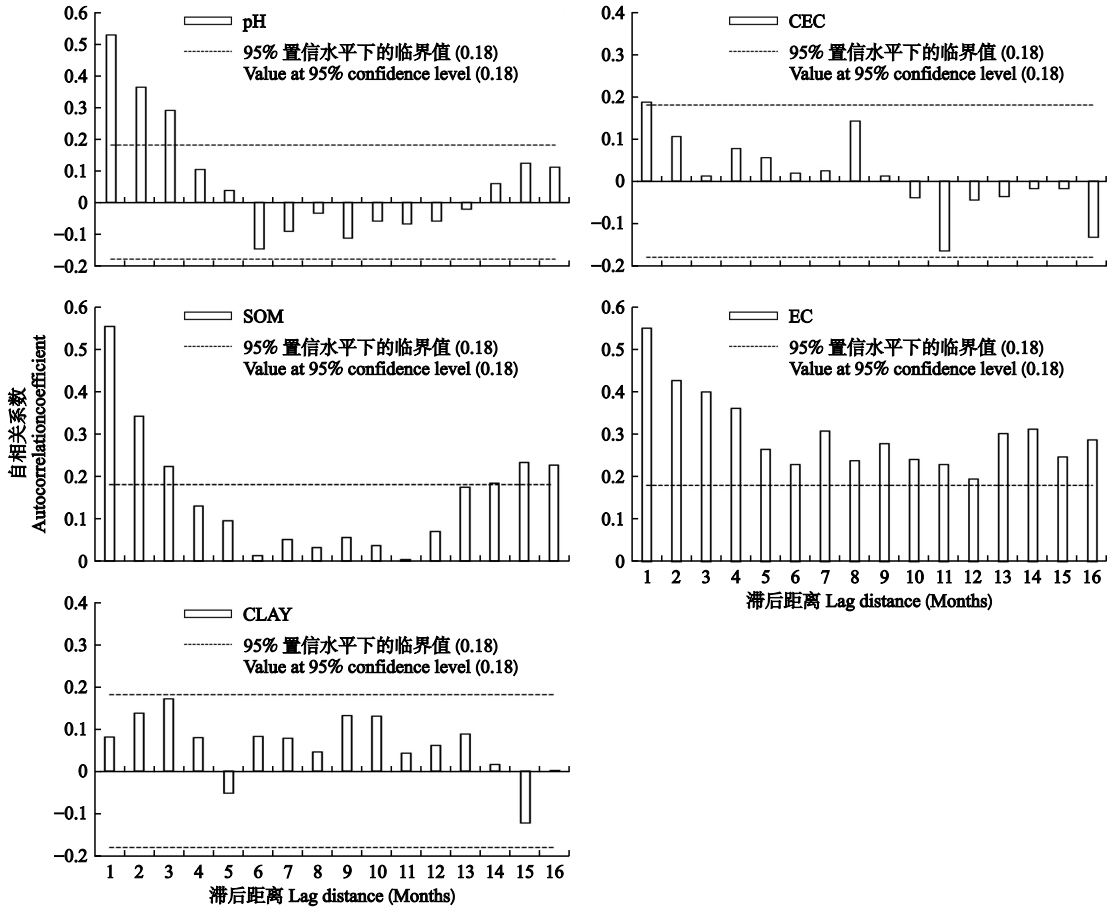


图 2 土壤 CEC 及其影响因素的空间自相关
Fig.2 Autocorrelation for soil CEC and its influencing factors

土壤 CEC 半变异函数的总趋势:间隔距离增加半方差值也增加,当间隔距离达到某一数值时所对应的半方差值趋于稳定,这说明间隔距离大于这个距离时,土壤 CEC 样点之间不再具有相关关系。由表 2 可以看出,土壤 CEC 含量的最佳拟合模型为指数模型。其中 $C_0/(C_0+C)$ 值反映变量空间变异来源中结构性和随机性因素所占的比重,其值越大说明施肥、耕作等随机性因素在主导地位,反之则地形、气候等结构性因素占主导地位(王幼奇等,2017)。本文中 $C_0/(C_0+C)$ 值为 89%,由 $C_0/(C_0+C)$ 值<25%时样点之间具有强烈的空间自相关性,25% < $C_0/(C_0+C)$ 值<75%时样点之间具有中等的空间自相关性,$C_0/(C_0+C)$ 值>75%时样点之间的空间自相关性较差可知(董悦等,2018),土壤 CEC 样点之间的空间自相关性较低,也证实了在自相关分析中土壤 CEC 仅在滞后距离 1 m 上自相关系数微大于 95%置信水平下的临界值 0.18。即中宁枸杞地土壤 CEC 的空间变异以施肥、耕作等随机性因素为主要变异来源。拟合精度较高的半方差函数模型是土壤属性空间变异结构分析的关键。在半方差函数拟合中指数模型决定系数(R^2)和残差(RSS)分别为 0.229和 0.077,模型的拟合效果较好。

空间自相关分析是判断某一地理现象的属性值在空间分布上是否具有相关性,特定属性值在某一

表 2 土壤 CEC 的半方差函数理论模型及相关参数
Table 2 Parameters and theoretical variogram models of soil ECE

	理论模型 Theoretical model	块金值 C_0	基台值 C_0+C	块金值/基台值 $C_0/(C_0+C)$ (%)	变程(m) Range	决定系数 R^2	残差 Residual
CEC	指数模型 Index model	0.257	2.417	0.894	144	0.229	0.077
	球状模型 Spherical model	0.073	2.411	0.970	154	0.213	0.078
	高斯模型 Gaussian model	0.276	2.410	0.885	128	0.213	0.078

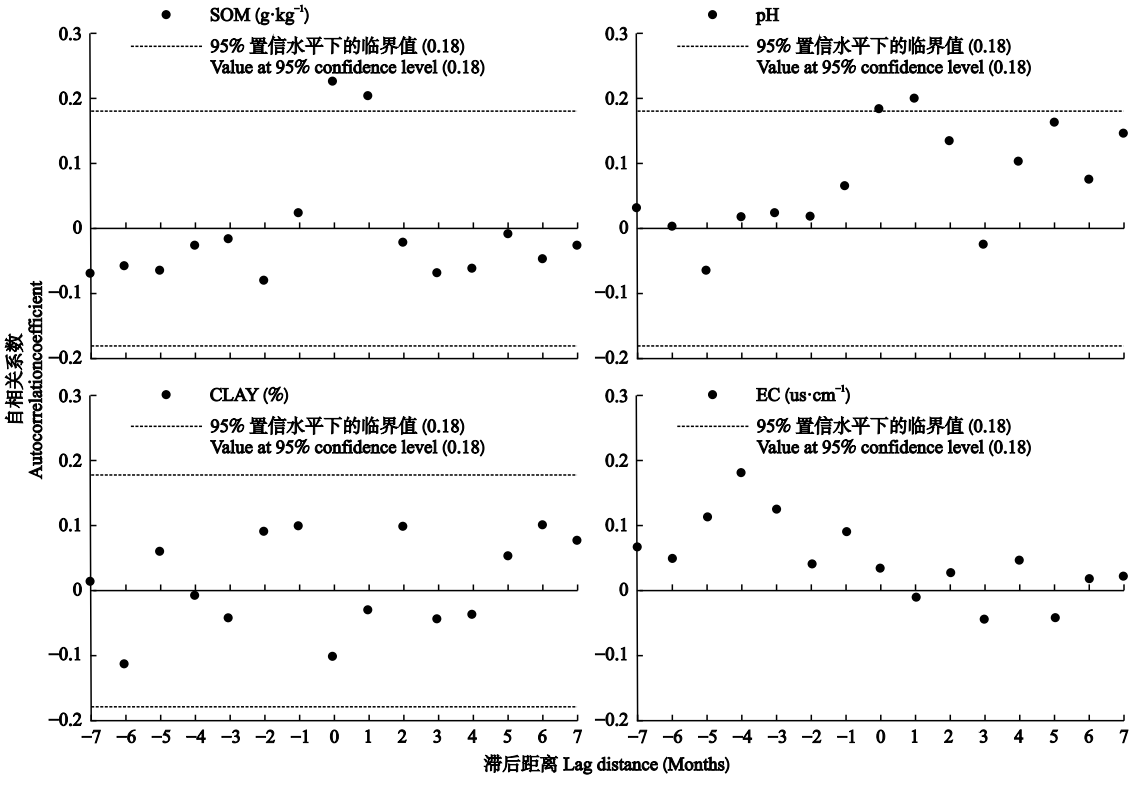


图 3 土壤 CEC 与其影响因素的空间交互相关
Fig.3 Cross-correlation between soil CEC and its influencing factors

空间位置上的变异是否由相邻空间位置的变异引起的空间统计方法(齐雁冰等,2017)。与传统的空间分析方法相比,空间自相关分析可以反映土壤 CEC 空间分布的相似性和依赖性,揭示其空间分布规律及其影响土壤 CEC 含量的随机变量。图 5 显示,土壤 CEC 的 Moran *I* 系数随间隔距离(步长)增大的变化趋势。土壤 CEC 在不同的间隔距离呈现出不同的相关性,间隔距离分别为119.74、329.06、530.15 m 时呈现出空间正相关,在 230.58、422.84、631.47 m 均为空间负相关。这说明土壤 CEC 更容易受到相邻区域耕作方式、施肥等人为措施的影响。

2.4 CoKriging、OLS 和 GWR 模型制图效果及精度比较

根据 EC、pH、CLAY 和 SOM 等土壤属性与 CEC 的相关性分析,对土壤 CEC 的空间分布进行 CoKriging 插值,得出土壤 CEC 的空间分布预测图(图 6)。从图 6 可以发现,土壤 CEC 在空间分布上表现出差异性,土壤CEC图斑边界两侧的破碎程度

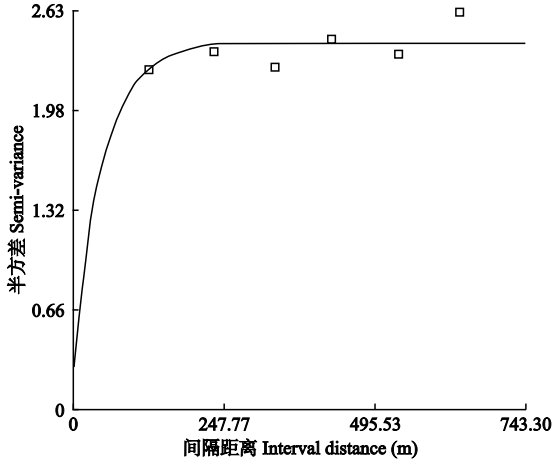


图 4 土壤 CEC 半方差函数图
Fig.4 The semi-variogram of soil CEC

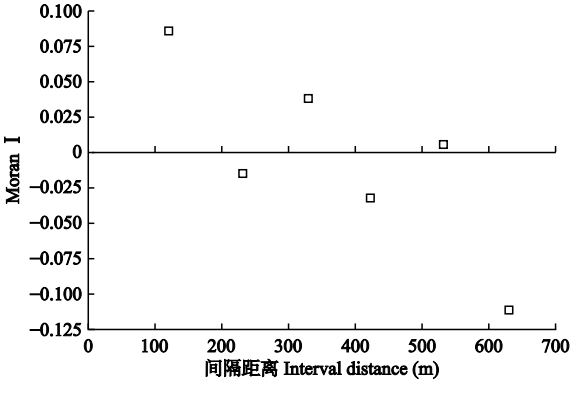


图 5 土壤 CEC Moran *I* 系数图
Fig.5 The Moran *I* of soil CEC

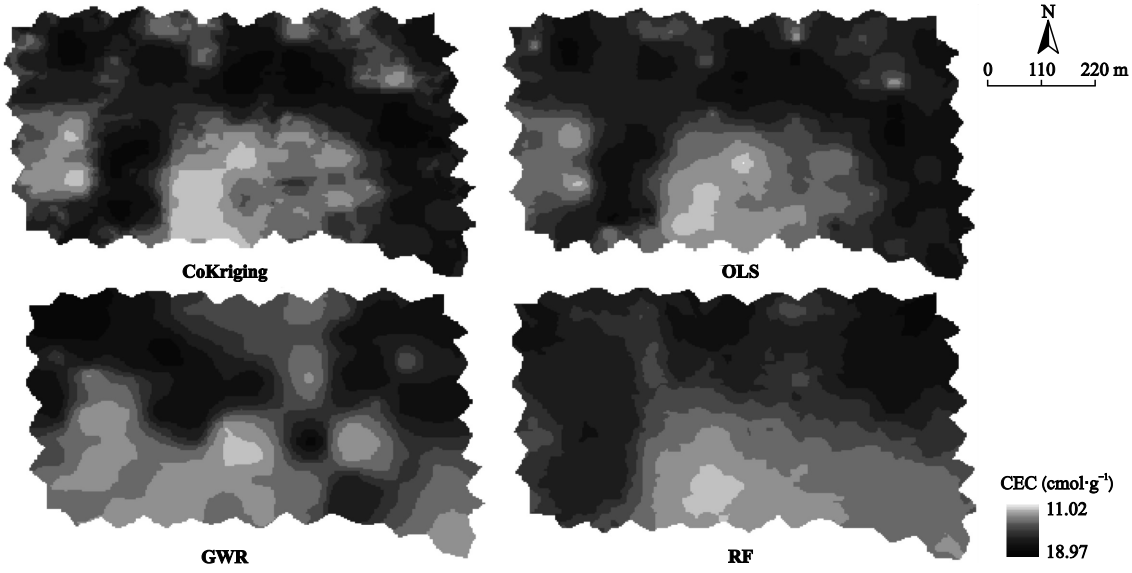


图 6 土壤 CEC 的 CoKriging、OLS、GWR、RF 插值图
Fig.6 Spatial distribution of soil CEC in CoKriging, OLS, GWR and RF

较大、突变明显,既 CoKriging 插值揭示土壤 CEC 空间变化的效果欠佳。从 OLS 模型预测图发现,土壤 CEC 在采样区域呈岛状和条带状相结合分布的特点。与 CoKriging 插值预测图相比其预测图减小了土壤 CEC 图斑边界两侧的破碎程度和突变,但减小程度有限,因而用 OLS 模型的插值预测图作为枸杞地改良依据时也存在一定的局限性。从 GWR 模型的土壤 CEC 预测图可以看出,土壤 CEC 在空间变化上表现为较自然、平缓的过渡,RF 模型的土壤 CEC 预测图 CEC 在空间变化比 GWR 模型的土壤 CEC 预测图更加自然、平缓的过渡,既 RF 模型的土壤 CEC 预测图有效避免了 CoKriging、OLS 和 GWR 模型的插值预测图中土壤 CEC 图斑边界两侧的破碎程度大、突变明显的缺陷。因此 RF 模型的土壤 CEC 预测图可以作为中宁枸杞地土壤施肥、改良的主要依据。

以土壤 pH、SOM、EC 和 CLAY 为控制因子,选择 CoKriging、OLS、GWR 和 RF 模型对土壤 CEC 进行回归分析。结果表明(表 3),CoKriging、OLS、GWR 和 RF 模型决定系数(R^2)分别为 0.11、0.25、0.48和 0.52,说明 RF 模型对土壤 CEC 的影响因子的解释力为 52%,相比 CoKriging、OLS 和 GWR 模型分别提高 78.84%、51.92%和 7.69%。不同预测模型中 RMSE 和 SE 值越小说明模型得到了显著改善(马冉等,2019)。本研究中 RF 模型的 RMSE 值分别比 CoKriging、OLS 和 GWR 模型减少 33.82%、20.55%、19.81%,说明 RF 模型相比 CoKriging、OLS 和 GWR 模型有了明显改善。同时 RF 模型也大大降低了回归的标准误差,故 RF 模型对土壤 CEC 的评估效果较佳且提高了预测精度。

表 3 CoKriging、OLS、GWR 及 RF 回归得到的相关参数
Table 3 Related parameters by using CoKriging, OLS, GWR and RF model respectively

模型 Models	决定系数 R^2	均方根误差 RMSE	标准误差 SE
CoKriging	0.11	3.243	1.55
OLS	0.25	3.014	1.52
GWR	0.48	3.013	1.51
RF	0.52	2.416	0.63

3 讨 论

根据土壤 CEC 含量及其相关属性的描述性统计可知,中宁枸杞林地粉壤土属中等肥力。李云翔等(2016)也发现,中宁枸杞粉壤土有机碳含量较高,且高于宁夏其他植被条件下土壤有机碳含量。说明枸杞林地有机肥施用强度较高(马冉等,2019)。在自相关和交互相关分析中,土壤 CEC 与 pH、SOM、CLAY、EC 具有不同的空间相关性,相对于 Pearson 相关,自相关和交互相关可以反映不同变量在不同滞后距离下的相关性。易小波等(2017)也发现,土壤属性的相关性随滞后距离的变化而不同。为有效反映样点空间结构特征,对土壤 CEC 进行半变异方差函数分析,发现研究区土壤

CEC 的 $C_0/(C_0+C)$ 值为 89.4%属较弱的空间自相关性,气候、地形、土壤等结构性因素在枸杞土壤 CEC 空间变异过程中不起决定性作用,而施肥、种植密度耕、作制度等随机性因素导致土壤 CEC 的空间自相关性减弱,并朝向均一化方向发展。这与李锦芬等(2017)研究江西进贤县土壤 CEC 空间结构所得结论一致。半方差及 Moran *I* 系数也说明粉壤土 CEC 在不同的空间距离下具有不同的相关性且相关性不强,也是由于随机性因素造成的。王幼奇等(2017)也发现,枸杞林地粉壤土电导率等相关属性主要受施肥、种植和耕作方式等随机因素影响。

依据建模样点和验证样点的插值分布,采用目前评价模型精确度的决定系数(R^2)、均方根误差(RMSE)和(RE)指标进行预测精度评价,发现 RF 模型 RMSE 值分别比 CoKriging、OLS 和 GWR 模型减少 33.82%、20.55% 和 19.81%,调节决定系数(R^2)分别为提高 8.84%、51.92% 和 7.69%。即 RF 模型相对于 CoKriging、OLS 和 GWR 方法在局部空间回归方面存在明显优势,能够准确反映局部地区土壤 CEC 实际空间分布状况、提高空间预测插值图可信度。一般而言,CoKriging 插值可以通过对误差进行估计来优化区域差异分析,但预测时忽略了环境要素的影响(徐占军等,2018),且预测精度受样点数的影响大,若样点数少,土壤属性空间自相关减弱,易出现纯块金效应,这导致预测结果不符合实际(赵宣等,2017;吴子豪等,2018)。OLS 模型是建立在平均值之上的回归方法,相比 CoKriging 插值,其具有回归分析时不受地理位置的限制且对数据要求不高的优点(谢恩泽等,2018),但是因变量和自变量之间需为线性相关,然而土壤 CEC 与辅助变量间并不是简单的线性关系,因此 OLS 模型分析土壤与辅助变量间复杂的非线性关系时难以达到理想效果,其适用性虽高于 CoKriging 但依然受分析要素之间线性关系的制约(姜赛平等,2018)。GWR 模型充分考虑了样点的空间位置及其相似性,以距离已知点越近的待测点权重越高为准则,但由于预测时受到带宽限制,待测点权重随预测点距离增大而减小,尤其在施肥、灌溉的农田中预测精度往往与实际差异较大,预测图也存在一定程度上的破碎,导致模型适用性也受到了一定限制(赵彦锋等,2018)。

RF 模型预测精度明显优于 CoKriging、OLS 和 GWR,原因在于 RF 模型采用集成算法可以处理非线性数据,能与各土壤因素参与大规模反复的训练,以及按照袋外预测误差最小原则确定节点分裂,反复设置生成树的数量,选择最优解有关(王卓然等,2016)。杨煜岑等(2018)发现,RF 模型比 CoKriging 模型更能精确地反映陕西周至县局部土壤有机质含量的空间部分状况;齐雁冰等(2017)、任丽等(2018)的结果也表明,RF 模型提高了插值精度并在捕捉空间异质性方面具有优越性。以上研究均与本研究结果相似。RF 模型有效克服了传统模型拟合过程中的不稳定性,较好地结合了 CoKriging、OLS 和 GWR 模型的优点,避免了从单一方面来进行预测,从而提高预测的精度及准确刻画了变量间的非线性关系(姜赛平等,2018)。因此,RF 模型也被广泛应用于土壤重金属源解析(刘斌等,2019)、土壤质地空间预测制图(Rad *et al.*,2018)和土壤盐度预测(王飞等,2018)等。可以看出,RF 模型在土壤属性空间预测中预测精度高、制图效果好、与实际情况较为一致。本研究只针对中宁粉壤土 CEC,后续研究还需比较、分析 RF 等不同预测方法,对不同质地土壤性质空间分布预测的适应性和优缺点,为区域土壤属性空间预测提供思路和方法参考。

4 结 论

土壤 CEC 平均值为 13.12 cmol · kg^{-1},达到全国第二次土壤普查 4 级标准属中等肥力,且 CEC、SOM、CLAY 均属中等变异程度。

地统计结果表明,中宁枸杞地粉壤土 CEC 块金系数为 89%,其与土壤 pH、SOM、CLAY 和 EC 在不同滞后距离上存在不同的空间交互相关关系并且各样点之间表现出空间自相关性较弱,主要受灌溉、施肥等人为因素的影响。

RF 模型有效揭示相关土壤属性对土壤 CEC 的空间非平稳影响,其 R^2 分别比 CoKriging、OLS 和 GWR 模型提高 8.84%、51.92% 和 7.69%,RMSE 值分别减少 33.82%、20.55% 和 19.81%,且预测图在空间变化上表现为自然、平缓过渡。

参考文献

董 悦, 张永清, 刘彩彩. 2018. 晋中市土壤养分空间分布与影响因子相关性研究. 生态学报, **38**(23): 1-8. [Dong Y, Zhang YQ, Liu CC. 2018. Study on soil nutrient spatial patterns and their driving factors in Jinzhong. *Acta Ecologica Sinica*, **38**(23): 1-8.]

姜赛平, 张怀志, 张认连, 等. 2018. 基于三种空间预测模型的海南岛土壤有机质空间分布研究. 土壤学报, **55**(4):

227-237. [Jiang SP, Zhang HZ, Zhang RL, *et al*. 2018. Research on spatial distribution of soil organic matter in Hainan Island based on three spatial prediction models. *Acta Pedologica Sinica*, **55**(4): 227-237.]

李锦芬, 瞿明凯, 黄 标, 等. 2017. 区域土壤 CEC 与相关控制因子的空间非平稳关系评估. 土壤学报, **54**(3): 637-645. [Li JF, Qu MK, Huang B, *et al*. 2017. Spatially non-stationary relationships between cation exchange capacity and related control factors. *Acta Pedologica Sinica*, **54**(3): 638-646.]

李云翔, 王少东, 柯 英, 等. 2016. 宁夏主要枸杞产区施肥现状与土壤养分特征. 干旱地区农业研究, **34**(2): 113-118. [Li YX, Wang SD, Ke Y, *et al*. 2016. Characteristics of soil nutrients and present situation of fertilization in the major wolfberry producing areas of Ningxia. *Agricultural Research in the Arid Areas*, **34**(2): 113-118.]

廖凯华, 徐绍辉, 程桂福, 等. 2010. 土壤 CEC 的影响因子及 Cokriging 空间插值分析——以青岛市大沽河流域为例. 土壤学报, **47**(1): 26-32. [Liao KH, Xu SH, Cheng GF, *et al*. 2010. Influencing factors and Cokriging spatial interpolation analysis of soil cation exchange capacity: A case study of Dagu river basin, Qingdao city. *Acta Pedologica Sinica*, **47**(1): 26-32.]

刘 斌, 郭 星, 朱宇恩. 2019. 基于随机森林模型的土壤重金属源解析——以晋中盆地为例. 干旱区资源与环境, **33**(1): 108-113. [Liu B, Guo X, Zhu YE. 2019. Analysis of soil heavy metal sources in Jinzhong basin based on random forest model. *Journal of Arid Land Resources and Environment*, **33**(1): 108-113.]

卢宏亮, 赵明松, 刘斌寅, 等. 2019. 基于随机森林模型的安徽省土壤属性空间分布预测. 土壤, **51**(3): 602-608. [Lu HL, Zhao MS, Liu BY, *et al*. 2019. Spatial prediction of soil properties based on random forest model in Anhui Province. *Soils*, **51**(3): 602-608.]

马 坤, 张 颖, 唐素贤, 等. 2016. 若尔盖高寒湿地土壤全氮空间分布特征. 生态学杂志, **35**(8): 1988-1995. [Ma K, Zhang Y, Tang SX, *et al*. 2016. Characteristics of spatial distribution of soil total nitrogen in Zoige alpine wetland. *Chinese Journal of Ecology*, **35**(8): 1988-1995.]

马 冉, 刘洪斌, 武 伟. 2019. 三峡库区草堂河流域土壤 pH 空间分布预测制图. 长江流域资源与环境, **28**(3): 691-699. [Ma R, Liu HB, Wu W. 2019. Spatial distribution prediction and mapping of soil pH of Caotang river basin in the Three Gorges Reservoir area. *Resources and Environment in the Yangtze Basin*, **28**(3): 691-699.]

齐雁冰, 王茵茵, 陈 洋, 等. 2017. 基于遥感与随机森林算法的陕西省土壤有机质空间预测. 自然资源学报, **32**(6): 1074-1086. [Qi YB, Wang YY, Chen Y, *et al*. 2017. Soil organic matter prediction based on remote sensing data and random forest model in Shaanxi province. *Journal of Natural Resources*, **32**: 1074-1086.]

任 丽, 杨联安, 王 辉, 等. 2018. 基于随机森林的苹果区土壤有机质空间预测. 干旱区资源与环境, **32**(8): 144-149. [Ren L, Yang LA, Wang H, *et al*. 2018. Spatial prediction of soil organic matter in apple region based on random forest. *Journal of Arid Land Resources and Environment*, **32**(8): 144-149.]

孙钰森, 王维芳, 李国春. 2019. 基于地理加权回归克里格模型的帽儿山地区森林碳储量空间分布. 应用生态学报, **30**(5): 1642-1650. [Song YL, Wang WF, Li GC, *et al*. 2019. Spatial distribution of forest carbon storage in Maoershan region, Northeast China based on geographically weighted regression kriging model. *Chinese Journal of Applied Ecology*, **30**(5): 1642-1650.]

田圣贤, 冯 盼, 杨 山, 等. 2018. 东北阔叶红松林腐殖质层土壤阳离子交换性能及其主要影响因素. 生态学杂志, **37**(9): 2549-2558. [Tian SX, Fang P, Yang S, *et al*. 2018. Soil cation exchange capacity and its main impact factors in the humic layer of broadleaved Korean pine forest in northeast China. *Chinese Journal of Ecology*, **37**(9): 2549-2558.]

王 飞, 杨胜天, 魏 阳, 等. 2018. 基于 RF 和 SGT 算法的子区优先建模对绿洲尺度土壤盐度预测精度的影响. 中国农业科学, **51**(24): 4659-4676. [Wang F, Yang ST, Wei Y, *et al*. 2018. Influence of sub-region priority modeling constructed by random forest and stochastic gradient treeboost on the accuracy of soil salinity. *Scientia Agricultura Sinica*, **51**(24): 4659-4676.]

王幼奇, 赵云鹏, 张 兴, 等. 2017. 引黄灌区枸杞地土壤电导率空间格局及条件模拟. 中国农村水利水电, **32**(6): 32-37. [Wang YQ, Zhao YP, Zhao X, *et al*. 2017. Spatial distribution and conditional simulation of soil electrical spatial distribution and conditional simulation of soil electrical. *China Rural Water and Hydropower*, **32**(6): 32-37.]

王卓然, 赵庚星, 高明秀, 等. 2016. 黄河三角洲垦利县夏季土壤水盐空间变异及土壤盐分微域特征. 生态学报, **36**(4): 1040-1049. [Wang ZR, Zhao GX, Gao MX, *et al*. 2016. Spatial variation of soil water and salt and microscopic variation of soil salinity in summer in typical area of the Yellow River Delta in Kenli County. *Acta Ecologica Sinica*, **36**(4): 1040-1049.]

吴子豪, 刘艳芳, 陈奕云, 等. 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 ZH, Liu YF, Chen YY, *et al*. 2018. Spatial interpolation model of soil organic carbon density considering land-use and spatial heterogeneity. *Chinese Journal of Applied Ecology*, **29**(1): 238-246.]

谢恩泽, 赵永存, 陆访仪, 等. 2018. 不同方法预测苏南农田土壤有机质空间分布对比研究. 土壤学报, **55**(5): 14-24. [Xie EZ, Zhao YC, Lu FY, *et al*. 2018. Comparison analysis of methods for prediction of spatial distribution of soil organic contents in farmlands south Jiangsu, China. *Acta Pedologica Sinica*, **55**(5): 14-24.]

徐占军, 张 媛, 张绍良, 等. 2018. 基于 GIS 与分区 Kriging 的采煤沉陷区土壤有机碳含量空间预测. 农业工程学报, **34**(10): 253-259. [Xu ZJ, Zhao Y, Zhao SL, *et al*. 2018. Spatial prediction of soil organic carbon content in coal mining subsidence area based on GIS and partition

Kriging. *Transactions of the Chinese Society of Agricultural Engineering*, **34**(10): 253-259.]

杨煜岑, 杨联安, 任 丽, 等. 2018. 基于随机森林的农耕区土壤有机质空间分布预测. 浙江农业学报, **30**(7): 1211-1217. [Yang YQ, Yang LA, Ren L, *et al*. 2018. Prediction for spatial distribution of soil organic matter based on random forest model in cultivated area. *Acta Agriculturae Zhejiangensis*, **30**(7): 1211-1217.]

易小波, 邵明安, 赵春雷, 等. 2017. 黄土高原南北样带不同土层土壤容重变异分析与模拟. 农业机械学报, **48**(4): 198-205.[Yi XB, Shao MA, Zhao CL, *et al*. 2017. Variation analysis and simulation of soil bulk density within different soil depths along south to north transect of loess plateau. *Transactions of the Chinese Society for Agricultural Machinery*, **48**(4): 198-205.]

张继舟, 吕 品, 王立民, 等. 2015. 大兴安岭森林土壤重金属含量空间变异与污染评价. 生态学杂志, **34**(3): 810-819. [Zhao JZ, Lu P, Wang LM, *et al*. 2015. Spatial variability of heavy metal contents and contamination assessment in forest soils of Daxing´an Mountains. *Chinese Journal of Ecology*, **34**(3): 810-819.]

赵 宣, 郝起礼, 孙婴婴. 2017. 典型毛乌素沙漠-黄土高原过渡带土壤盐渍化空间异质性及其影响因素. 应用生态学报, **28**(6): 1761-1768. [Zhao X, Hao QL, Song YY. 2017. Spatial heterogeneity of soil salinization and its influencing factors in the typical region of the Mu Us Desert-Loess Plateau transitional zone, Northwest China. *Chinese Journal of Applied Ecology*, **28**(6): 1761-1768.]

赵彦锋, 李豪杰, 陈 杰, 等. 2018. 基于土壤变异解释力的几种土壤制图方法的对比研究——以南阳市 1 m 土体土壤有机碳密度制图为例. 土壤学报, **55**(1): 43-53. [Zhao YF, Li HJ, Chen J, *et al*. 2018. Comparison between soil mapping approaches based on their ability explaining soil variability—A case of mapping soil organic carbon density of soil (1 m) in Nanyang district. *Acta Pedologica Sinica*, **55**(1): 43-53.]

Dharumarajan S, Hegde R, Singh SK. 2017. Spatial prediction of major soil properties using random forest techniques: A case study in semi-arid tropics of South India. *Geoderma Regional*, **10**: 154-162.

Rad MRP, Akbarimoghaddam A. 2018. Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *Catena*, **160**: 275-281.

作者简介 包维斌,男,1991 年生,硕士研究生,主要从事自然资源开发与持续利用研究。E-mail: baoweibin12@163.com

责任编辑 魏中青