

基于大尺度低密度样点的东北土壤 全氮空间插值方法比较

陈昌华, 陈锡云, 徐 英

(北京师范大学 地理学与遥感科学学院, 北京 100875)

摘 要: 基于全国第二次土壤普查东北地区土壤数据, 以 ArcGIS 和 GS⁺ 软件为支撑, 对比分析了反距离加权法(IDW)、径向基函数(RBF)、普通克里金(OK)和回归克里金(RK)4 种地统计空间插值方法在 7 个不同样本容量下土壤全氮含量(STNC)的空间插值效果。结果表明, 由普查数据得到的东北地区 STNC 在 0.08~21.48 g/kg 之间, 数据变异性较大; STNC 空间结构表现出中强度空间自相关性, 空间自相关范围大于同区域的小尺度采样研究; 样本容量<171 时, STNC 空间变异性发生变化, 空间结构特征和精度检验水平难以确信。4 种空间插值方法对 STNC 空间趋势表达均呈现从东北向西南方向递减规律, 空间趋势预测效果为: RK>OK>RBF>IDW。RK 方法通过线性回归分析添加了阳离子交换容量(CEC)、年均温(MAT)、土层厚度(*d*)和 pH 值等辅助信息, 比 IDW, RBF 和 OK 方法的插值精度分别提高了 19.40%, 18.50%和 16.15%; 在不同样本容量下 RK 方法的插值精度较为稳定且对无样点区 STNC 的空间趋势预测也体现出了更多细节信息, 因此对于大尺度低密度采样的土壤属性空间插值可重点考虑 RK 方法。

关键词: 空间插值; 土壤制图; 地统计; 土壤全氮; 东北地区

文献标识码: B

文章编号: 1000-288X(2014)06-0153-09

中图分类号: S153.6, S159.2

DOI:10.13961/j.cnki.stbctb.2014.06.036

Comparison of Spatial Interpolation Methods for Soil Total Nitrogen Content at Large Scale Using Low Density Soil Survey Data in Northeast China

CHEN Chang-hua, CHEN Xi-yun, XU Ying

(School of Geography, Beijing Normal University, Beijing 100875, China)

Abstract: Based on the relatively low density sampling data from China's second national soil survey in northeast China, we compared the performance of four spatial interpolation methods, inverse distance weighting(IDW), radial basis function(RBF), ordinary Kriging(OK) and regression Kriging(RK), under seven sample sizes in generating digital map of soil total nitrogen content at large scale based on the software ArcGIS and GS⁺. Results showed that soil total nitrogen content was in the range of 0.08~21.48 g/kg with great variability. The Nugget effect showed a medium strong spatial autocorrelation of soil total nitrogen content in the region, and the range of spatial autocorrelation was greater than research on smaller scale in the same region; Spatial variability of soil total nitrogen content changed when sample size was less than 171, in this case, spatial structure and accuracy test of interpolation were unbelievable. All of the four compared methods predicted the spatial pattern of soil total nitrogen content well with decreasing from northeast to southwest. The accuracy of interpolation changed in the order of RK>OK>RBF>IDW. With incorporated auxiliary variables of soil cation exchange capacity, soil depth, soil pH value and annual mean air temperature, the RK improved accuracy by 19.40%, 18.50% and 16.15% than IDW, RBF and OK, respectively. It also exhibited more details on soil total nitrogen content variation at the areas with sparse sample points. It suggested that the RK is a potential spatial interpolation method to improve the soil mapping accuracy at

收稿日期:2013-11-09

修回日期:2013-12-25

资助项目:国家自然科学基金项目“三江平原农业活动胁迫下的区域生态环境与安全调控研究”(40930740); 国家自然科学基金项目(41171360); 中央高校基本科研业务费专项(2009SD-5)

作者简介:陈昌华(1988—),女(汉族),山东省莒县人,硕士,研究方向为生物地球化学循环及生态安全评价。E-mail:chenchanghua.ok@163.com。

通信作者:陈锡云(1972—),男(汉族),甘肃省武威市人,博士,副教授,主要从事环境地理、流域水文生物地化过程与模拟的教学和科研工作。E-mail: chen_xiyun@bnu.edu.cn。

large area with low density sample sites.

Keywords: spatial interpolation method; soil mapping; geostatistics; soil total nitrogen; Northeast China

土壤属性的空间变异性研究对于实现土壤资源的科学管理与高效利用至关重要。土壤属性空间分布图是土壤属性空间变异最直观的表达形式,然而结合土壤剖面数据与土壤类型图制作土壤属性分布图的传统方法带有较强主观性,难以定量反映土壤属性的空间变异。空间统计是基于数据所代表的属性具有空间自相关特性,根据已知空间数据来估计未知空间数据的数学方法^[1]。20世纪70年代后期美国率先利用地理信息系统(GIS)的空间统计分析功能,开展了土壤普查、土壤属性数字制图及土壤变异性方面的研究,推动了利用地统计学方法研究土壤属性空间变异的快速发展,使得反距离加权、普通克里金等方法在该领域的应用趋于成熟。近年来,关于地统计插值方法比较及混合插值(hybrid interpolation)新模型^[2]的发展和验证等方面的研究表明,以目标变量的显著相关性环境因子作为辅助信息(auxiliary information)的混合插值对提高不同区域的土壤属性插值精度具有良好效果^[3]。常用的辅助信息有地形因子^[4]、气候因子^[5]及基于遥感卫星影像的土地利用类型^[6]、土壤类型等多分类变量^[7]。因辅助信息相对容易获得,在一定程度上降低了野外工作量、提高了空间插值的预测能力,使得利用有限野外采样点数据获得更高精度的土壤属性空间分布图成为可能。

我国在20世纪80年代以来也有不少土壤学学者开始这方面的研究,相关文献报道主要集中在小尺度下确定性内插和地统计插值方法比较^[8-9]、土壤属性插值效果评价及未采样区土壤属性预测等方面。对于大尺度低密度采样的土壤普查数据,由于采样不均匀且样点分布较为稀疏,土壤属性空间变异的随机性因素增加,使类似方法的应用受限。但从理论上讲,进行空间插值的样点数据满足以下两点^[10]即可:一是空间样点之间具有一定的相关性,二是空间变异具有连续性。因此,为充分发挥大尺度低密度采样下土壤普查数据的潜在价值,利用全国第二次土壤普查东北地区的土壤全氮含量(STNC)数据,比较不同插值方法下土壤属性的空间插值效果,同时利用土壤普查数据体系中显著相关的环境因子数据作为辅助信息,探索提高土壤普查数据空间插值精度的有效途径。为尽量避免土壤普查数据的采样设计不合理及数据内在变异性,比较了不同样本容量下插值结果以充分认识不同插值方法的空间趋势预测稳定性,希望能为今后土壤普查工作的土壤属性制图提供指导。

1 材料与方法

1.1 研究区域和数据来源

研究区域包括黑龙江、吉林、辽宁和内蒙古东部的赤峰市、通辽市、兴安盟和呼伦贝尔市(图1),位于 $115^{\circ}52'—135^{\circ}09'E$, $38^{\circ}72'—53^{\circ}55'N$ 之间,总面积约为 $1.24 \times 10^6 \text{ km}^2$,是世界三大黑土带之一和我国广义的黑土分布区。土壤数据来源于中国第二次土壤普查资料(1979—1992年),排除典型剖面中缺少经纬度坐标和STNC值的记录,共417个有效样点(图1)。每个样点还整理了详细的地理位置、海拔高度、年均温、年降水量等环境特征数据和土种类型、所属土类、剖面厚度、pH、土壤阳离子交换量等土壤理化数据,形成完整的辅助信息因子数据库。

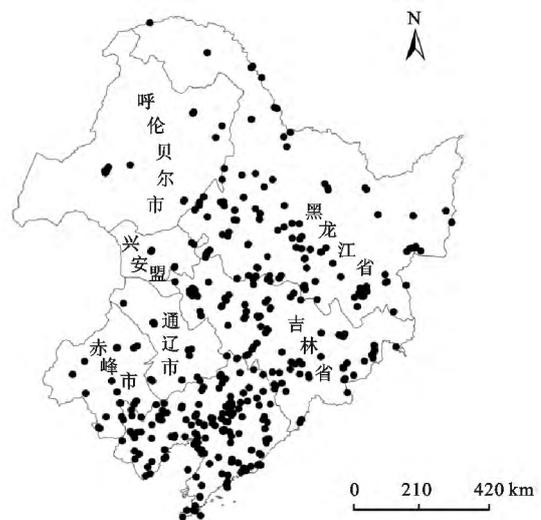


图1 研究区域和有效样点分布

1.2 数据处理

1.2.1 特异值处理与统计分析 由于特异值的存在会造成插值变量空间连续表面中断,采用“样本平均值3倍的标准差”来界定特异值,在该区间以外的数据均定为特异值,并分别用正常的最大值和最小值代替特异值^[11]。利用SPSS 18.0软件对特异值处理后的STNC数据进行单样本K-S(Kolmogorov—Smirnov)检验和描述统计分析,以检验数据分布情况。最后基于ArcGIS 10.0软件平台形成土壤氮素空间属性数据库,并对STNC数据作对数转换计算,以满足空间插值需要。

1.2.2 空间插值 采用确定性插值和地统计插值方法中较为常用且具有代表性的反距离加权法(inverse

distance weighting, IDW)、径向基函数(radial basis functions, RBF)、普通克里金(ordinary kriging, OK)和回归克里金(regression kriging, RK)4种方法对 STNC 进行空间插值,前3种方法的相关报道较多^[9-11],其理论依据这里不再赘述。回归克里金(RK)方法是一种可使用连续变量又可使用类别变量作为辅助信息,同时结合 Kriging 插值方法来预测土壤属性的混合插值技术。可用方程式(1)–(2)表示^[12]:

$$f(x) = a_0 + a_1 m(x) \quad (1)$$

$$Z(x) = f(x) + \varepsilon(x) \quad (2)$$

式中: $f(x)$ ——地理要素空间预测趋势项,用一个随辅助信息平稳变化的线性函数表示; a_0, a_1 ——系数项,使预测模型适宜任何偏离率; $m(x)$ ——影响空间预测趋势的辅助信息因子; $Z(x)$ ——土壤属性预测值; $\varepsilon(x)$ —— $f(x)$ 趋势预测土壤属性值的残差,模拟空间分布趋势及随机因素导致的不确定性。

本文采用趋势项 $f(x)$ 值与残差项的 Kriging 预测值之和作为 RK 方法的预测结果,即 STNC 与环境变量之间的线性回归预测值和回归残差值的 OK

插值之和作为 RK 方法的预测值^[13]。其中,参与线性回归分析的辅助信息因子共 8 个:经度(longitude, Lo)、纬度(latitude, La)、海拔(H)、年均温(mean annual temperature, MAT)、年均降水量(mean annual precipitation, MAP)、土层厚度(d)、pH 值和阳离子交换容量(CEC)。

参考目前研究较为成熟的县域尺度空间插值模型及本研究中不同样本容量下样点尺度的变化,在 ArcGIS 10.0 和 GS+ Version 7.0 中,对 4 种空间插值方法在大尺度低密度采样下的函数模型和参数设定详见表 1。OK 和 RK 方法中涉及到 Kriging 插值均首先在 GS+ 软件中进行相应变量的半方差分析,再基于以上所得最优半方差模型参数,在 ArcGIS 软件平台下完成空间插值。相关研究表明:通过软件得到的各向同性半方差模型实际上是对各向异性半方差函数模型均化的结果^[11],其差别不会对 Kriging 插值产生明显影响,因此在 OK 和 RK 半方差分析中均采用各向同性半方差模型。IDW 和 RBF 两种方法由于拟合模型参数调试相对方便,直接在 ArcGIS 平台上完成。

表 1 不同插值方法模型基本参数设定情况

插值方法	函数模型	设定参数	
		参数值	参数值
IDW	反距离加权模型	幂指数 P	2
		搜索的近邻样点个数 n	最大近邻个数 15 最小近邻个数 10
		搜索区域形状	圆形
RBF	比较完全规则样条、张力样条函数、高次曲面函数 3 种径向基函数拟合结果,以平均误差(ME)和均方根误差(RMSE)均较小为标准,选择最优模型	搜索的近邻样点个数 n	最大近邻个数 15 最小近邻个数 10
		搜索区域形状	圆形
		步长分组个数	15
OK 和 RK	比较球面模型、指数模型、高斯模型 3 种半方差模型,以模型拟合优度(R^2)较大、拟合模型残差(RSS)较小为标准,选择最优模型	搜索的近邻样点个数 n	最大近邻个数 5 最小近邻个数 3
		搜索区域形状	圆形

1.2.3 样本容量设定 为充分认识土壤调查数据 STNC 的空间变异性特征,以 110^[14]为最小个数设计了 7 个不同容量的样点系列。具体操作为:利用 ArcGIS 中的 Create Subsets 功能,随机抽取 417 条样点记录中 80% 的记录作为样本 A_1 ,抽取 A_1 中 80% 的记录作为样本 A_2 ,从 A_2 中抽取 80% 的记录作为样本 A_3, \dots ,依次按比例抽取形成 6 个不同容量的样本数据,再加上全部的 417 条样点记录,共形成 7 个样本系列。由于 STNC 原始数据中部分参与回归分析的环境因子数据缺失,RK 方法初始样本容量为 337,并单独记为样本 B,之后从 A_3 开始计算,尽量保证 4 种空间插值方法使用统一样点系列的 STNC 值进行

空间趋势预测。

1.2.4 插值结果精度检验与比较 选择交叉验证方法中的平均误差(ME, mean error)、均方根误差(RMSE, root mean square error)和预测值与观测值之间的相关系数 R 作为不同空间插值模型精度检验指标。为了消除不同样本容量下 STNC 平均值和标准差的波动性,将不同样本容量下计算的 ME 和 RMSE 分别除以各自的平均值(mean)和标准差(SD)作标准化处理。标准化后的 ME/mean 和 RMSE/SD 值越小, R 值越大则说明空间插值精度越高。同时参照相关文献^[15]引入相对提高指数(RI, Relative Improvement)来定量化表示 RK 方法对 RMSE 精度的

提高度,其计算公式为:

$$RI = \frac{RMSE_i' - RMSE_{RK}'}{RMSE_i'} \times 100\% \quad (3)$$

式中:RMSE'——标准化 RMSE,即对应的 RMSE/SD 值; i ——IDW, RBF, OK 方法。

RI 越大则说明 RK 方法的精度提高程度越高,即辅助信息的添加越有效。

2 结果与讨论

2.1 STNC 统计特征

417 条有效样点数据的统计分析(表 2)表明,STNC 原始取值范围为 0.08~21.48 g/kg,剔除特异值后的取值范围为 0.08~5.31 g/kg。

剔除特异值前后 STNC 变异系数均大于 80%,表现出较强的变异性,这与研究区域自然条件复杂,土壤类型和土地利用类型多样有关。但剔除特异值降低了数据的偏度和峰度,且使偏度和峰度更为接

近。剔除特异值后的 STNC 进行单样本 $K-S$ 检验表明,对数转换之前 STNC 属于偏正态分布($p < 0.001$);经过对数转换后 STNC 的 Z 值由对数转换之前的 4.12 减小为 1.03,且 P 值变为 0.244(> 0.01)。说明经对数转换后 STNC 数据呈正态分布,空间趋势预测中需要对有效样点的 STNC 数据进行对数转换,以保证空间插值数据满足正态分布。不同样点系列之间,随样本容量减小,样点最大间距减小,而最小间距的变化幅度不大(表 2)。样本系列 A_5 以后最大采样间距变化较大,但空间插值范围仍保持 A_1 时的 90%以上,对 STNC 空间变异性在大尺度上的分辨不会存在影响。不同样本容量下 STNC 基本描述统计量波动不大,偏度系数在 2.57~2.89 之间;峰度系数在 7.86~9.73 之间,说明各样本系列具有相似的统计特征,不同容量的样点系列设计比较合理。空间插值过程中不再针对每个样本进行正态分布检验,空间预测完成后将对数转换后的数据计算结果进行还原。

表 2 STNC 数据描述性统计特征及不同样本系列数据特征比较

数据项	样本容量	样点间距/km		含量值/(g·kg ⁻¹)		平均值/(g·kg ⁻¹)	变异系数/%	偏度系数	峰度系数
		最大值	最小值	最大值	最小值				
STNC 实测值	417	1 627.47	1.40	0.08	21.48	1.07	150.47	7.70	79.02
	A_{1-417}	1 627.47	1.40	0.08	5.31	0.98	92.80	2.89	9.73
	A_{2-333}	1 473.73	1.40	0.08	5.31	0.98	89.51	2.81	9.51
	A_{3-267}	1 473.73	1.88	0.08	5.31	1.01	98.51	2.65	8.22
STNC 剔除特异值	A_{4-214}	1 473.73	2.31	0.08	5.31	1.03	88.97	2.64	8.03
	A_{5-171}	1 473.73	2.31	0.08	5.31	0.99	88.06	2.65	8.56
	A_{6-137}	1 353.16	2.31	0.08	5.31	1.01	87.45	2.60	8.41
	A_{7-110}	1 333.60	2.31	0.08	5.31	1.04	90.70	2.57	7.86
	B_{337}	1 473.73	1.40	0.08	5.31	1.01	89.91	2.76	8.99

2.2 辅助信息因子的分析与筛查

根据相关性分析和回归分析结果得到 STNC 回归预测值及回归残差值,是采用 RK 方法分离趋势项和残差项的重要计算过程。由于进行逐步回归时因子变量间的共线性会影响结果的真实性,首先对 8 个环境变量进行共线性判别(表 3),以容忍度(tolerance, T)大于 0.1 和因子膨胀系数(variance inflation factor,

VIF)小于 10 作为因子非共线性的判断标准^[16]。

结果表明,8 个因子之间均满足非共线性,可以全部作为自变量参与到 STNC 的逐步线性回归分析中。通过相关性分析得到 STNC 与 Lo, La, H, CEC 呈显著正相关,与 MAT, d, pH 呈显著负相关,而与 MAP 相关性不显著。这可能是由于人类耕作过程中灌溉措施在一定程度上降低了降水对土壤养分的影响。

表 3 STNC 与影响因子的相关性及变量共线性诊断分析

分析方法	指标项	$Lo(^{\circ})$	$La(^{\circ})$	H/m	$MAT(^{\circ}C)$	MAP/mm	d/cm	pH 值	CEC
共线性诊断	容忍度 T	0.443	0.113	0.493	0.199	0.514	0.897	0.606	0.845
	因子膨胀系数 VIF	2.256	8.817	2.030	8.078	1.945	1.115	1.650	1.183
相关性分析	相关性系数 R	0.253**	0.245**	0.156**	-0.348**	0.033	-0.105**	-0.273**	0.328**

注:Lo 为经度,La 为纬度, H 为海拔高度, MAT 为年均温, MAP 为年均降水量, d 为土层厚度,CEC 为阳离子交换容量。**表示在 0.01 水平上显著相关。

通过逐步线性回归,设定因子变量进入方程的概率水平为 0.05,剔除的概率水平为 0.10,得到 STNC 预测的最优因子为 CEC, MAT, d 和 pH, 最优线性回归方程式如公式(4)。

考虑到本研究空间尺度较大且自然因素和土地利用变化较为复杂,同时对比其他研究尺度下 RK 方法回归方程的决定系数^[3,15,17],该回归拟合方程的决定系数($R^2=0.328$)处于可接受范围。以此计算得到的 STNC 回归预测残差值呈近似正态分布(图 2),故用 OK 方法进行回归预测残差值的空间插值时不再做数据转换处理。

$$\begin{aligned} \text{STNC} = & 0.033\text{CEC} - 0.072\text{AMT} - \\ & 0.005\text{D} - 0.103\text{pH} + 2.206 \end{aligned} \quad (4)$$

$(R^2 = 0.328)$

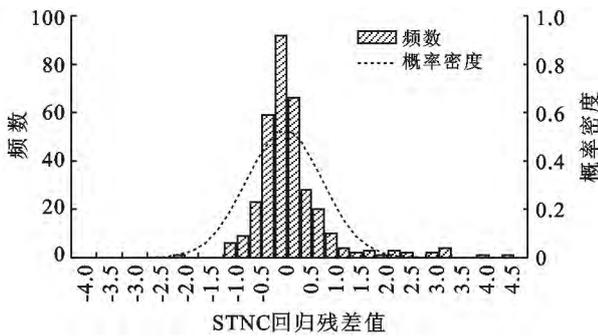


图 2 STNC 回归预测值的残差分布情况

注:样本数为 337,中值为 0.08,标准误差为 0.75,置信度为 95%。

2.3 不同样本容量下 STNC 空间变异性分析

利用最优半方差函数模型定量化描述土壤属性空间变异性,是利用地统计方法进行土壤属性空间插值的基础。通过比较不同模型拟合的 R^2 值和 RSS 值大小,STNC 除样本 A_3 满足球状模型(Sh)外(表 4),其他样本容量下均满足高斯模型(Gs),说明样本容量变化对拟合空间结构的半方差函数模型影响不大。在实际参数设定过程中,有效滞后距和步长的乘积为最大采样间距的 1/2,由于样本容量的减少对样点最大间距的影响,有效滞后距有所下降。块金效应(Nugget/Sill, N/S)是耕作措施、种植制度等随机因素引起的空间异质性占总变异的比,当样本容量 ≥ 171 时,STNC 的 N/S 值为 31.81%~49.93%,表现出中等水平的空间自相关性,且随着样本容量的减少 N/S 值基本呈现上升趋势,说明随机性因素影响比重上升造成了空间自相关性的减弱。当样本容量 < 171 时, N/S 值均小于 25%,此时半方差模型对空间变异性的辨识能力下降使随机性因素被掩盖,空间变异结构发生变化,得到的空间分析结果难以确信。STNC 回归预测残差值的 N/S 取值范围基本落在 STNC 块金效应波动范围以内,说明回归预测残差值保留了 STNC 固有空间结构。STNC 变程为 1 643.72~3 635.57 km,当样本容量 ≥ 171 时,STNC 样点最大间距均落在空间自相关范围内,此时各样本容量下的平均变程为 2 040.53 km。不同样本容量下 STNC 回归预测残差值的空间自相关范围也均大于 2 000 km。

表 4 STNC 半方差模型及拟合参数

插值变量	样本容量	最优模型	有效滞后距/km	步长/km	块金值	基台值	块金效应/%	变程/km	R^2	RSS
STNC	A1—417	Gs	818	55	0.347	1.091	31.81	2 083.66	0.973	2.98 E-03
	A2—333	Gs	815	54	0.347	0.935	37.11	1 882.74	0.961	3.70 E-03
	A3—267	Sh	815	54	0.365	0.747	48.86	2 088.00	0.924	4.53 E-03
	A4—214	Gs	749	50	0.382	1.168	32.71	2 504.55	0.920	5.36 E-03
	A5—171	Gs	749	50	0.379	0.759	49.93	1 643.72	0.766	0.013 7
	A6—137	Gs	749	50	0.384	1.821	21.09	3 635.57	0.624	0.024 1
	A7—110	Gs	749	50	0.429	1.781	24.09	3 635.57	0.504	0.034 8
STNC 回归预测残差值	B—377	Sh	818	55	0.342	0.685	49.93	2 099.00	0.723	0.015 8
	A3—267	Gs	815	54	0.384	1.182	32.49	2 767.82	0.651	0.025 6
	A4—214	Gs	815	54	0.394	1.939	20.32	3 638.65	0.781	0.017 1
	A5—171	Ep	749	50	0.310	0.621	49.92	2 928.00	0.548	0.026 4
	A6—137	Ep	749	50	0.264	0.529	49.91	2 922.00	0.565	0.016 9
	A7—110	Ep	749	50	0.309	0.620	49.92	3 780.00	0.315	0.043 9

注:Gs 为高斯模型,Ep 为指数模型,Sh 为球状模型。

目前针对大尺度低密度采样的土壤属性空间插值的报道很少,对比东北地区已有的相关研究,在小

流域尺度(992 hm^2)上 292 个样点的 STNC 空间插值的 N/S 块金效应为 89.2%,变程为 1 899 m^[18]。县

域尺度上,张素梅等^[3]对自采的吉林省农安县 471 个样点的 STNC 空间插值的 N/S 值为 36.7%,变程为 109.3 km。Liu 等^[19]利用吉林省德惠县的 354 个土壤肥力普查样点数据所作的 SOCC 空间插值的 N/S 值为 16.1%,变程达到 632 km。区域尺度上,Zhang 等^[20]利用 995 个自采样点对黑龙江和吉林两省黑土区的 STNC 空间插值的 N/S 值为 5.97%,变程为 1 037.5 km。上述结果不难发现,不同空间尺度下不同样本容量的土壤碳或氮的块金效应较为一致地表现出中强度空间相关性,且变程随着各研究尺度的增大而增大。本研究区域为整个东北地区,研究数据的样点间距远大于上述同类研究,得到的空间自相关范围也较大。

2.4 不同方法的空间插值结果

比较 4 种插值方法的空间趋势预测结果(图 3),整体来看各空间插值方法均能较直观的反映出该区 STNC 空间分布的基本态势,STNC 均呈现从东北向

西南逐渐递减规律。高值区(>1.10 g/kg)呈倒“V”形条带状贯穿于研究区域的东北部地区,包括黑龙江省大部分地区和内蒙古东北部的呼伦贝尔市及三江平原地区。所对应的基本是暗棕壤、漂灰土、黑钙林区及沼泽湿地等理论上均有利于有机质积累的区域。西北部中值区($0.63\sim 1.10$ g/kg)包括松嫩平原在内的吉林省和内蒙古的兴安盟、赤峰市以及通辽市的西北部地区。该区域主要为农业耕作区和草甸草原,其中受人类活动影响的松嫩平原多发育一些草甸白浆土、白浆化暗棕壤等过渡土壤类型,STNC 水平适中。低值区(<0.63 g/kg)则呈斑块状分布在通辽市的东南部和辽宁省的西北部沿海地区,是沙漠、旱地和水田区,土壤类型以风沙土、潮土和滨海盐土等有机质积累较弱的土壤类型为主。此插值结果分布趋势与 Yang 等^[21]使用全国第二次土壤普查数据得到的土壤有机碳密度的空间插值结果在该区的分布态势保持很好的一致性。

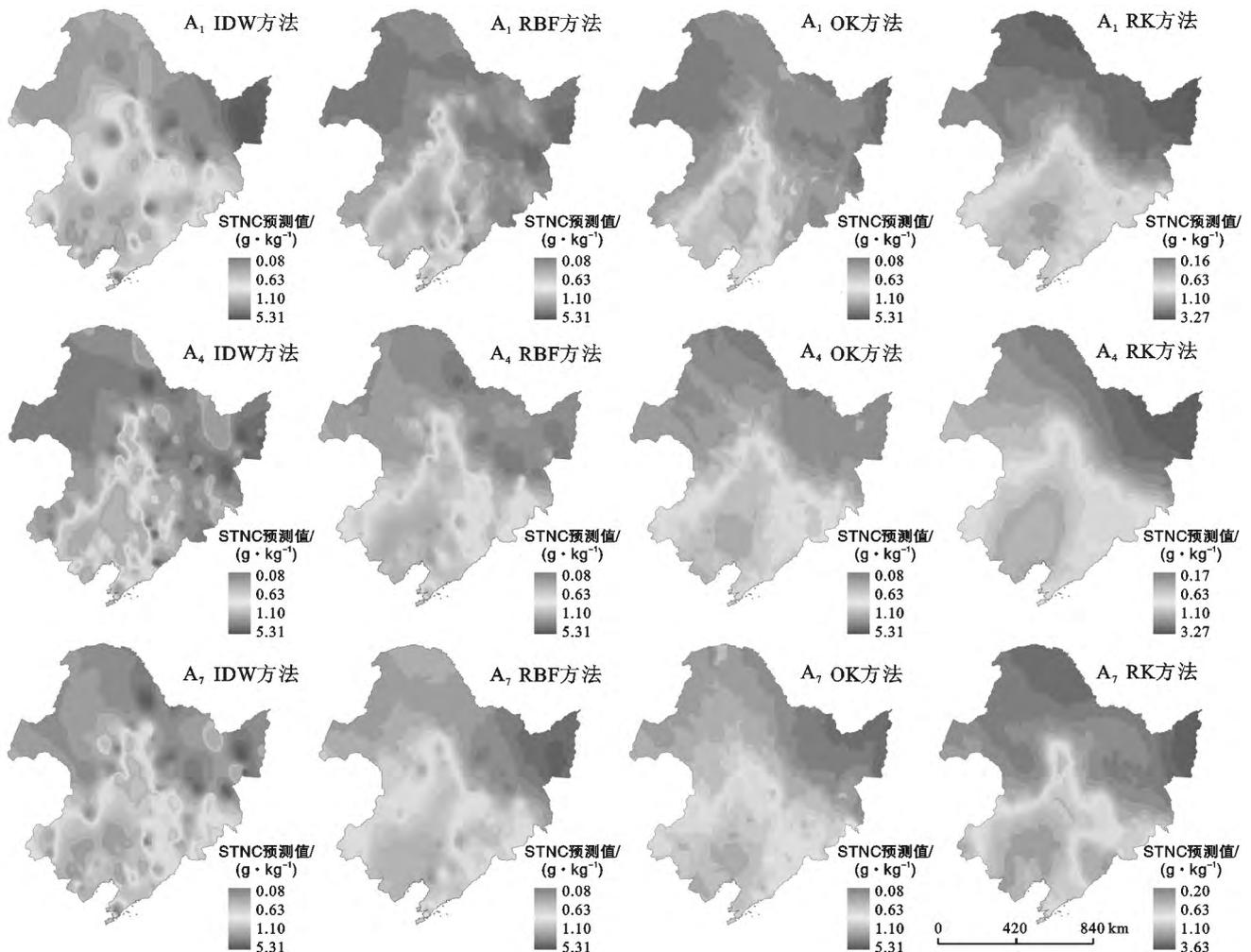


图 3 STNC 不同空间插值方法结果比较

注:RBF 方法中,A₁ 采用的最优模型为张力样条函数,A₄、A₇ 采用的最优模型为规则完全样条函数。

对比不同样本容量下插值结果(A₁, A₄, A₇ 3 个样本系列,图 3)可以发现,随着样本容量降低,4 种方法插值结果的中、低值区范围扩大,高值区缩小且有了明显的带状化表现,尤其是采样点比较稀疏的东北部三江平原地区、内蒙古东北和黑龙江西北部,过分的带状化削弱了空间变化特征。在空间分布的细节表达上,相同样本容量下 IDW 和 RBF 插值图中出现了特异值斑块或“牛眼”现象,造成空间趋势表达不连续,OK 和 RK 方法由于 Kriging 插值模型对空间特异性具有一定的平滑作用,STNC 空间变化规律明显。同一种空间插值方法下,随着样本容量的减少,整体变化规律一致但趋势变化的连续性变差。对比 RK 方法下 A₁ 和 A₄ 两个样点系列的插值含量水平和斑块面积变化,在样本容量较少(A₄)情况下 RK 方法在空间趋势预测上仍较为稳定并在西北部和东北部的样点稀疏区体现出更多的细节变化,说明辅助信息的加入对这些区域的空间预测具有一定帮助。

2.5 不同空间插值方法的精度分析

插值结果的精度比较是评价不同插值方法有效性及适用性的重要手段,不同插值方法对土壤属性插

值精度在不同地区相关研究中的表现不一,部分研究表明 Kriging 的效果优于 IDW^[22] 和 RBF 中的 Spline 函数模型^[23],也有研究表明 IDW 和 Kriging 插值方法均有较好的表现^[8]。仅针对不同 Kriging 插值方法的研究多体现为 RK 方法或其他添加辅助信息的 Kriging 插值方法和模型在提高插值精度方面表现较好^[24]。比较本研究中 4 种空间插值方法的标准化 ME,标准化 RMSE, R 三个精度指标可见(图 4),其空间预测效果顺序为:RK>OK>RBF>IDW。不同空间插值方法下 3 个精度指标值随着样本容量的变化规律不明显,计算不同样本容量下各精度指标平均值,得到 RK 方法的 ME 值同 OK, RBF 和 IDW 方法相比平均降低 0.18, 0.17 和 0.16, RMSE 值平均分别降低了 0.15, 0.18 和 0.19, R 平均分别提高了 0.40, 0.42 和 0.43。

从插值精度的变化幅度来看,随着样本容量的变化, RK 和 OK 方法对不同样本系列表达效果较为稳定。由上述分析可知,当样本容量<171 时, STNC 空间变异结构发生变化,此时 A₆, A₇ 两个样本容量下反映出的精度变化水平也不可信。

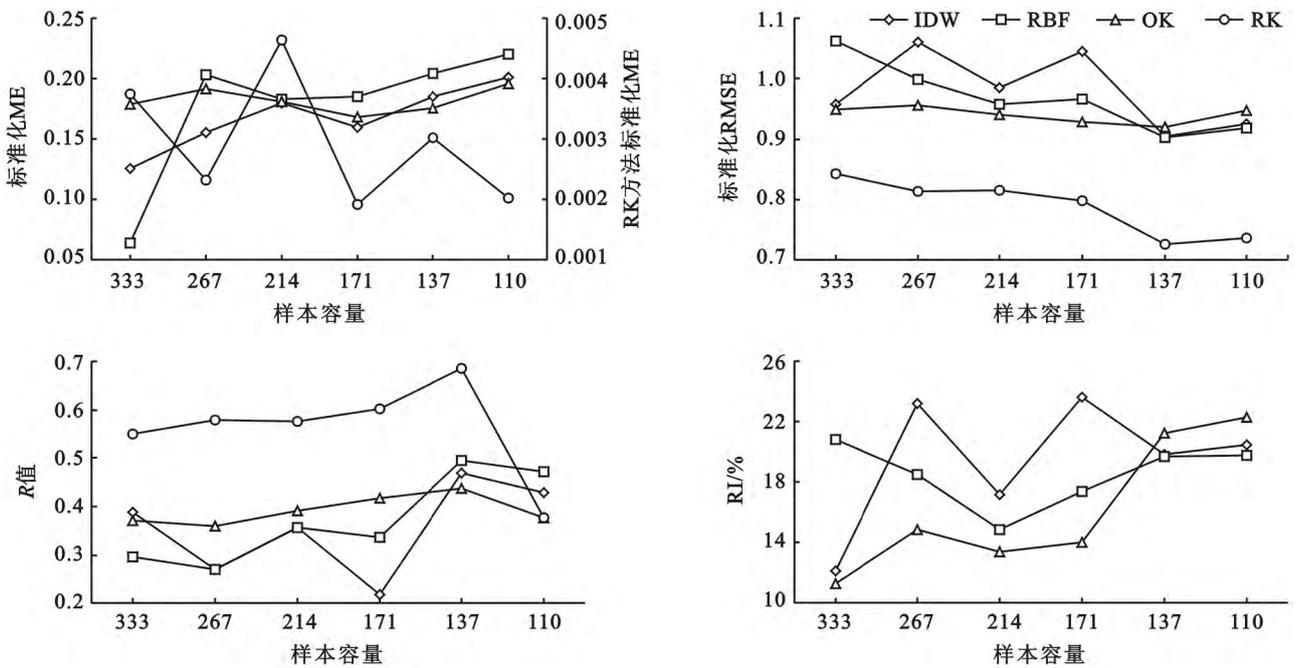


图 4 不同方法 STNC 空间插值精度随样本容量的变化

通过计算各样本容量的平均 RI 值(图 4), RK 方法相对于其他空间插值方法将预测精度分别提高了 16.15%(OK), 18.50%(RBF)和 19.40%(IDW), 表明利用环境因子作为辅助信息使该区空间插值精度得到了有效提高。相关研究中, Simbahan^[4] 等利用地表反射率、相对高程及土壤电导率 3 个环境因子作

为辅助性的 RK 方法提高了 C_s 元素制图精度的 16%。Zhang^[7] 等利用水流功率指数、湿润指数及坡度 3 个环境变量作为辅助信息的 RK 方法对土壤有机质空间插值提高了 13.93% 的精度。Sumfleth^[15] 等利用相对高程和河网高程 2 个环境变量作为辅助信息的 RK 方法对 STNC 空间预测精度提高了

13%。本研究使用的辅助信息是土壤普查工作中必须测定的项目或者是相应土壤点位上配套的气象站点数据,同上述研究相比,在辅助信息易获得性以及土壤属性的时空匹配性上具有明显的优势,这对于今后提高土壤普查数据的空间插值精度价值巨大。

3 结论

(1) 我国第二次土壤普查期间东北地区 STNC 在 0.08~21.48 g/kg 之间,剔除特异值前后 STNC 数值水平具有较大变异性。大尺度低密度采样下 STNC 表现出中等水平空间自相关性,但变程大于同区域的小尺度研究结果。目前针对大尺度(>1:100 000)低密度采样的土壤属性空间插值方法的研究较少,本研究得到的 STNC 空间变异性特征只能与同区域小尺度上的相关研究进行类比分析,得到的变程值对于 STNC 空间自相关范围的指示意义难以进行有效评价,但空间插值方法是利用土壤普查数据进行土壤属性制图的有效途径,因此对于大尺度低密度采样的土壤属性空间插值研究需要进一步的加强。

(2) 添加了辅助信息的回归 Kriging(RK)是对经度(Lo)、纬度(La)、海拔(H)、年均温(MAT)、年降水量(MAP)、土层厚度(d)、pH 值、土壤阳离子交换量(CEC)8 个环境变量与 STNC 值进行线性回归拟合,将得到的回归拟合残差值运用普通 Kriging(OK)进行 OK 空间插值模拟,由 OK 模拟后的残差值与回归拟合的预测值之和得到 RK 方法的空间预测值。共线性判别表明,8 个因子之间均满足非共线性,可以全部作为自变量参与到 STNC 的逐步线性回归分析中;通过逐步线性回归得到本研究 RK 方法的辅助信息为 CEC, MAT, d 和 pH 值,其中,STNC 与 CEC 呈显著正相关,与 MAT, d, pH 值呈显著负相关关系。

(3) 样本容量变化对不同插值方法的 STNC 空间趋势预测影响不明显。但在细节表达上,相同样本容量下 OK 和 RK 方法的 STNC 空间插值结果空间变化特征更鲜明。随着样本容量的减少,同一种空间插值方法的空间表达更简略。RK 方法对无样点区的空间趋势预测体现出更多的细节变化。此外,在以后的土壤普查工作中,应该着重加强本研究中氮含量插值结果的高、中、低值区的过渡地带以及内蒙古东北部的风沙地、黑龙江东部与东北部的多年冻土区和吉林东南沿海的盐渍地等生态敏感地区的布点设计。

(4) 不同样本容量下 4 种方法空间插值精度总体表现为:RK>OK>RBF>IDW,大尺度低密度采样下地统计插值方法要优于确定性内插。RK 方法

相对于其他空间插值方法将预测精度提高程度(RI)分别为 16.15%(OK),18.50%(RBF)和 19.40%(IDW)。土壤普查过程中所收集的环境因子作为辅助信息即可使该区土壤氮含量的空间插值精度得到有效地提高,而且在低样本容量下 RK 方法在空间趋势预测上仍较为稳定,预示着辅助信息的添加对大尺度低密度采样下 STNC 的空间表达具有较好的效果。

[参 考 文 献]

- [1] Lam N S. Spatial Interpolation[M]// International Encyclopedia of Human Geography. Elsevier Science Ltd., 2009.
- [2] Takata Y, Funakawa S, Akshalov K, et al. Spatial prediction of soil organic matter in northern Kazakhstan based on topographic and vegetation information[J]. Soil Science and Plant Nutrition, 2007, 53(3):289-299.
- [3] 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.
- [4] Simbahan G C, Dobermann A, Goovaerts P, et al. Fine-resolution mapping of soil organic carbon based on multivariate secondary data[J]. Geoderma, 2006, 132(3/4):471-489.
- [5] McKenzie N J, Ryan P J. Spatial prediction of soil properties using environmental correlation[J]. Geoderma, 1999, 89(1/2):67-94.
- [6] Bui E N, Moran C J. A strategy to fill gaps in soil survey over large spatial extents an example from the Murray-Darling basin of Australia[J]. Geoderma, 2003, 111(1/2):21-44.
- [7] Zhang Shiwen, Huang Yuanfang, Shen Chonggan, et al. Spatial prediction of soil organic matter using terrain indices and categorical variables as auxiliary information[J]. Geoderma, 2012, 171(S):35-43.
- [8] Schloeder C A, Zimmerman N E, Jacobs M J. Comparison of methods for interpolating soil properties using limited data[J]. Soil Science Society of America Journal, 2001, 65(2):470-479.
- [9] 石小华,杨联安,张蕾.土壤速效钾养分含量空间插值方法比较研究[J].水土保持学报,2006,20(2):68-72.
- [10] Webster R. Soil resources and their assessment[J]. Phil. Trans. R. London; B, 1997, 352(1356):963-973.
- [11] 刘付程,史学正,于东升,等.基于地统计学和 GIS 的太湖典型地区土壤属性制图研究:以土壤全氮制图为例[J].土壤学报,2004,41(1):20-27.
- [12] Stacey K F, Lark R M, Whitmore A P, et al. Using a process model and regression Kriging to improve predictions of nitrous oxide emissions from soil[J]. Geo-

- derma, 2006,135:107-117.
- [13] Odeh I O A, McBratney A B, Chittleborough D J. Further results on prediction of soil properties from terrain attributes; Heterotopic Cokriging and regression-Kriging[J]. Geoderma, 1995, 67(3-4):215-226.
- [14] Webster R, Oliver M A. Sample adequately to estimate variograms of soil properties[J]. Journal of Soil Science, 1992,43(1):177-192.
- [15] 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.
- [16] Menard S W. Applied logistic regression analysis[M]. Thousand Oaks; CA: SAGE Publication, Inc., 1995.
- [17] Ziadat F M. Analyzing digital terrain attributes to predict soil attributes for a relatively area[J]. Soil Science Society of America Journal, 2005,69(5):1590-1599.
- [18] Wei Jianbing, Xiao Duning, Zhang Xingyi, et al. Topography and land use effects on the spatial variation of soil organic carbon: A case study in a typical small watershed of the black soil region in Northeast China [J]. Euras Soil Sci., 2008,41(1):39-47.
- [19] Liu Dianwei, Wang Zongming, Zhang Bai, et al. Spatial distribution of soil organic carbon and analysis of related factors in croplands of the black soil region, Northeast China[J]. Agriculture, Ecosystems and Environment, 2006,113(1/4):73-81.
- [20] Zhang Xingyi, Sui Yueyu, Zhang Xudong, et al. Spatial variability of nutrient properties in black soil of Northeast China[J]. Pedosphere, 2007,17(1):19-29.
- [21] Yang Yuanhe, Mohammad A, Feng Jianming, et al. Storage, patterns and environmental controls of soil organic carbon in China[J]. Biogeochemistry, 2007,84(2):131-141.
- [22] Kravchenko A, Bullock D G. A comparative study of interpolation methods for mapping soil properties [J]. Agronomy Journal, 1999,91(3):393-400.
- [23] Panagopoulos T, Jesus J, Antunes M D C, et al. Analysis of spatial interpolation for optimising management of a salinized field cultivated with lettuce[J]. European Journal of Agronomy, 2006,24(1):1-10.
- [24] Hengl T, Heuvelink G, Alfred S. A generic framework for spatial prediction of soil variables based on regression-Kriging[J]. Geoderma, 2004,120(1/2):75-93.

~~~~~

(上接第 152 页)

- [5] 李惠茹. 陕北县南沟流域景观格局及动态研究[D]. 陕西杨凌:西北农林科技大学,2011.
- [6] 王继军,权松安,郭满才. 退耕还林还草中建立生态系统与经济系统“弹性资源”初论[J]. 水土保持通报,2004,24(5):95-98.
- [7] 唐华俊,李哲敏. 基于中国居民平衡膳食模式的人均粮食需求量研究[J]. 中国农业科学,2012,45(11):2315-2327.
- [8] 赵丹,李锋,王如松. 基于生态绿当量的城市土地利用结构优化:以宁国市为例[J]. 生态学报,2011,31(20):6242-6250.
- [9] 毛文永. 生态环境影响评价概论[M]. 北京:中国环境科学出版社,2003.
- [10] Perry J H, Ryan R J. The effect of urban leaf area on summertime urban surface kinetic temperatures: A Terre Haute case study[J]. Urban Forestry & Urban Greening, 2007,6(2):63-72.
- [11] 吴钦孝. 黄土高原的林草资源和适宜覆盖率[J]. 林业科学,2000,36(6):6-7.
- [12] 黄怀雄,赵红艳. 长株潭地区森林固碳释氧功能价值评价[J]. 林业调查规划,2010,35(2):136-138.
- [13] 马长欣,刘建军,康博文,等. 1999—2003年陕西省森林生态系统固碳释氧服务功能价值评估[J]. 生态学报,2010,30(6):1412-1422.
- [14] 刘忠琛,刘正梅. 肉羊舍饲高效养殖技术要点[J]. 养殖与饲料,2003(5):15-17.
- [15] 圈养舍饲绒山羊调查研究课题组. 绒山羊舍饲圈养技术研究:基于陕西省榆林市推行绒山羊舍饲圈养的实证分析[J]. 中国畜牧杂志,2008(S):184-191.
- [16] 刘吉利,程序,谢光辉,等. 收获时间对玉米秸秆产量与燃料品质的影响[J]. 中国农业科学,2009,42(6):2229-2236.
- [17] 折凤霞,郝明德,臧逸飞. 黄土高原沟壑区苜蓿生产力及养分特性的研究[J]. 草业学报,2013,22(2):313-317.
- [18] 高阳,程积民,刘伟. 黄土高原地区不同类型天然草地群落学特征[J]. 草业科学,2011,28(6):1066-1069.