

DOI: 10.13758/j.cnki.tr.2024.04.024

赵萍, 阮旭东, 刘亚风, 等. 基于遗传算法和 BP 神经网络的矿区土壤重金属含量空间分布预测. 土壤, 2024, 56(4): 889–896.

基于遗传算法和 BP 神经网络的矿区土壤重金属含量空间分布预测^①

赵萍¹, 阮旭东¹, 刘亚风^{2*}, 赵思逸¹, 孙雨¹, 常杰¹, 周俊¹

(1 合肥工业大学资源与环境工程学院, 合肥 230009; 2 安庆师范大学资源环境学院, 安徽安庆 246133)

摘要: 本研究提出了一种基于遗传算法(Genetic algorithm, GA)和 BP 神经网络(Back propagation neural network, BPNN)的复合模型——GABP 模型, 以安徽省池州市某矿区及其周边为研究区, 预测了土壤中 pH 和 7 种重金属元素(Cd、Pb、Cr、Cu、Ni、Hg、As)含量的空间分布, 并与 BPNN 和反比距离权重法(Inverse distance weighting, IDW)进行了比较。研究结果表明: 受采矿活动影响, 研究区土壤 pH 和重金属含量呈显著的空间分异性; GABP 复合模型的数据扩增能够有效弥补 BPNN 对样本数量的依赖, 同时结合了地理位置和高程属性, 精度评价结果显示 GABP 模型的平均 R^2 、 r 、RMSE、MAE 分别是 IDW 和 BPNN 的 3.03 倍、2.56 倍、2.93 倍、2.39 倍、0.85 倍、0.61 倍、0.79 倍、0.62 倍, 预测精度更高。模型解决了传统空间插值方法结果中可能出现负值和边界无法插值的问题, 为土壤重金属含量空间分布预测提供了一种新方法。

关键词: 遗传算法; BP 神经网络; GABP 模型; 空间分布预测; 重金属含量

中图分类号: X53; P934 **文献标志码:** A

Prediction of Spatial Distribution of Soil Heavy Metal Contents in Mining Areas Based on Genetic Algorithm and BP Neural Network

ZHAO Ping¹, RUAN Xudong¹, LIU Yafeng^{2*}, ZHAO Siyi¹, SUN Yu¹, CHANG Jie¹, ZHOU Jun¹

(1 School of Resources and Environmental Engineering, Hefei University of Technology, Hefei 230009, China; 2 School of Resources and Environment, Anqing Normal University, Anqing, Anhui 246133, China)

Abstract: Based on genetic algorithm (GA) and back propagation neural network (BPNN), this study proposed a composite model: GABP model. Focusing on a mining area and its surroundings in Chizhou City, Anhui Province, the spatial distribution of soil pH value and the concentrations of seven heavy metals (Cd, Pb, Cr, Cu, Ni, Hg and As) were predicted by GABP model, and the prediction results were compared with those of BPNN and inverse distance weighting (IDW) method. The results showed that, influenced by mining activities, there was significant spatial heterogeneity in soil pH value and heavy metal concentrations in the study area. The data augmentation of GABP model effectively compensated for the dependency of BPNN on the sample size, and simultaneously incorporated geographical location and elevation attributes. The precision evaluation results indicated that the average R^2 , r , RMSE and MAE of GABP model was 3.03 times and 2.56 times, 2.93 times and 2.39 times, 0.85 times and 0.61 times, 0.79 times and 0.62 times higher than those of IDW and BPNN, respectively, indicating a higher predictive accuracy. The proposed model can solve the issues in traditional spatial interpolation methods where negative values and boundary interpolation difficulties may occur, and provides a novel approach for predicting the spatial distribution of soil heavy metal contents.

Key words: Genetic algorithm; BP neural network; GABP model; Spatial distribution prediction; Heavy metal content

土壤是维持生态系统平衡与稳定的最主要媒介。矿山在开采、选矿、冶炼过程中产生的废弃物不仅污染企业用地, 而且对周边耕地土壤带来生态风险^[1]。

矿区土壤重金属污染具有高含量、空间异质性、伴生性等特征, 在直接暴露或食物链途径等作用下, 易对生态系统和人体健康产生潜在危害^[2-3]。

①基金项目: 国家自然科学基金项目(41972304)资助。

* 通讯作者(119817110@qq.com)

作者简介: 赵萍(1975—), 女, 安徽肥东人, 博士, 副教授, 主要研究方向为资源环境遥感与 GIS 应用。E-mail: njuzhp@sina.com

传统的土壤重金属空间分布研究是通过实地调查、采样和测试获取不同采样点的土壤重金属含量来分析。由于采样条件、研究区实际情况和研究成本等因素的限制,大量进行采样、测试分析并不现实。因此运用合理方法,通过样点数据预测掌握整个研究区土壤重金属空间分布具有重要意义。

利用已知点的数据对未知点重金属元素含量进行空间插值可以快速得到整个研究区域的重金属分布状况。目前空间插值的方法主要有确定性插值方法,如反比距离权重法(Inverse distance weighting, IDW)、径向基函数(Radial basis function, RBF)、全局多项式插值(Global polynomial interpolation, GPI)等和地统计插值方法(如克里金法)两类^[4]。曾星等^[5]和张龙等^[6]分别采用 IDW 方法和结合环境因子的克里金法对农田土壤指标的空间分布进行了分析。总体来看,确定性插值方法计算效率高,对于数据量充足且均匀分布的样本点,插值结果一般较为平滑,但对于非均匀分布的点,插值结果可能不够准确。地统计插值方法能够考虑空间数据的统计特性,更好地反映数据的空间变异性,提供插值结果的不确定性评估。但两种方法对于边界问题和异常值处理比较复杂,可能无法进行准确的插值^[7]。

近年来,机器学习算法凭借其计算速度快与准确度高特点,在土壤重金属污染状况评估、土壤重金属元素含量预测方面得到了广泛应用^[8]。由于土壤是一个空间连续性变异的非均质体,各样点的重金属含量与其所在点位置和污染等级之间存在复杂的随机非线性映射关系^[9]。人工神经网络(Artificial neural network, ANN)是一种仿生机制数学模型,能模仿动物神经连接节点分布式并行处理信息的行为^[10]。ANN 能较好地映射样点重金属含量与其空间位置的非线性关系,提供一批相互对应的输入和输出数据,分析其内在联系并预先设计预测模式,输入新数据即可推算输出结果,使得 ANN 能够对空间各点处的土壤重金属含量进行预测和评价^[11]。

BP 神经网络(Back propagation neural network, BPNN)模型建立过程需要大量训练样本进行学习,并使用测试样本进行验证。然而由于客观条件限制,有时很难获取足够的样本数据。样本数据不足可能会影响模型的准确性和可靠性。因此,在实际应用中,需要寻找其他有效方法来解决样本数据不足的问题,例如采用数据增强技术、合理选择特征等。Wang 等^[12]针对小样本数据进行表层土壤湿度反演存在的过拟合和目标任务欠拟合的问题,提出利用 α 近似方

法对实测表层土壤湿度数据进行扩充,为机器学习模型提供更多的训练数据。尹光彩等^[13-14]利用遗传算法(Genetic algorithm, GA)耦合 ANN 产生新个体实现空间插值,模型具有良好的预测性能,但其产生的位置具有随机性。戴倩倩等^[15]和龚仓等^[16]通过研究和收集关于地理探测器在土壤重金属空间分异领域的研究,发现高程差异对土壤 pH 和重金属空间分布和含量有显著影响。高程会影响土壤的湿度、类型、植被覆盖和雨水淋溶等,从而影响土壤重金属空间分布和含量。目前大多数关于 GA 和 BPNN 结合的研究只是利用 GA 可以收敛到全局最优解,来提高神经网络收敛速度、全局优化能力和学习能力。然而利用 GA 对小样本数据进行数据扩增,再耦合 BPNN,结合平面坐标与高程因素对指定点来进行预测插值方面的研究鲜有报道。

因此,本研究在前人研究的基础上,提出了基于 GA 和 BPNN 的复合模型——GABP。以安徽省池州市某矿区及其周边为研究区,首先利用 GA 对土壤采样点数据进行扩增,然后用遗传算法优化的 BP 神经网络(Genetic algorithm-BP evaluation network, GA-EN)对新个体数据进行评价;最后,使用遗传算法优化的 BP 预测网络(Genetic algorithm-BP prediction network, GA-PN)对研究区土壤 pH 和重金属含量空间分布进行预测,并与传统的 IDW 法进行对比。该模型不仅考虑采样点的地理位置,还考虑了高程对土壤重金属分布和含量的影响,数据扩增部分考虑了元素间的伴生性,克服了 BP 神经网络和传统插值方法需要大量数据及传统插值方法边界无法准确插值、出现负值的缺陷,使模型预测结果有较高的准确性与稳定性。研究方法为区域土壤重金属元素含量空间分布预测提供了一种新的参考。

1 材料与方法

1.1 研究区概况与采样点布置

本文以安徽省池州市“某矿区及周边区域土壤环境现状调查”项目数据为依托开展研究。该矿区位于长江中下游南岸、皖南山区北。原矿山地下分为东西两个开采区,东矿区为硫铁矿,西矿区为金矿。现西矿区已停止开采,研究区包含矿区和居民区、农田等,总面积为 4.42 km²(图 1)。分别于 2020 年 11—12 月和 2021 年 11 月在研究区进行布点采样。以国家场地调查和场地环境监测的相关技术导则要求为前提,结合前期踏勘与资料分析结果,对开采区及周边敏感目标土壤污染情况进行调查;以东矿区周边 1 km 为监

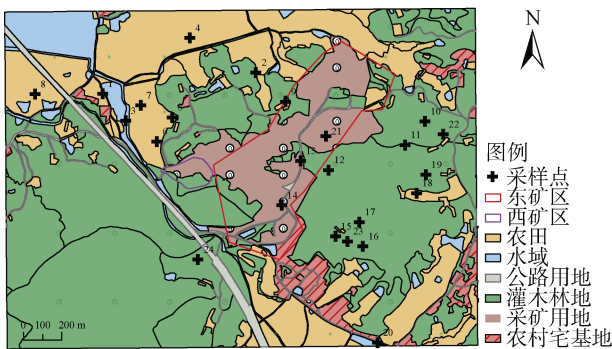


图 1 研究区位置及采样点示意图
Fig. 1 Location and sampling points in study area

测区域，参照监测区域用地类型、淋溶水流向等进行布点；基于二分法原理对监测半径进行划分，在矿区排污口下游、贫液池上下游和矿区排污口等有污染源

的地区加密布点；确保采样点的布置能够满足判别矿区及周边地块污染情况的要求，同时在现场环境条件不具备采样条件时，对点位进行调整，做到现场勘察与采样相结合^[17]，确定新的采样点位，用 GPS 记录坐标。鉴于前期调查发现西部地区整体地势较高，未受到本次项目区内淋溶水等污染的影响，因此在该区域未设置采样点。最终共采集表层(0~20 cm)土壤样品 24 份，点位分布见图 1。

1.2 土壤测定方法与数据来源

pH 采用 pH 计玻璃电极法测定，Cd、Pb 采用石墨炉原子吸收分光光度法测定，Cr、Cu、Ni 采用火焰原子吸收分光光度法测定，Hg、As 采用原子荧光法测定，检测结果见表 1。数字高程数据采用的是 GDEM V3 30M，精度为 30 m，下载自地理空间数据云网站(<https://www.gscloud.cn/>)。

表 1 土壤 pH 和重金属含量(mg/kg)
Table 1 Soil pH and heavy metal contents

土壤样品序号	pH	Cd	Pb	Cr	Cu	Ni	Hg	As
1	6.17	0.37	36.7	1.5	34	21	0.298	10.9
2	5.26	1.03	40.1	2.2	107	23	0.353	12.3
3	5.13	0.99	81.6	1.6	78	24	0.608	109
4	4.65	0.21	41.3	2.1	26	21	0.270	13.3
5	4.60	0.25	102	0	81	19	0.186	33.7
6	4.57	0.80	158	0	197	22	0.664	283
7	5.39	0.40	66.1	0	34	16	0.156	30.2
8	6.32	0.25	20.3	108	146	7	0.448	14.0
9	5.77	0.33	14.2	120	139	7	0.392	14.6
10	6.69	0.25	65.0	94	135	7	3.03	163
11	7.85	0.36	253	96	143	7	3.42	271
12	3.38	2.93	317	259	131	21	16.4	1 030
13	6.23	3.08	807	256	134	20	42.0	2 720
14	5.99	0.38	39.9	261	132	21	0.773	105
15	5.80	2.98	403	263	44	21	5.16	1 310
16	5.48	1.07	42.8	79	44	21	6.77	132
17	5.42	0.26	77.7	76	38	7	5.72	198
18	5.39	0.41	14.2	268	40	7	2.95	32.1
19	5.31	0.33	82.0	68	38	7	2.51	53.3
20	5.08	0.38	64.5	77	38	14	0.41	85.6
21	5.81	0.31	40.7	106	133	14	0.616	14.2
22	7.03	0.4	606	146	34	74	4.86	203
23	6.61	0.75	119	151	35	70	7.95	381
24	6.12	3.01	774	95	63	70	6.32	489

1.3 研究方法

本研究主要基于 GA 和 BPNN，构建了一种预测土壤 pH 和重金属含量空间分布的复合模型——GABP。

模型预测的具体流程(图 2)为：①使用 K-均值聚类算法，根据采样点坐标计算距离，对 24 个采样点进行分类，再利用 GA 的繁殖机制，生成一系列待检验的

新个体;②将新个体同样本数据一同作为训练集输入到 GA-EN 中,样本点的 11 个已知特征值作为输入,预测值作为输出,计算样本点预测值和真实值之间的均方根误差(RMSE);③重复前面两步,仅保留加入训练集后,计算的 RMSE 不比上一轮的值大的新个

体,确保只保留对模型预测精度提升有贡献的个体;④达到迭代次数后,输出所有保留的新个体,完成数据扩增;⑤构建 GA-PN 网络,将扩增后的数据和样本数据作为新的数据集,训练 GA-PN,预测研究区内待求点的含量值。

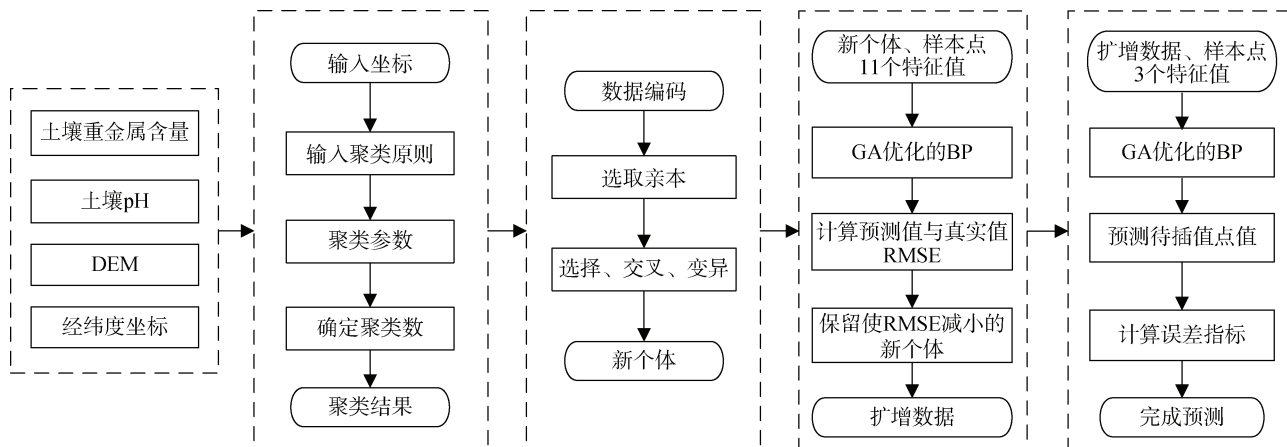


图 2 GABP 模型设计流程图

Fig. 2 GABP model design workflow diagram

1.3.1 K-均值聚类 K-均值聚类是一种基于距离度量的无监督学习方法,用于将一个数据集分成 k 个不同的组或类。其主要思想是找到一些中心点,将每个点分配给离它最近的中心点所对应的类^[18]。本研究中两点间的距离采用欧氏距离,种群中心 C_i 的横坐标用公式(1)计算。聚类的类别数 k 依据下面 3 个指标选取:①每个类内部采样点地理位置的标准差(Average within cluster std, AWCS)尽量小;②不同类之间采样点地理位置的标准差(Between-cluster std, BCS)尽量大;③每个类内采样点数量(Average cluster size, ACS)尽量多^[14]。相关计算公式如下:

$$C_{ix} = \frac{\sum_{i=1}^n X_i}{n} \quad (1)$$

$$\text{Std} = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}}, \quad \text{AWCS} = \frac{\sum_{i=1}^k \text{Std}_i}{k} \quad (2)$$

$$\text{BCS} = \sqrt{\frac{\sum_{i=1}^k (C_i - \bar{C})^2}{k}} \quad (3)$$

$$\text{ACS} = \frac{N}{k} \quad (4)$$

式中: C_{ix} 代表种群 C_i 的新中心点的横坐标, n 是每个种群内点的数量, X_i 是种群 C_i 中的第 i 个点的横坐标, \bar{X} 表示采样点横坐标的均值, \bar{C} 为所有种群中心坐标的平均值, N 为个体总数。

1.3.2 遗传算法(Genetic algorithms, GA) GA 是一种随机搜索最优解的智能算法,模拟了生物种群中个体优胜劣汰的自然规律^[5]。它以目标问题的不同解作为不同的个体生成种群,然后对种群进行选择、交叉、变异来寻求最优个体以实现对目标问题的求解^[19]。为了使 GA 可以更加有效地搜索适应度函数的最优解空间,且避免出现异常值,在算法中设置新个体的上界为种群内各元素对应的最大值,下界为种群内各元素对应的最小值。

1.3.3 BP 神经网络(Back propagation neural network, BPNN) BPNN 是一种非线性模型,它以数据驱动的方式进行建模和预测,可以避免人为设置权重的影响^[20],具有强大的拟合能力,能够处理多变量输入以及非参数建模等优点。土壤中重金属元素含量通常受多个因素的影响,如土壤性质和人为活动等,这些因素之间的关系是非线性的,BPNN 可以同时将它们作为输入变量,从而更全面地预测土壤中重金属含量。本研究分别构建 GA-EN 和 GA-PN 两种不同结构和功能的 BPNN。

1) GA-EN 评价网络。GA-EN 的作用是在遗传算法每一轮产生新个体后计算新个体 RMSE,决定是否保留。合理设计 BPNN 结构可以使其具有较好的泛化能力,该神经网络结构(图 3)为:输入层和输出层各 11 个神经元,代表采样点的 11 个特征值;2 个隐含层采用 sigmoid 函数;输出层采用的是正线性激活函数 Poslin,在输入值不小于 0 时,输出值与输入

相同,在输入值小于0时,输出0。在模型中,先将GA产生的新个体添加到数据集中,形成新的数据集;

然后训练GA-EN,用GA-EN进行预测,得到预测结果;计算RMSE并根据其大小取舍新个体。

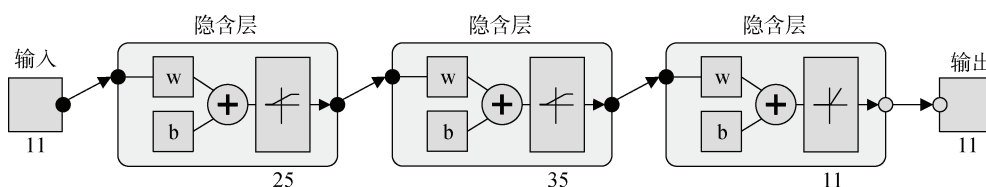


图3 GA-EN网络结构

Fig. 3 GA-EN network architecture

2)GA-PN 预测网络。在训练集中,GA-PN 模型采用 k 折交叉验证法来选择神经网络的参数,本研究采用了8折交叉验证($k=8$)。通过这种方式,可以更准确地评估模型的性能,选择较好的参数组合。GA-PN 模型基于前期数据扩增结果构建,用于预测

研究区内其他位置土壤pH和重金属元素含量。该神经网络结构(图4)为:输入层3个神经元,分别表示待预测点位的经度、纬度、高程;3个隐含层采用sigmoid函数;输出层1个神经元,表示预测点的某一种元素含量的预测值。

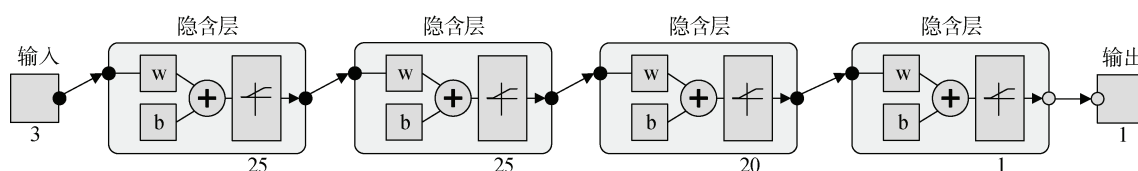


图4 GA-PN网络结构

Fig. 4 GA-PN network architecture

2 结果与讨论

2.1 基于GABP模型的预测结果

2.1.1 K-均值聚类结果 随着种群数量逐渐增加,观察到3个指标的变化逐渐减小或趋于稳定(图5)。得出以下结论:当 $k=5$,①种群内标准差逐渐减小;②种群间标准差逐渐增大;③种群内平均个

体数量趋于稳定。因此,可以认为当 $k=5$ 时,聚类效果最佳。

2.1.2 数据扩增与GA-EN评价 在初始样本数据构成的5个种群内分别使用GA进化繁衍,再对新个体用GA-EN进行评价。最终,样本点的数量从24个扩增到5019个,满足BPNN需要大量样本数据的要求。

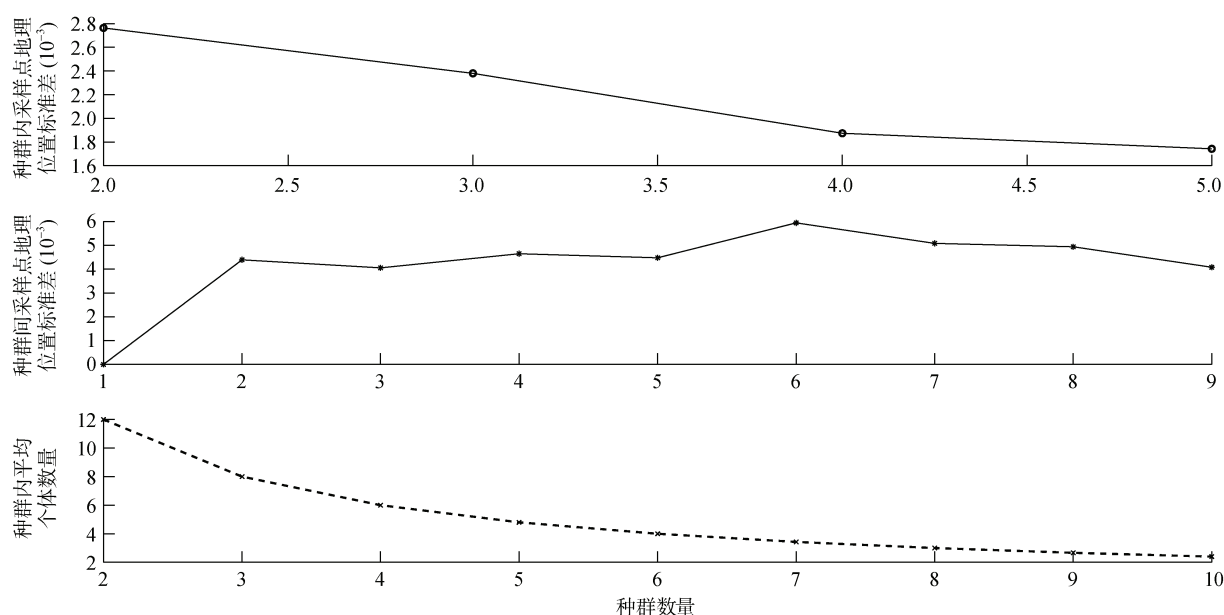


图5 K-均值聚类指标折线

Fig. 5 Line plot of K-means clustering metrics

2.1.3 GABP 模型预测 GABP 模型对 pH 和 7 种重金属的含量及分布进行了预测,模型的均方根误差 RMSE、决定系数 R^2 、相关系数 r 和平均绝对误差 MAE 见表 2。

表 2 三种模型精度评价
Table 2 Accuracy evaluation of three models

参数	模型	pH	Cd	Pb	Cr	Cu	Ni	Hg	As
R^2	GABP	0.219 5	0.156 0	0.390 2	0.295 8	0.269 3	0.114 3	0.554 5	0.133 7
	IDW	0.095 2	0.025 8	0.006 1	0.350 3	0.101 4	0.070 9	0.033 5	0.021 0
	BP	0.011 9	0.059 8	0.012 5	0.315 4	0.318 4	0.003 6	0.105 2	0.005 7
r	GABP	0.468 5	0.395 0	0.624 7	0.543 8	0.518 9	0.338 1	0.744 6	0.365 7
	IDW	0.308 6	0.160 8	-0.078 4	0.591 9	0.318 5	-0.266 3	0.183 0	0.144 8
	BP	0.109 2	0.244 4	-0.111 7	0.561 6	0.564 3	0.059 7	0.324 5	-0.075 4
RMSE	GABP	0.982 1	1.217 0	180.323 8	80.992 0	44.594 9	20.461 8	5.920 2	551.183 2
	IDW	0.879 6	0.965 1	253.953 5	75.876 0	48.515 8	23.494 8	9.010 1	628.079 1
	BP	1.345 6	1.266 7	345.283 8	95.081 1	45.687 0	26.114 3	8.626 0	927.215 1
MAE	GABP	0.743 7	0.904 7	124.707 1	56.204 6	33.418 1	13.614 5	3.661 0	299.185 7
	IDW	0.692 4	0.772 3	178.516 4	60.605 1	38.102 9	15.888 8	4.523 7	377.666 2
	BP	1.005 7	0.796 7	239.482 2	67.683 6	31.272 9	17.708 6	4.070 4	501.547 0

表 2 为 3 种模型交叉验证的参数值。结果显示 GABP 模型在预测 pH 和 7 种土壤重金属元素含量方面表现出了相对较好的精度和拟合效果。按照 R^2 从大到小进行排序,可以看出: Hg 最高,其余依次是 Pb、Cr、Cu、pH、Cd、As 和 Ni。GABP 相较于 BP 网络和 IDW 而言,除 Cr 外,对所有特征预测的 R^2 和 r 均有明显提升,且 RMSE 和 MAE 较小,表明较少样本经过 GA 数据扩增后,训练出的模型能够较好地解释数据变异,预测结果与实际值拟合程度较好,准确度明显提升。

2.2 GABP 模型与 IDW 预测方法的对比

2.2.1 两种方法空间分布预测结果对比 使用 ArcGIS 软件将预测值和矢量网格单元进行关联,再赋值和显示后,得到 GABP 空间分布预测结果,并将其与 IDW 空间插值结果进行对比。图 6 为 GABP 和 IDW 预测结果。

图 6 表明, GABP 复合模型预测结果与 IDW 插值结果在空间分布上总体上保持一致。特别是 Hg、Cu 和 pH 的空间分布与 IDW 结果非常接近,高值区和低值区的分布位置基本一致。其中,两种方法预测的 Cd、Pb、As 元素的高值区都分布在采矿用地的西南部,原因为研究区东矿区为硫铁矿,西矿区为废弃金矿,研究区外西南部存在一处金矿。而硫铁矿和金矿中都可能含有一定量的 Cd、Pb、As,在开采加工过程中可能导致其释放到土壤中,因此这些区域出现高值。Cr 元素高值区主要分布在东矿区南部和研究区东部。Ni 元素的高值区主要在研究区的西南部,

东部有一小范围高值区。Cu 元素从西部的高值区开始沿着东矿区逐渐降低直至研究区东北部。硫铁矿的形成除了与 S、Fe 成矿元素供给与沉积密切相关外,通常还会伴生如 Cu 等元素,矿山开采容易导致这些元素在土壤中累积。Hg 元素含量的高值区域都显示为一部分从研究区中部沿西南方向延伸,另一部分在研究区东部,两种方法预测的 Hg 最大值最小值区间完全一致,预测效果最佳,与精度评价的结果相同。两种方法预测 pH 的空间分布也基本一致, GABP 预测的结果相对更连续,而 IDW 预测结果显示研究区内重金属含量比较复杂,高值和低值区域相互交错,细节表现更好。

2.2.2 GABP 模型优越性 GABP 模型相较于传统的 BPNN 和常规 IDW 法具有以下优势:

1)利用 GA 进行数据扩增,大幅增加了样本数据量,提高了模型的泛化能力,改善了模型的稳定性,减少了误差。从交叉验证的结果可以看出, GABP 模型的 R^2 和 r 明显高于其他两种方法, RMSE 和 MAE 值也相对较低,表明其优于其他两种方法。

2)在 GA 内部设置新个体特征值的上下限来确保不会出现异常值,并在 BPNN 输出层采用 Poslin 函数,确保插值结果不会出现负值。从图 6 可以看出, GABP 预测的每一特征值均在 IDW 预测区间内,解决了一些插值方法边界可能出现负值的情况。

3)模型不仅考虑了平面地理位置,还综合考虑了高程因素,提高了非线性预测能力,使得结果更加准确。

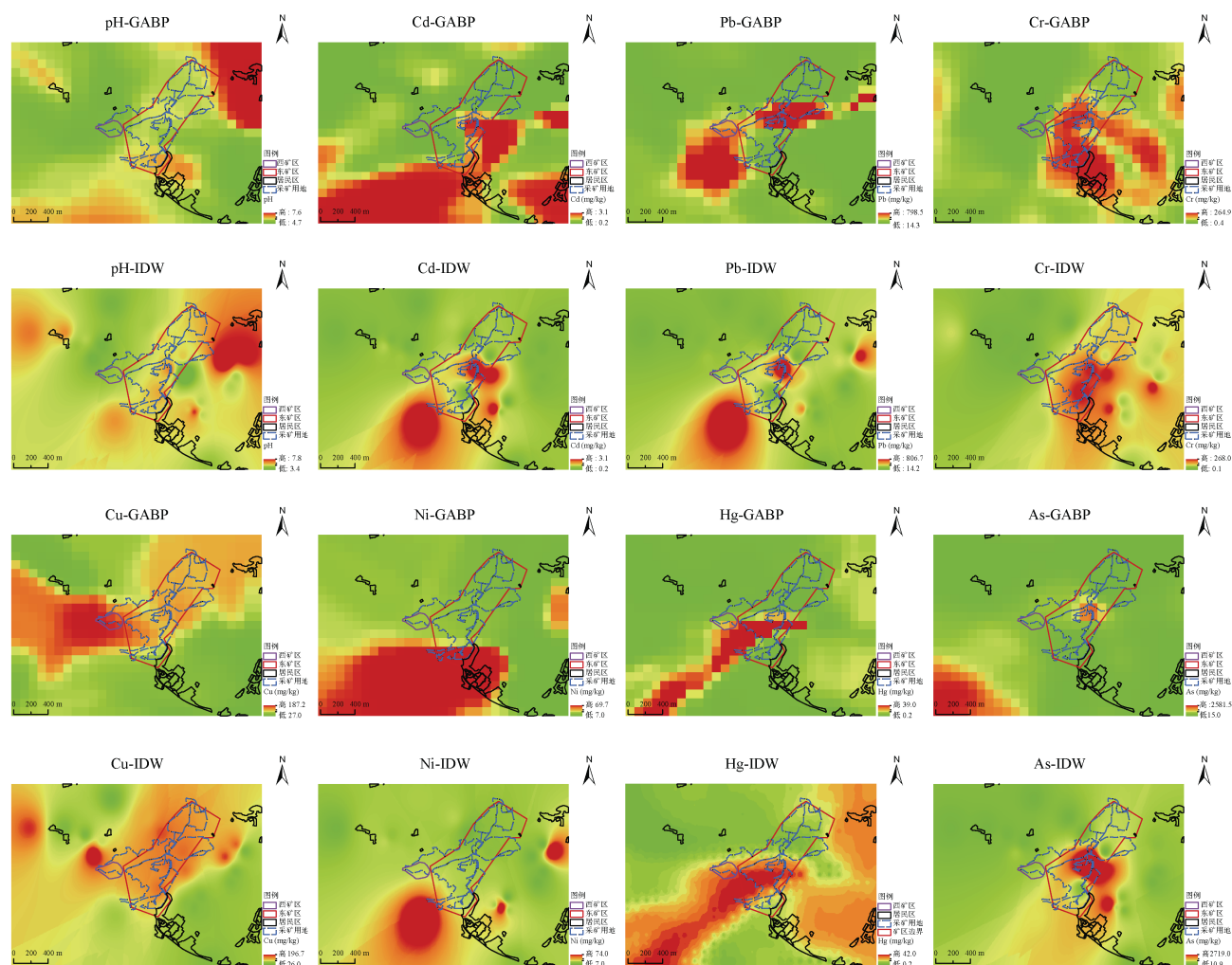


图6 GABP与IDW预测结果
Fig. 6 GABP and IDW prediction results

4)从图6中看出,在预测极大值与极小值时,GABP虽然没有IDW边缘光滑,但其不会有明显的牛眼效应。

3 结论

1)研究区土壤pH和7种重金属的分布趋势基本一致,高值区主要出现在研究区的西南部和东部,呈现出显著的空间分异性。

2)相较于传统的BPNN和IDW,GABP具有较高的决定系数和较低的误差,表明该模型弥补了传统方法需要大量实测样本数据支撑训练的缺点。与IDW法相比,模型的预测结果与其基本一致,说明GABP模型是可靠的。

3)模型不仅考虑了经纬度,还考虑了高程因素,模型对pH和7种重金属含量具有较好的预测结果,表明该模型在小尺度区域土壤重金属含量空间分布预测是可行的。

本文的研究方法对改进传统的空间预测方法和土壤重金属、水分等预测提供了新思路,对土壤环境评价具有参考意义。本文在预测部分考虑了地理位置和高程属性,但土壤重金属元素可能还受元素间伴生性、土地利用类型等因素的影响。因此,后续研究可以综合考虑这些因素,对输入特征进行优化。

致谢:感谢广东工业大学尹光彩老师为本研究提供的帮助!

参考文献:

- [1] Zhang Y M, Li S, Chen Z, et al. A systemic ecological risk assessment based on spatial distribution and source apportionment in the abandoned lead acid battery plant zone, China[J]. Journal of Hazardous Materials, 2018, 354: 170-179.
- [2] Liu M, Zhang A B, Liao Y J, et al. The environment quality of heavy metals in sediments from the central Bohai Sea[J].

- Marine Pollution Bulletin, 2015, 100(1): 534–543.
- [3] 成晓梦, 孙彬彬, 吴超, 等. 浙中典型硫铁矿区农田土壤重金属含量特征及健康风险[J]. 环境科学, 2022, 43(1): 442–453.
- [4] 汤国安, 杨昕. ArcGIS 地理信息系统空间分析实验教程[M]. 2 版. 北京: 科学出版社, 2012.
- [5] 曾星, 张华. 基于 GIS 的碧流河上游农田土壤性状空间分布特征[J]. 土壤, 2023, 55(6): 1354–1362.
- [6] 张龙, 张忠启, 蔡何青, 等. 贵州毕节植烟区土壤 pH 空间分布特征及对施肥的影响[J]. 土壤, 2023, 55(1): 85–93.
- [7] 尹光彩, 陶琳, 陈幸玲, 等. 一种耦合遗传算法和神经网络的土壤污染物含量插值方法: CN113159219B[P]. 2022-04-08.
- [8] 何云山. 区域土壤重金属污染预测模型研究与应用[D]. 北京: 北京信息科技大学, 2021.
- [9] 胡大伟, 卞新民, 许泉. 基于 ANN 的土壤重金属分布和污染评价研究[J]. 长江流域资源与环境, 2006, 15(4): 475–479.
- [10] 谢虹. 农田土壤重金属污染评价常用研究方法[J]. 山东化工, 2018, 47(10): 194–195, 200.
- [11] 陈晓燕, 范成五, 瞿飞, 等. 土壤重金属污染评价方法概述[J]. 浙江农业科学, 2017, 58(10): 1801–1804, 1810.
- [12] Wang Y L, Zhao J H, Guo Z W, et al. Soil moisture inversion based on data augmentation method using multi-source remote sensing data[J]. Remote Sensing, 2023, 15(7): 1899.
- [13] 尹光彩, 宋小旺, 陶琳, 等. 一种基于遗传算法的土壤重金属含量空间插值方法: CN111209679B[P]. 2023.09.29.
- [14] Yin G C, Chen X L, Zhu H H, et al. A novel interpolation method to predict soil heavy metals based on a genetic algorithm and neural network model[J]. The Science of the Total Environment, 2022, 825: 153948.
- [15] 戴倩倩, 徐梦洁, 庄舜尧, 等. 基于地理探测器的封丘县农田土壤重金属分布影响因素研究[J]. 土壤, 2022, 54(3): 564–571.
- [16] 龚仓, 王顺祥, 陆海川, 等. 基于地理探测器的土壤重金属空间分异及其影响因素分析研究进展[J]. 环境科学, 2023, 44(5): 2799–2816.
- [17] 谢芳, 黄海英. 重污染地区土壤调查点位布设方法[J]. 科技资讯, 2010, 8(23): 164.
- [18] 王政. 基于贝叶斯地统计学方法的区域土壤重金属空间预测研究[D]. 济南: 山东师范大学, 2022.
- [19] 曹文琪, 张聪. 深度复合模型下的土壤重金属含量预测[J]. 计算机工程与设计, 2021, 42(4): 1128–1134.
- [20] 李向, 管涛, 徐清. 基于 BP 神经网络的土壤重金属污染评价方法——以包头土壤环境质量评价为例[J]. 中国农学通报, 2012, 28(2): 250–256.