季佳运, 肖霄, 杨品璐, 等. GA-BP 神经网络在精准刻画场地地下水污染物扩散范围的应用研究[J]. 岩矿测试, 2025, 44(3): 406−419. DOI: 10.15898/j.ykcs.202409280204.
JI Jiayun, XIAO Xiao, YANG Pinlu, et al. Application of a GA-BP Neural Network in Accurately Characterizing the Diffusion Range of Groundwater Pollutants[J]. Rock and Mineral Analysis, 2025, 44(3): 406−419. DOI: 10.15898/j.ykcs.202409280204.

GA-BP 神经网络在精准刻画场地地下水污染物扩散范围的应用研究

季佳运[1,2]，肖霄[1]，杨品璐[3]，刘洋[1]，周亚红[1,4,5*]
(1. 河北地质大学水资源与环境学院，河北 石家庄 050031；
2. 河北地质大学城市地质与工程学院，河北 石家庄 050031；
3. 唐山市自来水有限公司，河北 唐山 063000；
4. 河北省水资源可持续利用与产业结构优化协同创新中心，河北 石家庄 050031；
5. 河北省水资源可持续利用与开发重点实验室，河北 石家庄 050031)

摘要：自 2021 年最新生态环境损害鉴定评估指南发布实施以来，对地下水中污染物 (如铬、铅、铁、锰等污染物) 的扩散范围刻画的精度要求越来越高。受研究区场地条件限制，采样点无法完全分布均匀，现有插值方法难以解决采样点分布不均而导致扩散范围刻画不准确的问题。本文通过 ArcGIS 空间插值图展示某化工园区地下水溶质的空间分布，发现 Mn^{2+}离子分布与其形成机制规律相差较大，且尝试使用 GIS 多种插值方法 (如克里金法、反距离权重法、样条函数等插值方法) 效果均不理想，其扩散方向与研究区地下水流向及形成机理不符，可能是由于其监测点位分布不均。因此以重金属 Mn^{2+}为例，使用 GA-BP 神经网络与标准 BP 神经网络对园区各点位 Mn^{2+}浓度进行回归预测，建立其浓度与空间分布的神经网络模型，选取拟合程度较好的神经网络模型对监测点位缺失区域进行浓度预测，并结合空间插值圈定化工园区中心 Mn^{2+}的扩散范围，同时用 Mn^{2+}的产生机制对扩散范围进行验证。结果表明：GA-BP 神经网络的 Mn^{2+}浓度预测效果最好，使用其补充监测点缺失位置的 Mn^{2+}浓度并重新绘制 Mn^{2+}浓度分布图，新 Mn^{2+}分布图显示化工园区中心 Mn^{2+}扩散范围为 $1.70\times10^{6}m^{2}$，超出化工园区面积为 $2.13\times10^{5}m^{2}$。与优化前的扩散范围相比，校正后的扩散范围符合 Mn^{2+}产生和运移规律。GA-BP 神经网络对场地地下水污染物扩散范围的精确圈定有较好的辅助效果，可为环境污染评估提供更加科学有效的方法支持。

关键词：地下水；化工园区；GA-BP 神经网络；扩散范围

要点：

(1) 研究区地下水 Mn^{2+}浓度分布不符合其 Mn^{2+}产生及运移规律，需要进行校正。

(2) GA-BP 神经网络预测补全采样点分布较少区域的 Mn^{2+}浓度，校正 Mn^{2+}分布。

(3) 经验证，校正后 Mn^{2+}浓度分布符合场地 Mn^{2+}产生机制及地下水动力学条件。

中图分类号：X523 **文献标识码：**A

收稿日期：2024−09−28；**修回日期：**2025−01−03；**接受日期：**2025−01−10；**网络出版日期：**2025−02−22
基金项目：河北省省级科技计划项目 (236Z4204G)；河北省自然科学基金项目 (D2022403016)；河北省教育厅科学研究项目 (ZD2022119)；河北地质大学第二十届学生科研项目 (KAG202402)
第一作者：季佳运，硕士，主要从事污染场地自然衰减方向的研究。E-mail：17805215951@163.com。
通信作者：周亚红，博士，副教授，主要从事地下水污染及治理研究。E-mail：zhyh327@163.com。

清洁安全的地下水是社会和人类可持续发展的基本保障。世界上大多数城市都面临着地下水枯竭和水质恶化导致的双重水资源压力[1]。随着中国城市化和工业化的不断推进, 工业场地地下水污染问题较为严重[2-3]。且随着中国化工企业的发展, 有机物及其降解产生的 Mn^{2+}二次污染物对人体的危害也逐渐受到重视[4]。为保护人体健康和环境的安全, 需要精细刻画出地下水污染物 (如铬、铅、铁、锰等污染物) 扩散范围[5], 并将二次污染纳入损害评估体系[6], 这对控制地下水污染十分重要[7]。

《生态环境损害鉴定评估技术指南 环境要素第 1 部分: 土壤和地下水》(GB/T 39792.1—2020) 指出, 损害鉴定需要明确地下水的当前损害范围和评估时间范围内的可能损害范围[8], 计算可能受损的地下水面积[9], 以便更好地评估环境损害。但研究者采样时, 受研究区条件限制, 采样点常常无法均匀分布。时而发生现有 ArcGIS 插值方法的插值结果与真实情况有较大差别的情况。王兴等[10]分析了不同插值方法对海水盐度插值适用性, 并指出无论采用哪种空间插值方法, 都会引入较大的误差, 适度加密监测站位是必要的, 尤其是在插值误差较大的区域。Mohammed 等[11]指出适当使用机器学习方法对插值精度有一定提升, BP 神经网络用于拟合时间较早, 是深度学习的一个关键分支, 它能拟合变量之间的非线性关系, 多名研究者在水文地质方面将其用于浓度预测, 方法较为成熟[12]。

近年来, 优化 BP 神经网络的算法不断迭代[13], 经过验证, 遗传算法优化的 BP 神经网络优化效果较好。且由于研究区监测点位分布不均, Mn^{2+}离子浓度分布与其产生规律及水动力条件不符, 研究区地下水 Mn^{2+}浓度受多因素影响, 采样点 Mn^{2+}浓度和极角与极径有较强的非线性关系, 因此本文采用遗传算法优化的 BP 神经网络 (GA-BP) 来生成缺失点位的 Mn^{2+}浓度数据, 从而对其扩散范围进行校正, 减少采样点分布不均对扩散范围圈定的影响。同时, 通过 Mn^{2+}成因机制结合其分布特征, 验证校正后的 Mn^{2+}扩散范围[14]。校正后的 Mn^{2+}浓度符合场地水动力条件及 Mn^{2+}产生规律, 校正效果较好, 相较优化前刻画的扩散范围更为准确。

1　研究区概况

研究区为河北省某工业园区, 场地面积约 5.25km^2, 主要生产化学原料、化学制品、医药、橡胶和纸制品。该园区周边皆为农田与果园, 以旱地农作物为主, 主要种植小麦、玉米、鸭梨、葡萄和山楂等。地貌单元为滹沱河冲洪积扇水文地质亚区。第四系发育较为齐全, 地层由新至老分为全新统 (Q4)、上更新统 (Q3)、中更新统 (Q2) 和下更新统 (Q1)。岩性由黄土状粉质黏土和冲积粉土组成, 并夹有砂层, 厚度约 130m 左右。研究区年平均降水量 456mm, 主要集中在夏季。浅层地下水位埋深总趋势是自西向东、由北往南由浅变深, 埋深由北部 30 ~ 35m 变至南部 50 ~ 60m, 地下水位年内变幅一般为 2m 左右。场地地下水排泄途径主要为人工开采, 其次是侧向流出和潜水蒸发, 其主要补给来源为大气降水与河渠渗漏。场地内企业产生的污染物均经过污水处理厂等设施处理后排放。

2　研究方法

2.1　数据收集

为了探究该化工园区地下水系统有机物对锰元素迁移规律的影响, 本研究结合场地水化学指标分布特征, 针对地下水中有机污染区域进行地下水采集, 共采集 12 组地下水水样。水样主要取自监测井, 依据园区地下水水动力条件以及前期获取的地下水污染源分布情况确定地下水取样深度, 本研究采集深度为地下水水位线 0.5m 以下。地下水水样的采样、密封、保存以及检测均按照《地下水环境监测技术规范》要求进行。本次地下水检测指标包括锰、硝酸盐 (以 N 计)、亚硝酸盐 (以 N 计)、苯酚、丙酮、硝基苯、4-硝基苯胺、石油烃 (C_{10} ~ C_{40})、石油类、甲苯、乙苯、间二甲苯+对二甲苯、邻二甲苯。参照“35+N”的原则确定化工园区地下水检测指标, “35”为《地下水质量标准》(GB/T 14848—2017) 扣除微生物指标和放射性指标后的 35 项常规指标, “N”为场地特征污染物指标。具体测试方法列于表 1, 在每个平行样品采样点位采集 3 份平行样品, 其中 2 份地下水测试样品送河北省地质实验测试中心, 另 1 份地下水质控测试样品送华测检测认证集团股份有限公司。所得数据均经过了准确度检验。

2.2　BP 神经网络模型构建

BP 神经网络是一类多层的前馈神经网络, 最早由 Rumelhart 等在 1986 年提出, 结构精简, 可调参数多, 训练算法众多, 操作性优良, 因此, BP 神经网络应用范围广泛, 接近 90% 的神经网络模型均为 BP 神经网络或其改进版本。BP 网络是前向网络的核心部分, 是神经网络的精华部分[15]。然而, 它也存在网络结构难以确定、学习收敛速度慢、不能够保证

收敛至全局最小点等缺点。除此之外，网络训练在很大程度上受网络结构、初始连接权值和阈值选择影响，因此本研究使用遗传算法对神经网络进行优化[16]。BP 神经网络的遗传算法流程图如图 1 所示。

2.2.1　模型构建及主要参数设置

以污染源中心点作为极坐标的原点，以东方向为极轴。将极角和极径为输入值，Mn^{2+}浓度设置为输出值。设置输入层的节点数为 i，隐含层节点数为 j，输出层节点数为 k，输入层到隐含层的权值为 w_{ij}，隐含层到输出层的权值为 w_{jk}，输入层到隐含层的阈值为 b_j[17]，隐含层到输出层的阈值为 c_k，学习速率为 η。其中 q 为 0~10 之间的任意数，其中 k 为期望输出值。

表 1　地下水样品分析测试方法

Table 1　Analysis methods for groundwater sample.

检测项目 Analytical items	分析测试方法 Analysis methods	方法检出限 Method detection limit
硝酸盐 Nitrate	离子色谱法 Ion chromatography	0.004mg/L
锰 Manganese	电感耦合等离子体质谱法 Inductively coupled plasma mass spectrometer	0.00012mg/L
石油类 Oil	红外分光光度法 Infrared spectrophotometry	0.06mg/L
石油烃 (C_{10} ~ C_{40}) Petroleum hydrocarbons (C_{10} ~ C_{40})	气相色谱法 Gas chromatography	0.004mg/L
丙酮 Acetone	气相色谱法 Gas chromatography	0.2mg/L
4-硝基苯胺 4-Nitroaniline	气相色谱-质谱法 Gas chromatography-mass spectrometry	4.6μg/L
苯酚 Phenol	气相色谱-质谱法 Gas chromatography-mass spectrometry	0.1μg/L
甲苯 Toluene	气相色谱-质谱法 Gas chromatography-mass spectrometry	0.3μg/L
乙苯 Ethylbenzene	气相色谱-质谱法 Gas chromatography-mass spectrometry	1.2μg/L
间二甲苯+对二甲苯 *m*-Xylene+*p*-Xylene	气相色谱-质谱法 Gas chromatography-mass spectrometry	1.2μg/L
邻二甲苯 *o*-Xylene	气相色谱-质谱法 Gas chromatography-mass spectrometry	1.2μg/L

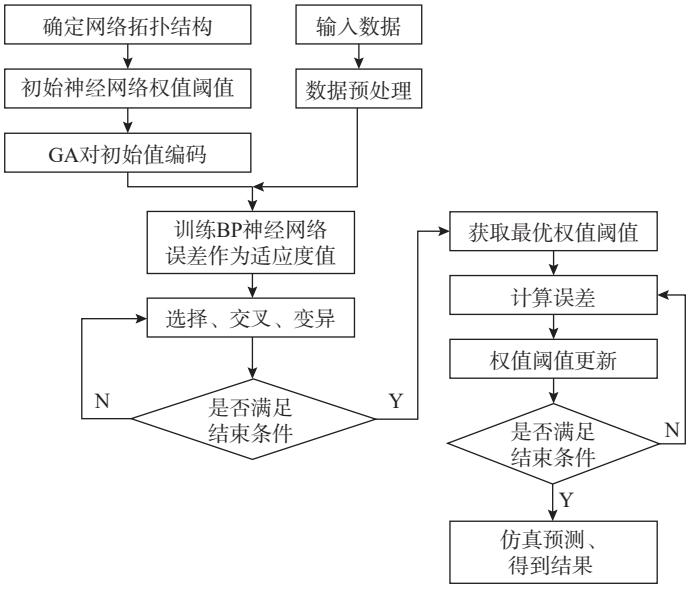

图1　GA-BP 神经网络拓扑图

Fig. 1　Topology diagram of GA-BP neural network.

隐含层节点数:

$$j = \sqrt{i+k}+q \tag{1}$$

隐含层输出:

$$a_j = f\left(\sum_{i=1}^{l} w_{ij}x_i + b_j\right) \tag{2}$$

输出层输出:

$$y_k = f\left(\sum_{j=1}^{l} w_{jk}a_j + c_k\right) \tag{3}$$

误差计算:

$$E = \frac{1}{2}\sum_{k=1}^{m}(Y_k - y_k)^2 \tag{4}$$

2.2.2　遗传算法优化 BP 神经网络

遗传算法存在共享机制, 整个种群具有一致的向最优值移动速度, GA-BP 神经网络的算法流程如下[18]。

步骤一: 对数据进行预处理, 确定 BP 神经网络输入层、隐含层以及输出层的节点数, 初始化权值及阈值。

步骤二: 设置初始种群规模为 200, 最大进化代数为 100, 交叉概率为 0.8, 变异概率为 0.2。遗传算法通过计算适应度函数, 执行选择、交叉、变异运算优化权值及阈值[19], 当父代个体的适应度满足终止条件或进化代数大于设定代数得到最优解时输出结果。

步骤三: 使用遗传算法得到的最佳权值和阈值, 训练 BP 神经网络, 并输出最终结果。

2.2.3　选择点位预测并进行插值

第一步: 选取与监测点距离相近点位, 对其 Mn^{2+}浓度进行预测: 对于训练与测试数据相近的点, 预测数值精度更高。

第二步: 选取能与实测点位一起均匀分布在园区及其附近的点位, 对其 Mn^{2+}浓度进行预测: 点位分布均匀, 扩散范围刻画得越为准确。

第三步: 将这些点位与实测点位 Mn^{2+}数据一起进行插值, 对扩散范围进行刻画。

3　结果与讨论

3.1　实测点位 Mn^{2+}污染晕及研究区流场

采用反距离权重插值法绘制场地各监测点 Mn^{2+}浓度空间分布, 如图 2 所示。该化工园区地下水中 Mn^{2+}主要呈现出局部性点状富集特征, 化工园区中部偏西区域的 Mn^{2+}浓度最高, 为 0.74mg/L, 整体上 Mn^{2+}浓度自西向东呈降低趋势。但由于采样点分布不均, 扩散范围刻画的精确度较低, M07 与 M12 点之间扩散范围向西北方向延伸, 而水流方向为西北向东南, 与 Mn^{2+}扩散机制不符。其扩散方向与地下水流向相反, 且 Mn^{2+}的产生与地下水中的有机物有关, 在还原条件下, NO_3^-、Fe^{3+}和 Mn^{4+}先后与有机物反应, 生成 Mn^{2+}。地下水中有机物越多, 被还原的离子越多。研究区监测数据见表 2。

3.2　有机物污染物的空间分布特征

根据不同监测井检出的 10 种有机污染物的浓度, 绘制出极坐标热力图 (图 3), 可以看出, 研究区丙酮、4-硝基苯胺、石油烃 (C_{10} ~ C_{40})、石油类在地下水

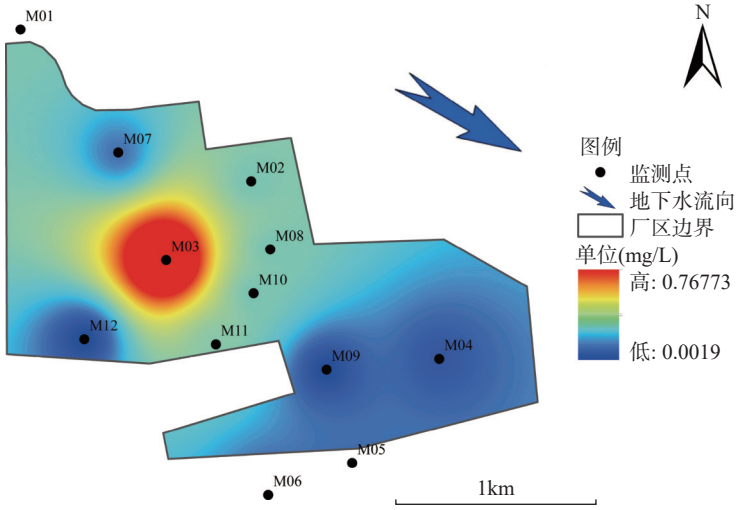

图2　Mn^{2+}浓度空间分布图

Fig. 2　Spatial distribution map of Mn^{2+} concentration.

表 2　研究区监测数据

Table 2　Monitoring data of the study area.

点位 Point positions	锰 Manganese (mg/L)	苯酚 Phenol (μg/L)	丙酮 Acetone (μg/L)	硝基苯 Nitrobenzene (μg/L)	4-硝基苯胺 4-Nitroaniline (μg/L)	石油烃 (C_{10} ~ C_{40}) Petroleum hydrocarbons (C_{10}-C_{40}) (mg/L)	石油类 Oil (mg/L)	甲苯 Toluene (μg/L)	乙苯 Ethylbenzene (μg/L)	间-二甲苯+对-二甲苯 *m*-Xylene+ *p*-Xylene (μg/L)	邻-二甲苯 *o*-Xylene (μg/L)
M01	0.184	0.3	—	—	—	0.24	0.45	—	—	—	—
M02	0.18	0.2	—	—	—	2.37	0.29	2.4	—	—	—
M03	0.768	0.1	—	—	—	5.46	0.98	—	0.4	1.3	0.6
M04	0.031	1.1	—	0.00864	39.7	34.8	1.64	18.2	0.4	1.2	0.4
M05	0.0658	—	—	0.00552	6.3	0.62	0.11	2.3	—	—	—
M06	0.0452	—	—	—	—	0.14	0.11	—	—	—	—
M07	0.0752	—	—	—	—	2.78	0.17	1	—	—	—
M08	0.182	0.4	—	—	—	23.5	2.52	—	—	—	—
M09	0.0024	—	—	—	—	0.55	0.17	—	—	—	—
M10	0.18	0.5	—	—	—	27.6	2.95	—	—	—	—
M11	0.194	—	58	—	—	8.95	8.98	—	—	—	—
M12	0.0018	—	—	—	—	0.02	0.03	—	—	—	—

注：“—”表示未检出。

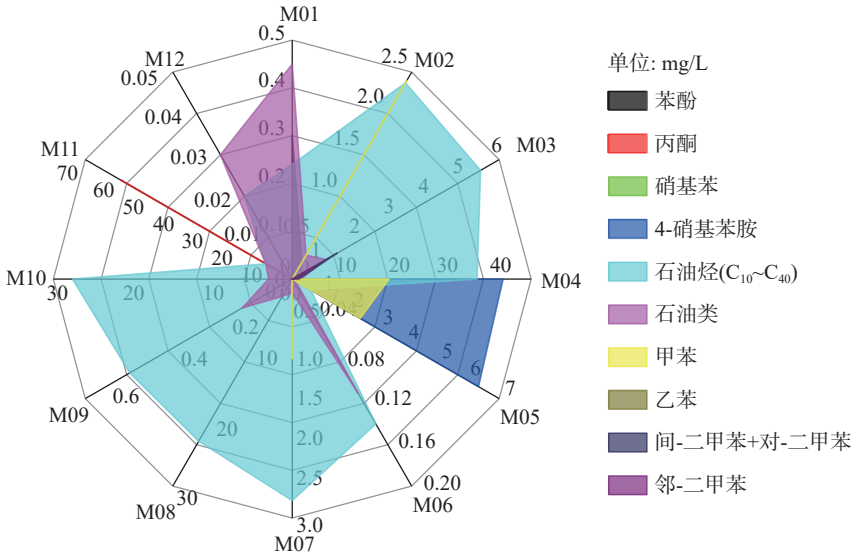

图3　各监测井有机指标极坐标热力图

Fig. 3　Polar heatmap of organic indicators in each monitoring well.

中浓度相对较大，分布点位相对较多，而硝基苯仅在 M11 监测井含量显著，甲苯、乙苯仅在监测井 M04 含量较高，其余点位含量较少，苯酚、间二甲苯+对二甲苯、邻二甲苯在各检测点位含量均较少。对 4 种典型有机指标 [丙酮、4-硝基苯胺、石油烃 (C_{10} ~ C_{40})、石油类] 进行空间插值 (图 4)。

丙酮、石油类浓度总体空间分布相似，4-硝基苯胺、石油烃 (C_{10} ~ C_{40}) 浓度空间分布相似 (图 4)。丙酮、石油类空间分布不均匀，高值区主要分布在场地中下部位置的 M11 监测井，最高浓度分别为 57.98mg/L、9mg/L，在横向上呈现点源污染特性，随着与污染源距离增加，丙酮与石油类浓度均降低，但丙酮浓度降低速率比石油类浓度大得多。4-硝基苯胺、石油烃 (C_{10} ~ C_{40}) 在厂区范围内分布不均，浓度呈现由东南部向西北部逐渐减少的特征，浓度最高的监测井为 M04，最高浓度分别为 30.58mg/L、25.82mg/L。

3.3　Mn^{2+}扩散范围圈定

3.3.1　神经网络训练结果

为确定标准 BP 神经网络最佳隐含层节点数，利用经验公式确定标准 BP 神经网络隐含层节点的数

量范围[20], 通过均方根误差确定最佳隐含层节点数量为 6, 相应的均方根误差为 0.031883(表 3)。

同时训练遗传算法优化的 BP 神经网络, 其拟合效果见图 5。训练集 r=0.99740, R^2=0.99481, 测试集 r=0.99576, R^2=0.99154, R^2 均大于 0.9, 这证明训练集拟合良好, 具有较强的可靠性。GA-BP、BP 神经网络预测结果与误差见表 4。

根据表 5 可看出, 优化后的 BP 神经网络误差比没有优化的 BP 神经网络 mae、mse、rmse、mape 更小, 拟合效果更好。训练集与测试集拟合效果均较好, 且最近隐含点层数按照均方根误差获得, 训练集与测试集样本点种类分布均匀, 设计合理, 没有发生过拟合[21]。

实测值与预测值对比见图 6, GA-BP 实测值与预测值相较于标准 BP 神经网络更为拟合, 且 GA-BP 神经网络误差均接近于 0, GA-BP 神经网络预测效果明显优于 BP 神经网络。研究区数据点不多, 因此将数据分成了训练集与测试集两个集。三个测试集点分布在污染源中心的污染羽及下游位置, 且训练集具有污染源上游、污染羽以及污染源下游三种类型的点位, 训练的模型结果拟合较好。测试集结果验证通过后, 模型可靠[22]。

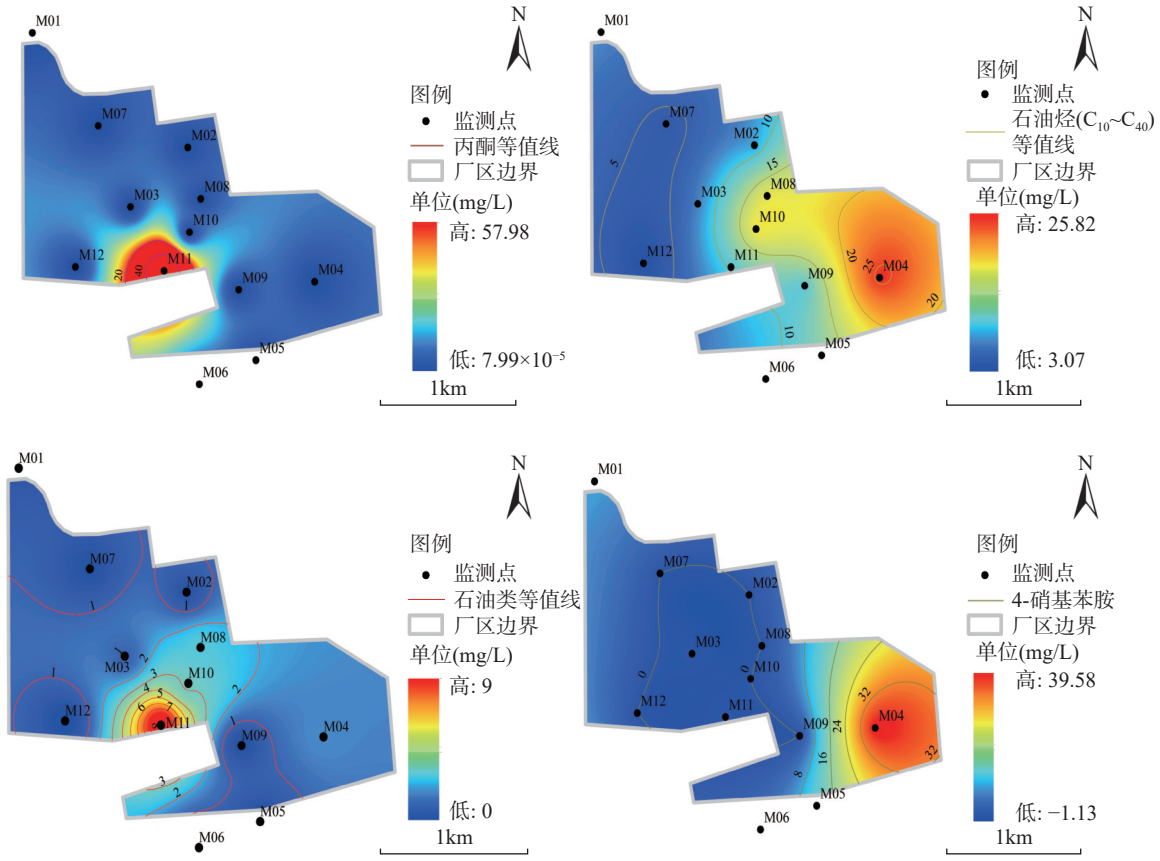

图4　四种典型有机指标空间分布图

Fig. 4　Spatial distribution maps of four typical organic indicators.

表 3　隐含层节点的确定过程

Table 3　Determination process of hidden layer nodes.

隐藏层节点 Hidden layer node	训练集的均方误差 Mean square error of training set	隐藏层节点 Hidden layer node	训练集的均方误差 Mean square error of training set
2	0.12601	8	0.67715
3	0.27314	9	0.32133
4	0.11288	10	0.2487
5	0.20273	11	0.06521
6	**0.031883**	12	2.0155
7	0.069991		

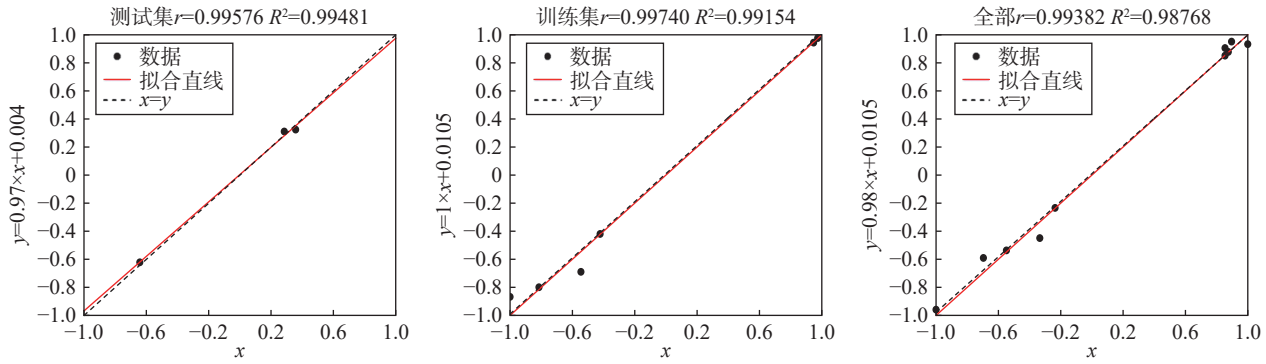

图5　神经网络拟合效果

Fig. 5　Fitting performance of neural network.

表 4　GA-BP、BP 神经网络预测结果与误差

Table 4　Prediction results and errors of GA-BP and BP neural network.

样本序号 Sample serial number	Mn^{2+}实测值 Mn^{2+} measured value (mg/L)	BP 预测值 (mg/L)	GA-BP 预测值 (mg/L)	BP 误差 (mg/L)	GA-BP 误差 (mg/L)
1	0.1800	−0.1366	0.1849	−0.3166	0.0049
2	0.1940	−0.1669	0.1876	−0.3609	−0.0064
3	0.0018	−0.1581	0.0057	−0.1599	0.0039

表 5　BP 神经网络误差

Table 5　Error of BP neural network.

BP 神经网络种类 Error in types of BP neural network	误差			
	mae	mse	rmse	mape
标准的 BP 神经网络模型 Standard BP neural network model	0.279	0.085	0.292	3082.097%
遗传算法优化的 BP 神经网络模型 Genetic algorithm optimized BP neural network	0.005	2.694×10^{6}	0.005	75.066%

选取与监测点距离相近或能够较为均匀地将园区 Mn^{2+}源中心紧密围绕的点位，对 Mn^{2+}浓度进行预测。结合实测数据，将化工园区中心 Mn^{2+}扩散范围精确刻画出来。预测点、监测点以及校正后 Mn^{2+}扩散范围图见图 7。

按照国家水质标准，Mn^{2+}浓度低于 0.05mg/L 即为一类水，因此本研究将地下水 Mn^{2+}浓度在 0.05mg/L 以下的区域视为没有受到 Mn^{2+}污染。由图 7 可知，园区 Mn^{2+}沿着水流方向从中心向东南方向迁移，距离中心 924m 处其浓度降至 0.05mg/L，园区中心 Mn^{2+}影响消失。将园区中心 Mn^{2+}扩散范围进行大致圈定，扩散范围约为 $1.7\times10^{6}m^{2}$，园区内扩散范围约为 $1.49\times10^{6}m^{2}$，园区外扩散范围约为 $2.13\times10^{5}m^{2}$。扩散范围延伸方向与水流方向一致。

3.3.2　现象机理验证

将精细刻画后的化工园区中心 Mn^{2+}污染范围与有机物指标分布进行对比，研究表明有机物在自然生物降解过程中，铁锰还原过程比硫酸盐过程发生更早[23]，有机物的降解过程中会增加土壤中锰的迁移性[24]，将原有的锰的氧化物转化成 Mn^{2+}，使

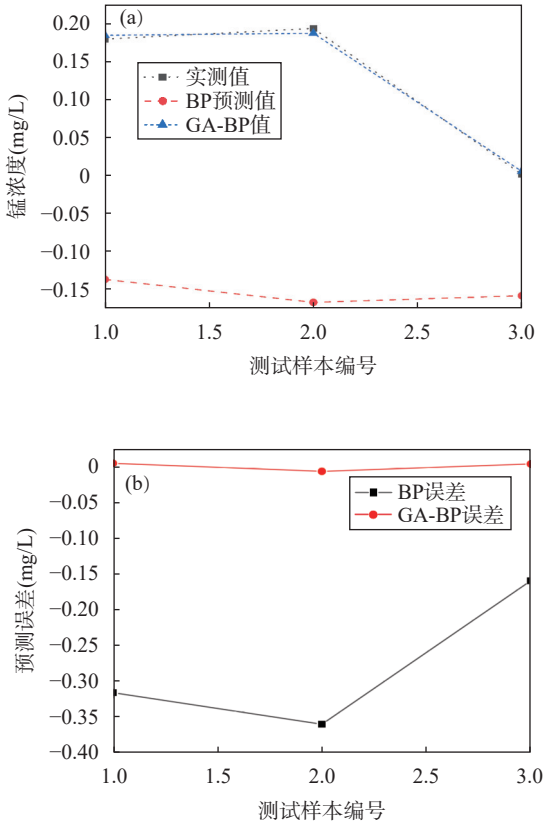

图6　预测效果对比：(a)锰浓度；(b)预测误差

Fig. 6　Comparison of prediction effect：(a) Concentration of Mn；(b) Prediction error.

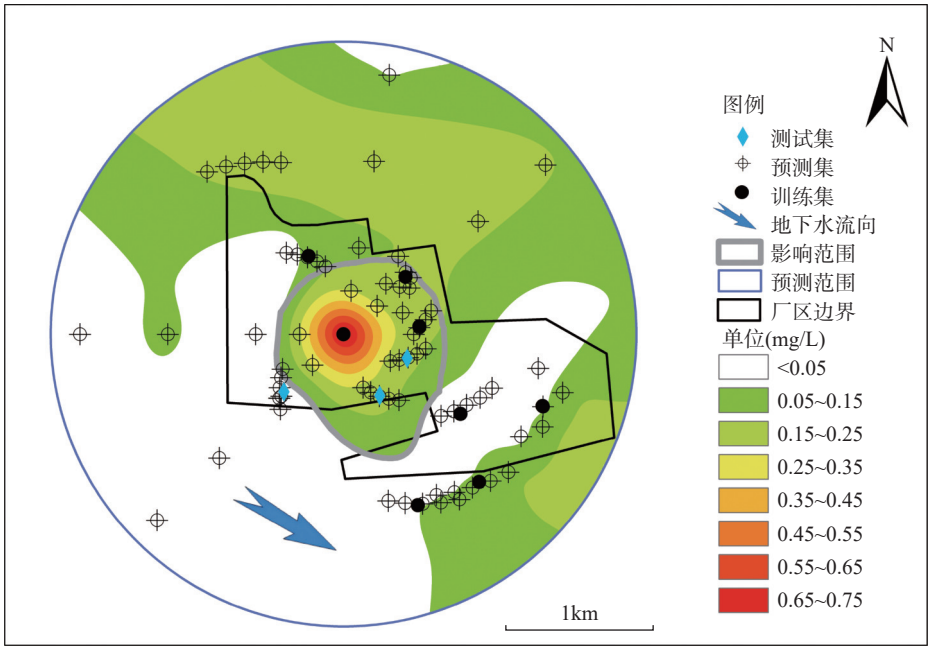

图7　校正后 Mn^{2+}点位分布图

Fig. 7　The distribution map of corrected Mn^{2+} point.

原有系统中土壤里的 Mn^{2+}减少, 水中 Mn^{2+}增加[25]。化工园区中心 Mn^{2+}扩散范围与石油类分布特征相一致, 即化工园区中心 Mn^{2+}与石油类扩散范围均分别向东北与东南方向延伸。石油类污染物从园区中心向东北方向延伸 M07 点位石油类浓度较低, 仅为 0.17mg/L, M03、M04、M08、M10、M11 等点均大于 1mg/L 或接近 1mg/L, M11 点浓度最大为 8.98mg/L。丙酮仅在 M11 处检出, 且此处也为有机物浓度最高点。丙酮与大部分石油类污染物相比, 分子量很小, 自然降解速度较快, 同时微生物降解能力受石油烃的物理状态影响显著[26], 溶解态、乳化态以及分散态是石油类在水中的主要存在形态, 其生物降解速率比大分子量的石油烃与 4-硝基苯胺大得多[27]。同时, 硝基具有的亲电子性、苯环结构具有的对称性[28], 这决定了 4-硝基苯胺的低生物降解能力[29]。因此, 4-硝基苯胺生物降解对地下水中锰元素的影响可以忽略不计。锰在自然条件下通常浓度较低, 厂区浓度较高可能是因为石油类有机物自然降解将岩层中的高价锰元素还原为低价态的 Mn^{2+}迁移至地下水中[29], 使得化工园区中心 Mn^{2+}扩散范围向东北方向延伸。石油类扩散范围与园区中心 Mn^{2+}扩散范围向东南方向延伸, 则是因为此区域地下水水流方向为自西北向东南[30]。Mn^{2+}扩散范围受石油类浓度分布与水流方向控制[31]。

选取已圈定扩散范围内的采样点, 绘制 Mn^{2+}浓度与石油类和硝酸根浓度关系图, 石油类与硝酸根的范围是通过现测打井数据插值, 并通过地表检测的泄漏点分布验证的, 且符合水动力条件流向。当硝酸根浓度 (mg/L) 最低时, Mn^{2+}浓度最高。圈画的园区中心 Mn^{2+}扩散范围内, 监测点硝酸根浓度均小于 1mg/L, 园区中心的 Mn^{2+}浓度在 M03 点附近最高。原因是为微生物降解过程中电子受体的反应顺序是 O_2>NO_3^->Mn^{4+}, 只有当硝酸根反应基本完全[32], 石油类降解才会产生 Mn^{2+}[33], 因此 M03 点 Mn^{2+}浓度最高, 园区中心 Mn^{2+}扩散范围内硝酸根浓度均低于 1mg/L[34]。

本文分析了各点位硝酸根浓度与 Mn^{2+}和石油类浓度关系, 由石油类、NO_3^-和 Mn^{2+}浓度分布图(图 8 和图 9)可知, 石油类的范围主要受 M07 与 M09 处硝酸根点源污染的扩散范围控制, 在研究区中心呈条状分布, 西北方向低浓度石油类小区域呈面状分布。Mn^{2+}基本产生于硝酸根浓度为 0 且具有一定石油类浓度的区域, 硝酸根浓度为 0 区域共有三个区域: 化工园区中心条区域、化工园区西北角以及化工园区东南角。将此三个区域内具有一定石油类浓度的位置视为 Mn^{2+}产出区域, 将区域沿着水流方向拉伸 (西北向东南), 即与 Mn^{2+}浓度分布基本一致。西北方向石油类降解的 Mn^{2+}沿水流方向扩散, 与园区中心 Mn^{2+}扩散范围相接。校正后的扩散范围符合 Mn^{2+}产生和运移规律。

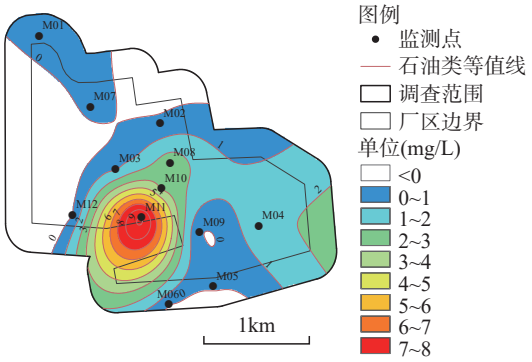

图8　石油类浓度分布图

Fig. 8　Concentration distribution of petroleum.

4　结论

采用 GA-BP 神经网络对化工园区地下水中 Mn^{2+}浓度与其空间位置的关系进行了拟合分析，取得较好的拟合效果，并对 Mn^{2+}的扩散范围进行了修正。结果表明，修正后的 Mn^{2+}扩散范围从园区中心向东南方向延伸，最远可达 924m，扩散面积达到 $1.70\times10^{6}m^{2}$，且超出园区范围 $2.13\times10^{5}m^{2}$。修正后，扩散方向与地下水流向一致，且扩散范围符合 Mn^{2+}产生机制。GA-BP 神经网络对场地地下水污染物扩散范围的精确圈定有较好的辅助效果。

该方法为污染范围刻画不准确的问题提供了一种有效的解决方式，也为地下水污染治理提供了新的技术手段。但实际应用中仍存在一定局限性，如模型参数敏感性和输入数据的精确性等问题。未来研究中，可以结合更多现场监测数据和不同区域的环境特征，进一步优化模型参数，提升预测的准确性和适用性。

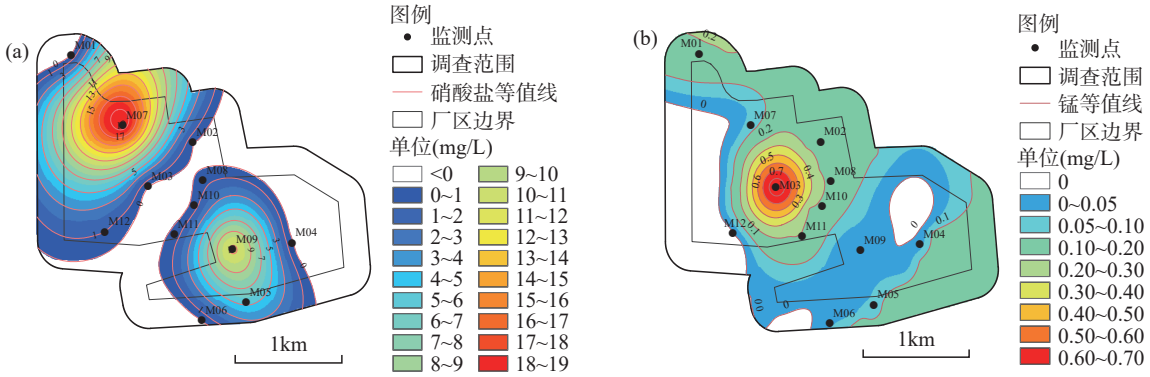

图9　硝酸根 (a) 和 Mn^{2+}(b) 浓度分布图

Fig. 9　Concentration distribution of nitrate (a) and Mn^{2+} (b).

Application of a GA-BP Neural Network in Accurately Characterizing the Diffusion Range of Groundwater Pollutants

JI Jiayun[1,2]，*XIAO Xiao*[1]，*YANG Pinlu*[3]，*LIU Yang*[1]，*ZHOU Yahong*[1,4,5*]

（1. College of Water Resources and Environment, Hebei GEO University, Shijiazhuang 050031, China;

2. College of Urban Geology and Engineering, Hebei GEO University, Shijiazhuang 050031, China;

3. Tangshan Water Supply Company Limited, Tangshan 063000, China;

4. Hebei Collaborative Innovation Center for Sustainable Utilization of Water Resources and Optimization of Industrial Structure, Shijiazhuang 050031, China;

5. Hebei Key Laboratory of Sustainable Utilization and Development of Water Resources, Shijiazhuang 050031, China）

HIGHLIGHTS

(1) The distribution of groundwater Mn^{2+} concentration in the study area does not conform to the law of Mn^{2+} production and migration, necessitating correction.

(2) The GA-BP neural network was used to predict the concentration of Mn^{2+} in the area with sparse sampling points and to correct the distribution of Mn^{2+}.

(3) It is verified that the corrected Mn^{2+} concentration distribution conforms to the Mn^{2+} production mechanism and groundwater dynamics.

ABSTRACT: This study addresses the issue of unevenly distributed sampling points, which leads to inaccurate characterization of pollutant diffusion ranges. Using ArcGIS spatial interpolation, the distribution of Mn^{2+} ions in a chemical park was analyzed, revealing discrepancies due to uneven sampling. To overcome this, two neural network models—GA-BP and standard BP—were applied to predict Mn^{2+} concentrations at unsampled locations. The GA-BP neural network, optimized with a Genetic Algorithm, showed the best performance, filling gaps in data and allowing for a more accurate concentration distribution map. This revised map was used to delineate the Mn^{2+} diffusion range, which was further validated with the known production and migration mechanisms of Mn^{2+}. The results demonstrate that the GA-BP model significantly improves the accuracy of pollutant diffusion mapping and offers a more reliable method for environmental pollution assessment, especially in areas with limited sampling data. The BRIEF REPORT is available for this paper at http://www.ykcs.ac.cn/en/article/doi/10.15898/j.ykcs.202409280204.

KEY WORDS: groundwater; chemical industrial park; GA-BP neural network; influence range

BRIEF REPORT

Significance: Groundwater pollution is a common environmental issue in industrial areas, in which heavy metal pollution (e.g., Mn^{2+}, chromium, lead, iron, etc.) poses a serious threat to ecosystems and human health. The "Technical Guideline for Identification and Assessment of Environmental Damage—Environmental Elements—Part 1: Soil and Groundwater" (GB/T 39792.1—2020) sets more stringent standards for the accuracy of characterizing groundwater pollutant diffusion. However, due to the limitations of the research area, sampling points are often unevenly distributed. Traditional spatial interpolation methods (such as the Kriging method, inverse distance weighting, and spline functions) introduce significant errors in predicting pollutant diffusion, making it difficult to accurately capture their migration patterns.

The GA-BP neural network method was used to construct a predictive model for the groundwater Mn^{2+} concentration in the study area. The model predicted Mn^{2+} concentration in areas with fewer monitoring points and near the monitoring points. By combining the actual monitoring data, a new Mn^{2+} distribution map was created. The results were validated through the principles of Mn^{2+} generation and the regional flow field of the study area. The results show that the GA-BP method significantly improves the correction of the Mn^{2+} ion distribution range, providing a new approach for accurately characterizing the extent of groundwater pollution.

Methods: Using Mn^{2+} pollution in a chemical industry park as a case study, ArcGIS spatial interpolation analysis was utilized to reveal a large deviation in the distribution trend of Mn^{2+} concentration and its formation mechanism, and multiple interpolation methods for correction were explored, yet the results still did not meet the required accuracy. In view of this, the back propagation neural network (GA-BP) optimized by genetic algorithm was compared with the standard back propagation neural network (BP) to optimize the prediction of pollutant concentration and improve the characterization accuracy of the diffusion range.

First, a spatial interpolation of Mn^{2+} concentrations using GIS interpolation methods (Kriging method, inverse

distance weighting, spline functions, etc.) showed that the predictions of these methods did not correspond with the direction of groundwater in the study area and to the mechanism of Mn^{2+} formation, indicating the low applicability of traditional interpolation methods when monitoring points were unevenly distributed. Therefore, a GA-BP neural network was used for concentration prediction, and the model with the best fit for Mn^{2+} concentration prediction at unmonitored points was chosen. In combination with ArcGIS spatial interpolation, the diffusion range of Mn^{2+} was delimited and verified in accordance with the mechanism of Mn^{2+} production.

Twelve groundwater monitoring sites were deployed in the chemical industry park, and the data were divided into a training set and a testing set. The training set included points upstream, contamination plume and downstream of pollution sources to ensure that the model learns pollution migration characteristics under different hydrodynamic conditions. In the test set, three monitoring points were selected, located in the contamination plume and downstream, to test the prediction ability of the model.

Data and Results: Data relevant to the Mn^{2+} ion concentration and organic indicators in the groundwater from a chemical industrial park in Shijiazhuang was used, and the distribution map was created. According to previous regional groundwater data, the groundwater flow direction was from the northeast to the southwest. However, the Mn^{2+} interpolation map shows that the diffusion direction of Mn^{2+} is opposite to the groundwater flow direction, which does not align with its migration patterns. After reviewing the literature, it was found that the electron acceptor response in microbial degradation follows the order: O_2>NO_3^->Mn^{4+}[32]. That is, petroleum degradation only releases Mn^{2+} after nitrate degradation has been substantially completed. It was found that the concentration of Mn^{2+} was highest when the concentration of nitrate was lowest, and the nitrate concentration at the monitoring points was less than 1mg/L within the diffusion area of Mn^{2+}, which was consistent with the theoretical mechanism. Among them, the highest concentration of Mn^{2+} at the M03 point indicated that the degradation of nitrate in the region was largely complete, leading to a large release of Mn^{2+}. This finding confirms that Mn^{2+} migration is influenced by the degradation of petroleum, and that the removal of nitrate is a key premise for Mn^{2+} release. The BP and GA-BP neural network model was established using groundwater Mn^{2+} concentration and its coordinates.

Since there were few data points in the study area, the data was divided into a training set and a testing set. The three testing points were located in the pollution plume and downstream of the pollution source, while the training set included points from upstream of the pollution source, the pollution plume, and downstream of the pollution source. The correlation coefficient (R) for the training set was 0.9921, and R^2=0.99747, which was greater than 0.9. The correlation coefficient (R) for the testing set was 0.9987, and R^2=0.9974, proving that the training set fitted well and had strong credibility. Compared to the standard BP neural network, the GA-BP model's actual values were more closely aligned with the predicted values, and the GA-BP neural network's errors were close to zero. The prediction performance of the GA-BP neural network was significantly better than that of the BP neural network, with the training model providing a good fit.

Using a GA-BP neural network to supplement the Mn^{2+} concentration data for the missing regions of the monitoring points and to replot the Mn^{2+} concentration distribution, the result shows that the dispersion of Mn^{2+} in the center of the chemical park is $1.70\times10^6 m^2$, of which $2.13\times10^5 m^2$ extends beyond the chemical park. Furthermore, the migration of Mn^{2+} is mainly influenced by the combination of the degradation of petroleum pollutants and the flow of groundwater. In areas where the nitrate concentration is near zero and the concentration of petroleum pollutants is moderate, the concentration of Mn^{2+} is highest and exhibits migration patterns following groundwater flow directions, forming three Mn^{2+}-enriched areas: the center sector, the northwest corner, and the southeast corner of the industrial park. The revised Mn^{2+} diffusion range is more consistent with known migration laws and further validates the accuracy of the GA-BP neural network predictions.

参考文献

[1] Gao Y, Qian H, Ren W, et al. Hydrogeochemical characterization and quality assessment of ground-water based on integrated-weight water quality index in a concentrated urban area[J]. Journal of Cleaner Production, 2020, 260: 121006.

[2] 张保会, 王林芳, 郭宏, 等. 我国污染场地修复决策系统研究进展[J]. 环境与可持续发展, 2021, 46(2): 138−143.
Zhang B H, Wang L F, Guo H, et al. Research progress of decision-making system for remediation of contaminated sites in China[J]. Environment and Sustainable Development, 2021, 46(2): 138−143.

[3] 陈卫平, 谢天, 李笑诺, 等. 欧美发达国家场地土壤污染防治技术体系概述[J]. 土壤学报, 2018, 55(3): 527−542.
Chen W P, Xie T, Li X N, et al. Generalization of technical systems for soil pollution prevention and control in developed countries[J]. Acta Pedologica Sinica, 2018, 55(3): 527−542.

[4] 陈吉吉, 陶蕾, 刘保献, 等. 北京市平原区地下水铁锰分布特征及成因分析[J]. 水文地质工程地质, 2024, 51(6): 198−207.
Chen J J, Tao L, Liu B X, et al. Distribution characteristics and genesis analysis of groundwater iron and manganese in Beijing plain area[J]. Hydrogeological Engineering Geology, 2024, 51(6): 198−207.

[5] Jiang W J, Sheng Y Z, Wang G C, et al. Cl, Br, B, Li, and noble gases isotopes to study the origin and evolution of deep groundwater in sedimentary basins, a review[J]. Environmental Chemistry Letters, 2022, 20: 1497−1528.

[6] 於方, 赵丹, 王膑, 等.《生态环境损害鉴定评估技术指南土壤与地下水》解读[J]. 环境保护, 2019, 47(5): 19−23.
Yu F, Zhao D, Wang B, et al. Interpretation of technical guidelines for identification and assessment of eco-environmental damage to soil and groundwater[J]. Environmental Protection, 2019, 47(5): 19−23.

[7] 刘宏伟. 中宁石空工业园区地下水锰污染评价及其水环境影响预测研究[D]. 西安: 长安大学, 2013.
Liu H W. Evaluation of groundwater manganese pollution and prediction of water environment impact in Zhongning Shikong industrial park [D]. Xi'an: Chang'an University, 2013.

[8] Jiang W J, Meng L S, Liu F T, et al. Distribution, source investigation, and risk assessment of topsoil heavy metals in areas with intensive anthropogenic activities using the positive matrix factorization (PMF) model coupled with self-organizing map (SOM)[J]. Environmental Geochemistry and Health, 2023, 45: 6353−6370.

[9] 李元杰, 王森杰, 张敏, 等. 土壤和地下水污染的监控自然衰减修复技术研究进展[J]. 中国环境科学, 2018, 38(3): 1185−1193.
Li Y J, Wang S J, Zhang M, et al. Research progress of monitored natural attenuation remediation technology for soil and groundwater pollution[J]. China Environmental Science, 2018, 38(3): 1185−1193.

[10] 王兴, 刘莹, 王春晖, 等. 海洋盐度分布的插值方法应用与对比研究[J]. 海洋通报, 2016, 35(3): 324−330.
Wang X, Liu Y, Wang C H, et al. Application and comparative study of interpolation methods for ocean salinity distribution[J]. Ocean Bulletin, 2016, 35(3): 324−330.

[11] Mohammed A, Paraskevas T, Gaetano P, et al. Application of multiple spatial interpolation approaches to annual rainfall data in the Wadi Cheliff Basin (North Algeria)[J]. Ain Shams Engineering Journal, 2024, 15(3): 10257.

[12] 段宁, 杨思言, 魏婉婷. 基于 BP 神经网络的铅酸蓄电池厂地下水重金属浓度预测[J]. 环境科学与技术, 2016, 39(1): 194−198.
Duan N, Yang S Y, Wei W T. Prediction of heavy metal concentrations of the groundwater from a lead-acid battery factory based on BP neural network[J]. Environmental Science & Technology, 2016, 39(1): 194−198.

[13] 刘伟韬, 李蓓蓓, 杜衍辉, 等. 基于改进的 SSA-BP 神经

网络的矿井突水水源识别模型研究[J]. 工矿自动化, 2024, 50(2): 98−105, 115.
Liu W T, Li B B, Du Y H, et al. Research on the recognition model of mine water inrush source based on improved SSA-BP neural network[J]. Industry and Mine Automation, 2024, 50(2): 98−105, 115.

[14] Robertson B, Dam-Bates V P, Gansell O. Halton iterative partitioning master frames[J]. Environmental and Ecological Statistics, 2021, 29(1): 1−22.

[15] 李海斌. 基于遗传算法的 BP 神经网络在柴油机故障诊断中的应用[J]. 科技视界, 2014, 2(13): 206, 211.
Li H B. Application of BP neural network based on genetic algorithm in diesel engine fault diagnosis[J]. Science & Technology Vision, 2014, 2(13): 206, 211.

[16] 王玉雯, 陈颖辉, 师庭飞. 改进的 BP 网络在深基坑变形预报中的应用[J]. 科学技术与工程, 2010, 10(15): 3791−3794.
Wang Y W, Chen Y H, Shi T F. Application of improved BP network in deformation prediction of deep foundation pit[J]. Science, Technology and Engineering, 2010, 10(15): 3791−3794.

[17] 徐杰, 但斌斌, 容芷君, 等. BP 神经网络在铁液预处理脱硫率预测中的应用[J]. 铸造技术, 2017, 38(9): 2183−2187, 2192.
Xu J, Dan B B, Rong Z J, et al. Application of BP neural network in prediction of desulfurization rate of molten iron pretreatment[J]. Casting Technology, 2017, 38(9): 2183−2187, 2192.

[18] 杨帆. 基于 BP 神经网络的 CO_2 通量预测模型研究[D]. 哈尔滨: 东北林业大学, 2017.
Yang F. Research on CO_2 flux prediction model based on BP neural network [D]. Harbin: Northeast Forestry University, 2017.

[19] 孙卫鹏, 徐合力, 高岚. 基于 GA-BP 的船舶同步发电机定转子绕组匝间短路故障诊断研究[J]. 中国修船, 2020, 33(4): 48−54.
Sun W P, Xu H L, Gao L. Research on fault diagnosis of inter-turn short circuit of stator and rotor windings of ship synchronous generator based on GA-BP[J]. China Ship Repair, 2020, 33(4): 48−54.

[20] 王嵘冰, 徐红艳, 李波, 等. BP 神经网络隐含层节点数确定方法研究[J]. 计算机技术与发展, 2018, 28(4): 31−35.
Wang R B, Xu H Y, Li B, et al. Research on the determination method of hidden layer nodes in BP neural network[J]. Computer Technology and Development, 2018, 28(4): 31−35.

[21] 陈克根. 统计机器学习中的过拟合问题[J]. 文渊(中学版), 2019(11): 755.
Chen K G. Overfitting problem in statistical machine learning[J]. Wen Yuan (Middle School Edition), 2019(11): 755.

[22] 王蒙. 机器学习中样本筛选方法的研究与应用[D]. 成都: 电子科技大学, 2017.
Wang M. Research and application of sample screening methods in machine learning [D]. Chengdu: University of Electronic Science and Technology of China, 2017.

[23] Zheng Z X, Zhang Y L, Su X S, et al. Responses of hydrochemical parameters, community structures, and microbial activities to the natural biodegradation of petroleum hydrocarbons in a groundwater-soil environment[J]. Environmental Earth Sciences, 2016, 75(21): 1.

[24] Klinchuch L, Delfino T. Reductive dissolution and precipitation of manganese due to biodegradation of petroleum hydrocarbons[J]. Environmental Geosciences, 2000, 7(4): 211−212.

[25] Ning Z, Zhang M, He Z, et al. Spatial pattern of bacterial community diversity formed in different groundwater field corresponding to electron donors and acceptors distributions at a petroleum-contaminated site[J]. Water, 2018, 10(7): 842.

[26] 温晶. 石油污染含水层石油相和水相有机质组分研究[D]. 北京: 中国石油大学 (北京), 2023.
Wen J. Study on organic matter composition of petroleum phase and aqueous phase in petroleum contaminated aquifer[D]. Beijing: China University of Petroleum (Beijing), 2023.

[27] 丁智晖, 董子萱, 于水利. 难降解有机物质的生物降解技术分析[J]. 当代化工研究, 2018(1): 49−51.

Ding Z H, Dong Z X, Yu S L. The analysis of biodegradation technology in the field of refractory organic matters[J]. Modern Chemical Research, 2018(1): 49−51.

[28] 盛连喜, 李明堂, 徐镜波. 硝基苯类化合物微生物降解研究进展[J]. 应用生态学报, 2007, 18(7): 1654−1660.
Sheng L X, Li M T, Xu J B. Research advances in microbial degradation of nitrobenzene and its substituted compounds[J]. Chinese Journal of Applied Ecology, 2007, 18(7): 1654−1660.

[29] 豆俊峰, 刘翔. 苯系化合物在硝酸盐还原条件下的生物降解性能[J]. 环境科学, 2006, 27(9): 1846−1852.
Dou J F, Liu X. Biodegradability of benzene series compounds under nitrate reduction conditions[J]. Environmental Sciences, 2006, 27(9): 1846−1852.

[30] 林广宇. 地下水位变动带石油烃污染物的迁移转化规律研究[D]. 长春: 吉林大学, 2014.
Lin G Y. Study on the migration and transformation of petroleum hydrocarbon pollutants in groundwater level fluctuation zone [D]. Changchun: Jilin University, 2014.

[31] 路莹. 浅层地下水系统石油类污染物的生物降解机制研究[D]. 长春: 吉林大学, 2013.
Lu Y. Study on biodegradation mechanism of petroleum pollutants in shallow groundwater system [D]. Changchun: Jilin University, 2013.

[32] 李烨, 潘涛, 刘菲, 等. 四氯乙烯在不同地下水环境的生物共代谢降解[J]. 岩矿测试, 2012, 31(4): 682−688.
Li Y, Pan T, Liu F, et al. Biological co-metabolic degradation of tetrachloroethylene in different groundwater environments[J]. Rock and Mineral Analysis, 2012, 31(4): 682−688.

[33] 宁卓, 郭彩娟, 蔡萍萍, 等. 某石油污染含水层降解能力地球化学评估[J]. 中国环境科学, 2018, 38(11): 4068−4074.
Ning Z, Guo C J, Cai P P, et al. Geochemical assessment of the degradation capacity of an oil-contaminated aquifer[J]. Environmental Sciences of China, 2018, 38(11): 4068−4074.

[34] 张敏, 张巍, 郭彩娟, 著. 污染场地自然衰减修复的原理与实践[M]. 北京: 科学出版社, 2019:128−130.
Zhang M, Zhang W, Guo C J, et al. Remediation by natural attenuation at contaminated sites: Principles and practice [M]. Beijing: Science Press, 2019:128−130.