doi:10.6041/j.issn.1000-1298.2018.09.032

基于 KHA 优化 BP 神经网络的地下水水质综合评价方法

刘 东^{1,2} 李 帅¹ 付 强^{1,2} 刘春雷¹

(1. 东北农业大学水利与土木工程学院,哈尔滨 150030; 2. 农业部农业水资源高效利用重点实验室,哈尔滨 150030)

摘要:为提高区域地下水水质评价精度,将磷虾群算法(Krill herd algorithm,KHA)引入到 BP 神经网络连接权值与 阈值的优化过程中,构建了 KHA - BP 地下水水质综合评价模型。以黑龙江省农垦建三江管理局为研究对象,运用 所建模型对其下辖 15 个农场进行地下水水质综合评价,并对造成地下水水质污染的主要原因进行辨识。为验证 本文所建模型的适用性,引入区分度法与序号总和理论分别分析了 KHA - BP 模型、PSO - BP 模型以及 BP 模型的 可靠性与稳定性。结果表明:各农场地下水水质良好,且存在一定的空间分布规律,I 类水质主要集中在管理局西 南位置,Ⅱ类水质主要集中在北部和南部,Ⅲ类水质主要分布于中东部和中西部。Fe、Mn、COD_{Mn}、NH₃-N 以及 NO₃⁻-N 是造成地下水水质污染的主要因素。其中 Fe、Mn 是当地原生危害,COD_{Mn}、NH₃-N、NO₃⁻-N 含量超标主要与 大量施用化肥、农药有关。KHA - BP 模型的区分度为 1. 107 0,Spearman 等级相关系数为 0. 928 6,与 PSO - BP 模 型、BP 模型相比优势明显。研究成果可为粮食生产核心区的地下水资源科学管理及水生态文明建设提供科学依据。 关键词:地下水水质;建三江管理局;磷虾群算法; BP 神经网络

中图分类号: X824 文献标识码: A 文章编号: 1000-1298(2018)09-0275-10

Comprehensive Evaluation Method of Groundwater Quality Based on BP Network Optimized by Krill Herd Algorithm

LIU Dong^{1,2} LI Shuai¹ FU Qiang^{1,2} LIU Chunlei¹

College of Water Conservancy and Civil Engineering, Northeast Agricultural University, Harbin 150030, China
 Key Laboratory of Effective Utilization of Agricultural Water Resources, Ministry of Agriculture, Harbin 150030, China)

Abstract: A new BP network model was developed to improve the accuracy and assess the groundwater quality. For this purpose, the krill herd algorithm (KHA) was established with the optimization process of the connection weights and thresholds of the BP neural network. Totally 15 farms were selected to evaluate the groundwater quality and identify the main causes of groundwater pollution in Jiansanjiang Administration. In addition, to verify the applicability of the model, the distinction degree method and the theory of serial number summation were used to analyze the reliability and stability of KHA - BP model, PSO - BP model and BP model, respectively. The results exhibited a good agreement of groundwater quality in each farm and there was a certain spatial distribution pattern such as the water quality of grade I was mainly concentrated in the southwest position, grade II was distributed in the north and south, while the grade III was located in the mid-west and mid-east of the administration. Fe, Mn, COD_{Mn} , NH_3 -N and NO_3^- -N were the main factors caused groundwater pollution. Fe and Mn were local primary hazard but excessive amounts of COD_{Mn}, NH₃-N and NO₃⁻-N were mainly related to use of a large number of fertilizers and pesticides. The distinction degree of KHA - BP was 1. 107 0 and Spearman's rank coefficient was 0.928 6, which was better than those of PSO - BP and BP. In conclusion, this research could provide a scientific basis for the comprehensive management of groundwater resources and construction of water ecological civilization in the core areas of food production.

Key words: groundwater quality; Jiansanjiang Administration; krill herd algorithm; BP neural network

收稿日期:2018-03-27 修回日期:2018-04-22

基金项目:国家自然科学基金项目(51579044、41071053、51479032)、国家重点研发计划项目(2017YFC0406002)、黑龙江省自然科学基金 项目(E2017007)和黑龙江省水利科技项目(201319、201501、201503)

作者简介: 刘东(1972—),男,教授,博士生导师,主要从事水土资源优化利用与管理研究,E-mail: liudong@ neau. edu. cn

通信作者:付强(1973—),男,教授,博士生导师,主要从事农业水土资源高效利用理论、方法及应用研究,E-mail: fuqiang@ neau. edu. cn

0 引言

地下水是水资源系统中的主要组成部分,作为 地球上的重要水体,以其稳定的供水条件、良好的水 质,成为农业、工业以及城市生活用水的重要水源, 是人类社会必不可少的资源^[1]。地下水资源具有 可再生性,但同时具有脆弱性^[2],循环周期较长,一 旦赋存环境遭到破坏,再生能力将会减弱。随着人 类社会的飞速发展,长期不合理、无规则的开采使地 下水水质恶化、水位持续下降,地下水系统遭受破坏 难以避免。因此,开展地下水水质评价可以掌握地 下水污染现状,对合理开采地下水资源、保护水环境 具有十分重要的意义。

迄今为止,已有众多国内外专家学者研究过地 下水水质评价问题。KIURSKI-MILPSEVIC 等^[3]运 用模糊逻辑法对塞尔维亚 Zrenjanin 市的地下水水 质进行了评价; ZHANG 等^[4] 运用灰关联分析法 (GRA)评价了中国北方滹沱河冲积扇区域的地下 水水质;BOUDERBALA^[5]将水质指数法(WQI)应用 于阿尔及利亚北部 Low - Isser 平原地区的地下水水 质评价:WU 等^[6]应用基于熵权的密切值法(OVM) 对银川北郊水源的地下水水质进行了评价。虽然上 述方法在水质评价方面取得了一定成果,但是也存 在一些不足之处:模糊逻辑法在获得评价集时采用 线性加权平均的方法,容易出现失真、失效、跳跃等 现象,导致等级评价不准确^[7-8]:传统灰关联分析法 面对复杂情况时存在评价因子权重影响程度低等缺 点[9-10]:水质指数法缺陷在于评价结果会出现失真 或物理意义不明确的现象[11-12]:密切值法进行多样 本评价时,对各指标权重的处理较为随意,导致评价 结果的误差较大^[13]。BP神经网络是一种基于人工 智能的非线性动力学系统,具有较强的自学习和自 适应能力,能够通过已有数据信息自动调整网络的 阈值和权值,从而减小主观因素的影响,使得评价结 果更加符合客观实际,已被国内外学者广泛应 用^[14-15]。因此,本文拟构建基于 BP 神经网络的地 下水水质评价模型。然而,传统 BP 神经网络各层 的连接权值和阈值均由随机初始化获得,收敛速度 慢,训练时间长且易陷入局部最优。为了提高模型 评价精度,必须改进 BP 神经网络。

GANDOMI 等^[16]于 2012 年提出的 KHA 算法是 一种新型的智能优化算法,该算法用于优化支持向 量机(SVM)的目标威胁估计模型^[17],整体性能要优 于 GSO(Group search optimization)和 PSO(Particle swarm optimization);用于入侵检测系统的构建^[18], 性能优于 PSO、GA(Genetic algorithm);应用于 SVDD 的参数优化^[19],性能优于 SA(Simulated annealing)、ACO(Ant colony optimization);应用于汽 轮机 热耗率预测问题^[20],性能优于 PSO、BBO (Biogeography-based optimization)。以上研究表明, KHA 具有较强的收敛速度和搜索能力,简单易用、 鲁棒性强,是一种性能优越的智能优化算法。因此,针对 BP 神经网络的缺陷,本文拟将磷虾群算法引 入到网络连接权值和阈值的优化过程中,改进 BP 神经网络。以建三江管理局为例,构建基于 KHA 优化 BP 神经网络的地下水水质综合评价模型,分析 区域地下水水质等级空间分布特征及造成地下水水质污染的可能成因,验证 KHA – BP 模型在区域地下水水质评价中的可行性与优越性。

1 材料与方法

1.1 研究区概况

黑龙江省农垦建三江管理局位于黑龙江省东北 部,地处三江平原腹地,位于黑龙江、松花江和乌苏 里江汇流的河间地带。地理坐标 132°31′38″ ~ 134°32′19″E,46°49′47″ ~48°12′58″N。区内水系发 达,河网纵横交错,地表水资源丰富^[21]。作为重要 的商品粮基地,建三江管理局粮食总产占黑龙江垦 区粮食总产的 1/3,占黑龙江省的 1/9,区内农业生 产主要以水稻种植为主^[22]。建三江管理局下辖 15 个国营农场,具体行政分区见图 1。



图 1 建二江省建向1]政方区 Fig. 1 Administrative division map of Jiansanjiang Administration

长期以来,建三江地区由于工程调蓄能力弱,地 表水利用率不高,农业灌溉主要以开发地下水为 主^[23]。由于大规模抽取地下水灌溉水田,同时大量 施用化肥农药,灌溉水携带污染物质渗入地下,对地 下水体造成污染,严重危害了饮用水安全和生态系 统安全。因此开展建三江管理局地下水水质评价和 污染分析,对确保该地区饮用水安全、合理提出地下 水资源保护措施等均具有重要的理论和现实意义。

1.2 数据来源

于 2016 年 5 月泡田期,选取 15 个代表性灌溉 水井作为水质采样点,每个采样点取 3 个水样。每 个水样测定 pH 值、COD_{Mn}、NH₃-N、NO₃⁻-N、Cl⁻、 F⁻、SO₄²⁻、Fe、Mn 含量共 9 个指标,其中 pH 值利用 PHS – 3C 型 pH 计测定,其余各指标利用美国 Hach DR2800 型便携式分光光度计测定。以每个采样点 的各指标浓度均值作为代表值,绘制各农场地下水 水质指标质量浓度图,如图 2 所示。





Fig. 2 Concentration map of groundwater quality index

从2017年《黑龙江垦区统计年鉴》收集整理建 三江管理局耕地总面积、顷均化肥施用总量、顷均氮 肥施用量、顷均磷肥施用量和顷均农药施用量数据, 用于分析建三江地下水水质污染的原因。

1.3 BP 神经网络

BP 网络是一种单向传播的多层前馈神经网络,标准 BP 网络算法包括前向传播和反向传递两个方面^[14-15,24-25]。在前向传播中,输入信号经过输入层、隐含层处理后,最终到达输出层。当输出层得不到期望输出时,就会进行反向传递,通过计算输出层所有神经元的期望输出与实际值之间的总误差对网络权值和阈值进行调整,从而使神经网络实际输出不断逼近期望输出。BP 神经网络结构如图 3 所示。





1.4 磷虾群算法

KHA 算法是对南极海洋中磷虾群觅食活动的 模拟^[16]。磷虾群中的磷虾个体在觅食活动中受到 食物和附近磷虾的综合影响,在这种综合作用下向 着食物移动^[26-27]。具体来说,磷虾位置变化主要受 3 个因素影响:其他磷虾个体引起的运动、觅食行动 以及随机扩散。

由上所知,每个磷虾个体 *i* 的位置移动由 3 部 分构成

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = N_i + F_i + D_i \tag{1}$$

式中 X_i——第 i 只磷虾个体的位置

N_i——第 i 只磷虾的诱导速度

F_i——第*i*只磷虾的觅食速度

D_i——第 i 只磷虾的扩散速度

(1)其他磷虾个体影响引起的运动

磷虾的游动方向受到种群排斥效应、附近磷虾 个体以及种群位置最优个体的影响,具体表示为

$$N_{i,new} = N_{max}(\alpha_{i,local} + \alpha_{i,larget}) + w_n N_{i,old}$$
 (2)
式中 $N_{i,new}$ 新的位置变化量
 N_{max} 最大诱导速度,通常取 0.01 m/s
 $N_{i,old}$ 世次产生的位置变化量
 w_n 惯性权值,取值范围[0,1]
 $\alpha_{i,local}$ 你近个体的诱导向量
 $\alpha_{i,larget}$ 最优个体提供的诱导向量
(2)觅食行为
磷虾个体的觅食运动可以表达为
 $F_{i,new} = V_f(\beta_{i,food} + \beta_{i,best}) + w_f F_{i,old}$ (3)
式中 $F_{i,new}$ — 新的觅食位置变化量
 $F_{i,old}$ 上一次觅食位置变化量
 V_f 觅食速度,通常取 0.02 m/s
 $\beta_{i,food}$ 像物源对于个体吸引的向量
 $\beta_{i,best}$ 第 i 个个体最优的目标向量
 w_f 惯性权值,取值范围[0,1]
(3)随机扩散

磷虾群的随机扩散一般由一个随机的方向向量 和磷虾群的最大扩散速度决定,即

$$D_i = D_{\max} \boldsymbol{\delta} \tag{4}$$

式中 D_{max} ——个体最大扩散速度,一般取 0.005 m/s δ ——随机向量, $\delta \in [-1,1]$

随着迭代次数的增加,磷虾群逐渐靠近最优位置,相应的其随机扩散也应该逐渐减弱。因此需要 在式(4)中加入新的变量使得随机扩散部分随迭代 线性递减,即

$$D_i = D_{\max} \left(1 - \frac{I}{I_{\max}} \right) \boldsymbol{\delta}$$
 (5)

式中 1——当前迭代次数

I_{max}——最大迭代次数

经过以上3种行为之后,在t到 $t + \Delta t$ 的时间间隔内磷虾个体的位置为

$$X_i(t + \Delta t) = X_i(t) + \Delta t \frac{\mathrm{d}X_i}{\mathrm{d}t} \tag{6}$$

其中
$$\Delta t = C_t \sum_{j=1}^{N_v} (U_j - L_j)$$
(7)

式中
$$\Delta t$$
——时间间隔常量,取决于搜索空间
 $X_i(t + \Delta t) X_i(t)$ —— $t + \Delta t t$ 时刻的磷虾个
体位置

$$U_j$$
——第 j 个变量的上限
 L_j ——第 j 个变量的下限
 N_v ——变量总数,通过经验发现

広 小 山 山 四

C,——[0,2]区间内的一个常数,其值越小, 算法的搜索步长越小

(4)遗传操作

在 KHA 算法中,遗传操作分为交叉和变异操作。交叉操作是指通过替换重组生成一个新个体的操作,即

$$x_{i,m} = \begin{cases} x_{r,m} & (\operatorname{rand} < C_r) \\ x_{i,m} & (\operatorname{rand} \ge C_r) \end{cases}$$
(8)

式中 $x_{i,m}$ — x_i 第 m 个参量 $x_{r,m}$ — 异于 $x_{i,m}$ 的个体的第 m 个参量, $r \in$ $[1,2,\dots,N_n] 且 r \neq i$

变异操作是对个体元素作变动调整的操作,即

$$x_{i,m} = \begin{cases} x_{gbest,m} + \mu(x_{p,m} - x_{q,m}) & (\operatorname{rand} < M_u) \\ x_{i,m} & (\operatorname{rand} \ge M_u) \end{cases}$$
(9)

式中 $x_{gbest,m}$ 当前全局最优个体的第 m 个参量 $x_{p,m}, x_{q,m}$ 异于 $x_{i,m}$ 的两个个体的第 m 个 参量, $p,q \in [1,2,\dots,i-1,i+1,i+2,\dots,N_p]$ μ [0,1]之间的数

M.——变异概率

KHA-BP 算法的具体步骤如下:

(1)设定算法初始参数,确定种群大小 N_p 、最大 迭代次数 M_i 、最大诱导速度 N_{max} 、觅食速度 V_f 以及 个体最大扩散速度 D_{max} 等。

(2)初始化搜索空间范围内的一组种群,每只 磷虾个体代表优化问题的一个可行解。

(3) 计算个体适应度值, 输入训练样本对神经 网络进行训练, 并用测试样本进行预测, 从而得到测 试样本的期望输出 *X* = (*x*₁, *x*₂, …, *x*_n) 及预测输出

$$Y = (y_1, y_2, \dots, y_n)$$
,则个体适应度计算公式为
 $F = ||x_i - y_i||$ (10)

本文构建评价模型的训练样本由地下水质量分 类标准数据获得,输出的预测则为具体的水质类别, 训练样本与测试样本的获取方法详见2.2.1节。

(4) 计算 3 个影响因素引起的磷虾位置变化 量, 加入遗传操作后重新计算磷虾个体的位置。

(5)返回步骤(3)计算个体适应度直到满足终止条件。

(6)输出最优个体位置,即最优权值和阈值,构 建最终评价模型,得出评价结果。

KHA-BP评价模型的构建流程如图4所示。





2 结果与分析

2.1 水质等级评价标准

依据 GB/T 14848—2017《地下水质量标准》选取 9 个具有代表性的指标作为水质评价因子,根据 GB/T 14848—2017 的分类标准,将水质等级划分为 5 类,各等级标准指标浓度见表 1。

2.2 评价模型的实现

2.2.1 训练样本集和测试样本集的生成

由表1共得到I~V类水的5个标准等级区间, 根据表1中数值可以得到用于神经网络训练的几组 输入和输出数据。然而,仅仅利用几组数据对 BP 神经网络进行训练无法实现精确建模,得到的水质 模型也无法精确评价水质。因此,本文对标准水质 评价指标采用内插取值方法对每类水质分别生成

	Tab. 1 Groundwater quality classification standard											
水质	ць	质量浓度 <i>p</i> /(mg·L ⁻¹)										
类别	рн ш –	$\mathrm{COD}_{\mathrm{Mn}}$	NH3-N	$NO_3^ N$	Cl -	F -	SO_4^2 -	Fe	Mn			
Ι	6.5 ~ 8.5	≤1	≤0.02	≤2	≤50	≤1	≤50	≤0.1	≤0.05			
П	6.5 ~ 8.5	≤2	≤0.10	≤5	≤150	≤1	≤150	≤0.2	≤0.05			
Ш	6.5 ~ 8.5	≤3	≤0.50	≤20	≤250	≤1	≤250	≤0.3	≤0.1			
IV	5.5~6.5,8.5~9	≤10	≤1.50	≤30	≤350	≤2	≤350	≤2.0	≤1.5			
V	< 5.5, >9	>10	> 1.50	> 30	>350	>2	> 350	> 2.0	> 1.5			

表 1 地下水质量分类标准 Tab. 1 Groundwater quality classification standard

500 组样本数据,取前 400 组为训练样本,后 100 组 为测试样本,以满足水质建模的需求,样本生成规则 以Ⅲ类水质标准为例,如表 2 所示。所选指标除 pH 值为适度指标外,其他指标均为越小越优型指标。 为消除各指标量纲和统一指标值的变化范围,采用 归一化方法处理上述生成的样本数据^[28]。

表 2 神经网络训练样本生成规则	表 2	神经网络训练样本生成规则
------------------	-----	--------------

Fab. 2	Generation	rules	of	training	sample	of	neutral	network
--------	------------	-------	----	----------	--------	----	---------	---------

		质量浓度 p/(mg・L ⁻¹)								扣 钥絵山	쿄;□4± 甩
рнш	$\mathrm{COD}_{\mathrm{Mn}}$	NH_3 -N	NO_3^- -N	Cl -	F ⁻	SO_4^2 -	Fe	Mn	则幼性平	别呈制币	班 以'4本
6.5 ~ 8.5	2 ~ 3	0.1~0.5	5 ~ 20	150 ~250	$0 \sim 1$	150 ~ 250	0.2~0.3	$0.05 \sim 0.1$	400	3	Ш

~	$\left\{ \frac{x_{\max} - x}{x_{\max} - x_{\min}} \right\}$	(越小越优型指标)
<i>x</i> = 1	$\left[1 - \frac{10.5(x_{up} + x_{down}) - x }{x_{max} - x_{min}}\right]$	(适度指标)

(11)

式中 *x̂*——经过标准化处理的数据 *x*——原始数据 *x_{max}、<i>x_{min}*——数据序列中的上、下限 *x_{up}、<i>x_{down}*——适度区间的上、下限

2.2.2 网络结构的确定

本文选取9个水质指标进行水质评价,因此 神经网络的输入层神经元个数为9;由于输出层 的期望输出是水质评价等级,故输出层有一个神 经元。采用单层隐含层的 BP 神经网络,在隐含 层节点数的选取上,目前并没有统一的计算方 法,一般采取试凑法进行选定。通过反复训练, 最终确定隐含层节点数为 15。选取指数型 S 函数,即 logsig 函数作为 KHA – BP 模型的激励函数^[29],以 BP 神经网络中权值和阈值为待优化变量,利用 KHA 优化 BP 神经网络,构建 9 – 15 – 1 的 KHA – BP 模型。参照文献[16],对 KHA 中重要参数进行设定,种群数量 N_p = 30,最大迭代次数 M_i = 200,最大诱导速度 N_{max} = 0.01 m/s,觅食速度 V_f = 0.02 m/s,个体最大扩散速度 D_{max} = 0.005 m/s。考虑算法随机性的问题,多次运行KHA – BP 算法,选取具有较高适应度的最优权值和阈值构建评价模型。

2.2.3 等级模拟区间确定

利用式(11),将表1各等级临界指标浓度值归 一化,归一化后数据代入上述所建的 KHA - BP 评 价模型中,得到临界值模拟结果和各等级模拟区间, 结果如表3所示。

表 3 KH-BP 地下水水质评价模型的等级模拟区间

Tab. 3	Hierarchical	simulation	interval	of	model	of	groundwater	quality	assessment	based	on	KHA	. – E	3P
--------	--------------	------------	----------	----	-------	----	-------------	---------	------------	-------	----	-----	-------	----

等级	Ι	П	Ш	IV	V
等级模拟区间	(0.8684,1.2981]	(1.2981,2.3864]	(2.3864,3.5943]	(3.5943,4.1612]	(4.1612,11.2623]

2.3 地下水水质等级空间分布

将图 2 所示建三江管理局各农场地下水水质指标值按式(11) 归一化, 归一化后数据代入所建的 KHA - BP 地下水水质评价模型中,得到各农场地下 水水质模拟结果和相应等级,结果如表 4 所示。

由表4可以看出,建三江管理局各农场地下水 水质总体较好,所有农场水质均处于Ⅲ类以上(包 括Ⅲ类)。胜利、红卫、创业、七星4个农场水质等 级为Ⅰ类,从水质模拟结果来看,地下水水质污染程 度由大到小为:创业、红卫、七星、胜利,红卫、创业、 七星3个农场水质模拟值接近所处等级区间上限, 可以认为这3个农场地下水水质有向Ⅱ类水质转化 的趋势,胜利农场距离Ⅰ类水质上下限均有一定距 离,说明水质较为稳定;前哨、二道河、前锋、鸭绿河、

Tab. 4 Simulation results and grades of groundwater quality in each farm

农场	模拟结果	水质等级	农场	模拟结果	水质等级
前哨	1.9222	Π	青龙山	1.6132	Π
二道河	1.4042	Π	前进	2.5107	Ш
八五九	2.5651	Ш	红卫	1.2253	Ι
胜利	0.9586	Ι	创业	1.2620	Ι
前锋	2.0490	Π	七星	1.2007	Ι
鸭绿河	1.7377	II	大兴	1.3432	Ш
洪河	1.3020	Π	勤得利	1.5693	Ш
浓江	1.8946	Π			

洪河、浓江、青龙山、大兴、勤得利9个农场水质均为 Ⅱ类,根据水质模拟结果,地下水水质污染程度由大 到小为:前锋、前哨、浓江、鸭绿河、青龙山、勤得利、 二道河、大兴、洪河,其中前锋、前哨、浓江、鸭绿河、 青龙山、勤得利6个农场水质模拟结果与Ⅱ类水质 上下限有一定距离,说明水质等级稳定,二道河、大 兴、洪河3个农场水质等级模拟结果距离Ⅱ类水质 下限很近,可以认为这3个农场水质有向Ⅰ类水质 改善的趋势;八五九、前进2个农场的水质等级为Ⅲ 类,根据水质模拟结果,八五九地下水污染程度高于 前进,其中前进农场的水质模拟结果距离Ⅲ类水质 下限很近,说明有向Ⅱ类水质改善的趋势,八五九农 场水质模拟结果与Ⅲ类水质上下限均有一定距离, 可以认为八五九农场地下水水质可以基本稳定在Ⅲ 类。

根据表 4 评价结果绘制建三江管理局地下水 水质等级空间分布图,见图 5。可以看出,建三江管 理局地下水等级在空间分布上呈一定规律性,水质 等级为 I 类的 4 个农场分布较为集中,主要分布于 管理局西南位置;水质等级为 II 类的除大兴农场在



图 5 建三江管理局地下水水质等级空间分布图 Fig. 5 Spatial distribution map of groundwater quality of Jiansanjiang Administration

管理局南部,其他农场均集中分布于北部;水质等级 为Ⅲ类的前进农场和八五九农场则分布于管理局中 部偏西和中部偏东位置,较为分散。

2.4 水质评价结果分析

参照 GB/T 14848—2017《地下水质量标准》中 分类对比,发现 pH 值、Fe、Mn、COD_{Mn}、NH₃-N、NO₃⁻-N 含量6种指标存在超标现象。为辨识区域地下水水 质污染的主要影响因子,计算 KHA – BP 评价结果 中各指标所占权重,如表5 所示。

表 5 评价结果中各指标权重

Tab. 5 Weight of each index in evaluation results

指标	权重
pH 值	0. 092 4
COD _{Mn} 含量	0. 170 1
NH3-N 含量	0. 210 8
NO3-N含量	0. 209 8
Cl-含量	0.0322
F-含量	0.006 8
SO4- 含量	0.008 2
Fe 含量	0. 154 8
Mn 含量	0. 114 9

由表5可以看出,评价结果中 pH 值所占权重 较小,而 Fe、Mn、COD_{Mn}、NH₃-N、NO₃-N 含量这5种 评价指标所占权重较大,可以认为这5种指标是造 成该地区地下水污染的主要影响因素。

Fe、Mn 化学性质非常接近,它们在水体中往往 相互伴生。地下水中富含铁、锰主要受几个因素影 响:含水层的上覆盖岩土性质、地下水径流条件、地 下水的酸碱度以及微生物对有机质的代谢活动^[30]。 本研究采样的农用机井水来自第四系松散岩类孔隙 水含水层,含水层和上覆盖岩土富含铁锰化合物。 此外,该地区地势平缓,地下水坡度与地面坡降相 近,含水层中地下径流滞缓,上覆盖层富含有机质且 微生物较多,使地下水呈弱酸性,形成还原环境,达 到了 Fe、Mn 迁移、转化和富集的必要条件。再者, 建三江管理局地下水开发程度较高,地下水用水量 占总用水量的99.0%^[23],加速了Fe、Mn物质的溶 滤。由此可见,Fe、Mn 是本区域地下水的原生污染 物。COD_{Mn}、NH₃-N、NO₃⁻-N含量超标与农业活动密 切相关。作为重要的商品粮基地,经过长期开发,土 壤肥力已经不能满足高产需求,加之近年来暖冬天 气频繁出现,为本区域害虫的越冬繁育提供了有利 条件,造成农作物病虫害发生频繁。此种背景之下, 为保证耕地的高产,增加化肥、农药施用强度不可避 免。据 2017 年《黑龙江垦区统计年鉴》 整理数据显 示,2016年建三江管理局农药施用量达到6.0 kg/hm²,

耕地化肥施用量为177.5 kg/hm²(折纯),高于世界 平均施肥量(139.4 kg/hm²)^[31],其中氮肥施用量为 60.1 kg/hm²,磷肥施用量为45.9 kg/hm²,然而农药 利用率不到30%,氮肥利用率仅为33%,磷肥利用 率仅为43%^[32],残留的农药和化肥,在降水或灌溉 过程中下渗进入土壤,进而污染地下水体。此外,建 三江管理局畜牧业畜禽养殖污水处理不当入渗地 下,也是地下水中 COD_{Mn}、NH₃-N、NO₃⁻-N 含量超标 的部分原因。

3 讨论

为验证所建模型的稳定性和可靠性,本文构建 基于粒子群优化的 BP 神经网络模型(PSO - BP)和 BP 模型,对建三江管理局各农场地下水水质进行评价。两种模型均采用与 KHA - BP 模型相同的网络结构,即9个输入层神经元,15个隐含层神经元以及1个输出层神经元的9-15-1 网络结构,利用前文中的训练集训练 PSO - BP 模型与 BP 模型。其中PSO - BP 模型与 KHA - BP 模型同样进行多次运行以获得具有较高适应度值的最优权值和阈值。PSO - BP 模型、BP 模型、KHA - BP 模型的评价结果如表 6 所示。

由表6可以看出,BP模型、PSO-BP模型的评价结果与KHA-BP虽有差异,但总体来看建三江管理局各农场的地下水水质基本良好,均不低于Ⅲ 类水质。

		-					
农场	BP模型	PSO – BP 模型	KHA – BP 模型	农场	BP 模型	PSO – BP 模型	KHA – BP 模型
前哨	Ш	Ш	Π	青龙山	Ш	Π	Ш
二道河	П	Ι	П	前进	Ш	Ш	Ш
八五九	Ш	П	Ш	红卫	Ш	Ι	Ι
胜利	П	П	Ι	创业	Ш	Ι	Ι
前锋	П	П	П	七星	П	Ι	Ι
鸭绿河	П	П	П	大兴	Ш	П	П
洪河	П	Ι	П	勤得利	Ш	Ι	П
浓江	П	П	П				

Tab. 6 Comparison of water quality evaluation results under different methods

表6 不同方法下的水质评价结果对比

计算 BP 模型、PSO - BP 模型评价结果的各指标权重并与 KHA - BP 模型进行对比,如表 7 所示。由表 7 可以看出,3 种方法所得评价结果中指标权重较大的均为 Fe、Mn、COD_{Mn}、NH₃-N、NO₃⁻-N 含量这 5 种评价指标,KHA - BP 模型所得评价结果比较可信。

表 7 不同方法下的水质评价结果指标权重对比

Tab. 7	Compari	son of v	weight	of index i	in water
quality	evaluation	results	under	different	methods

长行		权重	
1日7小 —	BP 模型	PSO – BP 模型	KHA - BP 模型
pH 值	0.0910	0.074 8	0.0924
COD _{Mn} 含量	0. 148 1	0. 162 8	0.1701
NH3-N含量	0. 209 6	0. 234 4	0.2108
NO3-N含量	0. 197 2	0. 206 5	0. 209 8
Cl ⁻ 含量	0.0536	0.0679	0.0322
F-含量	0. 018 9	0. 031 4	0.0068
SO ₄ ²⁻ 含量	0.0095	0.0014	0.008 2
Fe 含量	0. 142 4	0. 118 2	0. 154 8
Mn 含量	0. 129 7	0. 102 6	0.1149

3.1 评价方法可靠性分析

采用区分度作为各种评价方法评价结果的可靠

性指标^[33]。区分度概念来自于考试,指试卷测试题 目对被测试者知识和能力水平的鉴别能力。本文区 分度是指各种评价方法评价结果对评价对象实际水 平的区别能力。

对于某一种评价方法的评价结果,假设有 m 个 评价对象,将其按分值 G 高低进行降序排列,每个 分值编号为 A,则函数 G = f(N)是单调递减函数,评 价结果分值最好的坐标为(G_1 ,1),分值最差的坐标 为(G_m ,m),则区分度定义为

$$D = \frac{\sum_{i=1}^{m-1} \sqrt{(G_{i+1} - G_i)^2 + (A_{i+1} - A_i)^2}}{\sqrt{(G_1 - G_m)^2 + (A_m - A_1)^2}} \quad (12)$$

$$\vec{x} \oplus \quad D \longrightarrow \vec{x} \text{ fr}$$

由于各评价方法结果极值范围不一,为使评价 结果具有可比性,需要对评价结果进行标准化,本文 将标准化处理后分值设定处于0~m之间,计算公 式为

$$G'_{i} = m \left(1 - \frac{|G_{i} - G_{1}|}{|G_{1} - G_{m}|} \right)$$
(13)

式中 G'----标准化后的评价分值

标准化后区分度的计算公式简化为

$$\frac{D}{\sum_{i=1}^{m-1} \sqrt{(G'_{i+1} - G'_i)^2 + 1^2}}{\sqrt{(m-0)^2 + (m-1)^2}} = \frac{\sum_{i=1}^{m-1} \sqrt{(G'_{i+1} - G'_i)^2 + 1}}{\sqrt{2m^2 - 2m + 1}}$$
(14)

_ ת

根据式(13)和式(14)分别计算 BP 模型、KHA -BP 模型、PSO - BP 模型区分度,如表 8 所示。对 表 8 中的区分度进行排序,得出各评价方法的可靠 性由大到小为:KHA - BP 模型、PSO - BP 模型、BP 模型。

表 8 各评价方法区分度 Tab. 8 Diversity of different evaluation methods

评价方法	BP 模型	KHA – BP 模型	PSO – BP 模型
区分度	1.0263	1.1070	1.0819

3.2 评价方法稳定性分析

进一步采用序号总和理论^[34]对3种评价方法 的稳定性进行分析。根据序号总和理论,对各种评 价方法得到的排序序号进行加和并重新排序,所得 到的序号总和排序即是相对合理的排序结果。因 此,哪种评价方法与相对合理排序结果的相关性更 高,就可以认为这种方法更合理,稳定性更好^[34]。 评价方法稳定性分析可参考文献[34-35]。3种评 价方法的排序评价结果与相对合理排序如表9所 示。

表9 各评价方法的排序评价结果与相对合理排序

Tab. 9 Ranking evaluation results and relative rational ranking of each evaluation method

	0			
农场	BP 模型	KHA – BP	PSO – BP	相对合理
		模型	模型	排序
前哨	12	7	4	8
二道河	13	15	9	14
八五九	3	3	1	1
胜利	1	8	15	10
前锋	14	6	3	7
鸭绿河	10	5	6	6
洪河	11	12	11	12
浓江	6	2	5	3
青龙山	4	4	7	4
前进	7	1	2	2
红卫	15	14	13	15
创业	9	11	12	11
七星	8	13	14	13
大兴	5	9	10	9
勤得利	2	10	8	5

计算各评价方法排序评价结果与相对合理排序的 Spearman 等级相关系数,结果如表 10 所示。

表 10 各评价方法与相对合理排序的等级相关系数

Tab. 10 Rank correlation coefficient of each evaluation

method between relative reasonable order

河位于社	DD 拱到	KHA – BP	PSO – BP
评价方法	BP 榠型	模型	模型
Spearman 等级相关系数	0. 542 9	0. 928 6	0.8000

一般认为,随着排序次数的增加,接近真实排序 结果的合理排序出现概率会更大,排序结果的稳定 性会增加^[36]。因此如果经过足够多次的排序,大多 数情况下与相对合理排序结果相关性更高的评价方 法,就可以认为这种方法相对于其他方法更稳定。 为获得多次 Spearman 等级相关系数,本文从 15 个 农场中随机选取 12 个农场重新计算各评价方法排 序评价结果与相对合理排序结果的 Spearman 等级 相关系数,得到 25 组数据,如表 11 所示。

表 11 25 次随机抽样各评价方法与相对合理排序的 等级相关系数

Tab. 11Rank correlation coefficient of each evaluationmethod between relative reasonable order in 25

random samples

		KHA – BP	PSO – BP
评价万法	BP 模型	模型	模型
	0. 489 5	0. 909 1	0.8322
	0.4301	0.9441	0.6643
	0.4441	0.9650	0.7203
	0.5664	0.9580	0.8392
	0.4615	0. 923 1	0.8252
	0. 423 1	0.8252	0. 594 4
	0.5664	0.9091	0.7692
	0. 545 5	0.8881	0.8392
	0. 559 4	0. 916 1	0.7832
	0. 531 5	0. 930 1	0.8741
	0.6853	0.9441	0.7972
a kke lat	0.7483	0.9510	0.8671
Spearman 等级	0.4615	0.9091	0.8322
相关系数	0. 531 5	0. 930 1	0.7762
	0.6084	0. 930 1	0.7972
	0. 524 5	0.8741	0.8531
	0. 573 4	0. 916 1	0.7762
	0.4196	0. 916 1	0.7413
	0.5664	0.9091	0.7483
	0. 524 5	0. 916 1	0.8252
	0.3566	0.8811	0.7343
	0.5105	0. 923 1	0.6993
	0.4825	0.9021	0.7692
	0. 489 5	0.8706	0.6084
	0. 531 5	0.9021	0.7343
求和	13.0315	22. 842 7	19.3007
平均值	0. 521 3	0.9137	0.7720

由表 11 可知, Spearman 等级相关系数之和分 别为: BP 模型 13.031 5, KHA - BP 模型 22.842 7, PSO - BP 模型 19.300 7。对比表 10、11 可以看出, 表 11 所得等级相关系数与表 10 所得等级相关系数 的平均值极为接近。25 次随机抽样的等级相关系 数区间分配情况如图 6 所示。



由图 6 可知,25 次随机抽样中,BP 模型的等级 分配系数主要集中于[0.3,0.6],处于[0.3,0.5]的 有 10 次,处于[0.5,0.6]有 12 次;PSO - BP 模型的 等级分配系数则主要处于[0.7,0.9]这一区间,处 于[0.7,0.8]的有 12 次,处于[0.8,0.9]的有 9 次; KHA - BP 模型等级分配系数全部处在[0.8,1]区 间,其中分布于[0.8,0.9]的有 5 次,处在[0.9,1] 的则多达 20 次。显然,KHA - BP 模型在区域地下 水水质评价中的稳定性更好。

4 结论

(1)利用 KHA 算法对 BP 神经网络的权值和 阈值进行优化,建立了 KHA - BP 地下水水质综 合评价模型,并以建三江管理局下辖的 15 个农 场为典型案例进行地下水水质评价,研究发现各 农场地下水水质基本良好,均处于Ⅲ类(包括Ⅲ 类)以上。

(2)建三江管理局地下水水质等级空间分布具 有一定规律,Ⅰ类水质主要集中在管理局西南位置, Ⅱ类水质主要集中在北部和南部,Ⅲ类水质主要分 布于中东部和中西部。Fe、Mn、COD_{Mn}、NH₃-N、 NO₃⁻-N 是造成该地区地下水水质污染的主要因素, 其中 Fe、Mn 主要是当地原生危害;COD_{Mn}、NH₃-N、 NO₃⁻-N 含量超标主要与大量施用化肥、农药有关。 针对上述问题,建议当地降低农业开发强度,减少化 肥、农药施用量,同时推广测土施肥技术和低毒高效 农药,推广生物固锰除锰技术进行原水处理,保证当 地饮用水安全。

(3)对比 KHA - BP 模型、PSO - BP 模型以及 BP 模型 3 种评价方法的可靠性与稳定性,结果显示 KHA - BP 模型可靠性、稳定性均优于其他两种评价 方法,更适用于区域地下水水质评价。

- 参考 文献
- 1 鞠玉梅. 绵竹市地下水资源评价[D]. 南京:河海大学,2005.
- JU Yumei. Apprasial of groundwater resources in the city of Mianzhu[D]. Nanjing: Hohai University, 2005. (in Chinese)
- 2 COLLIN M L, MELLOUL A J. Assessing groundwater vulnerability to pollution to promote sustainable urban and rural development [J]. Journal of Cleaner Production, 2003, 11(7): 727 - 736.
- 3 KIURSKI-MILOSEVIĆ J Ž, VOJINOVIĆ-MILORADOV M B, RALEVIĆ N M. Fuzzy model for determination and assessment of groundwater quality in the city of Zrenjanin, Serbia[J]. Hemijska Industrija, 2015, 69(1): 17-28.
- 4 ZHANG Q Q, WANG H W, WANG Y C, et al. Groundwater quality assessment and pollution source apportionment in an intensely exploited region of northern China[J]. Environmental Science and Pollution Research, 2017, 24(20): 16639 16650.
- 5 BOUDERBALA A. Assessment of groundwater quality and its suitability for domestic and agricultural uses in Low Isser plain, Boumedres, Algeria[J]. Arabian Journal of Geosciences, 2017, 10(15): 333-346.
- 6 WU H, CHEN J, QIAN H, et al. Chemical characteristics and quality assessment of groundwater of exploited aquifers in Beijiao water source of Yinchuan, China: a case study for drinking, irrigation and industrial purposes [J]. Journal of Chemistry, 2015: Article ID 726340, 14 pages.
- 7 KHESHTI M, KANG X N, JIARULA Y. Smooth integration of Gansu wind farm into the grid using the stator flux-oriented vector method and fuzzy logic control[J]. Arabian Journal for Science and Engineering, 2017, 42(12): 5059 - 5069.
- 8 ZOU Y, CHENG K W, CHEUNG N C, et al. A polymer-based air gap length prediction method with current injection and fuzzy logic observer[J]. IEEE Transactions on Magnetics, 2017, 53(11): 1-4.
- 9 KIRUBAKARAN B, ILANGKUMARAN M. Selection of optimum maintenance strategy based on FAHP integrated with GRA TOPSIS[J]. Annals of Operations Research, 2015, 245(1-2): 285-313.
- 10 TSAUR R C, CHEN I F, CHAN Y S. TFT LCD industry performance analysis and evaluation using GRA and DEA models [J]. International Journal of Production Research, 2017, 55(15): 4378 - 4791.
- 11 ŞENER Ş, ŞENER E, DAVRAZ A. Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW Turkey)[J]. Science of the Total Environment, 2017, 584 585: 131 144.
- 12 DE LA M C, FLORES L H, RUBIO A H, et al. Developing a water quality index (WQI) for an irrigation dam[J]. International Journal of Environmental Research and Public Health, 2017, 14(5): 439-448.
- 13 SUN Y, HU K L, FAN Z B, et al. Simulating the fate of nitrogen and optimizing water and nitrogen management of greenhouse tomato in North China using the EU - Rotate_N model[J]. Agricultural Water Management, 2013, 128: 72 - 84.
- 14 GAO D Y, KINOUCHI Y, ITO K. Neural networks for event detection from time series: a BP algorithm approach [J]. Lecture Notes in Computer Science, 2003, 2658: 784 – 793.

- 15 NAGASHINO H, ZHANG Q Y, AKUTAGAWA M, et al. Identification of number of brain signal sources using BP neural networks[J]. Lecture Notes in Computer Science, 2004, 3214: 1074 – 1080.
- 16 GANDOMI A H, ALAVI A H. Krill herd: a new bio-inspired optimization algorithm [J]. Communications in Nonlinear Science and Numerical Simulation, 2012, 17(12): 4831-4845.
- 17 黄璇,郭立红,李姜,等. 磷虾群算法优化支持向量机的威胁估计[J]. 光学精密工程,2016,24(6):1448-1455. HUANG Xuan, GUO Lihong, LI Jiang, et al. Threat assessment of support vector machine optimized by krill herd algorithm[J]. Optics and Precision Engineering, 2016, 24(6): 1448-1455. (in Chinese)
- 18 李丛, 胡文军, 丁勇, 等. 基于改进磷虾群优化的中心极大化 KFCM 算法在 IDS 的应用[J]. 计算机应用研究, 2016, 33(2):507-512.

LI Cong, HU Wenjun, DING Yong, et al. Application of center maximization KFCM algorithm based on improved krill herd in intrusion detection system[J]. Application Research of Computer, 2016, 33(2): 507-512. (in Chinese)

- 19 孔祥鑫,周炜,王晓丹,等. 基于改进磷虾群算法的 SVDD 参数优化[J]. 计算机工程与应用,2017,53(22):137-142,216. KONG Xiangxin, ZHOU Wei, WANG Xiaodan, et al. Parameter optimization for SVDD based on improved krill herd algorithm [J]. Computer Engineering and Applications, 2017, 53(22): 137-142, 216. (in Chinese)
- 20 牛培峰,陈科,马云鹏,等. 基于磷虾群算法的汽轮机热耗率建模应用[J]. 动力工程学报,2016,36(10):781-787. NIU Peifeng, CHEN Ke, MA Yunpeng, et al. Modelling of turbine heat rate based on krill herd algorithm and its application[J]. Journal of Chinese Society of Power Engineering, 2016, 36(10): 781-787. (in Chinese)
- 21 LIU M, LIU D, LIU L. Complexity research of regional groundwater depth series based on multiscale entropy: a case study of Jiangsanjiang Branch Bureau in China[J]. Environmental Earth Sciences, 2012, 70(1): 353 - 361.
- 22 FU Q, LIU D, LI T X, et al. EMD RBFNN coupling prediction model of complex regional groundwater depth series: a case study of the Jiansanjiang Administration of Heilongjiang land reclamation in China[J]. Water, 2016, 8(8): 340 – 355.
- 23 付强,李国良. 三江平原地下水埋深时间序列的混沌研究[J]. 水土保持研究,2008,15(3):31-34. FU Qiang, LI Guoliang. Chaos studies on series of groundwater depth in Sanjiang Plain [J]. Research of Soil and Water Conservation, 2008, 15(3): 31-34. (in Chinese)
- 24 YI J Q, WANG Q, ZHAO D B. BP neural network prediction-based variable-period sampling approach for networked control systems[J]. Applied Mathematics and Computation, 2007, 185(2): 976-988.
- 25 陈英义,程倩倩,成艳君,等. 基于 GA BP 神经网络的池塘养殖水温短期预测系统[J/OL]. 农业机械学报,2017,48(8): 172 - 178. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1& file_no = 20170819&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2017.08.019. CHEN Yingyi, CHENG Qianqian, CHENG Yanjun, et al. Short-term prediction system of water temperature in pond aquaculture based

on GA – BP neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(8): 172 – 178. (in Chinese) JENSI R, JIJI G W. An improved krill herd algorithm with global exploration capability for solving numerical function optimization

- 26 JENSI R, JIJI G W. An improved krill herd algorithm with global exploration capability for solving numerical function optimization problems and its application to data clustering[J]. Applied Soft Computing, 2016, 46: 230 – 245.
- 27 REN Y T, QI H, HUANG X, et al. Application of improved krill herd algorithms to inverse radiation problems [J]. International Journal of Thermal Sciences, 2016, 103: 24 – 34.
- 28 刘明寿,马峥,潘云涛,等. 学术类科技期刊影响力归一化法评判定量模型的构建[J]. 编辑学报,2004,16(6):405-406. LIU Mingshou, MA Zheng, PAN Yuntao, et al. Establishment of normalization model for quantitatively evaluating impact power of university's academic journals[J]. Acta Editologica, 2004, 16(6): 405-406. (in Chinese)
- 29 隋惠惠. 基于 BP 神经网络的短期电力负荷预测的研究[D]. 哈尔滨:哈尔滨工业大学,2015. SUI Huihui. Research on short-term electric load forecasting based on BP neural network [D]. Harbin: Harbin Institute of Technology, 2015. (in Chinese)
- 30 刘成才. 佳木斯地区高铁高锰地下水生物净化工艺效能与应用研究[D]. 哈尔滨:哈尔滨工业大学,2013. LIU Chengcai. Biological purification of a typical groundwater with high concentration of iron and manganese in Jiamusi and engineering application[D]. Harbin: Harbin Institute of Technology, 2013. (in Chinese)
- 31 王艳语,苗俊艳. 世界及我国化肥施用水平分析[J]. 磷肥与复肥,2016,31(4):22-23. WANG Yanyu, MIAO Junyan. Analysis on fertilizer application level in world and China[J]. Phosphate & Compound Fertilizer, 2016, 31(4): 22-23. (in Chinese)
- 32 刘加海. 三江平原农业非点源污染现状分析及防治对策研究[J]. 水利规划与设计,2012(4):5-8. LIU Jiahai. Study on agricultural non-point source pollution and countermeasures in Sanjiang Plain[J]. Water Resources Planning and Design, 2012(4):5-8. (in Chinese)
- 33 俞立平,潘云涛,武夷山. 比较不同评价方法评价结果的两个新指标:以《泰晤士报高等教育副刊》大学排名为例[J]. 南京师范大学学报(自然科学版),2008,31(3):135-140. YU Liping, PAN Yuntao, WU Yishan. Two new indicators to compare different evaluation methods effect—based on Times Higher-QS Word University Rankings[J]. Journal of Nanjing Normal University(Natural Science Edition), 2008, 31(3):135-140. (in Chinese)
- 34 吕筱萍. "序号总和理论"对现行工业企业经济效益综合指数的评价及改进[J]. 管理工程学报,1996,10(1):61-65. LÜ Xiaoping. The evaluation and improvement of the current composite index of economic returns in industrial enterprises according to "the theory of the total of orders" [J]. Journal of Industrial Engineering and Engineering Management, 1996, 10(1):61-65. (in Chinese)
- 35 郭晓魁. 基于序号总和理论的御河大桥交通量预测研究[J]. 物流科技,2011(4):11-12. GUO Xiaokui. Study on traffic volume forecasting of imperial river bridge based on the theory of serial number summation[J]. Logistics Sci-Tech, 2011(4):11-12. (in Chinese)
- 36 杭州商学院省自然科学基金课题组."序号总和理论"及其在综合经济效益排序中的应用[J].数量经济技术经济研究, 1996(1):59-62.

Research group of provincial natural science foundation in Hangzhou Institute of Commerce. Application of "the theory of the total of orders" in comprehensive economic benefit ranking [J]. The Journal of Quantitative & Technical Economics, 1996(1): 59 – 62. (in Chinese)