基于 PCA-WOA-ELM 的矿井突水水源识别模型研究

姜兰兰，鲁海峰

（安徽理工大学 地球与环境学院，安徽 淮南 232000）

摘要：准确快速识别突水水源对于煤矿水害防治尤为重要，极限学习机（ELM）被广泛应用于突水水源识别领域。为了弥补 ELM 的不足、提高识别精度与效率，文章提出了一种主成分分析（PCA）、鲸鱼优化算法（WOA）与 ELM 相融合的突水水源识别模型（PCA-WOA-ELM 模型）。以新集二矿为例，选择影响安全回采的推覆体寒灰水、太灰水以及奥灰水作为突水水源，以水样中六大常规离子作为识别指标，首先采用 PCA 降低数据维度，其次基于 Matlab 运用 WOA 寻优 ELM 参数对测试样本进行识别，最后将 PCA-WOA-ELM 模型识别结果与其他识别模型进行对比。结果表明，该模型能够提取数据的主要特征，弥补了 ELM 初始参数随机选择的不足，识别效率与精度优于 PCA-ELM，WOA-ELM，PCA-PSO-ELM 以及 PCA-GA-ELM 这 4 种模型。

关键词：矿井突水水源；主成分分析；鲸鱼优化算法；极限学习机

中图分类号：X824；TD745^{+}.21 **文献标志码**：A **文章编号**：1672－349X(2025)06－0044－09

DOI：10.16160/j.cnki.tsxyxb.2025.06.008

Research on a Mine Water Inrush Source Identification Model Based on PCA-WOA-ELM

JIANG Lanlan, LU Haifeng

(School of Earth and Environment, Anhui University of Science and Technology, Huainan 232000, China)

Abstract: Accurate and rapid identification of water inrush sources is critical for the prevention and control of water inrush in coal mines, and extreme learning machine (ELM) is widely used in this field. In order to make up for the shortcomings of ELM and improve the accuracy and efficiency of recognition, a water source identification model (PCA-WOA-ELM model) is proposed, which combines principal component analysis(PCA), whale optimization algorithm (WOA) and ELM. With Xinji No. 2 Mine as an example, three types of water inrush sources of nappe Cambrian Limestone Water, Taiyuan Limestone Water, and Ordovician Limestone Water that affect the safe mining are selected, and the six conventional ions in the water sample are used as identification indicators. PCA is first used to reduce dimensionality; Then WOA is employed via Matlab to optimize ELM parameters for testing

收稿日期：2024－10－14

基金项目：国家重点研发计划项目(2022YFF1303302)

作者简介：姜兰兰(2001－)，女，河南洛阳人，硕士研究生，研究方向为水文地球化学；
鲁海峰(1982－)，男，安徽合肥人，教授，博士，研究方向为矿山地质灾害与防治。

sample recognition. Finally, results from the PCA-WOA-ELM model are compared with other models. The results show that the model can extract the main features of the data, make up for the shortcomings of random selection of initial parameters of ELM, and the recognition efficiency and accuracy are better than those of PCA-ELM, WOA-ELM, PCA-PSO-ELM and PCA-GA-ELM.

Key Words: mine water inrush source; principal component analysis; whale optimization algorithm; extreme learning machine

0　引言

近年来，我国煤炭资源逐渐向深部开采，使得煤矿水害事故频发，呈现出“预测难、影响大、伤亡重”的特点。据统计，2001－2022 年我国共发生煤矿安全事故 33 546 起，死亡 56 224 人，其中水害事故共 1 103 起，死亡 4 667 人，造成了严重的经济损失和社会影响[1]。因此快速准确识别矿井突水水源对于煤矿防水治水工作以及保障生产安全尤为重要。

矿井突水水源识别方法有水温法、水化学分析法和数理分析法等[2]。其中，水温法是指通过对比突水点的水温和具有突水隐患的含水层的水温来对矿井突水水源进行初步预测。然而由于受地质构造的作用或者地下水补径排的影响，水温在空间上存在较大差异，单靠水温识别是否为矿井突水水源效果不佳，故该方法仅适用于水文地质较为简单的区域[3－4]。水化学分析法是对水中的常量离子和水质综合指标进行检测，并借助图示法对检测结果进行可视化分析，从而判断矿井水的来源[5]。水化学分析法对矿井突水水源识别有着重要意义，但这种方法只适用于水质类型简单的区域，对于水质类型复杂的区域则需要综合运用多种方法，以精确识别矿井突水来源[6]。

近年来，许多学者将水化学指标与数理分析法相结合建立矿井突水水源识别模型。比如，孙延辉等[7]构建了一种基于 Piper 三线图、主成分分析(Principal Component Analysis, PCA)和 Fisher 判别法的一种矿井突水水源判别模型，并将其应用于临涣矿区，判别准确率高达 100％。据棋定等[8]以潘二矿 24 个水样数据为研究对象，将 PCA 与 Bayes 判别法相结合构建矿井突水水源判别模型，判别准确率达 90.9％。同时，模糊数学评判法[9－10]、灰色关联度法[11]、神经网络法[12]、可拓识别法[13]、支持向量机[14]等也在矿井突水水源识别领域得到了广泛应用，但这些方法也存在不足，如模糊数学评判法需要人为划分隶属度，主观性较大；灰色关联度法因子权重较难确定；神经网络法需要大量数据作为支撑，而实际收集数据的量却难以达到神经网络的训练要求；可拓识别法理论体系尚不完善，方法论相对复杂，对数据质量依赖性较强；支持向量机在处理多分类问题上存在局限性。

基于此，以新集二矿为例，构建一种基于主成分分析(PCA)、鲸鱼优化算法(Whale Optimization Algorithm, WOA)和极限学习机(Extreme Learning Machine, ELM)的矿井突水水源识别模型(PCA-WOA-ELM 模型)，即在广泛应用于突水水源识别领域的 ELM 基础上，利用 PCA 降低数据间的相关性，对突水水源指标进行筛选；并利用 WOA 算法对隐藏层的连接权值和阈值进行寻优，避免模型陷入局部最优解。

1　PCA-WOA-ELM 模型的原理与构建

1.1　主成分分析

PCA 是一种广泛应用于数据降维的统计方法。它的核心思想是将一组可能存在相关性的高维变量，通过正交变换转换成一组线性不相关的低维变量，同时保留原始数据的大部分信息[15]。主要步骤如下：假设有一个数据集，包含 n 个样本，每个样本有 p 个特征，将数据集表示为一个 $n\times p$ 的矩阵 $\boldsymbol{X}$：

$$\boldsymbol{X}=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}。\tag{1}$$

由于不同特征的量纲不同，因此要对数据进行标准化处理，标准化公式如下：

$$X_{ij}=\frac{x_{ij}-\mu_{ij}}{\sigma_{ij}}。\quad(2)$$

式中，X_{ij} 为标准化后数据，μ_{ij} 和 σ_{ij} 分别为特征 x_{ij} 的均值和标准差，其中 $i=1,2,3,\cdots,n$，$j=1,2,3,\cdots,p$。

然后计算标准化后数据矩阵的协方差矩阵，求解特征值和特征向量，公式如下：

$$\boldsymbol{C}=\frac{1}{n-1}(\boldsymbol{X}_{std}{}^{\mathrm{T}}-\boldsymbol{X}_{std})。\quad(3)$$

式中，$\boldsymbol{X}_{std}$，$\boldsymbol{X}_{std}{}^{\mathrm{T}}$ 分别为标准化后数据矩阵及其转置，n 为样本数量，$\boldsymbol{C}$ 为一个 $p\times p$ 的矩阵。

将 $\boldsymbol{C}$ 的特征值按从大到小排列，根据累计方差贡献率≥85%的原则，选择前 k 个主成分，这些主成分对应的特征向量为 $\boldsymbol{\gamma}_i$，组成投影矩阵 $\mathbf{V}=(\boldsymbol{\gamma}_1,\boldsymbol{\gamma}_2,\cdots,\boldsymbol{\gamma}_k)$。

最后将标准化后的数据投影到选定的主成分上，得到降维后的数据，公式如下：

$$\boldsymbol{Z}=\boldsymbol{X}\mathbf{V}。\quad(4)$$

式中，$\boldsymbol{Z}$ 为降维后的数据矩阵。

1.2 鲸鱼优化算法

WOA 是由澳大利亚格里菲斯大学的 Mirjalili 等于 2016 年提出的一种元启发式优化算法[16]。发明该算法的灵感起源于自然界中座头鲸群体的狩猎行为，通过鲸鱼群的自组织和自适应来寻找最优解（猎物）。该算法结构简单、收敛速度快，主要由三部分动作组成：包围猎物、发泡网攻击猎物、随机搜索猎物。

1.2.1 包围猎物

WOA 将当前迭代的假设最优解作为目标猎物，鲸鱼在获取猎物位置后，会朝着猎物进行包围收缩。此时鲸鱼位置更新的公式如下：

$$\vec{D}=|\vec{C}\cdot\vec{X}^*(t)-\vec{X}(t)|；\quad(5)$$

$$\vec{X}(t+1)=\vec{X}^*(t)-\vec{A}\cdot\vec{D}。\quad(6)$$

式中，t 表示当前迭代次数，$\vec{D}$ 为鲸鱼与猎物之间的距离，$\vec{A}$ 和 $\vec{C}$ 表示向量系数，$\vec{X}^*(t)$ 表示当前最佳候选鲸鱼的位置向量，$\vec{X}(t)$ 表示当前迭代次数中鲸鱼的位置向量。

$\vec{A}$ 和 $\vec{C}$ 由式（7）和式（8）得出：

$$\vec{A}=2\vec{a}\,\vec{r}_1-\vec{a}；\quad(7)$$

$$\vec{C}=2\vec{r}_2。\quad(8)$$

式中，$\vec{r}_1$ 和 $\vec{r}_2$ 为（0，1）中的随机向量，$\vec{a}$ 的值从 2 线性下降到 0，$\vec{a}$ 的定义式如下：

$$\vec{a}=2-2\,\frac{t}{T_{\max}}。\quad(9)$$

式中，$T_{\max}$ 为初始设置的最大迭代次数。

1.2.2 发泡网攻击猎物

鲸鱼捕食时，采用螺旋运动的方式向猎物靠近，其数学公式如下：

$$\vec{X}(t+1)=\vec{D}'\cdot e^{bl}\cdot\cos(2\pi l)+\vec{X}^*(t)；\quad(10)$$

$$\vec{D}'=|\vec{X}^*(t)-\vec{X}(t)|。\quad(11)$$

式中，$\vec{D}'$ 表示鲸鱼与当前猎物之间的距离，b 是一个表示螺旋形状的参数，l 表示（−1，1）区间均匀分布的随机数。

在 WOA 中，公式（10）并非单独使用，而是与公式（6）通过概率随机切换，具体公式如下：

$$\vec{X}(t+1)=\begin{cases}\vec{X}^*(t)-\vec{A}\cdot\vec{D},p<0.5\\ \vec{D}'\cdot e^{bl}\cdot\cos(2\pi l)+\vec{X}^*(t),p\geqslant 0.5\end{cases}。\quad(12)$$

式中，p 为（0，1）之间的随机数。

1.2.3 随机搜索猎物

WOA 算法能够探索全局最优解，并跳出局部最优，其数学公式如下：

$$\vec{D}=|\vec{C}\cdot\overrightarrow{X_{rand}}(t)-\vec{X}(t)|；\quad(13)$$

$$\vec{X}(t+1)=\overrightarrow{X_{rand}}(t)-\vec{A}\cdot\vec{D}。\quad(14)$$

式中，$\overrightarrow{X_{rand}}$ 是随机选取的一只鲸鱼的位置，当 $|\vec{A}|\geqslant 1$ 时，算法采用随机搜索方式，其他鲸鱼均向该随机选取的鲸鱼个体的位置更新自身位置，算法通过这种随机搜索方式扩大探索范围，助力种群跳出局部最优；当 $|\vec{A}|<1$ 时，算法则切换为包围捕食方式，引导所有鲸鱼向当前迭代过程中的最优解位置靠拢，实现对潜在全局最优区域的聚焦开发。

1.3 极限学习机

ELM 是由南洋理工大学的黄广斌等在 2004 年提出的一种基于前馈神经网络构建的机器学习系统[17]，它的基本构架由三部分组

成,即输入层、隐藏层、输出层。相对于传统的单隐层前馈神经网络,ELM通过随机选择隐藏层的连接权值和阈值分析得到输出层权重,在保证精度的情况下拥有更快的学习速度,并能有效防止算法陷入局部最优问题[18]。ELM神经网络结构图如图1所示。

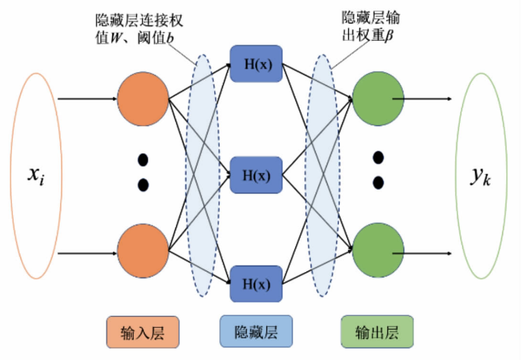


图1　ELM神经网络结构图

在ELM神经网络中,假设有任意样本函数(x_i,t_i),其中x_i和t_i的取值为:

$$x_i=[x_{i1},x_{i2},\cdots,x_{in}]^{\mathrm{T}}\in \mathrm{R}^n; \tag{15}$$

$$t_i=[x_{i1},x_{i2},\cdots,x_{im}]^{\mathrm{T}}\in \mathrm{R}^m。 \tag{16}$$

式中,x_i为第i个输入样本,t_i为第i个样本对应的目标输出,n为输入层神经元数,m为输出层神经元数。

对于一个有l个隐藏层神经元的ELM神经网络,其输出解可表示为:

$$y_{i,k}=\sum_{j=1}^{l}\beta_{j,k}\cdot g(\omega_j x_i+b_j)。 \tag{17}$$

式中,$\beta_{j,k}$为第j个隐藏层节点到第k个输出层节点的权重,ω_j为第j个隐藏层节点的权值向量,b_j为第j个隐藏层节点的阈值,$g(x)$为激活函数,可用Sigmoid函数进行激活。

将式(17)用矩阵表示为:

$$\boldsymbol{Y}=\boldsymbol{H\beta}。 \tag{18}$$

$\boldsymbol{H}$为隐藏层期望输出矩阵,$\boldsymbol{\beta}$为输出层权重矩阵,$\boldsymbol{Y}$为理想输出矩阵,其中$\boldsymbol{H}$的表达式如下:

$$\boldsymbol{H}=\begin{bmatrix} g(\omega_1 x_1+b_1) & \cdots & g(\omega_l x_1+b_l) \\ \vdots & \vdots & \vdots \\ g(\omega_1 x_n+b_1) & \cdots & g(\omega_l x_n+b_l) \end{bmatrix}_{n\times l}。 \tag{19}$$

根据Moore-Penrose广义逆矩阵定义,求解广义逆矩阵$\hat{\boldsymbol{\beta}}$,如下式:

$$\hat{\boldsymbol{\beta}}=\boldsymbol{H}^{+}\boldsymbol{Y}。 \tag{20}$$

式中,$\boldsymbol{H}^{+}$为$\boldsymbol{H}$的Moore-Penrose广义逆矩阵。

1.4　PCA-WOA-ELM模型构建流程

基于PCA-WOA-ELM的矿井突水水源识别模型算法流程如图2所示。

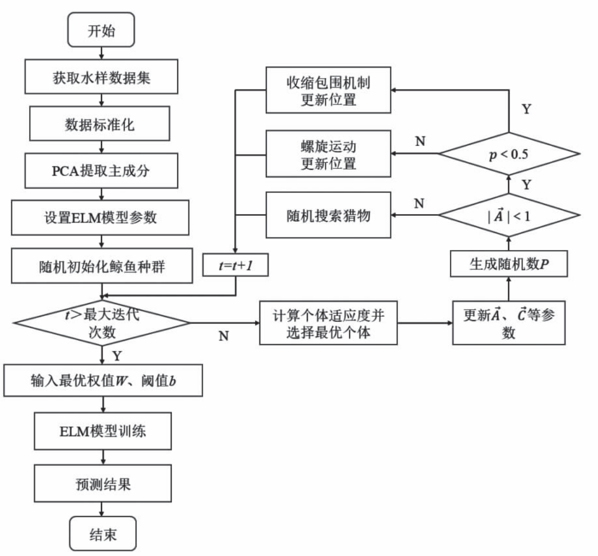


图2　矿井突水水源识别模型算法流程

2　PCA-WOA-ELM模型实际应用

2.1　研究区水文地质概况

新集二矿位于安徽省淮南市凤台县城西12 km处,地处淮南复向斜的谢桥向斜南翼、颖凤阜凤推覆构造的中段。井田含水层由上至下依次为新生界松散含水层、推覆体下元界片麻岩裂隙承压含水层、推覆体寒武系灰岩岩溶裂隙承压含水层、推覆体下夹片岩溶裂隙承压含水层、二叠系煤层顶板砂岩裂隙承压含水层,底部为石炭系太原组灰岩含水层及奥陶系岩溶裂隙承压含水层[19]。

矿井主采煤层为1煤组、6-1煤层、8煤层、11-2煤层、13-1煤层,被新生界松散含水层和推覆体所覆盖。井田内新生界松散含水层被外来阜凤推覆体阻隔,与含煤地层及下伏各含水层水力联系弱,接受垂直补给少,因此新生界中下伏含水层组不会对矿井充水产生威胁;推覆

体下元界片麻岩裂隙承压含水层富水性弱，在留设防水煤柱情况下，对 13-1 煤层、11-2 煤层开采影响程度一般；推覆体寒武系灰岩岩溶裂隙承压含水层因上部受风化作用、底部受断层影响，富水性弱至中等，对 13-1 煤层开采有影响；推覆体下夹片岩溶裂隙承压含水层富水性不均匀；二叠系煤层顶板砂岩裂隙承压含水层富水性不均匀，以淋渗水为主，少见集中突水，并且以静储量为主，对煤层开采影响不大；石炭系太原组灰岩含水层为 1 煤组底板直接充水含水层，在高水压、薄底板条件下极易引发底板突水；奥陶系岩溶裂隙承压含水层一旦发生突水，则会来势凶猛、出水量大且持续时间长，容易引发突水事故[20－21]。

2.2　水化学特征

基于实测数据，采用 Origin 软件对矿井影响较大的 3 个含水层——推覆体寒武系灰岩岩溶裂隙承压含水层（推覆体寒灰水）、石炭系太原组灰岩含水层（太灰水）、奥陶系岩溶裂隙承压含水层（奥灰水）进行水质分析，结果如图 3 所示。根据 Piper 三线图可以看出，阳离子分布较为集中，阴离子较为分散；3 类水样中阳离子 $Na^{+}+K^{+}$ 的含量较高，阴离子含量 $Cl^{-}>HCO_3^{-}>SO_4^{2-}$；推覆体寒灰水的水质类型主要有 Cl-Na，Cl-HCO_3-Na；太灰水的水质类型主要有 Cl-HCO_3-Na，Cl-Na，HCO_3-Cl-Na 等；奥灰水的水质类型主要有 Cl-Na，Cl-HCO_3-Na。另外，推覆体寒灰水和太灰水部分水样中还混有 SO_4^{2-}，这可能来源于硫酸盐沉积物的溶解。从以上可以看出，3 个含水层水样的水质类型非常相似，由此根据水化学特征识别水源类型较为困难。

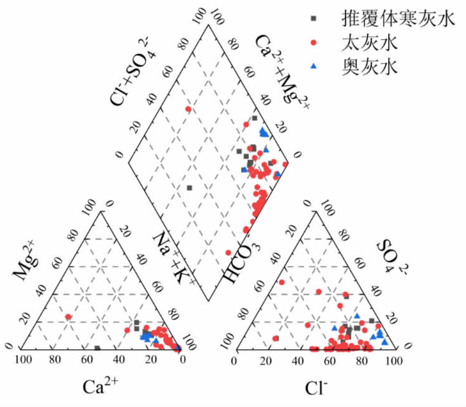


图 3　研究区水样 Piper 三线图

2.3　数据处理

选取新集二矿推覆体寒灰水水样、太灰水水样、奥灰水水样共 87 组实测数据，按 7∶3 的比例划分为训练样本（60 组）和测试样本（27 组），以六大常量离子（$Na^{+}+K^{+}$，Ca^{2+}，Mg^{2+}，Cl^{-}，SO_4^{2-}，HCO_3^{-}）作为识别指标。训练样本实测数据见表 1，测试样本实测数据见表 2。

表 1　新集二矿实测数据(训练样本)　mg/L

水样组	$Na^{+}+K^{+}$	Ca^{2+}	Mg^{2+}	Cl^{-}	SO_4^{2-}	HCO_3^{-}	类别
1	448.50	57.55	32.45	474.32	189.43	331.95	推覆体寒灰水
2	529.00	83.37	36.94	614.65	76.85	528.18	推覆体寒灰水
3	549.75	95.39	49.70	755.09	238.31	282.09	推覆体寒灰水
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	1 179.71	10.02	5.11	1 029.82	7.68	860.38	太灰水
12	1 013.00	16.03	8.75	932.34	30.74	848.18	太灰水
13	797.97	18.84	12.28	707.23	77.81	750.55	太灰水
14	664.50	99.40	73.39	1 063.50	23.05	433.24	太灰水
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
58	1 541.42	325.60	92.26	2 747.50	195.66	239.41	奥灰水
59	1 539.27	321.60	95.89	2 733.42	218.28	235.72	奥灰水
60	1 419.40	279.16	68.46	2 411.31	487.33	190.99	奥灰水

水化学指标相关性矩阵如表 3 所示。从中可以看出，6 种指标之间相关性明显。例如，$Na^{+}+K^{+}$ 与 Cl^{-} 存在显著的相关性，相关性系数为 0.840，说明二者有着相似的来源，可能来

源于岩盐溶解；Ca^{2+} 与 Mg^{2+}，Cl^{-}，SO_4^{2-} 的相关性系数依次为 0.902，0.854，0.593，其中 Ca^{2+} 与 Mg^{2+} 相关性系数接近于 1，存在极强的相关性，表明各含水层存在较明显的白云岩溶解作用，Ca^{2+} 与 Cl^{-}，SO_4^{2-} 相关性高可能是由于含钙矿物以及硫酸盐岩的溶解。由上可见，各指标之间存在信息重叠，如果直接采用 6 种指标进行水源识别，则会影响识别的准确性，因此数据降维非常有必要。

表 2　新集二矿实测数据(测试样本)　　mg/L

水样组	$Na^{+}+K^{+}$	Ca^{2+}	Mg^{2+}	Cl^{-}	SO_4^{2-}	HCO_3^{-}	类别
1	485.72	68.14	33.55	582.27	148.41	370.70	推覆体寒灰水
2	562.41	93.54	43.72	744.37	230.44	305.01	推覆体寒灰水
3	483.47	65.53	31.72	586.70	111.43	387.48	推覆体寒灰水
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6	121.00	2.40	0.15	58.49	51.49	39.66	太灰水
7	1 179.08	12.42	5.23	1 057.30	13.93	866.48	太灰水
8	97.18	4.03	2.91	58.91	66.25	24.84	太灰水
9	1 287.33	13.83	6.08	905.75	8.65	1 644.49	太灰水
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
25	1 558.10	316.33	95.28	2 803.26	206.92	246.16	奥灰水
26	1 361.83	254.18	79.44	2 202.47	356.92	252.41	奥灰水
27	1 465.94	267.56	92.51	2 377.25	393.77	265.27	奥灰水

表 3　水化学指标相关性矩阵

水化学指标	$Na^{+}+K^{+}$	Ca^{2+}	Mg^{2+}	Cl^{-}	SO_4^{2-}	HCO_3^{-}
$Na^{+}+K^{+}$	1					
Ca^{2+}	0.492**	1				
Mg^{2+}	0.406**	0.902**	1			
Cl^{-}	0.840**	0.854**	0.798**	1		
SO_4^{2-}	0.146	0.593**	0.549**	0.413**	1	
HCO_3^{-}	0.320**	−0.495**	−0.552**	−0.201	−0.543**	1

注：** 表示在 0.01 级别(双尾)相关性显著。

本文使用 SPSS27 对样本实测数据进行降维。经主成分分析提取样本特征值，得到各主成分方差贡献率占比，如图 4 所示。从中可知，前 3 种成分累计方差贡献率达 95.79%，从第 4 个成分开始，累计方差贡献率增长缓慢，因此选取前 3 种成分为主成分，可以解释 90% 以上的信息，同时实现有效降维。令 $Na^{+}+K^{+}$，Ca^{2+}，Mg^{2+}，Cl^{-}，SO_4^{2-}，HCO_3^{-} 标准化后数据分别为 X_1，X_2，X_3，X_4，X_5，X_6，根据主成分得分系数矩阵，可得到新识别指标 Y_1，Y_2，Y_3 和上述水化学指标之间的关联矩阵(即相关系数矩阵与标准化后的数据矩阵对应相乘)，其中 Y_1，Y_2，Y_3 代表新的主成分因子，表达式如下：

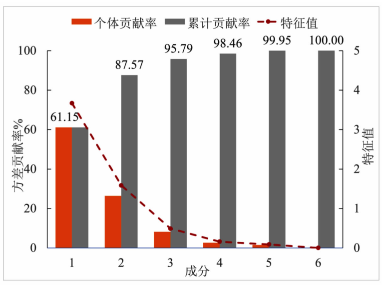


图 4　各主成分方差贡献率占比

$$\begin{cases} Y_1 = 0.158X_1 + 0.261X_2 + 0.253X_3 + \\ \quad 0.248X_4 + 0.185X_5 - 0.138X_6 \\ Y_2 = 0.498X_1 - 0.018X_2 - 0.068X_3 + \\ \quad 0.242X_4 - 0.255X_5 + 0.505X_6 \\ Y_3 = 0.244X_1 - 0.229X_2 - 0.453X_3 - \\ \quad 0.159X_4 + 1.238X_5 + 0.392X_6 \end{cases} \text{。}$$

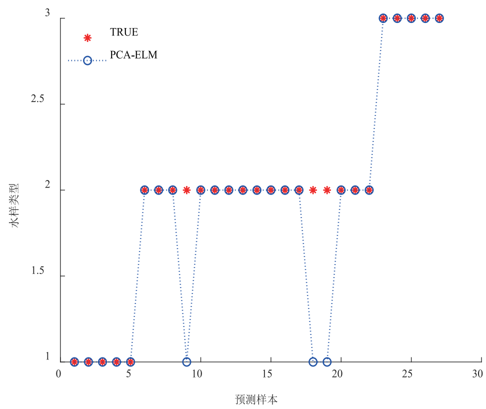
2.4 突水水源识别

设置 WOA 优化算法最大迭代次数为 100，初始种群规模为 30；设置 ELM 神经网络隐藏层神经元数为 20，激活函数选取 Sigmoid 函数；采用 Matlab R2021b 编写程序代码；将经主成分分析后得到的新指标 Y_1，Y_2，Y_3 作为输入向量，输出向量为水样类型，代入 PCA 降维后的数据，建立 PCA-WOA-ELM 模型并对模型进行训练；将测试样本代入训练好的模型中进行分类识别。同时，将未经 PCA 处理的标准化后的数据代入 WOA-ELM 模型中，并将其预测结果与 PCA-WOA-ELM 模型的预测结果进行对比。

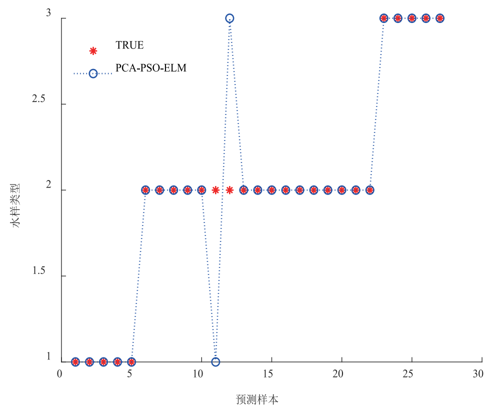
为了验证 WOA 优化算法在本模型中的优势，建立无优化算法模型(PCA-ELM)，以及利用粒子群优化算法(Particle Swarm Optimization，PSO)、遗传算法(Genetic Algorithms，GA)对 ELM 进行参数寻优分别建立 PCA-PSO-ELM 模型和 PCA-GA-ELM 模型，对同一组测试样本进行预测，同样将新指标 Y_1，Y_2，Y_3 作为输入向量，输出向量为水样类型，最后将这些模型的预测结果与 PCA-WOA-ELM 模型的预测结果进行对比。模型具体参数设置如表 4 所示，各模型的预测结果如图 5 所示。

表 4 各模型具体参数

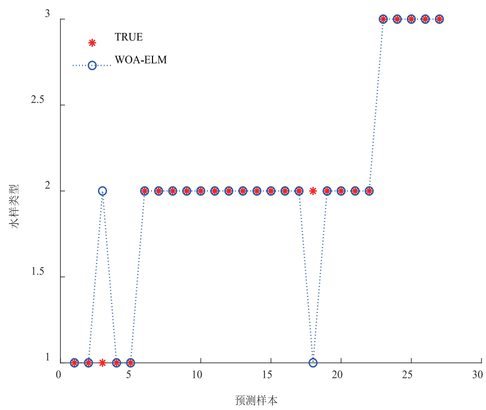
模型	参数设置
PCA-ELM	隐藏层节点数 18
PCA-PSO-ELM	迭代次数 100；粒子个数 30；隐藏层节点数 12；学习因子 $C_1=2$，$C_2=2$
PCA-GA-ELM	迭代次数 100；种群数量 30；隐藏层节点数 30；$Pc=0.7$，$Pm=0.01$
PCA-WOA-ELM(WOA-ELM)	迭代次数 100；种群数量 30；隐藏层节点数 20



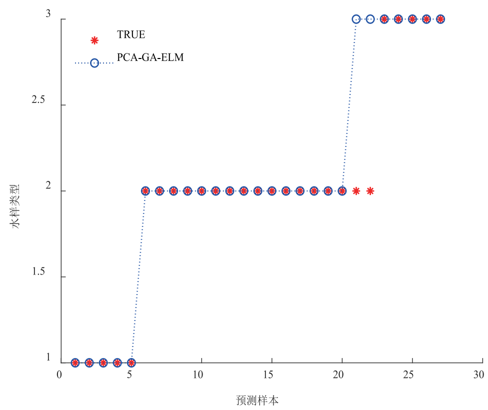
(a)PCA-ELM 测试集(准确率：88.889%)



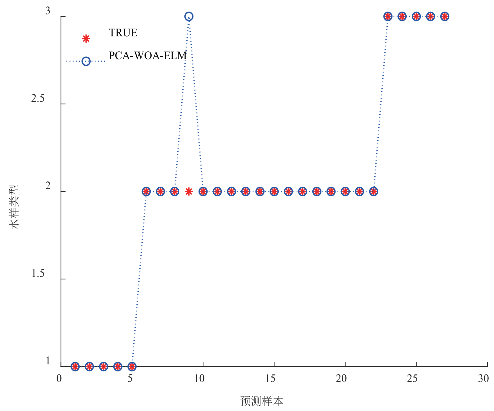
(c)PCA-PSO-ELM 测试集(准确率：92.593%)



(b)WOA-ELM 测试集(准确率：92.593%)



(d)PCA-GA-ELM 测试集(准确率：92.593%)



(e)PCA-WOA-ELM 测试集(准确率:96.296%)

图 5　各模型预测结果

对于 60 组训练样本，各模型训练效果均较好，训练集识别准确率分别为 91.667%，98.333%，95.000%，98.333%，98.333%；对于 27 组测试样本，由图 5 可知，PCA-WOA-ELM 模型预测准确率最高，为 96.296%，与传统模型 PCA-ELM 模型相比准确率提高了7.407%，与其他 3 种模型相比提高了 3.703%。对于测试集，PCA-ELM 模型误判了 3 个样本，分别为样本 9、样本 18 和样本 19；WOA-ELM 模型误判了 2 个样本，分别为样本 3 和样本 18；PCA-PSO-ELM 模型与 PCA-GA-ELM 模型均误判了 2 个样本；PCA-WOA-ELM 模型对样本 9 发生了误判。其中误判样本大多数是太灰水，其含水层构造条件复杂，裂隙发育不均一，水样的混合效应给水样识别带来了一定困难。

由以上可知，PCA-WOA-ELM 模型预测效果最好，原因为：PCA 可消除数据的冗余信息，提升识别速度与准确率；同时，针对 ELM 模型预测性能受隐藏层连接权值和阈值影响的问题，WOA 优化算法能对 ELM 分类器的初始权值和阈值进行寻优，使模型在训练过程中更加稳定，减少过拟合或欠拟合的风险，提高预测精度。

3　结语

(1)对新集二矿水样数据进行了主成分分析，提取出 3 个主成分(Y_1，Y_2，Y_3)，有效降低了数据集的维度，消除了冗余信息；将降维后的数据代入 PCA-WOA-ELM 模型，使模型更快地收敛到最优解，提高了识别效率。

(2)采用鲸鱼优化算法(WOA)对极限学习机(ELM)隐藏层的连接权值和阈值进行寻优，克服了极限学习机初始参数随机取值的缺点，提高了模型的稳定性和预测精度。

(3)将 PCA-WOA-ELM 模型与 PCA-ELM，WOA-ELM，PCA-PSO-ELM 以及 PCA-GA-ELM 模型进行对比，结果证明，该模型预测精度最高。

参考文献：

[1]　孙文洁，李文杰，宁殿艳，等. 我国煤矿水害事故现状、预测及防治建议[J]. 煤田地质与勘探，2023，51(12)：185－194.

[2]　崔中良，黄保胜，龙向前. 矿井突水水源判别研究现状及发展趋势[J]. 中国钨业，2018，33(4)：42－50.

[3]　马雷，钱家忠，赵卫东. 基于 GIS 和水质水温的矿井突水水源快速判别[J]. 煤田地质与勘探，2014，42(2)：49－53.

[4]　汪洋，左文喆，王斌海，等. 矿井突水水源判别方法研究进展[J]. 现代矿业，2018，34(1)：69－73.

[5]　贺亮亮，吕广罗，胡安焱，等. 基于水化学特征分析的矿井突水水源判别[J]. 中国煤炭地质，2022，34(6)：34－39.

[6]　JI Y，YU L，WEI Z，et al. Research progress on identification of mine water inrush sources：A visual analysis perspective[J]. Mine water and the environment，2025，44(1)：1－13.

[7]　孙延辉，胡文博，马红林，等. 基于 Piper-PCA-Fisher 的矿井突出水源判别模型构建及应用[J]. 化工矿物与加工，2023，52(9)：52－58.

[8]　琚棋定，胡友彪，张淑莹. 基于主成分分析与贝叶斯判别法的矿井突水水源识别方法研究[J]. 煤炭工程，2018，50(12)：90－94.

[9]　谢治刚. 模糊数学方法在矿井突水水源识别中的应用[J]. 地下水，2020，42(5)：54－57.

[10]　许延春，高玉兵，李卫民，等. 基于水化学特征的模糊评判法分析突水水源[J]. 煤

炭技术,2014,33(5):17－20.
[11] 朱赛君,姜春露,毕波,等.基于组合权一改进灰色关联度理论的矿井突水水源识别[J].煤炭科学技术,2022,50(4):165－172.
[12] 韩忠,王晓丽,施龙青.PCA-BP 神经网络在矿山岩溶突水水源判别中的应用研究[J].河南理工大学学报(自然科学版),2023,42(1):46－53.
[13] 张瑞钢,钱家忠,马雷,等.可拓识别方法在矿井突水水源判别中的应用[J].煤炭学报,2009,34(1):33－38.
[14] 陈祖云,张桂珍,邬长福,等.支持向量机在矿井突水水源识别中的应用[J].江西理工大学学报,2009,30(5):10－13.
[15] 赵茉溪,杨玉民,周传波,等.基于 MD-PCA-BP 模型的露天矿山爆破振动速度预测[J].爆破,2024,41(2):203－211.
[16] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. Advances in engineering software,2016,95:51－67.
[17] HUANG G B, ZHU Q, SIEW C. Extreme learning machine: Theory and applications[J]. Neurocomputing,2005,70(1):489－501.
[18] HUYNH H T, WON Y G, KIM J J. An improvement of extreme learning machine for compact single-hidden-layer feedforward neural networks[J]. International journal of neural systems,2008,18(5):433－441.
[19] 赵春永.淮南新集二矿井田水文地质条件分析及矿井涌水量预测[D].南京:南京大学,2020.
[20] 武佩萍.水化学特征分析法在出水水源判别中的应用[J].江西煤炭科技,2017(4):48－53.
[21] 万磊,孙茂如,赵吉诚,等.复杂地质构造下煤层底板灰岩水害防治技术研究[J].能源技术与管理,2021,46(6):89－91.

(责任编校:李秀荣)

(上接第 7 页)

(1)在给定结构尺寸前提下给出了计算消音器接管许用外载荷的计算公式,为接管设计提供了基础数据。

(2)采用所计算的接管许用外载荷,对消音器封头与接管连接处的 3 条评定路径进行了应力评定,结果显示,各应力分量满足设计应力强度限值,由此补齐了原设计计算中没有数据而导致的缺项部分。

(3)采用单点反应谱方法对消音器进行抗震分析,计算得到消音器 4 个支座所受到的地震载荷,以及考虑其自重条件下的支座支反力,这可为后续设备的设计及校核提供一定依据。

参考文献:

[1] 高远.城市环境噪声污染与监测技术[J].环境与发展,2020,32(5):154－156.
[2] 严索,高婷.国内消音器专利技术分析[J].河南科技,2021,40(24):53－55.
[3] 温华兵,李晓亮,吴晓佳.空压机排气放空复合消声器设计与性能分析[J].噪声与振动控制,2018,38(5):230－233.
[4] 郑津洋,桑芝富.过程设备设计[M].北京:化学工业出版社,2021:69－70.
[5] 陈冠锋,沈大园.吸气消音器结构优化及可靠性验证[J].日用电器,2022,203(11):76－79.
[6] HEI. Standards for closed feedwater heaters[S]. Cleveland: Heat Exchange Institute, Inc. ,2015:72－85.

(责任编校:冯兆娜)