

改进麻雀搜索算法的 RBF 神经网络水质预测^①



宋 健, 丛秋梅, 杨帅帅, 杨 健

(辽宁石油化工大学 信息与控制工程学院, 抚顺 113001)

通信作者: 丛秋梅, E-mail: cong_0828@163.com

摘 要: 针对污水处理过程中化学需氧量 (chemical oxygen demand, COD) 难以在线测量的问题, 提出了一种基于径向基函数 (radial basis function, RBF) 神经网络的软测量模型. 首先, 用污水处理厂实测数据挑选出与 COD 相关的过程变量作为输入变量; 其次, 基于 RBF 神经网络建立出水 COD 软测量模型, 利用自适应遗传算法改进的麻雀搜索算法 (adaptive genetic algorithm improved sparrow search algorithm, AGAISSA) 优化 RBF 神经网络的中心值、宽度值以及权值, 通过改进麻雀位置更新公式以及引入遗传算法中的自适应交叉和变异操作保证了软测量模型的精度; 最后, 将 RBF 神经网络的软测量模型应用于污水处理厂实测数据加以验证, 结果表明: AGAISSA 优化 RBF 神经网络模型能够对出水 COD 进行准确的预测, 具有较高的预测精度.

关键词: 污水处理; 麻雀搜索算法; 自适应遗传算法; RBF 神经网络

引用格式: 宋健, 丛秋梅, 杨帅帅, 杨健. 改进麻雀搜索算法的 RBF 神经网络水质预测. 计算机系统应用, 2023, 32(4): 255-261. <http://www.c-s-a.org.cn/1003-3254/9044.html>

RBF Neural Network Water Quality Prediction Based on Improved Sparrow Search Algorithm

SONG Jian, CONG Qiu-Mei, YANG Shuai-Shuai, YANG Jian

(School of Information and Control Engineering, Liaoning Petrochemical University, Fushun 113001, China)

Abstract: To address the problem that chemical oxygen demand (COD) is difficult to be measured on line during sewage treatment, this study proposes a soft sensing model based on a radial basis function (RBF) neural network. First, the process variables related to COD are selected as input variables by using the measured data of a sewage treatment plant. Second, the soft sensing model of COD in effluent is built on the basis of an RBF neural network. The center value, width value, and weight of the RBF neural network are optimized by an adaptive genetic algorithm improved sparrow search algorithm (AGAISSA). The accuracy of the soft sensing model is ensured by improving the sparrow position update formula and introducing the adaptive crossover and mutation operation in the genetic algorithm. Finally, the soft sensing model based on the RBF neural network is applied to the measured data of a sewage treatment plant for verification. The results show that the AGAISSA optimized RBF neural network model can accurately predict the COD in effluent and has high prediction accuracy.

Key words: sewage treatment; sparrow search algorithm; adaptive genetic algorithm; radial basis function (RBF) neural network

污水处理过程是一个复杂的生物化学反应过程, 其关键水质参数 COD 的检测是污水处理厂稳定、高效运行的保障^[1,2]. 现如今, COD 的检测方法可以分为具有强非线性、时滞性、耦合性、不确定性等特点,

① 基金项目: 国家自然科学基金 (61803191); 辽宁省自然科学基金 (2019-KF-03-05)

收稿时间: 2022-09-12; 修改时间: 2022-10-10; 采用时间: 2022-10-27; csa 在线出版时间: 2023-01-06

CNKI 网络首发时间: 2023-01-06

3类:水质离线化验方法、水质在线检测方法以及水质软测量方法。离线化验方法存在时滞性,在线检测装置普遍存在仪表安装和维护成本十分昂贵等问题,而现有的机理模型很难体现污水处理的特性,软测量精度还需要进一步提高。因此,建立可靠、精确的污水处理模型对关键水质参数进行软测量是污水处理过程领域需要解决的问题,具有重要的研究意义。针对以上问题,基于数据驱动的关键水质参数软测量模型得到广泛关注,神经网络具有非线性拟合能力,在污水水质参数软测量中应用广泛。何丹等把线性降低权系数和变异算子引入粒子群算法中,提出了改进粒子群算法的BP神经网络预测污水处理的COD和悬浮固体(suspended solids, SS)浓度,并与BP神经网络和遗传算法优化的BP神经网络进行对比,预测性能大大提高^[3]。张明涛等用多核LS-SVM建立污水处理过程中出水能耗和水质罚款模型,用分子动理论和人工记忆算法优化该模型,提高了该模型收敛速度和精确度^[4]。谢武明等用遗传算法优化的BP神经网络的权值和阈值,使权值和阈值在可选范围取得最佳值,并对污水处理电耗进行预测,与未经优化的BP神经网络相比,预测精度得到了极大的改善^[5]。李晋贤等用和声算法优化RBF神经网络的隐含层节点数预测污水出水COD指标,通过与标准的RBF神经网络相比具有较高的预测精度和良好的泛化能力^[6]。任忠铭等提出一种引力搜索算法的出水生化需氧量(biochemical oxygen demand, BOD)预测方法,将均方根误差作为适应度函数,根据特征变量转变为0-1组合问题来选出最优特征子集并作为RBF神经网络的输入,在保证预测精度的情况下,证明了算法的有效性^[7]。任东红等把轮盘赌策略以及交叉和变异引入到粒子群算法中,提出了混合粒子群算法的集成神经网络软测量模型,可以有效地预测污水处理过程中的关键水质参数,软测量精度较高^[8]。乔俊飞等提出了粒子群优化算法对回声状态神经网络权重训练,并对污水处理过程水质参数总磷进行预测,结果表明该模型具有较好非线性拟合能力和准确的预测能力^[9]。由于不同的智能优化算法具有各自的特性,因此,采用不同智能优化算法软测量模型得到的污水水质软测量精度不同。

自20世纪70年代遗传算法被提出以来,不少学者根据对大自然中生物种群及物理现象的认识分析,相继提出了许多智能优化算法,包括:细菌觅食优化算法^[10]、蛙跳算法^[11]、人工蜂群算法^[12]、鲸鱼优化算法^[13]、和

声搜索算法^[14]、布谷鸟算法^[15]、麻雀搜索算法(sparrow search algorithm, SSA)。其中麻雀搜索算法^[16]2020年被薛建凯提出,具有较强的全局寻优能力、参数少、易于实现、收敛快的优点。尽管如此,麻雀搜索算法也和其他智能优化算法一样,在解决实际问题时,容易陷入局部最优值,并且在迭代过程中种群多样性会有所减少,导致精度降低。针对麻雀搜索算法的不足,本文对麻雀搜索算法进行改进,引入改进麻雀搜索算法的位置更新公式以及遗传算法中的自适应交叉和变异操作优化RBF神经网络的中心值、宽度值以及权值,通过某大型污水处理厂的实测数据进行仿真实验,以均方根误差和平均绝对百分比误差验证了AGAISSA的性能优势。

1 模型的建立

1.1 辅助变量的选取

COD是衡量水体有机污染的一项重要指标,可以反映水体的受污染程度,如果想要高效准确地预测出水COD指标,辅助变量的选取也很重要。本文辅助变量通过机理分析选取,如表1所示。

表1 辅助变量

序号	辅助变量参数	序号	辅助变量参数
1	总进水悬浮固体浓度	10	曝气池氧化还原电位
2	总进水pH值	11	曝气气量
3	总进水氨态氮浓度	12	曝气氧量
4	计量槽2出水BOD ₅	13	曝气氧量
5	计量槽2出水COD	14	曝气池混合液悬浮固体浓度
6	计量槽2出水悬浮固体浓度	15	曝气池沉降比
7	分配流量	16	曝气池污泥容积指数
8	污泥流量	17	曝气池pH
9	曝气池氧化还原电位		

注:同一变量位于不同采样位置

1.2 污水处理水质软测量模型

本文提出了基于RBF神经网络^[17,18]污水处理过程水质软测量模型,RBF神经网络有很强的非线性逼近能力,可映射任意复杂的非线性关系,并且学习规则简单,便于设计实现。RBF神经网络的结构图如图1所示,为典型的3层前向神经网络,即输入层、隐含层、输出层。输入层选用表1作为输入变量,隐含层选用高斯函数作为激活函数,高斯函数是非线性函数,可以依据输入样本的变化不断学习,其表达式为:

$$h(x) = \exp\left(-\frac{1}{2\sigma_i^2}\|x - c_i\|^2\right), \quad i = 1, 2, 3, \dots, m \quad (1)$$

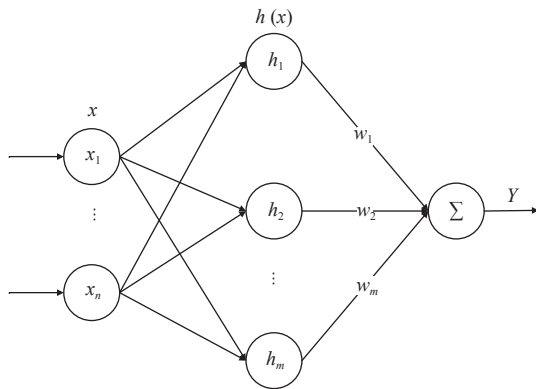


图1 RBF神经网络结构图

其中, x 是 n 维输入向量, 通过以上分析 $n = 17$; c_i 是第 i 个隐含层节点的圆心, 即中心值; σ_i 是第 i 个隐含层节点的宽度, 即宽度值; 输出层的输出为:

$$Y = \sum_{i=1}^m w_i \exp\left(-\frac{1}{2\sigma_i^2} \|x - c_i\|^2\right) \quad (2)$$

其中, w_i 为权值. 输出层选用水 COD 作为输出指标, m 为隐含层节点个数, 根据经验公式, 隐含层节点数表达式为:

$$m = M(n+1) \quad (3)$$

其中, M 为输出层节点数, 通过输入层的分析可知 $n = 17$, 故隐含层节点数为 18. RBF 神经网络中 σ_i 、 w_i 、 c_i 的取值直接影响 RBF 神经网络的预测性能, 因此, 本文采用 AGAISSA 优化 RBF 神经网络的 σ_i 、 w_i 、 c_i 参数以提高 RBF 神经网络的预测性能.

2 算法的实现

2.1 麻雀搜索算法

SSA 算法受麻雀觅食和反捕食行为而提出的一种新型群智能优化算法, 其原理如下.

麻雀在觅食的过程中由发现者、加入者、预警者 3 部分构成. 在 SSA 中, 适应度较高的发现者在搜索食物过程中会引导种群搜索和觅食, 加入者为获取更高的适应度会跟随发现者觅食, 当麻雀种群意识到有危险时, 则会发生反捕食行为.

发现者位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(\frac{-i}{\alpha \cdot N}\right), & \text{if } R_2 < ST \\ X_{i,j}^t + Q \cdot L, & \text{if } R_2 \geq ST \end{cases} \quad (4)$$

其中, t 代表迭代次数, $j = 1, 2, \dots, d$, d 表示待优化问题变量的维度, N 是个常数, 表示最大迭代次数, $X_{i,j}$ 表示

第 i 个麻雀在第 j 维中的位置信息, α 为 $(0, 1]$ 之间的随机数, ST 和 R_2 分别表示安全值和预警值, L 为 $1 \times d$ 且元素均为 1 的矩阵, Q 为服从正态分布的随机数. 当 $R_2 < ST$ 时, 表明麻雀种群觅食环境相对安全, 没有捕食者入侵, 发现者可以搜索更广泛的食物. 当 $R_2 \geq ST$ 时, 表明种群中的部分麻雀发现了捕食者的入侵, 同时向种群中的其他麻雀发出警告, 这时麻雀种群都会快速飞到安全的地方觅食.

加入者位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{X_{\text{worst}}^t - X_{i,j}^t}{i^2}\right), & \text{if } i > n/2 \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, & \text{if } i \leq n/2 \end{cases} \quad (5)$$

其中, X_p 是当前发现者占据的最优位置, X_{worst} 表示当前最差位置, A 为随机包含 -1 和 1 两个元素的 $1 \times d$ 矩阵, $A^+ = A^T(AA^T)^{-1}$, n 为麻雀总数, 当 $i > n/2$ 时, 表明第 i 个加入者没有觅食, 适应度较低, 正处于饥饿状态, 需要飞到其他地方觅食, 以获得更高的能量. 当 $i \leq n/2$ 时, 表明第 i 个加入者将在当前最优位置 X_p 附近范围找到合适的位置觅食.

预警者位置更新如下:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot |X_{i,j}^t - X_{\text{best}}^t|, & \text{if } f_i > f_{\text{best}} \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_{\text{worst}}) + \varepsilon}\right), & \text{if } f_i = f_{\text{best}} \end{cases} \quad (6)$$

其中, X_{best} 是当前最优位置, β 为服从均值为 0, 方差为 1 的正态分布随机数, f_i 为麻雀个体当前的适应度值, K 为 $[-1, 1]$ 之间的随机数, f_{best} 和 f_{worst} 为全局最优和最差的适应度值. ε 为一个极小的数, 避免分母为 0. 当 $f_i > f_{\text{best}}$ 时, 此时麻雀处于种群边缘觅食, 很容易遭到捕食者攻击; $f_i = f_{\text{best}}$ 时, 处于种群中间的麻雀会意识到危险, 则会飞往其他麻雀的位置觅食, 以减小被捕食的概率.

2.2 改进的麻雀搜索算法

(1) 预备知识

迭代函数 $\phi(x)$ 在 $[a, b]$ 上连续可导, 且满足:

(i) $x \in [a, b]$ 时, $a \leq \phi(x) \leq b$.

(ii) 存在一正数, 满足 $0 < L < 1$, 且任意 $x \in [a, b]$, 有 $|\phi'(x)| \leq L$.

对任意 $x \in [a, b]$, 迭代法 $x_{k+1} = \phi(x_k)$ 均收敛于 x^* .

证明: 设 $f(x) = x - \phi(x)$, 则 $f(x)$ 在 $[a, b]$ 上连续可导:

$$f(a) = a - \phi(a) \leq 0 \quad (7)$$

$$f(b) = b - \phi(b) \geq 0 \quad (8)$$

故 $f(x) = 0$ 在 $[a, b]$ 上至少有一个根.

由 $|\phi'(x)| \leq L < 1$:

$$f'(x) = 1 - \phi'(x) > 0 \quad (9)$$

则 $f(x)$ 在 $[a, b]$ 上单调递增, 有 $x = \phi(x)$ 在 $[a, b]$ 内有唯一解 x^* .

$$x_{k+1} = \phi(x_k) \quad (10)$$

由微分中值定理:

$$x_{k+1} - x^* = \phi(x_k) - \phi(x^*) = \phi'(\xi)(x_k - x^*) \quad (11)$$

$$x_{k+1} - x_k = \phi(x_k) - \phi(x_{k-1}) = \phi'(\xi)(x_k - x_{k-1}) \quad (12)$$

由于 $|\phi'(x)| \leq L$:

$$|x_{k+1} - x_k| \leq L|x_k - x_{k-1}| \quad (13)$$

$$\begin{aligned} |x_{k+1} - x^*| &\leq L|x_k - x^*| \\ &= L|x_{k+1} - x^* - (x_{k+1} - x_k)| \\ &\leq L|x_{k+1} - x^*| + L|x_{k+1} - x_k| \end{aligned} \quad (14)$$

$$|x_{k+1} - x^*| \leq \frac{L}{1-L}|x_{k+1} - x_k| \quad (15)$$

$$\begin{aligned} |x_k - x^*| &\leq \frac{L}{1-L}|x_k - x_{k-1}| \\ &\leq \frac{L^2}{1-L}|x_{k-1} - x_{k-2}| \\ &\vdots \\ &\leq \frac{L^k}{1-L}|x_1 - x_0| \end{aligned} \quad (16)$$

由于 $L < 1$, $\lim_{k \rightarrow \infty} (x_k - x^*) = 0$, 因此, 对任意初值 x_0 , 迭代法 $x_{k+1} = \phi(x_k)$ 均收敛于 x^* .

(2) 改进发现者位置

对发现者位置的改进, 引入正弦函数对其改进, 并去掉随机数 α , 改进后的公式:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \sin\left(\frac{-i}{N}\right), & \text{if } R_2 < ST \\ X_{i,j}^t + Q \cdot L, & \text{if } R_2 \geq ST \end{cases} \quad (17)$$

证明改进发现者位置公式的收敛性如下:

由预备知识 $x_{k+1} = \phi(x_k)$, 只需证明 $|\phi'(x)| < 1$ 即可.

$X_{i,j}^{t+1} = X_{i,j}^t \cdot \sin\left(\frac{-i}{N}\right)$, $\phi'(x) = \sin\left(\frac{-i}{N}\right)$, 由于 i 为第 i 个麻雀的位置信息, 故 i 为 1 并以 1 为单位增加到发现者数量为 7 的数, 故 $i_{\min} = 1$, $i_{\max} = 7$, $N = 30$, 故 $\frac{-i}{N} \neq$

$\frac{\pi}{2} \pm k\pi$, $k = 1, 2, \dots, m$, 因此 $|\phi'(x)| = \left| \sin\left(\frac{-i}{N}\right) \right| < 1$, 说明改进发现者位置公式收敛.

(3) 改进加入者位置

对加入者位置的改进, 引入正弦函数以及反比例函数对其改进, 改进后的公式:

$$X_{i,j}^{t+1} = \begin{cases} \frac{\exp(X_{\text{worst}}^t - X_{\text{best}}^t)}{i^2} + \frac{\sin(i/n)}{X_{i,j}^t}, & \text{if } i > n/2 \\ X_p^{t+1} + |X_{i,j}^t - X_p^{t+1}| \cdot A^+ \cdot L, & \text{if } i \leq n/2 \end{cases} \quad (18)$$

证明改进加入者位置公式的收敛性如下:

由预备知识 $x_{k+1} = \phi(x_k)$, 只需证明 $|\phi'(x)| < 1$ 即可.

$X_{i,j}^{t+1} = \frac{\exp(X_{\text{worst}}^t - X_{\text{best}}^t)}{i^2} + \frac{\sin(i/n)}{X_{i,j}^t}$, 因此 $\phi'(x) = -\frac{\sin(i/n)}{(X_{i,j}^t)^2}$, 由于 i 是由发现者数量+1 并以单位 1 增加到种群规模为 10 的数. 即 $i_{\min} = 8$, $i_{\max} = 10$, $i/n \neq \frac{\pi}{2} \pm k\pi$, $k = 1, 2, \dots, m$, 由于加入者在不断跟随发现者的位置以得到更多的资源, 故加入者每次更新的位置范围较大, 即 $|\phi'(x)| = \left| -\frac{\sin(i/n)}{(X_{i,j}^t)^2} \right| < 1$, 说明改进加入者位置公式收敛.

(4) 改进预警者位置

对预警者位置的改进, 引入正弦函数对其改进, 改进后的公式:

$$X_{i,j}^{t+1} = \begin{cases} X_{\text{best}}^t + \beta \cdot \sin(X_{i,j}^t - X_{\text{worst}}^t), & \text{if } f_i > f_{\text{best}} \\ X_{i,j}^t + K \cdot \left(\frac{|X_{i,j}^t - X_{\text{worst}}^t|}{(f_i - f_{\text{worst}}) + \varepsilon} \right), & \text{if } f_i = f_{\text{best}} \end{cases} \quad (19)$$

证明改进预警者位置公式的收敛性如下:

由预备知识 $x_{k+1} = \phi(x_k)$, 只需证明 $|\phi'(x)| < 1$ 即可.

$X_{i,j}^{t+1} = X_{\text{best}}^t + \beta \cdot \sin(X_{i,j}^t - X_{\text{worst}}^t)$, 对其求导 $\phi'(x) = \beta \cdot \cos(X_{i,j}^t - X_{\text{worst}}^t)$, 由于 β 为服从均值为 0, 方差为 1 的正态分布随机数, 即 $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ 在 0 处取得最大值为 $f(0) = \frac{1}{\sqrt{2\pi}} < 1$, β 的值域为 $\left(0, \frac{1}{\sqrt{2\pi}}\right]$, 余弦函数值域为 $[-1, 1]$, 故导数的绝对值小于 1, 说明改进预警者位置公式收敛.

2.3 自适应遗传算法改进的麻雀搜索算法

交叉操作与变异操作在遗传算法^[19,20]中起着关键性的作用, 对不同问题具有较好的鲁棒性, 能够避免种群陷入局部最优值. 在麻雀种群发现食物与跟随以及

侦察预警过程中,对麻雀个体的适应度排序,选出当前的最佳适应度麻雀个体进行交叉与变异操作,通过子代与父代适应度比较,淘汰适应度差的个体,这样可以增加麻雀种群的多样性,充分利用麻雀群体的优良特性,使其跳出局部最优值,从而精度得到进一步提高.算法公式如下.

① 交叉操作

选择种群中的麻雀个体使他们随机地进行相互配对,然后使配对好的麻雀个体通过选定的概率 P_c 发生交叉操作.对于交叉的两个麻雀 X_i, X_j ,其位置交叉的公式如下:

$$\begin{cases} X_i^{k+1} = \gamma \cdot X_i^k + (1-\gamma) \cdot X_j^k \\ X_j^{k+1} = (1-\gamma) \cdot X_i^k + \gamma \cdot X_j^k \end{cases} \quad (20)$$

其中, γ 是区间(0,1)之间的随机数.

② 变异操作

对麻雀个体的位置以选定的概率 P_m 发生变异.假设个体最优值 P_i 的第 d 维变量为 P_i^d ,对 P_i^d 施加随机扰动的策略进行变异操作.公式如下:

$$P_i^d = P_i^d \cdot (1 + \mu) \quad (21)$$

其中, μ 为随机扰动变量,服从均值为0,方差为1的正态分布.

其中,交叉操作 P_c 和变异操作 P_m 是影响算法性能指标的重要因素^[21,22],在实际应用中,很难确定最好的交叉操作 P_c 与变异操作 P_m 进而达到预期效果.本文提出了自适应交叉和变异操作,使 P_c, P_m 能够自动调节,自适应交叉和变异操作定义公式如下:

$$P_c = \begin{cases} 1 - (P_{c1}/(f_1 + f_3/f_2)), & \text{if } fitness(1) < f_1 \\ P_{c1}, & \text{if } fitness(1) \geq f_1 \end{cases} \quad (22)$$

$$P_m = \begin{cases} P_{m1}/(f_1 + f_3/f_2), & \text{if } fitness(1) < f_1 \\ P_{m1}, & \text{if } fitness(1) \geq f_1 \end{cases} \quad (23)$$

其中, P_{c1}, P_{m1} 为常数, f_1 为整个种群适应度的平均值, f_2 为整个种群适应度的最大值, f_3 为整个种群适应度的最小值, $fitness(1)$ 为全局最优适应度.

自适应遗传算法改进的麻雀搜索算法流程图如图2所示.

3 仿真实验

本次仿真实验采用污水处理厂365组数据为样本,选用345组数据作为训练样本,20组数据为测试样本,

初始种群规模为10,最大迭代次数为30,安全值为0.6,发现者的比例为0.7,预警值为0到1之间的随机数,用Matlab软件对RBF神经网络、GA-RBF神经网络、SSA-RBF神经网络、AGAISSA-RBF神经网络模型进行实验,最后用均方根误差(root mean square error, RMSE)和平均绝对百分比误差(mean absolute percentage error, MAPE)来评价预测值与实测值的偏差.

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (actual_j - predict_j)^2} \quad (24)$$

$$MAPE = \frac{1}{N} \sum_{j=1}^N \frac{|actual_j - predict_j|}{|actual_j|} \times 100\% \quad (25)$$

其中, $predict_j$ 为预测值, $actual_j$ 为实测值, N 为样本总数.

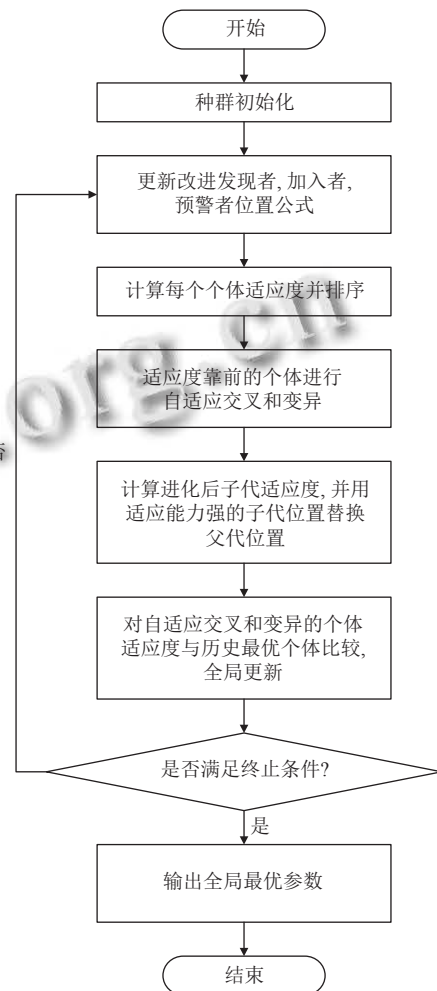


图2 自适应遗传算法改进的麻雀搜索算法流程图

预测值和实测值对比图如图3所示,由图3可知,RBF神经网络预测值远远低于实测值,GA-RBF神经网络和SSA-RBF神经网络的预测值会在实测值的较小范围内波动,而在引入AGAISSA时,AGAISSA-RBF神经网络部分预测值与实测值重合,其他未重合的数据点也非常接近实测值,说明麻雀种群在寻优过程中,引入自适应交叉和变异操作,会使模型精度更高。

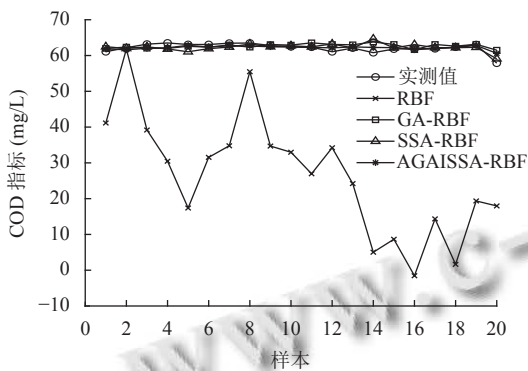


图3 预测值和实测值对比图

预测误差对比图如图4所示,从图4我们可以看出,RBF神经网络的预测误差变化范围很大,预测误差远小于0;GA-RBF神经网络和SSA-RBF神经网络的预测误差相对RBF神经网络的预测误差具有明显的改善;而AGAISSA-RBF神经网络的预测误差非常接近于0,说明,AGAISSA可以增加麻雀种群多样性,使其跳出局部最优值,从而使AGAISSA-RBF神经网络具有更高的预测精度。

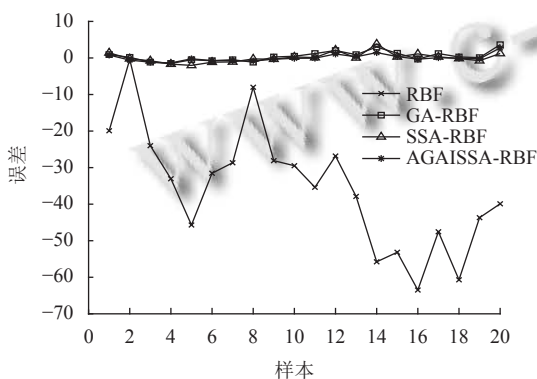


图4 预测误差对比图

迭代收敛曲线对比图如图5所示,从图5我们可以看出,GA在第5代开始收敛,SSA在第7代开始收敛,AGAISSA在第17代开始收敛,尽管AGAISSA收

敛速度较慢,但AGAISSA最佳适应度优于SSA和GA最佳适应度。

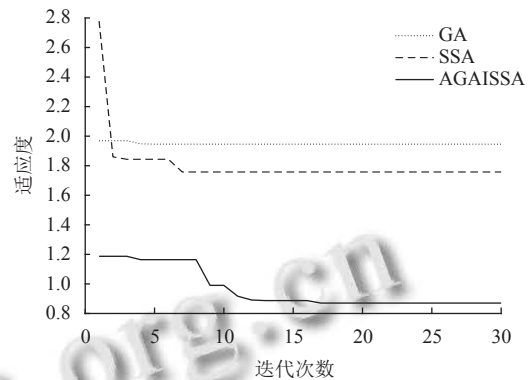


图5 迭代收敛曲线对比图

均方根误差对比图如图6所示,从图6我们可以看出RBF神经网络、GA-RBF神经网络、SSA-RBF神经网络、AGAISSA-RBF神经网络的均方根误差分别为39.06、1.39、1.32、0.93,依次呈递减趋势,说明引入自适应遗传算法可以使软测量模型的精度得到进一步的改善。

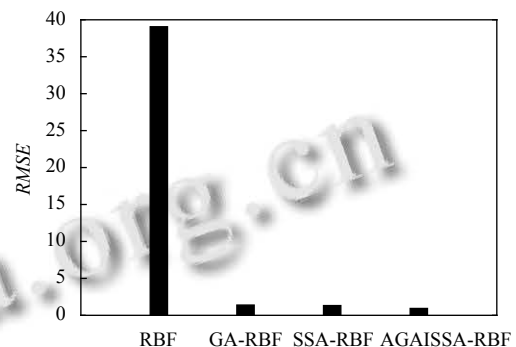


图6 均方根误差对比图

平均绝对百分比误差对比图如图7所示,从图中我们可以看出RBF神经网络的平均绝对百分比误差为57.49%、GA-RBF神经网络的平均绝对百分比误差为1.73%、SSA-RBF神经网络的平均绝对百分比误差为1.55%、AGAISSA-RBF神经网络的平均绝对百分比误差为1.11%,平均绝对百分比误差值越小,预测精度越高,因此可以说明AGAISSA-RBF神经网络的模型精度较高。

4 结论与展望

本文提出了解决污水处理水质参数软测量问题的

一种新型 AGAISSA 算法, 将遗传算法中的自适应交叉操作和变异操作引入改进麻雀搜索算法中, 有效地改变了麻雀个体的适应度, 避免了麻雀搜索算法陷入局部极值, 提高了算法的全局寻优能力. 通过对 AGAISSA 算法的性能指标分析, 验证了该算法的优势性. 运用 AGAISSA-RBF 神经网络模型预测污水出水 COD, 相对于 RBF 神经网络、GA-RBF 神经网络、SSA-RBF 神经网络找到了更优的预测模型, 有效降低了污水出水 COD 预测偏差. 由于麻雀搜索算法提出较短, 以后还可以结合其他智能优化算法进行相关改进, 并运用到复杂工业建模中, 提高算法的实用性.

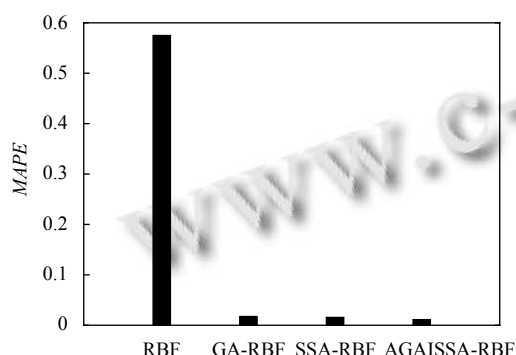


图7 平均绝对百分比误差对比图

参考文献

- 张涛. 污水处理厂污水深度处理工艺研究. 皮革制作与环保科技, 2021, 2(18): 12–13.
- 张明. 污水处理厂节能降耗措施分析. 化工设计通讯, 2021, 47(8): 181–182. [doi: 10.3969/j.issn.1003-6490.2021.08.087]
- 何丹, 林来鹏, 李小勇, 等. 基于改进粒子群算法优化 BP 神经网络的废水处理软测量模型. 华南师范大学学报 (自然科学版), 2021, 53(2): 114–120.
- 张明涛, 王瑞峰. 基于记忆分子动力学理论的污水处理过程优化. 计算机仿真, 2019, 36(3): 401–405. [doi: 10.3969/j.issn.1006-9348.2019.03.086]
- 谢武明, 李俊, 周峰平, 等. 基于 BP 神经网络和遗传算法的污水处理厂电耗预测. 水电能源科学, 2018, 36(8): 202–204.
- 李晋贤, 李少甫, 李洪. 和声改进 RBF 神经网络在污水出水 COD 软测量中的应用. 电子技术与软件工程, 2017, (5): 34–35.
- 任忠铭, 李文静, 乔俊飞. 基于 GSA 的递归 RBF 神经网络出水 BOD 预测模型. 第 38 届中国控制会议论文集 (7). 广州: 中国自动化学会控制理论专业委员会, 2019. 156–160.
- 任东红, 韩红桂, 乔俊飞. 基于 IHPSO 算法的污水处理过程集成神经网络软测量模型. 信息与控制, 2014, 43(1): 123–128.

- 乔俊飞, 李瑞祥, 韩红桂. 基于粒子群优化回声状态网络的污水总磷预测方法. 第 34 届中国控制会议论文集 (B 卷). 杭州: 中国自动化学会控制理论专业委员会, 2015. 244–248.
- Guo C, Tang H, Niu B. Evolutionary state-based novel multi-objective periodic bacterial foraging optimization algorithm for data clustering. Expert Systems, 2022, 39(1): e12812. [doi: 10.1111/exsy.12812]
- 徐逸夫, 段隆振. 基于蛙跳算法的无线传感器网络节点重部署. 计算机仿真, 2021, 38(10): 328–332. [doi: 10.3969/j.issn.1006-9348.2021.10.066]
- Yu XW, Li Y, Liu Y, et al. Inertial optimization MCL deep mine localization algorithm based on grey prediction and artificial bee colony. Wireless Networks, 2021, 27(4): 3053–3072. [doi: 10.1007/s11276-021-02633-y]
- Pan YK, Xia KW, Wang L, et al. A novel approach to oil layer recognition model using whale optimization algorithm and semi-supervised SVM. Symmetry, 2021, 13(5): 757. [doi: 10.3390/sym13050757]
- Kumar P, Jain M, Meena RK. Optimal control of fault tolerant machining system with reboot and recovery in fuzzy environment using harmony search algorithm. ISA Transactions, 2022, 119: 52–64. [doi: 10.1016/j.isatra.2021.02.027]
- 方园园. 基于布谷鸟算法的气味源识别方法研究. 矿业研究与开发, 2021, 41(2): 183–189. [doi: 10.13827/j.cnki.kyyk.2021.02.032]
- 薛建凯. 一种新型的群智能优化技术的研究与应用——麻雀搜索算法 [硕士学位论文]. 上海: 东华大学, 2020.
- 张雯雯. 基于 RBF 神经网络的固定时间控制器设计. 电子技术与软件工程, 2022, (9): 154–158.
- 张启凡, 王永忠, 马俊逸. 基于 RBF 神经网络的短时降水量预测方法研究. 计算机与数字工程, 2022, 50(4): 807–811. [doi: 10.3969/j.issn.1672-9722.2022.04.024]
- 徐逸扬, 潘兆一, 李锦键. 基于遗传算法优化 PID 的电磁制热系统温度控制方法. 电气传动自动化, 2022, 44(3): 1–5. [doi: 10.3969/j.issn.1005-7277.2022.03.001]
- 董佳琦, 许洪华. 基于遗传算法优化支持向量机的电梯交通模式识别. 工业控制计算机, 2022, 35(1): 101–103. [doi: 10.3969/j.issn.1001-182X.2022.01.036]
- 吴振宇. 基于改进遗传算法的配电网故障区段定位方法研究 [硕士学位论文]. 天津: 天津理工大学, 2022.
- 周勇. 改进遗传算法的 RBF 神经网络控制研究. 科技与创新, 2022, (2): 11–13, 21. [doi: 10.15913/j.cnki.kjycx.2022.02.004]

(校对责编: 牛欣悦)