

控制与决策

Control and Decision

基于隐性记忆的非平稳时变污水处理过程多目标运行优化

周平, 张天娇

引用本文:

周平, 张天娇. 基于隐性记忆的非平稳时变污水处理过程多目标运行优化[J]. *控制与决策*, 2023, 38(8): 2389–2400.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.2112>

您可能感兴趣的其他文章

Articles you may be interested in

基于复杂昂贵仿真的体系效能多目标优化

Complex and expensive simulation based multi-objective optimization to system-of-system effectiveness

控制与决策. 2021, 36(3): 589–598 <https://doi.org/10.13195/j.kzyjc.2019.0844>

基于R2指标和目标空间分解的高维多目标粒子群优化算法

R2 indicator and objective space partition based many-objective particle swarm optimizer

控制与决策. 2021, 36(9): 2085–2094 <https://doi.org/10.13195/j.kzyjc.2020.0113>

输入约束不确定系统的点对点迭代学习控制与优化

Point-to-point iterative learning control and optimization for uncertain systems with constrained input

控制与决策. 2021, 36(6): 1435–1441 <https://doi.org/10.13195/j.kzyjc.2019.0908>

基于数据驱动的浓密-压滤过程协调优化控制

Data driven coordinated optimization control of thickening-filter process

控制与决策. 2021, 36(5): 1095–1100 <https://doi.org/10.13195/j.kzyjc.2019.1151>

基于多种群分解预测的动态多目标引力搜索算法

Dynamic multi-objective gravitational searching algorithm based on multi-population decomposition prediction

控制与决策. 2021, 36(12): 2910–2918 <https://doi.org/10.13195/j.kzyjc.2020.1002>

基于隐性记忆的非平稳时变污水处理过程 多目标运行优化

周平, 张天娇

(东北大学 流程工业综合自动化国家重点实验室, 沈阳 110819)

摘要: 针对污水处理过程中运行能耗大和水质超标严重等问题, 基于隐性记忆策略, 提出一种适用于无规律且非平稳时变过程的多目标运行优化方法. 首先, 采用集成即时学习在线建模算法建立运行指标模型, 实现时变工况下运行优化目标的准确描述; 然后, 提出基于隐性记忆的多目标优化算法, 通过充分考虑复杂时变工况, 挖掘埋藏在历史优化信息中的结构化知识, 引导进化搜索过程, 并结合平均距离聚类指标进行均匀随机局部搜索, 提高算法收敛性、多样性和分布性; 最后, 利用基于集成即时学习的智能决策方法, 获取操作变量溶解氧浓度和硝态氮浓度的优化设定值并进行跟踪控制. 数据实验表明: 所提出方法能够实现污水处理过程的稳定运行, 在满足操作限制的情况下, 提高污染物去除性能, 降低运行成本.

关键词: 运行优化; 污水处理过程; 多目标优化; 隐性记忆; 局部搜索; 即时学习

中图分类号: TP18

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.2112

引用格式: 周平, 张天娇. 基于隐性记忆的非平稳时变污水处理过程多目标运行优化[J]. 控制与决策, 2023, 38(8): 2389-2400.

Implicit memory based multi-objective operation optimization of wastewater treatment process under nonstationary time-varying conditions

ZHOU Ping, ZHANG Tian-jiao

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China)

Abstract: A multi-objective operation optimization method based on an implicit memory strategy is proposed for irregular and nonstationary time-varying processes to solve the problems of excessive energy consumption and severe water quality in wastewater treatment process (WWTP). First, an ensemble just-in-time learning (E-JITL) algorithm is used to build an operation optimization model, which realizes an accurate description of the operation optimization objectives. Then, a multi-objective optimization algorithm based on the implicit memory (IM-NSGAI) is designed, considering the dynamic time-varying working conditions, mining the structured knowledge buried in the historical optimization information, guiding the evolutionary search process, and combining the average distance clustering index and uniform random local search, which improves the convergence, diversity and distribution of the algorithm. Finally, an intelligent decision-making method based on the E-JITL algorithm is used to obtain the optimal setpoints of dissolved oxygen concentration and nitrate concentration for tracking control. Experiments show that the proposed this method can achieve stable operation of the WWTP, improve pollutant removal performance, and reduce operating costs.

Keywords: operation optimization; wastewater treatment process; multi-objective optimization; implicit memory; local search; just-in-time learning

0 引言

随着社会经济的快速发展, 水污染问题日渐突出, 而污水处理也成为城市化进程必须解决的问题之

一^[1]. 污水进入被动且随机接受, 其处理过程包含众多受多重随机因素影响的微生物反应和复杂的物理化学反应, 是一个强耦合性、强不确定性和非平稳时

收稿日期: 2022-12-06; 录用日期: 2023-03-20.

基金项目: 国家重点研发计划项目(2022YFB3304903); 国家自然科学基金项目(61890934, U22A2049).

责任编辑: 杨涛.

†通讯作者. E-mail: zhouping@mail.neu.edu.cn.

变的复杂非线性系统. 目前, 如何实现城市污水处理过程的优化运行, 以及如何在满足操作限制的情况下提高污染物去除性能和降低运行能耗, 仍是一个亟待解决的重要学术与工程技术问题^[2]. 随着工业自动化与优化技术在人工智能带动下的快速发展, 以提质、增效和降耗为目标的运行优化方法和技术已在城市污水处理过程中得到了广泛应用. 通过建立特定的运行优化问题, 采用智能建模与智能优化技术设计合适的运行优化策略, 可实现工业过程整体运行性能的优化. 对于具有无规律、非平稳时变动态的污水处理过程, 其运行优化的实现面临两方面的挑战^[3], 即: 一是如何建立可精确捕捉污水处理过程复杂动态特性的运行指标优化模型; 二是如何实现多个操作变量设定值的高性能优化决策以实现多个运行指标的多目标优化.

针对上述第1个问题, 目前已有许多污水处理运行优化的数据驱动建模方法. 如: 文献[4]使用自适应神经模糊推理系统和广义线性模型回归技术, 完成了活性污泥污水处理过程的非线性系统识别, 建立了生化需氧量(BOD)、化学需氧量(COD)等关键出水水质变量的数据驱动模型. 文献[5]将溶解氧浓度作为污水能耗指标, 将污水含碳生化需氧量、总悬浮固体、总溶解磷和总悬浮磷等作为水质指标, 采用聚类分析、自适应多元回归分析等技术建立模型, 实验表明所建立的模型虽然计算简单但精度不足. 文献[6]针对实际污水处理过程出水指标离线化验时间滞后的问题, 提出一种类脑模块化神经网络建模方法. 该方法首先采用基于互信息和专家知识的方法筛选关键出水参数的相关变量, 获取各出水参数的辅助变量, 再通过神经网络构建软测量子模型, 具有较高的建模精度及效率. 但该方法仅包含单个指标, 难以反映污水处理的大部分动态特性. 因此, 如何根据实际污水处理过程非平稳时变动态特性, 构建动态、高精度、综合性的运行指标优化模型, 仍是一个具有挑战性的难题.

针对上述第2个问题, 文献[7]以污水运行能耗和出水水质为优化目标, 提出一种改进的基于分解的多目标进化算法(MOEA/D), 利用较少的迭代次数搜索来优化决策溶解氧浓度和硝态氮浓度的最优设定值. 文献[8]利用反向传播算法动态调整出水水质、运行能耗和关键操作变量之间的关系, 同时利用标准多目标优化方法平衡出水水质和成本. 文献[9]提出一种基于强化学习的粒子群算法来降低能耗, 其构建的精英粒子群集记录优化设定值及调整趋势, 训练精英网络作为策略函数对粒子速度进行预测, 提高了算

法的收敛速度. 文献[10]采用一种知识引导方法构建知识库, 精确储存和重用过去的解引导搜索, 并结合定向局部区域寻优以及随机全局寻优策略获得更高质量的设定值优化解. 以上方法较好地解决了污水处理过程的运行优化及控制问题, 但是, 当污水处理过程被动接受的污水进水的流量、组分等发生无规律、非平稳波动时, 优化算法的搜索效率和精度均会降低, 寻优性能有待进一步提升.

针对上述实际工程问题与现有方法的不足, 本文提出一种基于隐性记忆的污水处理过程多目标运行优化方法. 如图1所示, 所提出方法首先建立基于集成即时学习的出水水质模型和能耗模型, 实现关键运行指标的在线高精度建模; 然后, 提出一种隐性记忆策略, 在工况发生无规律、非平稳变化时, 采用单层去噪自动编码器挖掘埋藏在过去搜索经验中的结构化知识, 并将最优解储存在知识库中, 从而提高计算效率, 增强进化搜索效果. 同时, 采用基于平均距离聚类多样性指标, 基于聚类思想, 将平均距离内的多个个体划分为一个小种群, 应用于进化搜索中的选择、交叉、变异过程, 对平均距离聚类较小的小种群中个体进行均匀随机局部搜索, 提高Pareto解的分布性和多样性. 最后, 考虑当前工况与具体工艺要求, 设计基于集成即时学习的智能决策方法, 从而获得更适合当前工况的溶解氧浓度、硝态氮浓度设定值. 通过底层控制器对设定值的跟踪控制, 实现污水处理过程安全稳定运行, 达到“保质降耗”的目标.

1 过程简介与问题陈述

现代城市污水处理厂普遍采用活性污泥处理法^[11]处理废水, 其工艺流程如图1所示, 主要由曝气池和沉淀池两部分组成. 该过程利用活性污泥中的微生物群体吸附、氧化、分解污水中有机物, 并通过物理沉降功能实现泥水分离^[12], 其本质是一个动态特性极复杂、大滞后的生物化学反应过程. 污水处理过程的污水随机进入且被动接受, 受水文天气等多重不确定随机因素的影响. 此外, 由于活性污泥法微生物活性的复杂性, 即使进入系统的微小变化(例如流速、废水温度等变化)也会影响出水水质和运行能耗. 因此, 污水处理过程具有典型的无规律非平稳随机性、强耦合性和高度非线性等特征. 例如, 鼓风机向曝气池供应的空气允许氧气从空气中转移到液相(废水), 该氧转移就是一个复杂的过程, 其特点是大时间延迟以及强非线性和强耦合. 因此, 上述复杂特性大大增加了污水处理过程运行操作的难度.

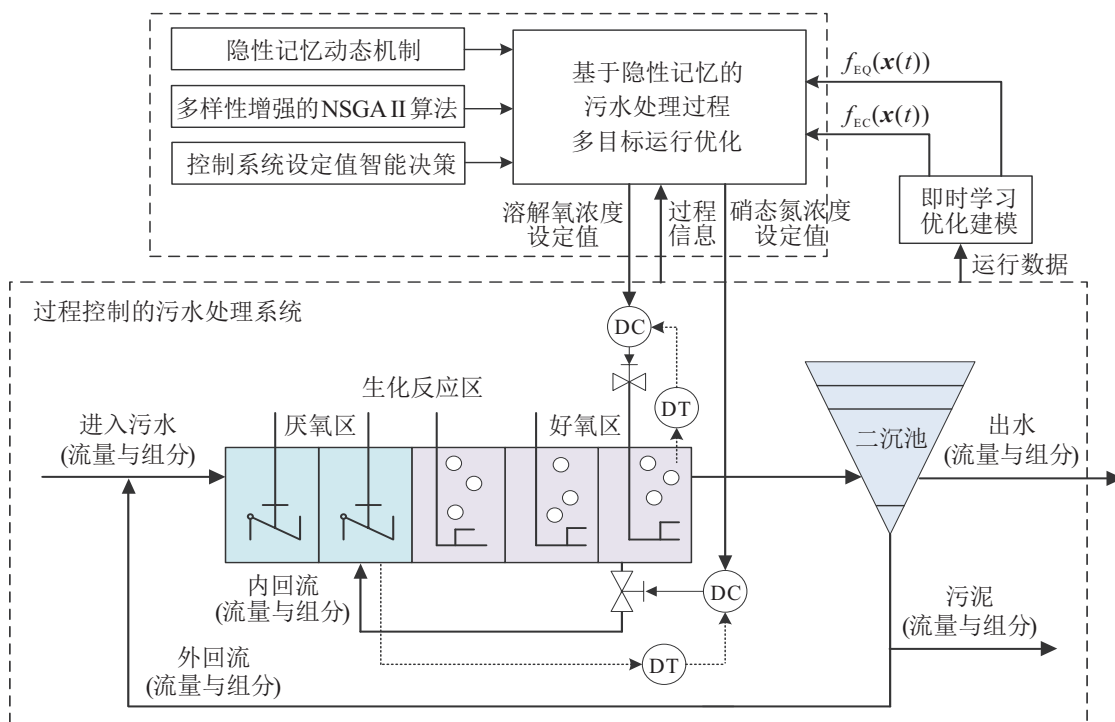


图1 污水处理过程多目标运行优化结构

图1所示污水处理包含多个反映过程运行性能的运行指标,如运行能耗和出水水质等,其运行状态随被动进入污水的流量与组分、外部环境以及生化反应过程的变化而变化,且各性能指标互相影响、相互冲突^[2]。此外,被动接受的污水组分与流量非平稳随机动态变化特性使得人们很难采用常规方法合理地决策底层第5分区溶解氧浓度(S_{O_5})和第2分区硝态氮浓度(S_{NO_2})等关键过程变量控制系统的设定值,从而难以保证最终出水水质达标和降低运行能耗与生产成本。随着我国城市化的发展和环保的要求越来越高,城市污水处理的要求也日益严格,在满足出水水质的同时降低运行成本已迫在眉睫,污水处理多目标优化控制为解决污水处理过程关键问题提供了解决方法。

2 运行指标集成即时学习优化建模

建立准确有效的水质(EC)与能耗(EQ)等运行指标模型是实现污水处理过程运行优化的前提。考虑到污水处理过程的非线性、时变以及非平稳动态特性,本文采用集成即时学习在线建模方法建立污水处理过程的运行优化模型。即时学习是一种擅长应对过程特性和状态变化的局部学习策略,其从历史数据库中根据相似度度量准则选取与当前查询样本最相关的样本子集建立局部模型^[13],实现在线学习建模。为了提高模型精度,将常规即时学习模型与基于

互信息^[13-14]的相似度衡量指标的即时学习模型加权集成,得到最终污水处理过程的运行优化模型,可满足污水运行指标优化模型高精度在线学习建模的要求。所提出的集成建模算法流程如图2所示,具体描述如下。

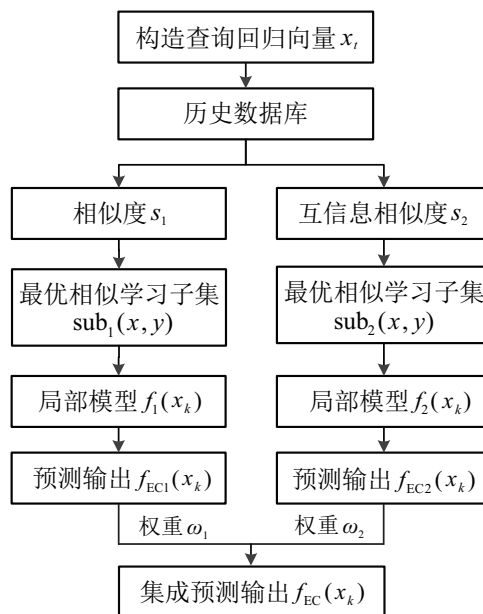


图2 集成即时学习优化建模方法流程

设 $\{(x_i, y_i) | x_i \in \mathbf{R}^m, y_i \in \mathbf{R}, i = 1, 2, \dots, n\}$ 。其中: x_i 为输入向量,即 S_{O_5} 和 S_{NO_2} 的设定值; y_i 为输出变量,分别为EC和EQ。对于当前工作点 x_t ,采用K-VNN(k-vector nearest neighbors)方法^[13]计算其与

历史数据的相似度值 $s_1(x_i, x_t)$, 即

$$s_1(x_i, x_t) = \gamma \cdot e^{-d(x_i, x_t)} + (1 - \gamma) \cdot \cos \beta(x_i, x_t),$$

$$\begin{cases} d(x_i, x_t) = \|x_i - x_t\|_2 = \\ \sqrt{\sum_{j=1}^n (x_{i,j} - x_{t,j})^2}, \\ \beta(x_i, x_t) = \cos^{-1} \left(\frac{x_i^T x_t}{\|x_i\|_2 \cdot \|x_t\|_2} \right). \end{cases} \quad (1)$$

其中: $d(x_i, x_t)$ 表示向量 x_i 与 x_t 的欧氏距离, $\beta(x_i, x_t)$ 为向量夹角, $\gamma \in [0, 1]$ 表示权重参数. 此外, 相似度 $s_1(x_i, x_t) \in [0, 1]$, 且 $s_1(x_i, x_t)$ 越接近于 1, x_i 与 x_t 越相似.

在历史数据库中搜索 $s_1(x_i, x_t)$ 最大的 k 组数据样本并降序排列, 以构造 x_t 的最大近邻学习集 Ω_k , 即

$$\Omega_k = \{(x_1, y_1), \dots, (x_k, y_k) | s(x_1, x_q) > \dots > s(x_k, x_q)\}. \quad (2)$$

这样, 运行指标建模问题转化为如下优化问题:

$$J(\theta) = \min_{(x_i, y_i) \in \Omega_k} (f_{EC1}(x_i, \theta) - f_1(x_i, \theta))^2 s_1(x_i, x_t). \quad (3)$$

进一步, 采用自回归模型作为局部模型结构, 极小化式(3)所示目标函数 $J(\theta)$ 的解可得当前工作点的模型参数 θ , 从而得到当前工作点的局部模型.

为了提高模型精度, 采用基于互信息的相似度指标, 将工作点的输入输出值组合在一起, 计算其与历史数据库中其他点的欧氏距离并进行升序排序, 输入输出向量间的互信息如下式所示:

$$\begin{cases} MI(x_d, y) = \\ \phi(K) - \frac{1}{N} \sum_{i=1}^N (\phi(nx_{id} + 1) + \\ \phi(ny_{id} + 1)) + \phi(N), \\ \phi(n+1) = \phi(n) + 1/n. \end{cases} \quad (4)$$

其中: $d = 1, 2$; N 为数据个数; $\phi(\cdot)$ 为 digamma 函数.

根据变量相关度决定各变量权重, 假设 $W = [w_1, w_2]^T$ 为输入权重向量, 第 d 个输入变量权重为

$$w_d = \frac{MI(x_d, y)}{\sum_{k=1}^2 MI_k(x_d, y)}, \quad d = 1, 2. \quad (5)$$

令 $x_i^w = x_i \text{diag}(W)$, $x_q^w = x_q \text{diag}(W)$, 则基于互信息的相似度指标 s_2 可表示为

$$\begin{cases} s_2(x_i^w, x_t^w) = \exp(-d(x_i^w, x_t^w)^2), \\ d(x_i^w, x_t^w) = \sqrt{\sum_{d=1}^m (x_i^w - x_t^w)^2}. \end{cases} \quad (6)$$

与常规即时学习计算一样, 可得当前工作点的局部模型. 根据信息熵法获取单一模型的权重 ω , 利用集成学习方法可得最终能耗与出水水质的性能指标模型如下:

$$\begin{cases} f_{EC}(x(t)) = \\ \omega_1 \sum_{k=1}^{k_1} f_1(x_i, \theta_1) s_1(x_i, x_t) + \\ \omega_2 \sum_{k=1}^{k_2} f_1(x_i, \theta_2) s_2(x_i, x_t), \\ f_{EQ}(x(t)) = \\ \omega_3 \sum_{k=1}^{k_1} f_2(x_i, \theta_3) s_3(x_i, x_t) + \\ \omega_4 \sum_{k=1}^{k_2} f_2(x_i, \theta_4) s_4(x_i, x_t). \end{cases} \quad (7)$$

3 基于隐性记忆的多目标运行优化与设定值智能决策

为了在操作受限情况下提高污染物去除性能, 降低运行成本, 需同时优化 EC 和 EQ, 从而决策过程控制系统的优化设定值. 因此, 建立以下两目标运行优化问题:

$$\begin{aligned} \min_{x_1^*, x_2^*} F(x(t)) &= \{f_{EC}(x(t)), f_{EQ}(x(t))\}; \\ \text{s.t. } f_{EC}(x(t)) &= \omega_1 \sum_{k=1}^{k_1} f_1(x_i, \theta_1) s_1(x_i, x_t) + \\ &\quad \omega_2 \sum_{k=1}^{k_2} f_1(x_i, \theta_2) s_2(x_i, x_t), \\ f_{EQ}(x(t)) &= \omega_3 \sum_{k=1}^{k_1} f_2(x_i, \theta_3) s_3(x_i, x_t) + \\ &\quad \omega_4 \sum_{k=1}^{k_2} f_2(x_i, \theta_4) s_4(x_i, x_t), \end{aligned}$$

$$x_{\min, i} \leq x_i \leq x_{\max, i}, \quad i = 1, 2. \quad (8)$$

其中: 将已建立的 t 时刻 EC 模型 $f_{EC}(x(t))$ 和 EQ 模型 $f_{EQ}(x(t))$ 作为优化目标函数; $x(t) = [x_1(t), x_2(t)]^T$ 为决策变量, 即 S_{O5} 和 S_{NO2} 的设定值.

为了解决上述多目标运行优化问题, 本文设计一种基于隐性记忆的多目标优化方法, 主要包括: 基于隐性记忆的动态机制、多样性增强的 NSGA II 算法和过程控制设定值智能决策等.

3.1 基于隐性记忆的动态机制

污水处理过程的污水被动接受且不规律非平稳变化,精确储存和重用过去的解无法满足该过程运行优化的需求. 因此,提出隐性记忆的方法,建立知识库,并将自动编码进化搜索^[15]应用于无规律非平稳动态时变工业过程的多目标运行优化问题,以充分利用历史信息,得到深埋在历史优化周期中的结构化知识,并与实际污水运行工况相结合,从而最大限度地利用污水处理的运行工况信息.

知识库通过保存污水处理过程中的历史问题信息及其最优解来指导新问题的求解,知识库中的信息表示为 $KL_i = (x_i, P_i)$. 其中: x_i 表示保存的第 i 个环境监测变量, P_i 表示第 i 个历史问题的最优解. 知识库的主要操作包括: 工况匹配相似度、相似历史数据的应用和知识库的更新等^[16]. 工况相似度检测由式(1)所示 K-VNN 方法执行. 若在知识库中搜索到大于相似度阈值的工况信息,则根据历史工况与当前工况的相似度确定历史解集引入当前初始种群的比例. 若在知识库中没有搜索到相似工况信息,则利用自动编码进化搜索模块进行信息提取和搜索引导,并将最后得到的优化解集存入知识库.

自动编码进化搜索从历史过程的种群信息中提取结构化知识,采用单层去噪自动编码器,构建映射关系 M . 然后,将知识库中与当前工况相似度最大的最优解与 M 相乘得到的隐性记忆解,通过自适应策略加入到当前种群中. 假设当前优化周期为 t , 周期 $t - 2$ 和周期 $t - 1$ 的非支配解集分别为 $NDS_{t-2} = \{p_1, p_2, \dots, p_N\}$ 和 $NDS_{t-1} = \{q_1, q_2, \dots, q_N\}$, 其中 N 表示种群规模. 这里将 NDS_{t-2} 视为 NDS_{t-1} 的带噪声的“破坏”版本. 采用如下单层映射 M 重构被“破坏”的信息之间的关系,其损失函数为

$$L(M) = \frac{1}{2N} \sum_{i=1}^N \|q_i - Mp_i\|^2. \quad (9)$$

采用最小二乘求解,得

$$M = (NDS_{t-1} NDS_{t-2}^T) (NDS_{t-2} NDS_{t-2}^T)^{-1}, \quad (10)$$

其中 M 为 NDS_{t-1} 与 NDS_{t-2} 之间的映射. 另外,由于过多围绕隐性记忆解的搜索可能会导致多样性的丢失,从而陷入局部最优. 隐性记忆动态机制采用每迭代 h 次引入部分隐形记忆解的方式引导搜索,引入隐性记忆解的过程由自适应系数 α 进行调节,有

$$\alpha_i = \alpha_{\max} - \log_{i_{\max}} i \times (\alpha_{\max} - \alpha_{\min}). \quad (11)$$

在进化搜索前期,自适应系数 α 保持较大值,有利于

加快知识引导进程. 随着迭代进行, α 迅速减小,加快算法收敛,如图3所示.

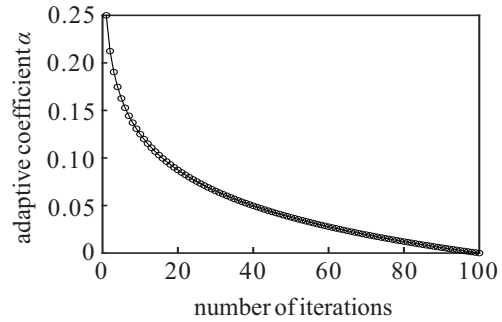


图3 自适应迭代系数

隐性记忆动态机制流程如图4所示. 对于给定新周期的优化问题,算法首先根据环境监测变量判断知识库中是否存在工况与当前工况匹配. 若存在,则将历史最优解按照相似度比例引入初始种群,引导搜索;若不存在,则激活自动编码进化搜索,对之前优化周期经验进行知识提取操作,引入当前周期种群搜索过程,加速算法收敛.

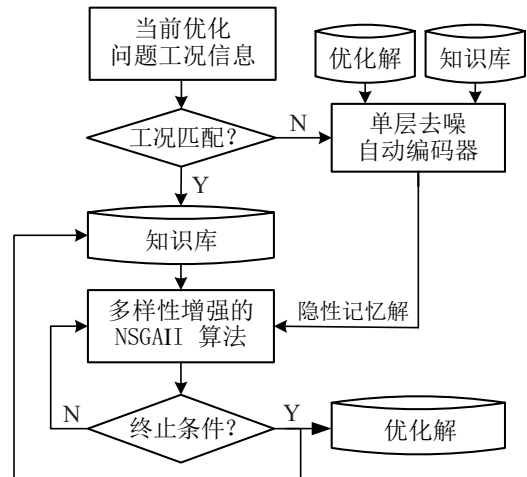


图4 隐性记忆动态机制流程

3.2 多样性增强的NSGA II算法

在满足操作限制的情况下,最大限度降低经济成本和环境影响一直是污水处理操作优化的关键问题. 为了维持和增强经典NSGA II算法^[17-18]进化种群的多样性和分布性,得到更高质量 SO_5 和 S_{NO2} 的优化设定值,本文采用一种平均距离聚类^[19]的多样性指标,并引入均匀随机局部搜索策略^[20],具体方法如下.

常规拥挤度距离指标主要利用个体的局部信息,因而可能导致算法陷入局部最优. 为此,本文采用聚类思想,基于平均距离聚类的多样性指标,把平均距离内的多个个体划分为一个小种群. 假设种群 $P = [x_1, x_2, \dots, x_N]$ 有 N 个个体, $f(x_i) = [f_1(x_i), f_2(x_i),$

$\dots, f_M(x_i)]$ 为个体 x_i 的 M 个目标函数, 则该指标可描述为

$$f_{j,step} = \frac{S \times (f_{j,max} - f_{j,min})}{2N}. \quad (12)$$

其中: $f_{j,max} = \max\{f_j(x_1), f_j(x_2), \dots, f_j(x_N)\}$ 和 $f_{j,min} = \min\{f_j(x_1), f_j(x_2), \dots, f_j(x_N)\}$ 分别表示种群 P 中第 j 个目标函数的最大值和最小值, S 表示每个小种群所含个体数量.

将上述平均距离聚类指标应用于 NSGA II 的选择过程: 首先基于指标将平均距离内的多个个体划分为一个小种群; 然后对小种群执行选择操作, 判断小种群内个体 a_j 与 a_i 之间的支配关系, 选择较优个体. 若个体 a_j 与 a_i 互不支配, 则比较平均距离内的个体数量, 选择较少个体; 若个体 a_j 与 a_i 平均距离内个体数量相同, 则随机选择个体.

交叉过程中, 引入小种群决策变量的平均距离, 即引入扰动操作, 增加种群的多样性. 如图 5 所示, 若 A 和 B 为随机选择的交叉父代, 则 A 与 B 的子代一定落在小矩形框里. 若引入小种群决策变量的平均距离, 则可以确定 A 和 B 的小种群范围, 通过随机选择 A 点所在区域的个体 A' 和 B 点所在区域的个体 B', A 与 B 交叉产生的子代范围将比标准交叉算子产生子代范围更大(大矩形框所在区域).

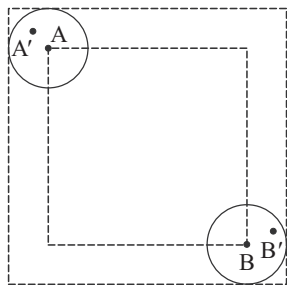


图 5 交叉算子示意

定义交叉运算算子 \otimes , 以及定义 $x^A = (x_1^A, x_2^A, \dots, x_D^A)$ 和 $x^B = (x_1^B, x_2^B, \dots, x_D^B)$ 为交叉运算的父代, 则交叉过程可描述为

$$x^{new} = (x^A + (r_1 \times x_{step}^{SPA})) \otimes (x^B + (r_2 \times x_{step}^{SPB})). \quad (13)$$

其中: x^{new} 表示产生的子代; $r_1, r_2 \in \text{rand}[-1, 1]$; $x_{step}^{SPA} = (x_{1,step}^{SPA}, x_{2,step}^{SPA}, \dots, x_{D,step}^{SPA})$ 为 A 点所在小种群 SP_A 的平均距离; $x_{step}^{SPB} = (x_{1,step}^{SPB}, x_{2,step}^{SPB}, \dots, x_{D,step}^{SPB})$ 为 B 点所在小种群 SP_B 的平均距离.

局部搜索策略可以有效提高 NSGA II 算法的分布性和多样性, 现已提出多种基于局部搜索策略的改进型 NSGA II 算法, 如文献 [21-23]. 然而, 由于局部搜

索过程产生大量局部解, 使得计算量成倍增加. 因此, 本文在平均距离指标最小的小种群内展开局部搜索, 在提升解的分布性和多样性的同时, 可大大减小计算量. 这里采用一种均匀随机局部搜索策略, 这是一种无导数局部搜索技术, 具体描述如下.

若 t 代变异操作之后平均距离指标最小的小种群内个体表示为 $M_t = \{y_{1,t}, y_{2,t}, \dots, y_{S,t}\}$, 基于平均距离聚类多样性指标展开局部搜索, 即

$$x_{i,t} = y_{i,t} + \Delta_t U, \quad (14)$$

其中 Δ_t 为步长参数, 计算公式为

$$\Delta_{t+1} = \frac{1}{K} \sum_{i=1}^K \gamma^i. \quad (15)$$

$\Omega = \{x_1, x_2, \dots, x_q\}$ 是在种群中随机选取 q 个不同的个体得到的集合, 计算其均值 $\bar{x} = (1/q) \sum_{i=1}^q x_i$. 比较每个 $x_i \in \Omega$ 与 \bar{x} 之间的欧氏距离, 选取其中 K 个最小的欧氏距离得到 $\{\gamma^1, \gamma^2, \dots, \gamma^K\}$. 当 $t = 0$ 时, $\Delta_0 = \tau \times \max\{u^j - l^j, j = 1, 2, \dots, n\}$, $\tau \in [0, 1]$. 利用 Δ_0 生成初始步长, 是为了加速搜索, 从一个大的步长开始, 快速遍历搜索空间. 此外, 式 (14) 中 $U = (U_1, U_2, \dots, U_n)^T$ 为随机分量的余弦, 计算公式如下:

$$U_j = R_j / (R_1^2 + \dots + R_n^2)^{\frac{1}{2}},$$

$$R_j \sim \text{Unif}([-1, 1]), j = 1, 2, \dots, n. \quad (16)$$

如果 $x_{i,t}$ 不在搜索空间内, 则

$$x_{i,t} = \begin{cases} y_{i,t} + \lambda(u - y_{i,t}), & x_{i,t} > u; \\ y_{i,t} + \lambda(y_{i,t} - l), & x_{i,t} < l. \end{cases} \quad (17)$$

其中 $\lambda \sim \text{Unif}([0, 1])$.

本文所提出的基于隐性记忆的 NSGA II (IM-NSGA II) 算法实现步骤如图 6 所示. 在初始化相关参数后, 根据环境监测变量检测知识库中是否存在当前优化周期的相似工况:

1) 若存在, 则根据相似度确定历史解集引入当前初始种群的比例, 剩余初始种群随机生成, 之后对当代种群进行快速非支配排序, 采用平均距离聚类的多样性指标将种群划分为多个小种群, 并基于该指标执行选择、交叉操作; 然后执行多项式变异, 在平均距离指标最小的小种群内展开局部搜索; 最后, 经环境选择更新种群个体, 重复以上步骤.

2) 若不存在, 则随机生成初始种群, 且启动自动编码进化搜索, 将隐性记忆解引入环境选择阶段, 引导搜索, 直到达到终止条件, 获得一组 Pareto 最优解.

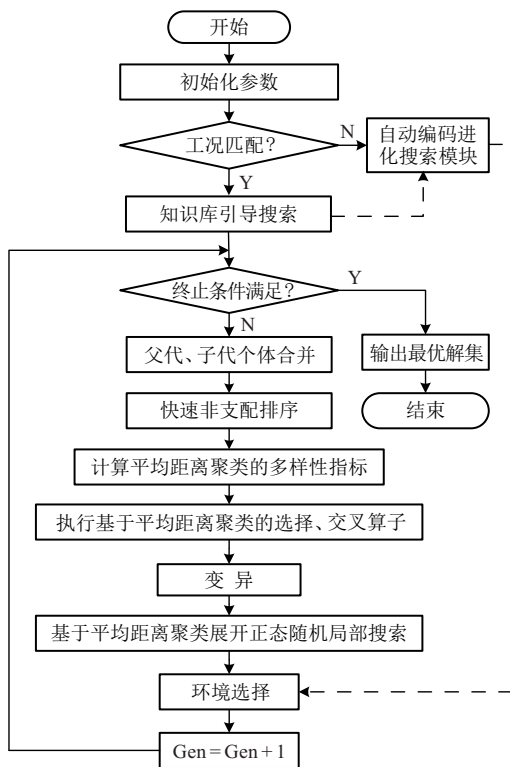


图6 所提出IM-NSGA II算法实现流程

3.3 控制系统设定值智能决策

多目标优化方法获得的 Pareto 最优解集具有众多等同的优秀解,如何从中选择出一组适合当前工况和特定工艺需求的解,使得 S_{O_5} 和 S_{NO_2} 的最优设定值能够在满足操作约束的前提下保证出水水质达标和降低运行能耗,是污水处理运行优化必须解决的另一关键问题. 实际污水处理过程是一个慢反应过程,从操作到产生结果有较大时延,想要保证出水水质达标,必须对其进行预测. 为此,本文提出一种基于集成即时学习预测的控制系统设定值智能决策方法.

首先,采用前述集成即时学习算法建立氨氮浓度 (S_{NH})、总氮浓度 (N_{tot})、生化需氧量 (BOD)、化学需氧量 (COD)、总悬浮物固体浓度 (TSS) 等出水水质指标的预测模型. 然后,判断优化所得 Pareto 解集的 5 种出水水质预测值是否满足排放标准. 若不满足,则从 Pareto 解集选择 EQ 最小值所对应的解作为最优设定值;若 5 种出水水质预测值全部达标,则需要在 Pareto 解集中选择出一组最适合当前工况的解. 由于出水水质和运行能耗两目标间存在不可公度性,本文利用非整秩次秩和比法 (rank-sum ratio, RSR)^[24] 对得到的 Pareto 解集进行综合评价,选取最优设定值. 非整秩次秩和比法将原始数据中的信息转化为秩次的形式,结合统计理论为目标信息排序,具体步骤如下:

对出水水质和运行能耗指标进行编秩操作,对于第 i 个解的第 j 个性能指标函数 f_{ij} , 其秩次表述为

$$R_{ij} = 1 + \frac{(n-1)[f_{ij} - \min(f_{1j}, \dots, f_{nj})]}{\max(f_{1j}, \dots, f_{nj}) - \min(f_{1j}, \dots, f_{nj})} \quad (18)$$

每组决策变量的评价指标秩和表示为

$$RSR_i = \frac{1}{mn} \sum_{j=1}^m R_{ij} \quad (19)$$

选取秩和最大的解作为 S_{O_5} 和 S_{NO_2} 的最优设定值用于底层跟踪控制. 这样,选取的 S_{O_5} 和 S_{NO_2} 的最优设定值能够在满足操作约束的前提下,平衡环境目标与经济影响. 最后,利用增量式 PID 控制方法对得到的优化设定值进行跟踪控制. 其中:第 1 个控制回路为 S_{O_5} 控制回路,其被控量为 S_{O_5} ,控制量为第 5 分区氧气转换系数 (K_{La5});第 2 个控制回路为 S_{NO_2} 控制回路,其被控量为 S_{NO_2} ,控制量为内回流量 (Q_a).

4 实验与分析

为了客观、公正地验证所提出方法的性能,采用国际污水处理基准仿真平台 (BSM1)^[3,25] 的雨天数据进行验证. 其中:采样间隔为 15 min,运行优化周期为 2 h,底层控制器的采样周期为 45 s. 实验中阴雨天气工况的前 8 天的进水情况与晴天基本一致,但是在第 9 天~第 11 天持续降雨,进水流量猛增,污染物浓度明显降低,直到第 11 天降雨结束,进水流量和污水组分浓度又回归到晴天平稳状态^[26-27]. 实验采用雨天工况下前 7 天的数据进行训练,第 8 天~第 14 天 (共 7 天) 的数据用于测试实验,测试所提出的基于隐性记忆的运行优化方法的效果.

首先测试本文建模方法建立的能耗与水质运行指标优化模型的性能. 为此,分别采用常规即时学习在线建模算法与本文集成即时学习在线建模 (E-JITL) 算法进行比较,如图 7 所示.

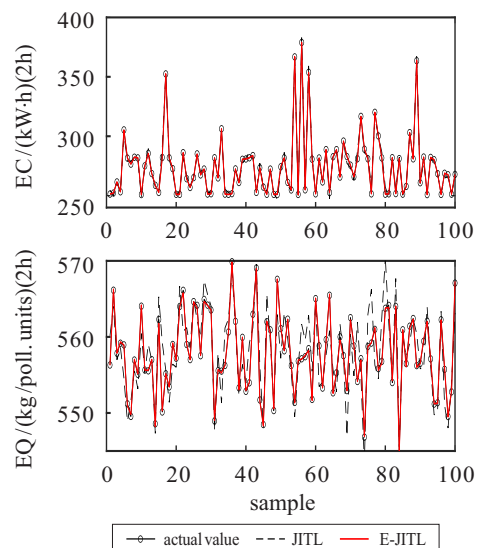


图7 运行指标模型建模效果

由图7可以看出:常规即时学习算法建立的优化模型在拐点处无法精确有效地跟踪实际值,尤其在水质EQ模型中出现了较大的偏差;而采用所提出的基于互信息相似度的集成即时学习算法得到了更好的预测效果.因此,本文E-JITL建模算法更适用于建立非平稳动态时变污水处理过程的运行优化模型.

在运行优化实验时,本文基于隐性记忆NSGA II (IM-NSGA II)的多目标运行优化方法的参数选取如下:1)种群规模 $N = 100$; 2)最大迭代次数 $i_{max} = 100$; 3)NSGA II相关参数:交叉概率 $p_c = 0.9$,变异概率 $p_m = 0.2$,交叉参数 $\eta_c = 20$,变异参数 $\eta_m = 20$; 4)基于记忆的动态机制参数:引入隐性记忆迭代次数 $h = 5$,自适应参数上界 $\alpha_{max} = 0.25$,自适应参数下界 $\alpha_{min} = 0$,相似度阈值 $\beta = 0.8$; 5)平均距离指标参数 $S = 1$.根据Friedman检验,表1给出了 S 取值分别为1~6时HV(hypervolume)指标^[28]秩均值排名(排名已修正).表1中的结果表明,当 $S = 1$ 时具有更好的算法性能.

表1 参数 S 取值测试结果

value of S	mean (HV)	ranking
1	0.985 1	2.8
2	0.985 0	3.2
3	0.980 0	3.8
4	0.982 5	3.8
5	0.980 6	3.6
6	0.982 8	3.8

此外,实验中环境检测周期与优化周期保持一致,环境监测变量选择进水流量(Q_0)和入水总氮浓度($N_{tot;n}$),这些变量的典型值以及对应的最优解将作为案例储存在知识库中.需要注意的是,隐性记忆动态机制在每次工况变化时仅运行一次,式(10)得出的 M 是一个封闭解.图8给出了新工况时,所提出隐性记忆动态机制运行过程示例.可以看出,当种群信息构成的变量为2维,种群数量为100时,得到它们之间的结构化知识矩阵 M 的大小为 2×2 ,因而不会为运行优化过程增加很大的计算量.

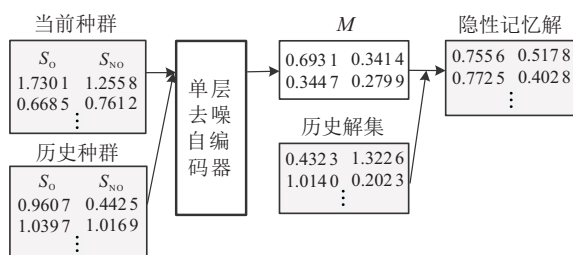


图8 隐性记忆动态机制示例

图9为雨天工况下第10天连续12次运行优化得出的Pareto前沿曲线.可以看出,所提出IM-NSGA II算法的Pareto前沿离原点更近,优化结果明显优于经典的NSGA II算法,优化得到的EC和EQ更优.另外,进一步采用HV(hypervolume)指标和SP(spacing)指标^[28-29]评价NSGA II算法和所提出IM-NSGA II算法的污水处理多目标运行优化效果.HV指标是唯一一种可以评价未知情况Pareto解集质量的综合性能指标,数值越大,算法的优化效果越好.

图10给出了HV指标单次收敛曲线,在迭代次数为50次时,所提出IM-NSGA II算法已经基本收敛,其效果明显优于经典NSGA II算法.这表明隐性记忆动态机制可以显著加快进化搜索进程,并提升优化效果.SP指标是一种评价算法分布性的指标,数值越小,算法分布性越好,其收敛曲线如图11所示.可以看出,所提出IM-NSGA II算法的SP数值明显更低,表明IM-NSGA II算法的分布性优于经典的NSGA II算法,可以得到更高质量的 S_{O_5} 和 S_{NO_2} 的优化设定值.

表2给出了以第10天为例的HV指标和SP指标的结果均值,这里采用威克逊秩和检验来验证经典NSGA II算法和所提出IM-NSGA II算法的运行优化结果在 $P = 5\%$ 显著水平下是否存在显著性差异.原假设 H_0 表述为:两种算法不存在显著性差异.当 $P < 5\%$ 时,拒绝原假设,表明两算法之间存在显著性差异.计算得出两种指标的检验结果均为 $P = 0.2\%$,表明所提出IM-NSGA II算法的寻优效果明显优于经典NSGA II算法.

在图9得到Pareto最优解集的基础上,利用所提出基于集成即时学习预测的控制系统设定值智能决策算法,进一步得出雨天工况下第10天经典NSGA II算法和所提出IM-NSGA II算法的设定值决策结果,如图12所示.可以看出,所提出IM-NSGA II算法得到优化解的出水水质和运行能耗比经典NSGA II算法小得多,从而进一步验证了所提出IM-NSGA II在污水处理过程运行优化的优越性能.在阴雨工况下,利用所提出IM-NSGA II算法对污水处理进行优化及控制,得到5种关键出水水质,即 S_{NH} 、 N_{tot} 、BOD、COD、TSS的变化情况,如图13所示.可以看出,各出水水质指标变化平稳,边界清晰,均达到国家排污标准.

为了进一步验证所提出的基于IM-NSGA II运行优化方法及相应控制方法的性能,将其与4种不同的运行优化控制策略进行对比,并根据Friedman检验、Holm方法对不同控制方法进行排序,结果如表3所示.

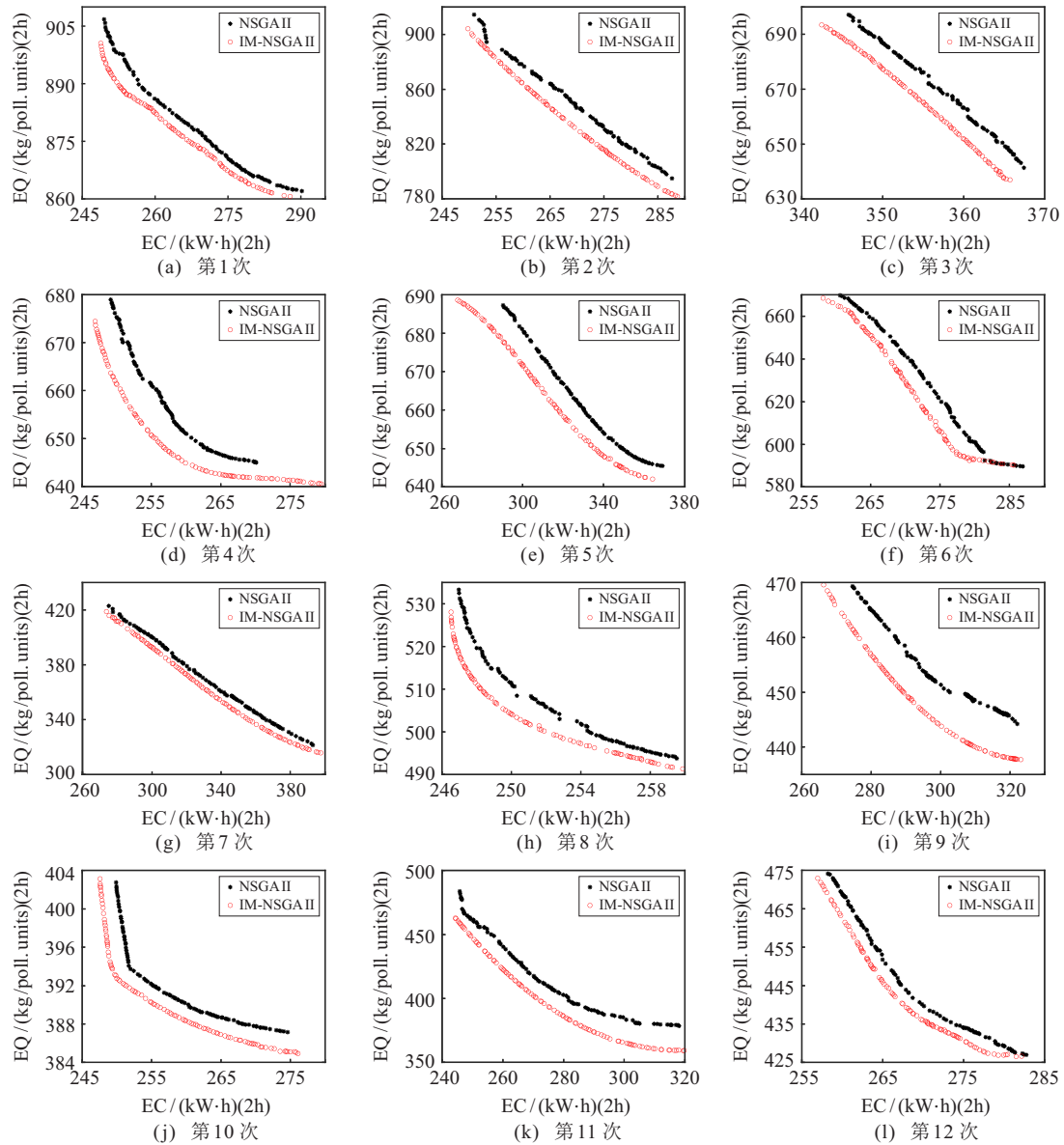


图9 连续12次运行优化的Pareto最优解

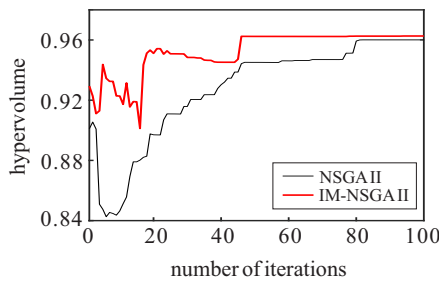


图10 两种算法HV指标收敛曲线

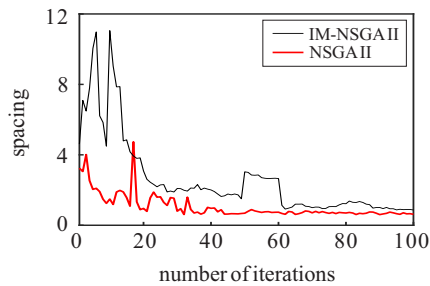


图11 两种算法Spacing指标收敛曲线

表2 两种算法评价指标值

	HV		SP	
	NSGA II	IM-NSGA II	NSGA II	IM-NSGA II
	0.898 0	0.898 7	0.308 0	0.297 2
	0.816 4	0.817 3	0.803 4	0.643 3
	0.807 2	0.809 7	0.350 6	0.241 8
	0.847 0	0.877 4	0.234 3	0.231 8
	0.946 1	0.980 8	0.539 6	0.537 9
	0.798 5	0.803 8	0.459 9	0.432 5
	0.845 0	0.845 6	0.995 5	0.823 1
	0.803 1	0.809 2	0.230 5	0.219 2
	0.900 2	0.909 4	0.407 9	0.368 2
	0.866 5	0.871 4	0.157 2	0.150 3
	0.805 6	0.806 6	0.584 2	0.550 1
	0.811 6	0.814 5	0.328 3	0.260 2

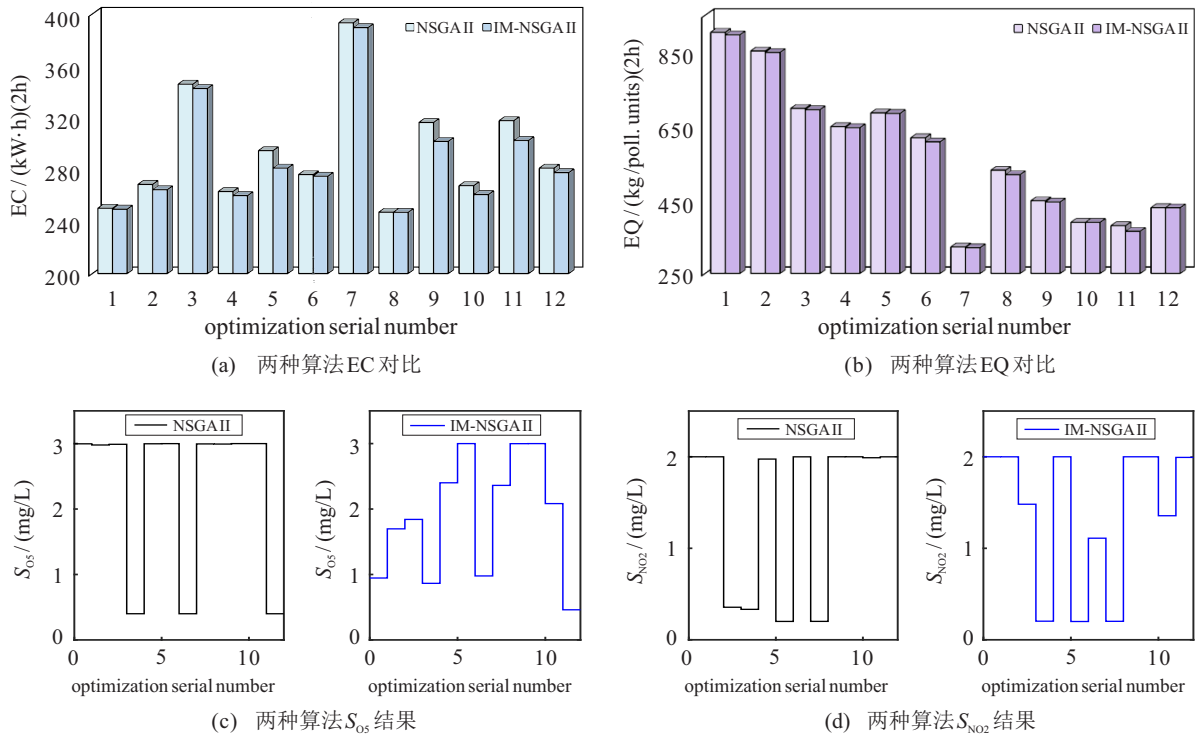


图 12 过程控制系统设定值智能决策结果

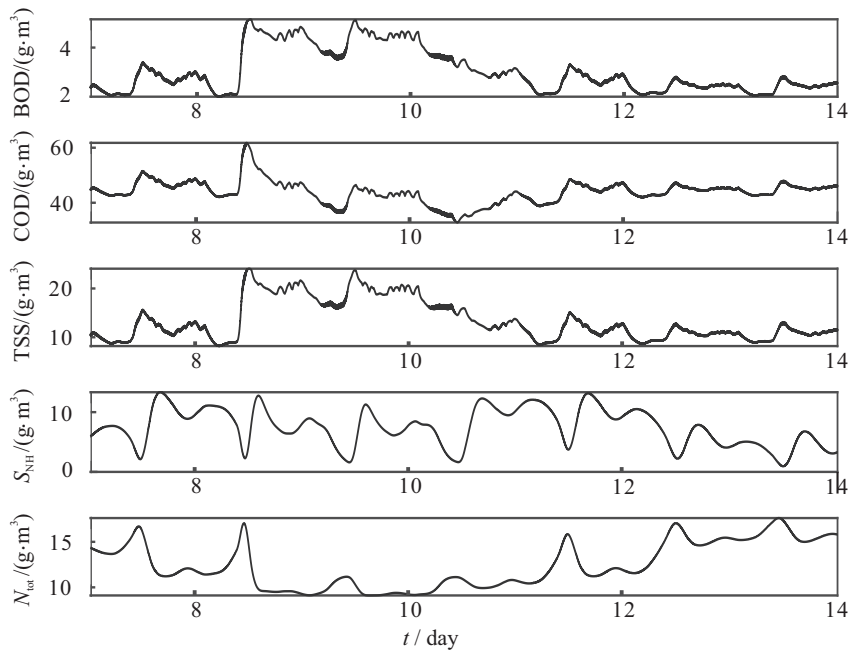


图 13 IM-NSGA II方法得到的5种主要出水水质实际值

表 3 不同运行优化方法性能比较

方法	EC				EQ			
	mean	ranking	Z	P-value	mean	ranking	Z	P-value
Open-loop	4810.33	5.0	5.656854	0	8871.33	3.9	3.959798	0
Closed-loop	3829.15	3.5	3.535534	0	8701.00	3.4	3.252691	0
NSGA II	3787.17	2.9	2.687006	0.000037	8300.95	2.3	1.697056	0.000164
SPEA2 ^[30]	3783.99	2.6	2.121320	0.000059	8984.44	4.3	4.525483	0
IM-NSGAII	3689.59	1.0	—	—	7566.84	1.1	—	—

由表3可以看出,所提出方法得到的污水处理运行能耗和出水水质均值都明显优于其他运行优化控制, Friedman 检验排序结果也验证了这一点. 综上,所提出IM-NSGA II算法及相应的运行优化方法是一种解决污水处理过程运行优化问题的有效方法,在出水水质达标的基础上可显著降低运行能耗,最大限度降低经济成本 and 环境影响.

5 结论

本文提出了一种基于隐性记忆NSGA II的多目标优化算法,用于解决无规律非平稳动态时变污水处理过程出水水质不达标、运行能耗高的运行优化问题. 首先,基于集成即时学习的运行指标优化模型能够根据污水处理过程非平稳时变工况进行在线建模. 然后,在经典NSGA II算法基础上引入隐性记忆动态机制,在无规律、非平稳变化的新工况发生时,利用单层去噪自动编码器提取历史解集的结构化知识,引导种群搜索过程,增强进化搜索效果,并将结果储存在知识库中. 结合平均距离聚类的多样性指标,在种群稀疏的位置展开均匀随机局部搜索策略,提高算法的多样性和分布性. 最后,设计基于集成即时学习的智能决策方法,得到更高质量的溶解氧浓度和硝态氮浓度控制系统的设定值. 数据实验表明,所提出算法在满足操作限制的情况下,提升了污染物去除性能,降低了运营成本,兼顾了环境影响与经济性能的平衡.

参考文献(References)

- [1] Shannon M A, Bohn P W, Elimelech M, et al. Science and technology for water purification in the coming decades[J]. *Nature*, 2008, 452(7185): 301-310.
- [2] Zhou P, Wang X, Chai T Y. Multiobjective operation optimization of wastewater treatment process based on reinforcement self-learning and knowledge guidance[J]. *IEEE Transactions on Cybernetics*, DOI: 10.1109/TCYB.2022.3164476.
- [3] Sun X Y, Zhou P, Ding J L, et al. Enhanced MPC based on unknown state estimation and control compensation[J]. *Journal of Process Control*, 2023, 121: 60-72.
- [4] Araromi D O, Majekodunmi O T, Adeniran J A, et al. Modeling of an activated sludge process for effluent prediction — A comparative study using ANFIS and GLM regression[J]. *Environmental Monitoring and Assessment*, 2018, 190(9): 495.
- [5] Asadi A, Verma A, Yang K, et al. Wastewater treatment aeration process optimization: A data mining approach[J]. *Journal of Environmental Management*, 2017, 203: 630-639.
- [6] Meng X, Qiao J F, Han H G. Soft measurement of key effluent parameters in wastewater treatment process using brain-like modular neural networks[J]. *Acta Automatica Sinica*, 2019, 45(5): 906-919.
- [7] Yang Z, Yang C L, Gu K, et al. Multi-objective evolutionary algorithm for wastewater treatment process optimization control[J]. *Control Theory & Applications*, 2020, 37(1): 169-175.
- [8] Zhang R, Xie W M, Yu H Q, et al. Optimizing municipal wastewater treatment plants using an improved multi-objective optimization method[J]. *Bioresource Technology*, 2014, 157: 161-165.
- [9] Lu L, Zheng H, Jie J, et al. Reinforcement learning-based particle swarm optimization for sewage treatment control[J]. *Complex & Intelligent Systems*, 2021, 7(5): 2199-2210.
- [10] Qiao J F, Han G T, Zhou H B. Knowledge-based intelligent optimal control for wastewater biochemical treatment process[J]. *Acta Automatica Sinica*, 2017, 43(6): 1038-1046.
- [11] Zhang S, Zhou P, Xie Y F, et al. Improved model-free adaptive predictive control method for direct data-driven control of a wastewater treatment process with high performance[J]. *Journal of Process Control*, 2022, 110: 11-23.
- [12] Du X J, Wang J L, Jegatheesan V, et al. Dissolved oxygen control in activated sludge process using a neural network-based adaptive PID algorithm[J]. *Applied Sciences*, 2018, 8(2): 261.
- [13] Zhou P, Chen W Q, Yi C M, et al. Fast just-in-time-learning recursive multi-output LSSVR for quality prediction and control of an ironmaking blast furnace[J]. *Engineering Applications of Artificial Intelligence*, 2021, 100: 104168.
- [14] Kuroda A, Takiguchi N, Gotanda T, et al. A simple method to release polyphosphate from activated sludge for phosphorus reuse and recycling[J]. *Biotechnology and Bioengineering*, 2002, 78(3): 333-338.
- [15] Feng L, Ong Y S, Jiang S W, et al. Autoencoding evolutionary search with learning across heterogeneous problems[J]. *IEEE Transactions on Evolutionary Computation*, 2017, 21(5): 760-772.
- [16] Zhang Y, Yang K, Hao G S, et al. Evolutionary optimization framework based on transfer learning of similar historical information[J]. *Acta Automatica Sinica*, 2021, 47(3): 652-665.
- [17] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. *IEEE Transactions on Evolutionary Computation*, 2002, 6(2): 182-197.
- [18] Wang F, Han M C, Zhao Y Y, et al. An improved NSGA-II algorithm for multi-objective resource-constrained

- project scheduling problem[J]. Control and Decision, 2021, 36(3): 669-676.
- [19] Cui Z H, Zhang M Q, Chang Y, et al. NSGA-II with average distance clustering[J]. Acta Automatica Sinica, 2021, 47(5): 1171-1182.
- [20] Sawyerr B A, Adewumi A O, Ali M M. Real-coded genetic algorithm with uniform random local search[J]. Applied Mathematics and Computation, 2014, 228: 589-597.
- [21] Qian X Y, Fang W. Opposition-based learning competitive particle swarm optimizer with local search[J]. Control and Decision, 2021, 36(4): 779-789.
- [22] Li S Y, Li W J, Qiao J F. A local search strategy based on density for NSGA2 algorithm[J]. Control and Decision, 2018, 33(1): 60-66.
- [23] Wang Y K, Chen X B. Improved multi-area search and asymptotic convergence PSO algorithm with independent local search mechanism[J]. Control and Decision, 2018, 33(8): 1382-1390.
- [24] Wu X, Shen S. Comprehensive evaluation of medical service efficiency in TCM hospitals based on data envelopment method and rank sum ratio method[C]. IEEE International Conference on Bioinformatics and Biomedicine (BIBM). San Diego, 2019: 2486-2492.
- [25] Jeppsson U, Pons M N, Nopens I, et al. Benchmark simulation model no 2: General protocol and exploratory case studies[J]. Water Science and Technology, 2007, 56(8): 67-78.
- [26] Jiang W K, Zhou P. Multi-condition fault diagnosis based on adaptive transfer and broad learning under imbalanced data[J]. Scientia Sinica Informationis, 2022, 52: 1-26.
- [27] Han G, Qiao J F, Han H G, et al. Optimal control for wastewater treatment process based on Hopfield neural network[J]. Control and Decision, 2014, 29(11): 2085-2088.
- [28] Bader J, Zitzler E. HypE: An algorithm for fast hypervolume-based many-objective optimization[J]. Evolutionary Computation, 2011, 19(1): 45-76.
- [29] Han H G, Liu Z, Lu W, et al. Dynamic MOPSO-based optimal control for wastewater treatment process[J]. IEEE Transactions on Cybernetics, 2021, 51(5): 2518-2528.
- [30] Zitzler E, Laumanns M, Thiele L. SPEA2: Improving the strength Pareto evolutionary algorithm for multiobjective optimization[C]. Proceedings of the 5th Conference on Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems. Athens, 2001: 95-100.

作者简介

周平(1980—),男,教授,博士生导师,从事复杂工业过程运行优化控制、数据驱动智能建模等研究, E-mail: zhouping@mail.neu.edu.cn;

张天娇(1998—),女,硕士生,从事复杂工业过程运行优化控制的研究, E-mail: Zhangtianjiao1225@163.com.



特邀专家 周平,分别于2003、2006和2013年本、硕和博毕业于东北大学,2010年留校任教,并分别于2012、2015和2018年评为讲师、副教授和教授,目前是东北大学的教授和博士生导师,主要从事工业自动化方面的教学与科研工作. 入选第四批国家“万人计划”青年拔尖人才、辽宁省“兴辽英才”和“百千万人才”百人层次. 兼任IFAC-MMM技术委员会委员、IEEE计算智能学会神经网络委员会“随机神经网络与学习系统”工作组成员、IEEE高级会员. 以第1(通讯)作者发表期刊论文100余篇,第1作者学术专著2部,第1发明人中美发明专利40余件;主持国家重点级科研项目4项,并荣获中国自动化学会自然科学一等奖(1/5)、教育部自然科学一等奖(3/4)和二等奖(2/4)、首届中国自动化学会优博、辽宁省优博等奖励.

专家寄语 作为本硕博学习和毕业工作一直在母校的东北大学人,我深切感谢母校的栽培,并亲身体会百年东大开拓奋进的点点滴滴. 值此东北大学百年华诞之际,祝愿母校积百年历史之厚蕴,继往开来,开拓创新,续写绚丽华章.