

DOI:10.16799/j.cnki.esdqyfh.2023.06.064

基于 IPSO-LSSVM 的污水 BOD 预测应用研究

于洋, 王萍

[上海市市政工程设计研究总院(集团)有限公司, 上海市 200092]

摘要: 生化需氧量(BOD)的快速准确测量对于污水处理过程的调控至关重要。针对污水处理过程中 BOD 浓度测量时效性较低等问题,选择最小二乘支持向量机(LSSVM)作为 BOD 浓度预测模型,并选用粒子群优化算法(PSO)优化回归性能参数,同时使用自适应惯性权重计算方法对 PSO 进行改进,建立了 IPSO-LSSVM 预测模型。运用预测模型对某污水厂数据进行仿真研究以及 3 种误差分析,结果表明该模型具有较好的预测精度。

关键词: 污水处理; BOD 预测; IPSO; LSSVM

中图分类号: X5

文献标志码: A

文章编号: 1009-7716(2023)06-0247-03

0 引言

随着城镇化和工业的快速发展,污水处理相关措施越来越受到社会的关注。现如今污水处理过程的重点是检测污水处理过程中各项指标数据,然后对污水处理过程进行控制以达到排放标准,其中生化需氧量(BOD)是评价水质的重要指标。目前我国污水处理厂检测 BOD 主要依靠人工化验离线分析,分析测定时间为 5 d,检测滞后性较大,直接影响污水处理实时监测与控制。

另一方面,水质预测技术目前已成为过程控制与检测学科方向的一大研究热点,该技术原理为利用工业过程中容易测得的参数来推导出难以直接测量或测量方法较难、测量成本较高的参数。施汉昌等在 IAWQ 模型基础上实现了对曝气池化学需氧量(COD)的实时预测;任敏等利用神经网络将氧化还原电位(ORP)、溶解氧浓度(DO)、pH 值作为输入层的 3 个神经元来估计水厂的进水水质^[1]。

本文针对污水处理过程中 BOD 指标在线检测较难等问题,选择最小二乘支持向量机(LSSVM)作为 BOD 浓度预测模型,并选用改进的粒子群优化算法(IPSO)进行优化,建立了 IPSO-LSSVM 预测模型。根据某污水处理厂实测数据进行训练的结果表明,IPSO-LSSVM 预测模型可以实现较高精度的 BOD 浓度预测。

收稿日期: 2022-09-16

作者简介: 于洋(1996—),男,硕士,工程师,主要从事自控及仪表设计工作。

1 基于 LSSVM 的污水 BOD 预测建模

1.1 辅助变量选取

目前国内外通常是通过建立易测得的过程变量与主导变量之间的函数关系来预测主导变量。根据工艺过程分析,进水 pH 值会影响污水处理工艺过程稳定性;DO 可以判断水中微生物耗氧情况;污水温度(T)会影响水中微生物活性;COD 反映了水样中需要被氧化的物质质量;污泥浓度(MLSS)反映了污泥中微生物相对数量。因此本文选取 pH 值、DO、T、COD、MLSS 作为 BOD 浓度预测模型的辅助变量^[2]。

1.2 最小二乘支持向量机

LSSVM 最早由 Suykens 等学者提出,该模型是在支持向量机(SVM)的基础上,与最小二乘线性系统相结合得到的。LSSVM 将 SVM 中的不等式约束替代成等式约束,对应的线性系统称为 KKT 系统^[3]。

假设有输入样本 $\{x_i, y_i\}$, $x_i \in R^n$ 且 $y_i \in R, i=1, \dots, l$ 。LSSVM 优化问题为:

$$\begin{cases} \min_{w, b, \zeta} \frac{1}{2} w w^T + \frac{1}{2} \gamma \sum_{i=1}^l \zeta_i^2 \\ \text{st. } y_i = w^T \phi(x_i) + b + \zeta_i \end{cases} \quad (1)$$

式中: γ 为正则化参数本; ζ 为误差变量。

构建拉格朗日函数:

$$L(w, b, \alpha, \zeta) = \frac{1}{2} w w^T + \frac{1}{2} \gamma \sum_{i=1}^l \zeta_i^2 - \sum_{i=1}^l \alpha_i [w^T \phi(x_i) + b + \zeta_i - y_i] \quad (2)$$

式中: $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$, 是 Lagrange 乘子。

根据 KKT 系统,对式(2)求偏微分得到如下线性系统:

$$\begin{pmatrix} 0 & y^T \\ y & Q+\gamma^{-1}I \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ e \end{pmatrix} \quad (3)$$

式中: e 为元素为 1 的向量; $y=[y_1, \dots, y_l]^T$; $\alpha=[\alpha_1, \dots, \alpha_l]^T$; $Q=y_i y_j K(x_i, x_j), i, j=1, \dots, l$ 。

求解出 LSSVM 模型如下:

$$y(x)=\sum_{i=1}^l \alpha_i K(x_i, x)+b \quad (4)$$

式中: $K(x_i, x)$ 为核函数。

利用 LSSVM 进行污水 BOD 浓度预测时,核函数对模型精确度有着重要影响。本文选取对非线性问题有较好处理能力的径向基(RBF)核函数,其表达式见式(5),其中 σ 为核函数的宽度^[4]。

$$K(x_i, x)=\exp\left(-\frac{\|x_i-x\|^2}{2\sigma^2}\right) \quad (5)$$

2 粒子群优化算法

利用 LSSVM 建模时,需确定核函数宽度 σ 、正则化参数 γ 选择集合以获得最佳性能。本文利用粒子群优化算法 PSO 对参数寻优,该算法调整参数少、不易陷入局部极小,适用求解非线性问题。假设在 D 维寻找空间里某个群体由 N 个粒子组成。第 i 个粒子表示为 $X_i=(x_{i1}, x_{i2}, \dots, x_{id})$,对应的粒子速度为 $V_i=(v_{i1}, v_{i2}, \dots, v_{id})$ ^[5]。该粒子搜寻经过最好位置为 p_{best} ,所有粒子搜寻经过最好位置为 g_{best} 。速度和位置更新公式为:

$$\begin{cases} v_{id} = \omega v_{id} + c_1 r_1 (p_{id} - x_{id}) + c_2 r_2 (p_{gd} - x_{id}) \\ x_{id} = x_{id} + v_{id} \end{cases} \quad (6)$$

式中: ω 为惯性权重; r_1, r_2 为 $[0, 1]$ 内的随机数; c_1, c_2 为学习因子。

粒子群优化算法 PSO 流程如下:

Step1:初始化粒子速度 v_i 、粒子位置 x_i 等。

Step2:确定适应度函数,计算每个粒子适应度。

Step3:每个粒子适应度与 p_{best} 、 g_{best} 比较,更新

p_{best} 、 g_{best} 。

Step4:更新粒子速度和位置。

Step5:未满足结束条件,则返回 Step2。

3 粒子群算法的改进

对于惯性权重 ω 而言, ω 取值较大则全局收敛能力较强; ω 取值较小则局部收敛能力较强。为了使 PSO 在迭代初期具有较强的全局收敛能力,在迭代后期具有较强的局部收敛能力,惯性权重应随着迭

代次数增加而不断减小,即利用自适应惯性权重对 PSO 进行改进,得到改进的粒子群优化算法 IPSO。 ω 计算见式(7):

$$\omega = (\omega_{\max} - \omega_{\min}) \exp\left(-\beta \left(\frac{t}{T_{\max}}\right)^2\right) + \omega_{\min} \quad (7)$$

式中: β 根据经验取值,一般 $\beta \in [15, 20]$; t 为迭代次数; T_{\max} 为最大迭代次数。

在改进的粒子群优化算法 IPSO 中,令 $\varphi_1=c_1 \cdot \text{rand}$, $\varphi_2=c_2 \cdot \text{rand}$ 。 ω 与 φ_1, φ_2 关系需满足式(8):

$$\omega > \frac{(\varphi_1 + \varphi_2)}{2} - 1 \quad (8)$$

4 BOD 浓度预测模型实现

通过 3σ 原则选取某污水处理厂 300 组污水处理数据,其中 250 组用来训练,50 组用来测试。部分数据如表 1 所示。同时将数据进行归一化处理。

表 1 某污水处理厂实测数据

pH 值	DO/ (mg·L ⁻¹)	T/°C	进水 COD/ (mg·L ⁻¹)	MLSS/ (mg·L ⁻¹)	出水 BOD/ (mg·L ⁻¹)
7.28	3.03	20.46	313.7	0.50	10.08
7.45	1.03	22.67	278.1	0.35	11.42
7.37	2.93	23.55	289.1	0.70	9.63
7.17	2.87	20.26	252.2	0.47	11.08
7.61	2.94	19.58	285.5	0.64	12.59
7.03	1.85	23.58	261.7	0.52	13.55
7.15	2.93	19.99	302.9	0.33	8.58
7.81	3.08	20.25	299.4	0.60	12.51
7.37	2.87	23.42	247.5	0.38	10.47
7.37	2.81	19.05	188.4	0.62	11.69

首先通过 Matlab 2014 建立 LSSVM 模型以实现污水 BOD 浓度的预测,再选用 PSO 优化 LSSVM 模型,即 PSO-LSSVM 模型。设置粒子群初始参数学习因子 $c_1=1.5, c_2=1.5$,最大迭代次数为 200。通过训练数据的学习,将得到的核函数宽度 σ 、正则化参数 γ 这 2 个参数运用到预测模型中,得到预测结果如图 1 所示,相对误差如图 2 所示。

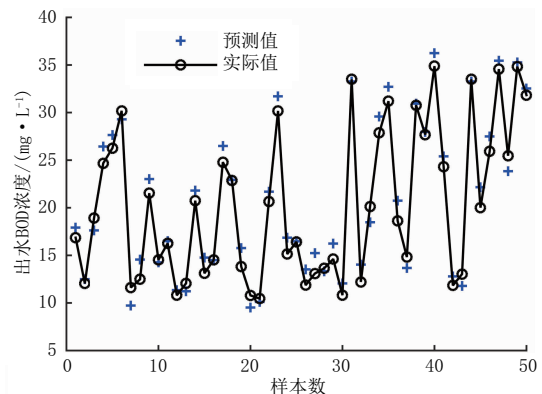


图 1 PSO-LSSVM 预测结果

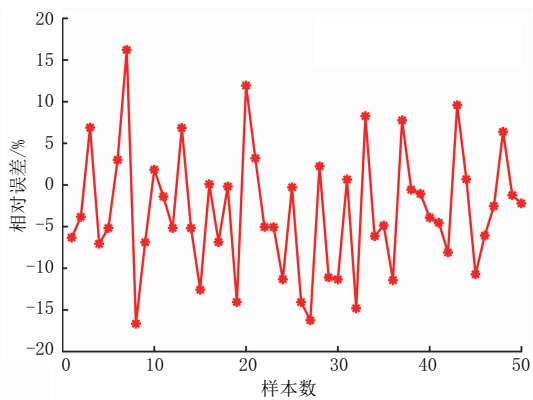


图2 PSO-LSSVM 预测误差

其次选用改进的 PSO 优化 LSSVM 模型,即 IP-SO-LSSVM 模型。设置 $\omega_{max}=0.9, \omega_{min}=0.5, \beta=18$ 。将 IPSO 得到的核函数宽度 σ 、正则化参数 γ 这 2 个参数再次运用到预测模型中,得到预测结果如图 3 所示,相对误差如图 4 所示。

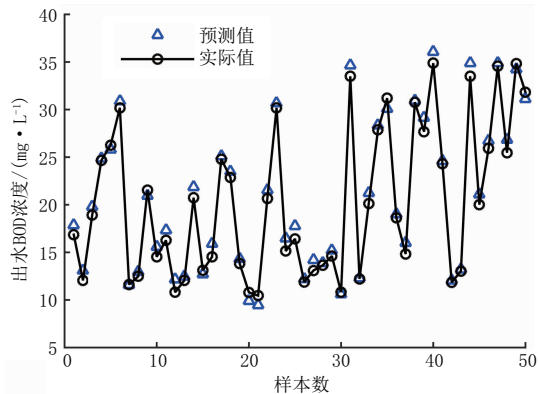


图3 IPSO-LSSVM 预测结果

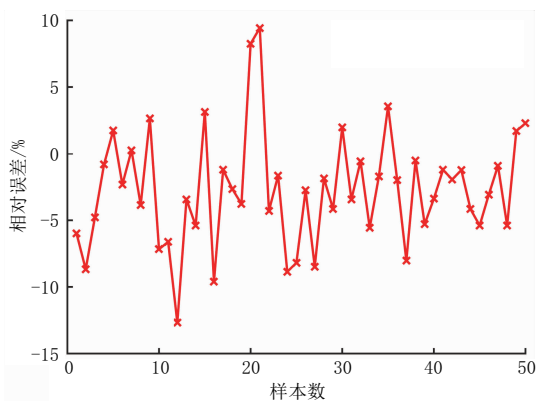


图4 IPSO-LSSVM 预测误差

通过以上仿真图可知,IPSO-LSSVM 模型的预估能力优于 PSO-LSSVM 模型,预测值与实测值相差较小。图 1 在第 1 到第 30 样本附近有着较大偏差,其余部分拟合效果较好;图 3 表明 IPSO-LSSVM 模型的预测效果明显较高。选取平均相对误差、最大相对误差、均方根相对误差 3 种误差计算方法进行分析,经过计算,3 种模型对应的预测误差如表 2 所示。

表2 3种预测模型误差对比

模型	平均相对误差 /%	最大相对误差 /%	均方根相对误差 /%
LSSVM	8.79	18.1	9.09
PSO-LSSVM	6.96	16.6	8.23
IPSO-LSSVM	4.48	10.7	5.39

由表 2 可知,IPSO-LSSVM 模型的 3 种误差均小于 PSO-LSSVM 模型以及 LSSVM 模型,其中平均绝对相对误差由 8.79% 下降为 4.48%,说明该模型预测精度较优。

5 结语

针对污水处理过程中 BOD 浓度测量实时性低等缺点,建立了预测 BOD 浓度的 IPSO-LSSVM 模型。经分析,优化后的模型预测准确度明显提高,证明了 IPSO 对 LSSVM 优化的有利影响。实验结果可为污水处理现场提供可靠的水质参考信息。

参考文献:

[1] 邵袁.基于 WEST 软件模拟的城市污水厂的优化运行研究[D].南京:东南大学,2019.
 [2] 杨壮,武利,乔俊飞.基于 GM-RBF 神经网络的污水环境预测[J].控制工程,2019,26(9):1728-1732.
 [3] 叶小岭,顾荣,邓华,等.基于 WRF 模式和 PSO-LSSVM 的风电场短期风速订正[J].电力系统保护与控制,2017,45(22):48-54.
 [4] 吴以文,杜坤,吴汉清,等.基于 LSSVM 交互预测的供水管网爆管检测[J].中国给水排水,2022,38(9):58-63.
 [5] 王语园,李嘉波,张福.基于粒子群算法的最小二乘支持向量机电池状态估计[J].储能科学与技术,2020,9(4):1153-1158.

(上接第 198 页)

参考。

参考文献:

[1] 宋文智,吴迪,刘建民,等.临近既有建筑物沉井不排水施工技术[C]//《施工技术》杂志社,亚太建设科技信息研究院有限公司.2020 年全

国土工程施工技术交流论文集(中册).《施工技术》杂志社,2020:55-59.

[2] 于得水.沉井不排水下沉施工技术控制[J].江西建材,2016(16).

[3] 詹赞贤.卵石及岩石地质条件下沉井施工实例探讨[J].中国科技信息,2006(22).