

doi:10.3969/j.issn.1673-9833.2016.01.011

基于 LSSVM 的污水处理过程建模

王 欣, 宋翼颀, 秦 斌, 彭小玉

(湖南工业大学 电气与信息工程学院, 湖南 株洲 412007)

摘 要: 针对基准仿真 1 号模型 (BSM1) 有耗时久、结构复杂、不利于研究和优化等缺点, 提出使用多输入多输出最小二乘支持向量机 (LSSVM) 对污水处理过程进行建模。先介绍 BSM1, 再在 BSM1 的基础上建立基于 LSSVM 的污水处理过程简化模型, 最后将 BSM1 的仿真结果与基于 LSSVM 的污水处理模型的仿真结果进行比较, 验证得出基于 LSSVM 的污水处理过程模型运行效率更高。

关键词: 最小二乘支持向量机; 污水处理过程; 基准仿真 1 号模型

中图分类号: TP15

文献标志码: A

文章编号: 1673-9833(2016)01-0059-05

Modeling of Sewage Treatment Process Based on MIMO-LSSVM

WANG Xin, SONG Yijie, QIN Bin, PENG Xiaoyu

(School of Electric and Information Engineering, Hunan University of Technology, Zhuzhou Hunan 412007, China)

Abstract: In view of shortcomings of Benchmark simulation model no.1 (BSM1), such as time consuming, complicated structure and hard to research and optimization, a model based on LSSVM (least squares support vector machine) for sewage treatment process is presented. Introduces BSM1 firstly, then establishes LSSVM-based simplified model of sewage treatment process on BSM1, finally compares the simulation results of BSM1 and LSSVM-based model. It indicates that the LSSVM-based model has higher operation efficiency.

Keywords : least squares support vector machine; sewage treatment process; benchmark simulation model no.1

0 引言

在大力推进生态文明建设的大背景下, 切实加大水污染防治的力度十分必要, 污水处理的研究也应运而生。目前, 污水生物处理的所有反应过程都可以在一个单纯的活性污泥系统内完成。由于微生物系统间存在着相互作用, 描述这些过程的数学模型较复杂^[1]。在污水处理过程综合模型研究的初期阶段, 研究人员不得不从事大量的计算, 并最大限度地从数学模型中寻求答案。国际水污染控制与研究协会提出了活性污泥 1 号模型, 该模型为污水处理过程

模型的深入研究奠定了基础, 同时也衍生出了许多其它污水处理过程模型^[2]。国际水质协会 (International Association on Water Quality, IAWQ) 和欧盟科学技术委员会 (Committee on Science and Technology, COST) 开发了基准仿真 1 号模型 (benchmark simulation model no.1, BSM1), 该模型成为了活性污泥废水处理过程控制方案的一个评估工具^[3]。但是 BSM1 在许多场合的工作效率低下, 因此本文提出基于最小二乘支持向量机 (least squares support vector machine, LSSVM) 的污水处理过程简化模型, 并进

收稿日期: 2015-11-03

作者简介: 王 欣 (1971-), 女, 湖南株洲人, 湖南工业大学教授, 博士, 主要研究方向为复杂过程建模, 机器学习,
E-mail: 1114084900@qq.com

通信作者: 宋翼颀 (1991-), 男, 湖南长沙人, 湖南工业大学硕士生, 主要研究方向为复杂过程建模控制与优化,
E-mail: 6569107@qq.com

行仿真验证。为了提高仿真结果的可接受性，在 MATLAB/SIMULINK 环境中分别建立了 BSM1 和基于 LSSVM 的污水处理过程简化模型，并把两者的仿真结果进行对比分析。

1 在 MATLAB/SIMULINK 中搭建 BSM1

BSM1 分为 2 个部分，其设备布局如图 1 所示。

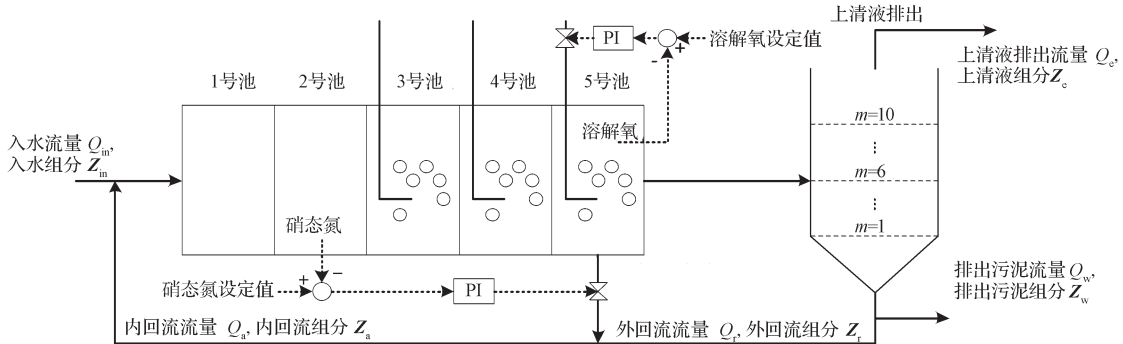


图 1 BSM1 设备布局

Fig. 1 BSM1 layout

图 1 中各个流量的单位均为 m^3/d ，各个组分为污水中 13 种物质的质量浓度组成的向量，单位均为 mg/L ^[4]。BSM1 有 2 个部分：一部分是基于 BSM1 的 5 个生化反应池模型，其中包括 2 个缺氧生化反应池模型和 3 个好氧反应池模型；另一部分是沉淀池模型，沉淀池模型使用了 Takács 的双指数沉淀速度模型来描述沉淀池的运作情况^[5]。

BSM1 中所有的微分方程、化学计量数和动力学参数在国际水质协会 (IAWQ) 的网站^[6]上均可查到。在仔细阅读 J. B. Copp 等编写的编程仿真手册后即可在 MATLAB 中建立起 BSM1^[5]。BSM1 的仿真过程就是解一个包含 145 个微分方程的方程组。为了使模型更容易理解和更具有可操作性，选择在 SIMULINK 环境下搭建 BSM1。考虑到求解速度，BSM1 中的生化反应池和沉淀池模块均使用 S 函数编写。

在国际水质协会的网站^[6]上，可以找到干燥天气、少雨天气和暴雨天气的入水流量和组分的数据。同时，为了方便用户对搭建起来的模型进行验证，还提供了这 145 个微分方程的初值以及在开环状态下第一天出水总悬浮颗粒质量浓度的变化曲线，如图 2 所示。

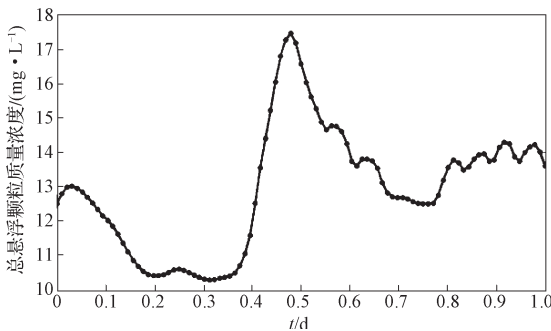


图 2 国际水质协会所提供的出水总悬浮颗粒质量浓度曲线
Fig. 2 The mass concentration curve of total suspended particles provided by IAWQ

搭建好 BSM1 后，使用干燥天气的入水数据，在开环情况下进行仿真，可以得到出水总悬浮颗粒质量浓度的变化曲线，如图 3 所示。

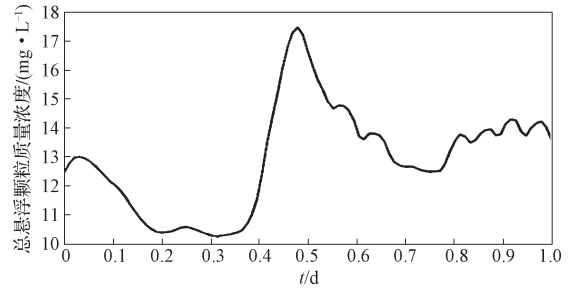


图 3 BSM1 开环仿真的出水总悬浮颗粒质量浓度曲线
Fig. 3 The mass concentration curve of total suspended particles made by BSM1 in open loop

由图 2 和图 3 可以看出，搭建起来的 BSM1 的开环仿真结果与国际水质协会所提供的结果完全一致，可进行下一步研究。

污水处理过程中最重要的就是生物脱氮部分。生物脱氮分为 2 个部分，硝化反应和反硝化反应。硝化反应在好氧区进行，反硝化反应在缺氧区进行。为了使硝化反应能够充分进行，好氧区的溶解氧就成了一个比较重要的变量。因此，国际水质协会提出了如下控制策略，采用 PI 控制器控制 5 号池的氧转移系数，从而使 5 号池中的溶解氧质量浓度达到稳定；同时采用 PI 控制器控制混合液回流量，从而使 2 号池出水的硝态氮质量浓度达到稳定。根据这一控制策略，可以在 SIMULINK 环境中搭建出如图 4 所示闭环状态下的 BSM1。图中 Q_0 为流入 1 号池的流量， Q_i 为第 i 号池的出水流量 ($i=1, 2, \dots, 5$)， k_i 为 i 号池的氧转移系数 ($i=1, 2, \dots, 5$)， Q_f 为流入二沉池的流量， Q_r^* 为外回流流量的给定值， Q_w^* 为排出污泥流

量的给定值, $\rho_2^*(\text{NO})$ 为 2 号池硝态氮质量浓度的给定值, $\rho_5^*(\text{O}_2)$ 为 5 号池溶解氧质量浓度的给定值, $\frac{1}{z}$ 为延迟因子。

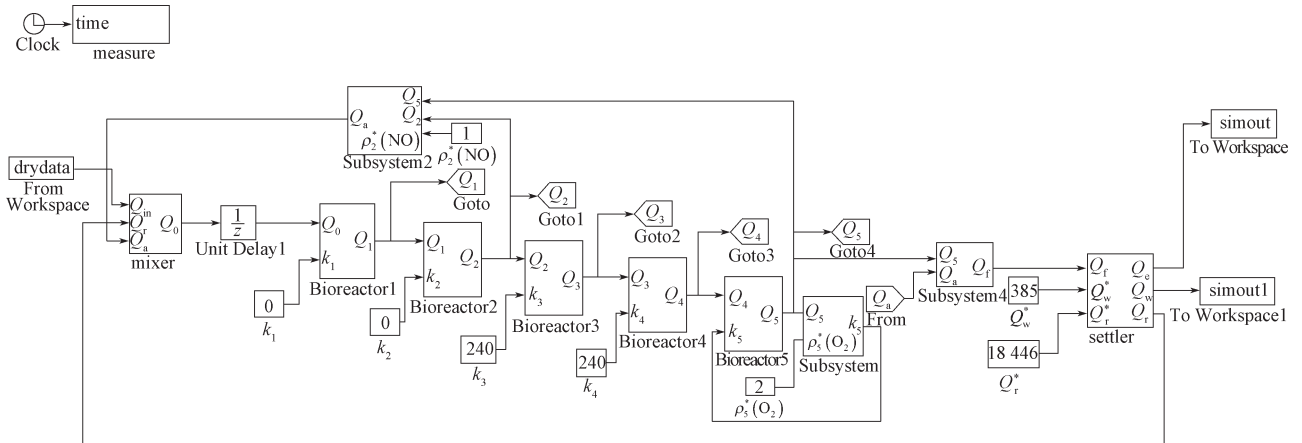


图4 在 SIMULINK 环境中搭建的 BSM1

Fig. 4 BSM1 model in SIMULINK

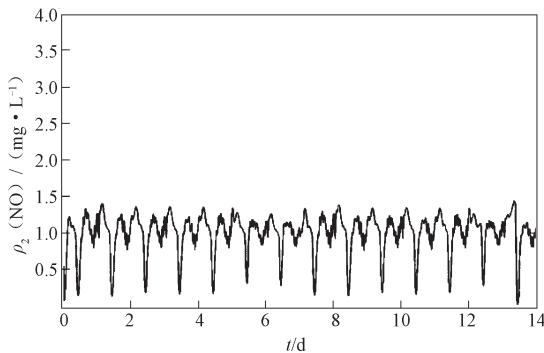


图5 BSM1 闭环仿真 $\rho_2(\text{NO})$ 的变化曲线

Fig. 5 The simulated $\rho_2(\text{NO})$ curve of BSM1 in closed loop

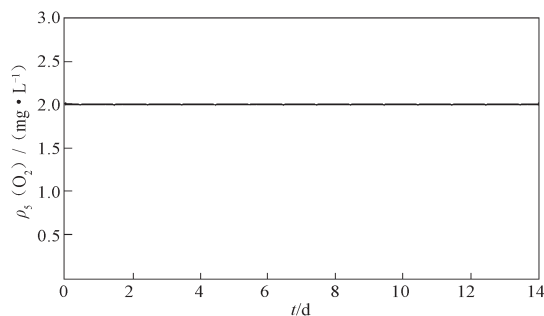


图6 BSM1 闭环模型仿真 $\rho_5(\text{O}_2)$ 的变化曲线

Fig. 6 The simulated $\rho_5(\text{O}_2)$ curve of BSM1 in closed loop

由图 5 和图 6 可知, 2 号池出水的硝态氮质量浓度在较快达到稳定后, 大致稳定在 0 到 1.3 之间, 波动较大, 这说明 2 号池出水的硝态氮质量浓度的控制器还有较大的优化空间。5 号池的溶解氧质量浓度有比较大的超调量, 但是很快达到了稳定状态, 稳态误差基本为 0。

将仿真时间设定为 [0, 14], 可以得到 2 号池出水的硝态氮质量浓度 ($\rho_2(\text{NO})$) 的变化曲线如图 5 所示, 5 号池中的溶解氧质量浓度 ($\rho_5(\text{O}_2)$) 的变化曲线如图 6 所示。

2 基于 MIMO-LSSVM 的污水处理过程简化模型

2.1 最小二乘支持向量机

统计学习理论解决了传统统计学中的一些难题, 例如“过学习”和“欠学习”的问题^[7]。支持向量机 (support vector machine, SVM) 是由 Vapnik 等在统计学习理论的基础上, 建立的一种机器学习方法。支持向量机在分类和回归上有比较广泛的应用, 但是支持向量机的解法存在耗时的二次规划问题, 因此又发展出了最小二乘支持向量机 (least square support vector machine, LSSVM)^[7-8]。下面介绍最小二乘支持向量机回归的数学理论。

LSSVM 在求解过程中用等式代替了 SVM 中的不等式, 从而避免了求解二次规划问题。因此 LSSVM 在求解速度上比 SVM 具有更大的优势。LSSVM 对回归问题的描述为给定样本集:

$$\begin{aligned} & (x_1, y_1), (x_2, y_2), \dots, (x_i, y_i) \dots, (x_l, y_l), \\ & x_i \in \mathbf{R}^d, y_i \in \mathbf{R}, i = 1, 2, \dots, l. \end{aligned} \quad (1)$$

式中: x_i 为输入向量数据样本;

y_i 为输出数据样本;

d 为输入向量的维度;

l 为样本个数。

最小二乘支持向量机的回归函数形式为

$$f(x) = \mathbf{w}^T \varphi(x) + b, \quad (2)$$

式中: $\varphi(\cdot)$ 为非线性映射;

$f(x)$ 为待确定的映射关系;

w 为权重向量;

w^T 为 w 的转置向量;

b 为偏置量。

最小二乘支持向量机的求解就是求解 w 和 b , 这样就可以使用结构风险最小化原则来描述 LSSVM 的求解:

$$\begin{cases} \min_{w,b,e} J^* = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^l e_i^2; \\ \text{s.t. } y_i = w^T \varphi(x_i) + b + e_i; i = 1, 2, \dots, l. \end{cases} \quad (3)$$

式中: e_i 为第 i 个样本的预测误差;

C 为正则化参数, C 的取值可以用来在泛化能力和模型训练误差之间进行折中。

式(3)中, 等式右边第一项 $\frac{1}{2} w^T w$ 是正则化风险, 代表模型的泛化能力和模型的平滑程度; 第二项 $\frac{1}{2} C \sum_{i=1}^l e_i^2$ 是经验风险, 代表模型的训练误差。

因此, 可以构建出拉格朗日函数:

$$L = J^* - \sum_{i=1}^l \alpha_i [w^T \varphi(x_i) + b + e_i - y_i], \quad (4)$$

式中 α_i 为拉格朗日乘子。

从式(3)可以看出, LSSVM 的求解最终化为一个线性方程组的求解。正是因为最后不需要再象 SVM 那样去解二次规划问题, 所以 LSSVM 在求解时效率更高。

2.2 简化模型的结构

BSM1 中有 2 个控制量和 2 个被控量, 这 2 个被控量同时还影响着出水水质。由于 BSM1 中的参数较多、数学模型较复杂, 所以在预测、优化等场合, BSM1 的运行效率较低。

因此, 本研究提出建立基于 LSSVM 的污水处理过程模型。以 2 号池出水硝态氮质量浓度 ($\rho_2(\text{NO})$) 和 5 号池出水的溶解氧质量浓度 ($\rho_5(\text{O}_2)$) 作为输出, 5 号池氧转移系数和混合液回流量为输入, 以入水流量和入水中各种物质质量浓度作为扰动, 建立污水处理过程简化模型。由于 BSM1 中的微分方程均为一阶, 且没有延迟环节, 所以在建立简化模型时可以认为当前输出与上一时刻的输出 $\rho_2(\text{NO})(k-1)$ 和 $\rho_5(\text{O}_2)(k-1)$ 相关。这样简化模型的输入还需要加入上一采样时刻的 2 号池硝态氮质量浓度 $\rho_2(\text{NO})(k-1)$, 上一采样时刻的 5 号池溶解氧质量浓度 $\rho_5(\text{O}_2)(k-1)$, 简化模型的结构如图 7 所示。

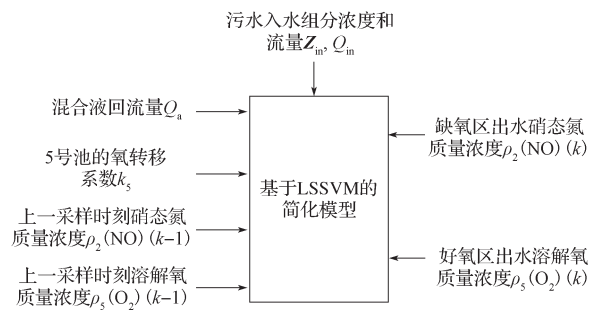


图 7 基于 LSSVM 的污水处理过程模型的结构

Fig. 7 The model structure for LSSVM-based sewage treatment

2.3 简化模型的仿真

将 BSM1 14 d 的开环仿真结果作为数据样本, 开环输入为参数在各自区间上的均匀分布的随机数, 采样时间定为 3 min, 这样采样结果将有 6 721 组数据。将其中 75% 的数据作为训练样本, 25% 的数据作为测试样本。对于同样的数据样本, 标准支持向量机和最小二乘支持向量机对测试样本的预测能力如表 1 所示。从表中可以看出, 标准支持向量机对测试样本的预测误差比最小二乘支持向量机的预测误差要高很多。

表 1 SVM 和 LSSVM 对测试样本的预测误差对比

Table 1 Prediction error comparison of SVM and LSSVM on test samples

向量机类型	质量浓度平均平方误差 / ($\text{mg} \cdot \text{L}^{-1}$)	
	$\rho_2(\text{NO})$	$\rho_5(\text{O}_2)$
标准支持向量机	0.363 0	0.018 1
最小二乘支持向量机	0.001 8	0.015 3

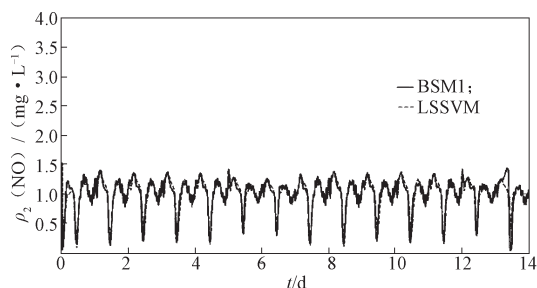
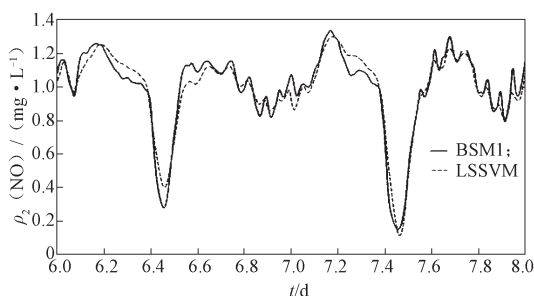
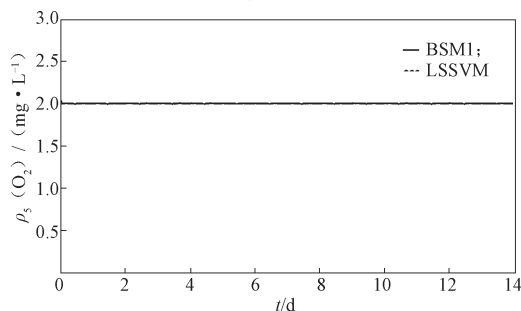
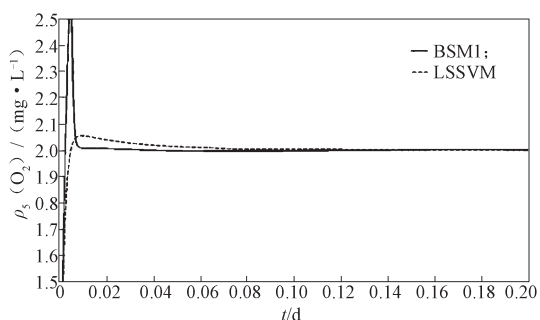
将仿真时间设为 $[0, 2]$, 入水数据为干燥天气的数据, BSM1 和基于 LSSVM 的污水处理过程模型各自的运行时间对比如表 2 所示。

表 2 BSM1 和基于 LSSVM 的模型的运行时间对比

Table 2 The runtime comparison of BSM1 and the model based on LSSVM

模型	基准仿真 1 号模型	基于最小二乘支持向量机的简化模型
运行时间 / s	663.930 3	6.118 9

最后, 为了验证模型, 将基于 LSSVM 的污水处理过程模型在 SIMULINK 环境下仿真, 仿真时间设为 $[0, 14]$, 使用干燥天气的入水数据, 仿真结果如图 8~11 中虚线所示。从图 8 和图 9 可以看出, 硝态氮质量浓度的预测曲线基本与 BSM1 的曲线保持一致。从图 10 可以看出, 溶解氧质量浓度的预测曲线在稳定状态下能够和 BSM1 的曲线保持一致; 但是从图 11 可以看出, 溶解氧的预测曲线在超调部分未能有较好的预测效果。

图8 $\rho_2(\text{NO})$ 的变化曲线对比Fig. 8 The $\rho_2(\text{NO})$ curve comparison图9 放大的 $\rho_2(\text{NO})$ 对比图Fig. 9 The amplified $\rho_2(\text{NO})$ curve comparison图10 $\rho_5(\text{O}_2)$ 的变化曲线对比Fig. 10 The $\rho_5(\text{O}_2)$ curve comparison图11 放大的 $\rho_5(\text{O}_2)$ 对比图Fig. 11 The amplified $\rho_5(\text{O}_2)$ curve comparison

3 结语

基于机理模型运行速度慢的缺点, 本文提出在BSM1的基础上建立基于LSSVM的简化模型, 这种简化模型忽略了机理模型中大量不需要关心的中间变量和参数, 改善了机理模型结构复杂、耗时长、中间变量多的缺点。研究人员只需要找出所关心的变

量的数据对模型进行训练, 即可获得一个预测精准的黑箱模型。从仿真结果可以看出, 基于LSSVM的污水处理过程模型可以在和BSM1的结果保持一致的基础上, 运行速度比BSM1更快。

参考文献:

- [1] 顾夏声. 废水生物处理数学模式[M]. 2版. 北京: 清华大学出版社, 1993: 123-124.
GU Xiasheng. Sewage Treatment Mathematical Model[M]. 2nd ed. Beijing: Tsinghua University Press, 1993: 123-124.
- [2] 余颖, 乔俊飞. 活性污泥法污水处理过程的建模与仿真技术的研究[J]. 信息与控制, 2004, 33(6): 709-713, 728.
YU Ying, QIAO Junfei. Modeling and Simulation Technology of Activated Sludge Method on Wastewater Treatment Process[J]. Information and Control, 2004, 33(6): 709-713, 728.
- [3] 黄明智. 废水处理系统水质特征动态分析的混合智能控制研究[D]. 广州: 华南理工大学, 2010.
HUANG Mingzhi. Study on Hybrid Intelligent Control for Water Quality Dynamic Parameters in Wastewater Treatment System[D]. Guangzhou: South China University of Technology, 2010.
- [4] 黄晓琪. 污水处理过程节能优化控制方法的研究[D]. 北京: 北京工业大学, 2013.
HUANG Xiaoqi. Research on Energy-Saving Optimal Control Method for Wastewater Treatment Process[D]. Beijing: Beijing University of Technology, 2013.
- [5] 刘大伟, 沈文浩. 用Matlab/Simulink建立废水处理仿真模型[J]. 环境科学与技术, 2007, 30(11): 86-89.
LIU Dawei, SHEN Wenhao. Set Up BSM1 with Matlab/Simulink[J]. Environmental Science and Technology, 2007, 30(11): 86-89.
- [6] COPP J B. Description of BSM1[EB/OL]. [2014-11-17]. <http://www.benchmarkwwtp.org/>.
- [7] 程萍. 基于最小二乘支持向量机的非线性广义预测控制[D]. 太原: 太原理工大学, 2010.
CHENG Ping. Generalized Predictive Control Algorithm Based on Least Squares Support Vector Machines[D]. Taiyuan: Taiyuan University of Technology, 2010.
- [8] 郭振凯, 宋召青, 毛剑琴. 基于最小二乘支持向量机的非线性广义预测控制[J]. 控制与决策, 2009, 24(4): 520-525, 531.
GUO Zhenkai, SONG Zhaoqing, MAO Jianqin. Nonlinear Generalized Predictive Control Based on Least Squares Support Vector Machine[J]. Control and Decision, 2009, 24(4): 520-525, 531.

(责任编辑: 邓光辉)