本文引用格式：金绍琴,唐莉丽,吴菁,等.基于时差-即时学习的相关向量机软测量建模研究[J].自动化与信息工程,2023,44(5):22-31.

JING Shaoqin, TANG Lili, WU Jing, et al. Research on soft sensor modeling of relevance vector machine based on time difference-just in time[J]. Automation & Information Engineering, 2023,44(5):22-31.

基于时差-即时学习的相关向量机软测量建模研究*

金绍琴[1]　唐莉丽[1]　吴菁[1,2]　李明珠[2]　黄道平[3]

（1.贵州民族大学数据科学与信息工程学院，贵州 贵阳 550025

2.海口经济学院腾竞依智网络学院，海南 海口 570203

3.华南理工大学自动化科学与工程学院，广东 广州 510640）

摘要：针对污水处理厂的过程数据时变性较大、非线性较强，传统的离线模型难以应对实际处理过程中的工况变化等问题，提出一种基于时差-即时学习的相关向量机（RVM）模型 TD-JIT-RVM。通过时间差分（TD）建模提取过程变量之间的关联关系，采用即时学习（JIT）解决时滞引起的模型退化问题。利用 TD-JIT-RVM 模型对仿真数据集和真实的工业数据集进行分析，结果表明，该模型在两个数据集中比 RVM 基础模型的 *RMSE* 分别提高了 94.59%和 82.26%。

关键词：污水处理；时间差分；即时学习；相关向量机；在线软测量

中图分类号：TP 277　　**文献标志码：**A　　**文章编号：**1674-2605(2023)05-0004-10

DOI：10.3969/j.issn.1674-2605.2023.05.004

Research on Soft Sensor Modeling of Relevance Vector Machine Based on Time Difference-Just in Time

JING Shaoqin[1]　TANG Lili[1]　WU Jing[1,2]　LI Mingzhu[2]　HUANG Daoping[3]

(1.College of Data Science and Information Engineering, Guizhou Minzu University, Guiyang 550025, China

2.TJ-YZ School of Network Science, Haikou University of Economics, Haikou 570203, China

3.School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: A relevance vector machine (RVM) model TD-JIT-RVM based on Time Difference-Just in Time is proposed to address the issues of large time-varying and strong nonlinearity of process data in sewage treatment plants, and traditional offline models are difficult to cope with changes in actual processing conditions. Extract the correlation between process variables through Time Difference (TD) modeling, and solve the model degradation problem caused by time delay using just in time (JIT). The TD-JIT-RVM model was used to analyze the simulation dataset and the real industrial dataset, and the results showed that the R*MSE* of the model increased by 94.59% and 82.26% compared to the RVM basic model in both datasets, respectively.

Keywords: wastewater treatment; time difference; just in time; relevance vector machine; online soft sensors

0　引言

软测量技术是通过统计模型、机器学习等方法建立易测变量与目标变量（难测变量）之间的数学关系，并从易测变量推断目标变量数值的一种技术[1]。软测量建模方法主要有机理建模和数据驱动建模。与机理建模相比，数据驱动建模先从数据中选取特征，再进行模型训练和优化，无需大量的专业知识和经验来开发模型，更适合具有多样性和多层次性的复杂工业过

* 基金项目：贵州省省级科技计划项目（黔科合基础[2020]1Y276）；贵州省教育厅自然科学研究项目（黔教技[2022]015 号，黔教技[2022]047 号）。

程建模[2-4]。目前，数据驱动软测量技术主要包括主成分回归（principal component regression, PCR）[5]、偏最小二乘回归（partial least squares regression, PLSR）[6]、人工神经网络（artificial neural network, ANN）[7]和支持向量机（support vector machine, SVM）[8]等。其中，SVM 是一种基于统计学习理论的结构风险最小化（structural risk minimization, SRM）原理的机器学习算法[9]，比 ANN 具有更好的泛化性能，广泛应用于模式分类和回归估计，但对数据噪声较敏感。1999 年，SUYKENS 等[10]提出了基于 SVM 的最小二乘支持向量机（least squares support vector machine, LSSVM），与 ANN 相比，它对非线性工艺数据具有更好的泛化性能[11]，常用于离线数据建模，但不适用于在线数据建模。LYU 等[12]提出的一种基于数据相似性的在线学习技术，可有效地筛选并实时调整输入数据，但该方法计算成本较高，不适合实时应用。

近几年，尽管针对 SVM 算法的研究有所进展，但也呈现出不适定的问题，如后验概率分布不可计算、惩罚因子 C 不可估计、核投影需要 Mercer 条件等。相关向量机（relevance vector machine, RVM）[13-14]是 SVM 的贝叶斯替代方案，具有相关向量稀疏、泛化性能好、具备统计解释性的预测结果等特点，明显缓解了不适定的问题[15]。RVM 算法不需要对超参数进行交叉验证，即可实现输出的概率分布计算，且核函数可以任意指定，不受正范式的约束。王春雷等[16]通过使用组合核函数来提高 RVM 算法的泛化性，并构建用于预测锂电池使用寿命的模型，但预测结果的拟合程度没有明显变化，并不适用于所有电池。吴菁等[17]提出一种优化多核 RVM 模型的方法，增强了 RVM 中多核函数的建模能力，克服了简单单核函数无法适应强非线性数据的局限性，提高了预测效率和精度，但该模型使用所有过程数据进行训练，无法精准捕捉过程数据的局部动态特征，且数据集中包含冗余信息，约束了模型精度的提高。谭承诚等[18]开发了一种基于集成 RVM 的水质在线预测模型，提高了模型的稳定性和可靠性，但计算时间较长。许玉格等[19]利用自优化理论对核函数进行参数优化，构建基于 RVM 的多属性核函数预测模型，能够有效地处理复杂的高维数据，且具有良好的准确性，但在线学习能力差，鲁棒性弱。

污水处理的过程数据具有多变量、强干扰、大时滞、强耦合等特点，导致其检测和采样不可避免地存在误差。传统的离线模型用于在线建模时，需实时更新参数以确保模型的有效性。为此，建模过程通常引入在线学习的方法，使软测量模型能够根据新样本实时更新参数，从而在不断变化的数据环境中保持良好的测量精度，同时有效缩减数据的存储成本[20]。目前，主流的在线软测量建模算法包括递归算法（recursion algorithm, RA）[4]、移动窗口（moving window, MW）算法[21]、时间差分（time difference, TD）算法[22]、即时学习（just in time, JIT）算法[23]等。其中，RA 和 MW 算法适用于数据漂移的情况；JIT 算法适用于采样延迟的情况；TD 算法适用于数据延迟和突变的情况[24]。相比于其他 3 种算法，TD 算法不仅能够有效地抵御输入输出变量的漂移，而且在运行过程中也不需要模型重构。RA、MW、TD 这 3 种算法都是根据时间相关性更新模型，属于时间自适应算法。JIT 算法[27-31]基于空间相关性更新和维护模型，属于空间自适应算法，可以更好地适应生产过程中的复杂现象，且该算法能够对每个样本建立局部模型，能很好地描述过程变量之间的非线性关系。QIN 等[25]提出一种递归偏最小二乘（partial least squares, PLS）算法可有效地解决软测量的退化问题。KANEKO 等[26]提出一种 MW 算法，使用最近获取的数据来更新模型，能够有效地跟踪过程的动态特征。

本文针对污水处理过程中存在的生化反应复杂、非线性强、运行状态分布不均匀、出水指标误差大等问题，将 TD 算法和 JIT 算法结合并引入 RVM 模型，提出基于时差-即时学习的相关向量机 TD-JIT-RVM 模型。该模型结合了 TD 算法和 JIT 算法的优点，可减少测量数据漂移对模型性能退化的影响，有效捕捉过程数据随时间的动态变化，并实时更新模型参数，保证模型的准确率，解决过程数据漂移、突变、强非线性等导致的模型预测精度随时间下降的问题。

1 理论基础

1.1 RVM 模型

RVM 模型是一种基于贝叶斯框架的机器学习模型，其通过最大化边际似然获得相关向量和权重[15]。

设 $\{\boldsymbol{x}\}_{u=1}^{N}$ 和 $\{\boldsymbol{t}\}_{u=1}^{N}$ 分别是输入向量和输出目标，N 为样本数，目标 $\boldsymbol{t}$ 采用回归模型：

$$\boldsymbol{t}=y(\boldsymbol{x})+\xi_n \tag{1}$$

式中：ξ_n 为零均值，方差 δ^2 的噪声；$y(\boldsymbol{x})$ 定义为

$$y(\boldsymbol{x})=\sum_{u=1}^{N}\boldsymbol{w}_u K(\boldsymbol{x},\boldsymbol{x}_u)+b_0 \tag{2}$$

式中：$\boldsymbol{w}_u$ 为权重向量，$K(\boldsymbol{x},\boldsymbol{x}_u)$ 为核函数，b_0 为偏差。

设目标 $\boldsymbol{t}$ 是独立的，其概率定义为

$$p(\boldsymbol{t}\,|\,\boldsymbol{w},\delta^2)=(2\pi\delta^2)^{-\frac{N}{2}}\exp\left\{-\frac{\|\boldsymbol{t}-\boldsymbol{\varphi}\boldsymbol{w}\|^2}{2\delta^2}\right\} \tag{3}$$

式中：$\boldsymbol{w}=\{w_i\}_{i=0}^{N}$ 为权重，δ^2 为方差，$\boldsymbol{t}=(t_1,t_2,\cdots,t_N)^{\mathrm{T}}$，$\boldsymbol{\varphi}$ 为 $N(N+1)$ 的矩阵。

公式(3)中 $\boldsymbol{w}$ 和 δ 的最大似然估计会导致过拟合，为约束参数，定义一个零均值高斯先验概率分布为

$$p(\boldsymbol{w}\,|\,\boldsymbol{\alpha})=\prod_{u=0}^{N}N(\boldsymbol{w}_u\,|\,0,\boldsymbol{\alpha}_u^{-1}) \tag{4}$$

式中：$\boldsymbol{\alpha}$ 是 N+1 维的超参数向量。

依据贝叶斯公式，未知参数的后验概率为

$$p(\boldsymbol{w},\boldsymbol{\alpha},\sigma^2\,|\,\boldsymbol{t})=p(\boldsymbol{w}\,|\,\boldsymbol{\alpha},\sigma^2,\boldsymbol{t})p(\boldsymbol{\alpha},\sigma^2\,|\,\boldsymbol{t}) \tag{5}$$

后验分布的权重为

$$p(\boldsymbol{w}\,|\,\boldsymbol{t},\boldsymbol{\alpha},\sigma^2)=(2\pi)^{-\frac{N+1}{2}}|\Sigma|^{-\frac{1}{N}}\exp\left\{-\frac{1}{2}(\boldsymbol{w}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{w}-\boldsymbol{\mu})\right\} \tag{6}$$

式中，后验均值 $\boldsymbol{\mu}=\sigma^{-2}\boldsymbol{\Sigma}\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{t}$，协方差 $\boldsymbol{\Sigma}=(\sigma^{-2}\boldsymbol{\varphi}^{\mathrm{T}}\boldsymbol{\varphi}+\boldsymbol{A})^{-1}$，$\boldsymbol{A}=diag(\alpha_0,\alpha_1,\cdots,\alpha_N)$。

为了实现统一的超参数，做出以下定义：

$$p(\boldsymbol{t}\,|\,\boldsymbol{\alpha},\sigma^2)=\int p(\boldsymbol{t}\,|\,\boldsymbol{w},\sigma^2)p(\boldsymbol{w},\boldsymbol{\alpha})\mathrm{d}\boldsymbol{w}=$$
$$(2\pi)^{-\frac{N}{2}}\left|\sigma^2\boldsymbol{I}+\varphi\boldsymbol{A}^{-1}\varphi^{\mathrm{T}}\right|\exp\left\{-\frac{1}{2}\boldsymbol{t}^{\mathrm{T}}\left(\sigma^2\boldsymbol{I}+\varphi\boldsymbol{A}^{-1}\varphi^{\mathrm{T}}\right)^{-1}\boldsymbol{t}\right\} \tag{7}$$

高斯径向基函数具有较强的非线性处理能力，被用作核函数，定义为

$$K(\boldsymbol{x},\boldsymbol{x}_u)=\exp\left(-\frac{(\boldsymbol{x}-\boldsymbol{x}_u)^2}{2\gamma^2}\right) \tag{8}$$

式中：γ 为宽度因子，对模型精度有较大影响，需要预先设定。

1.2 TD 算法

TD 算法是一种处理时序数据的算法，通过计算相邻数据之间的差异，获取数据的变化趋势，得到数据的动态变化规律。利用 TD 算法进行建模：

首先，计算在目标时间 t 的数据 $\boldsymbol{X}(t)$ 和 $\boldsymbol{Y}(t)$ 与之前的一段时间 i 内，$\boldsymbol{X}(t-i)$ 和 $\boldsymbol{Y}(t-i)$ 之间的时间差，时差变量表示公式为

$$\Delta\boldsymbol{X}_i(t)=\boldsymbol{X}(t)-\boldsymbol{X}(t-i) \tag{9}$$

$$\Delta\boldsymbol{Y}_i(t)=\boldsymbol{Y}(t)-\boldsymbol{Y}(t-i) \tag{10}$$

然后，使用回归模型建立 $\Delta\boldsymbol{X}_i(t)$ 和 $\Delta\boldsymbol{Y}_i(t)$ 之间的关系

$$\Delta\boldsymbol{Y}_i(t)=f\left(\Delta\boldsymbol{X}_i(t)\right)+\boldsymbol{e} \tag{11}$$

式中：f 表示回归模型，在文中为 RVM 模型；$\boldsymbol{e}$ 为计算误差向量。

当输入一个新的数据，先计算差分量：

$$\Delta\boldsymbol{X}_j(t)=\boldsymbol{X}(t)-\boldsymbol{X}(t-j) \tag{12}$$

式中：j 为新的时间间隔。

通过训练好的模型得到差分的预测目标值。

$$\Delta\boldsymbol{Y}_{j,pre}(t)=f(\Delta\boldsymbol{X}_j(t)) \tag{13}$$

最后，通过反 TD 得到真正的预测值。

$$\boldsymbol{Y}_{j,pre}(t)=\Delta\boldsymbol{Y}_{j,pre}(t)+\boldsymbol{Y}(t-j) \tag{14}$$

1.3 JIT 算法

JIT 算法作为一种局部建模算法，具有分析复杂非线性数据的能力，它可以根据新的数据不断地更新模型参数，并在实时数据处理中学习和优化模型参数，从而实现数据的动态建模和实时预测。JIT 算法流程如图 1 所示。

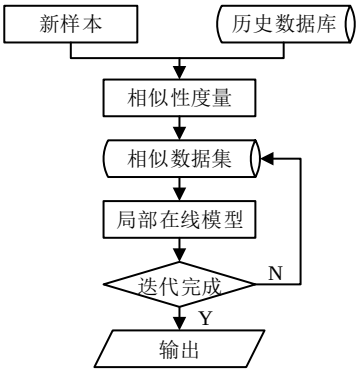


图 1 JIT 算法流程图

JIT 算法主要有 3 个步骤：

1） 建立历史数据库，采集可测量的输入输出数据，这些数据尽可能地覆盖所有的工况条件，并通过减少冗余、归一化等方法进行预处理后，存入历史数据库；

2） 确定相似数据集，获得局部模型，基于欧式距离函数，从历史数据库中选出与当前输入数据 $\boldsymbol{X}$ 相似度最高的样本，构建相似数据集（similar data），利用局部多项式拟合求得局部模型参数 $\boldsymbol{\Phi}$，计算当前输入数据对应的输出 $\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{\Phi}$；

欧式距离公式为

$$d_j=\left\|x(t),x_j\right\|^2, j=1,2,\ldots,n \tag{15}$$

3） 当新的输入数据出现时，转到步骤 2）。

2 TD-JIT-RVM 模型

在 TD-JIT-RVM 模型中，TD 算法将过程数据划分为一系列时间窗口，对每个时间窗口内的数据进行时间差分处理，可充分反映数据的动态变化，并作为 RVM 模型的输入；JIT 算法先离线采集大量的数据样本，再边建模边预测；RVM 模型在离线训练时表现出较好的预测效果。但由于时变问题，在在线应用过程中，RVM 模型的性能容易下降。为了保持 RVM 模型的高性能和及时跟踪状态，引入自适应算法，不断更新模型参数。TD-JIT-RVM 建模由 2 个阶段组成：

1） 离线训练阶段，先对历史数据进行预处理，利用 TD 算法将预处理后的数据转换为一阶差分向量，作为 RVM 模型的输入，训练好模型参数，并保存该模型参数用于后续的在线监测；

2） 在线监测阶段，JIT 算法通过历史数据的输出特征提取在线数据的特征，模型参数在同步在线建模状态下不断更新，以保证模型的正常运行和性能稳定。TD-JIT-RVM 建模过程和框架分别如表 1、图 2 所示。

表 1 TD-JIT-RVM 建模过程

Step1：离线准备，选定训练样本，进行标准化处理以及公式(9)、(10)通过 TD 运算获得历史数据库；
Step2: 离线训练，根据历史数据库训练 RVM 模型，并保存训练模型参数；
Step3：在线预测，将新的数据样本与历史数据库中的数据进行逐一对比，利用距离计算公式(15)，确定 2 个数据之间的相似性；将最相似的数据条目放置到新数据库的输入点中，这些条目的集合形成一个相关的数据集，用于模型训练，更新模型参数；
Step4：模型训练完成后，利用经过 TD 处理后的新样本进行预测，得到相应的 TD 输出；对 TD 输出执行反标准化和反时差处理，生成预测输出；
Step5：通过将最新的和实验室分析的数据合并到历史数据库，可以在线获取最新过程变化，同时保存更丰富的过程信息；
Step6：重复 Step3，直到查询数据结束，获得最后的预测输出。

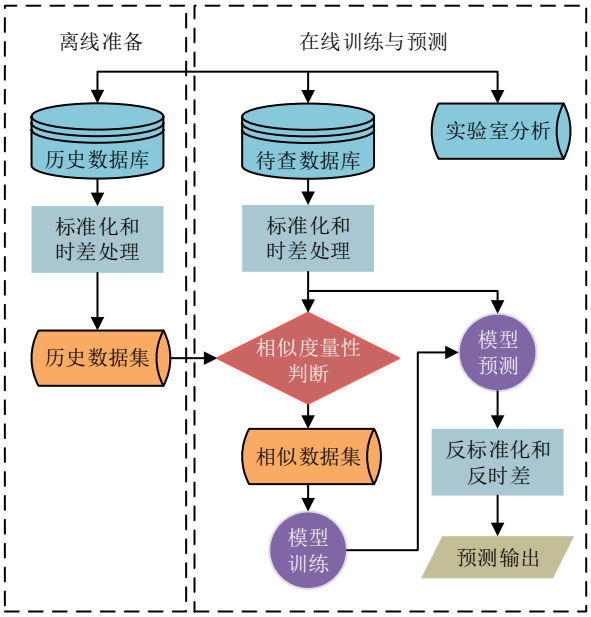


图 2　TD-JIT-RVM 框架图

3　案例分析

3.1　BSM1 案例分析

基准仿真 1 号模型（benchmark simulation model No.1, BSM1）是一个用于模拟和评估污水处理系统的基准平台[26]，其涵盖了一个完整的污水处理系统，包括入水、生物反应池、二次沉淀池、出水处理等环节。通过 BSM1 可以更加准确地模拟和评估不同污水处理系统的性能。BSM1 平台污水处理原理如图 3 所示。

BSM1 仿真数据集包含 14 天的 3 种（晴天、雨天、暴雨天）天气数据，每 15 s 采样一次，共有 1 344 组样本数据。其中，前 1 000 组样本数据用于模型训练，剩余 344 组样本数据用于在线测试。BSM1 仿真数据集中，前 7 天的数据较为平稳，BOD_5 含量变化较小，从第 7 天开始发生降雨，污水流量大幅增加，使相应的悬浮固体浓度等变量增加，导致出水变量 BOD_5 含量变化显著。这对在线建模提出了挑战，TD-JIT-RVM 模型需连续几天不断地更新状态，同时提供目标变量的准确预测。本文选择的辅助变量与文献[26]一致。工业现场通过传感器等仪器采集的数据易受生产环境、仪器精度、测量方法等影响，数据集维数存在差异，通过零均值可标准化这些数据，如公式(16)所示。

$$z_n = \frac{x_i - \bar{x}}{s} \tag{16}$$

式中：z_n 是标准化后的数据，x_i 是原始数据，$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}$ 是原始数据的标准差，$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$ 是原始数据的均值。

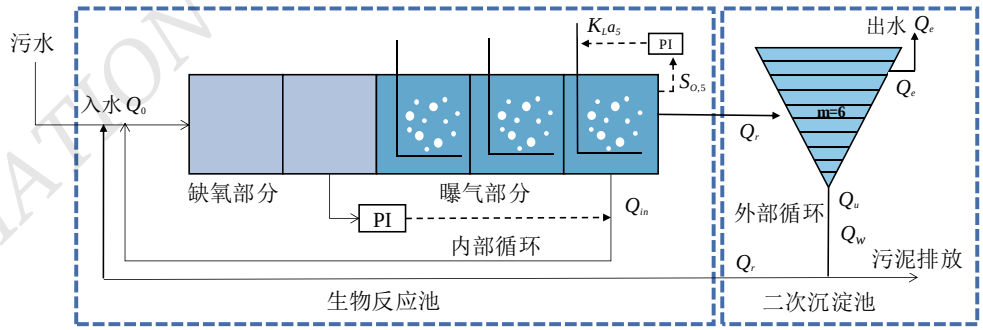


图 3　BSM1 平台污水处理原理图[26]

采用预测 BOD_5 的均方根误差 RMSE、预测值与实际 BOD_5 值的均方根误差以及相关系数 r 来评估模型性能。

$$R_{\mathrm{RMSE}} = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)}{N}} \tag{17}$$

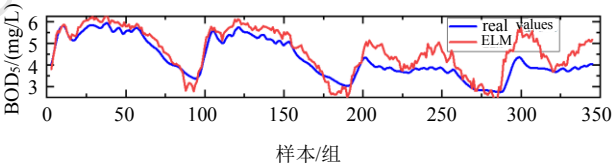
$$r=\frac{\operatorname{cov}(y_i,\hat{y}_i)}{\sqrt{\operatorname{var}(y_i)\operatorname{var}(\hat{y}_i)}}=\frac{N\left(\sum_{i=1}^{N}\hat{y}_i y_i\right)-\left(\sum_{i=1}^{N}\hat{y}_i\right)\left(\sum_{i=1}^{N}y_i\right)}{\sqrt{\left(N\sum_{i=1}^{N}\hat{y}_i^{\,2}-\left(\sum_{i=1}^{N}\hat{y}_i\right)^2\right)\left(N\sum_{i=1}^{N}y_i^{\,2}-\left(\sum_{i=1}^{N}y_i\right)^2\right)}} \tag{18}$$

式中：y_i 为真实值，$\hat{y}_i$ 为预测值，N 为样本数。

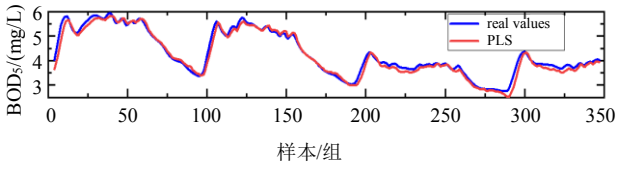
为验证TD-JIT-RVM模型的性能，本文选择PLS、极限学习机（Extreme Learning Machine, ELM）、SVM 3个基础模型作为对比模型，分别引入TD算法和JIT算法进行软测量建模性能测试分析，模型参数如表2所示（PLS无参数设置），预测结果如图4所示。

表2　对比模型参数设置

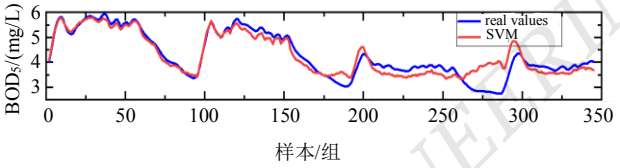
	模型	参数设置
基础模型	ELM	神经元 *N*=15，输入输出层激活函数 *sigmoid*
	SVM	惩罚系数 *c*=95 *gamma*=0.7　*RBF* 核函数
	RVM	核函数 *gauss*，*width*=5
在线软测量模型	TD+JIT+ELM	时间间隔 *i*=1，相似因子 *d*=10，神经元 *N*=15，输入输出层激活函数 *logsig*
	TD+JIT+SVM	惩罚系数 *c*=95 *gamma*=0.7　*RBF* 核函数
	TD+JIT+RVM	时间间隔 *i*=1，相似因子 *d*=10，核函数 *gauss*，*width*=5



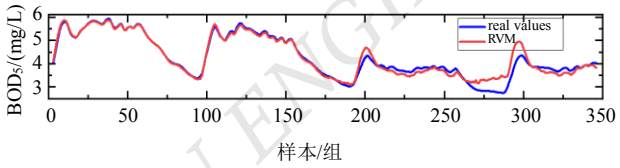
(a)　ELM



(b)　PLS



(c)　SVM

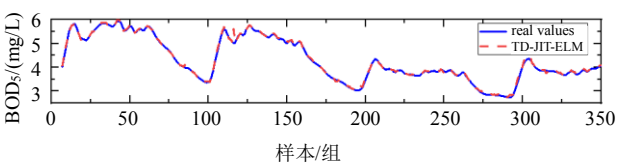


(d)　RVM

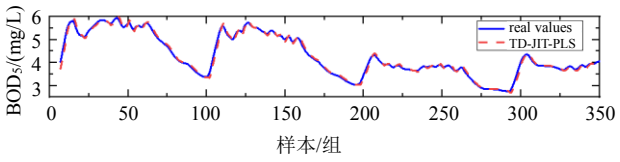
图4　对比预测结果图

由图4可以看出，PLS模型的预测效果较好，但在前期 BOD_5 波动较大时，RVM模型的预测效果比PLS模型好，说明RVM模型在非线性建模中比PLS模型具有更好的跟踪性能，PLS模型更适用于线性预测，这是因为PLS是一个线性模型，不能提取非线性数据特征；ELM是一个前馈神经网络，非线性处理能力比PLS好，但是模型泛化能力较弱，无法很好地控制模型的复杂度，对于小样本数据集，容易出现过拟合的情况，模型在新数据上的表现可能会受到影响；SVM在小样本数据集上泛化能力表现出色，但其性能受核函数的影响较大，且核函数必须符合Mercer条件；RVM无需满足Mercer条件，能够设立多种核函数，相对于SVM算法更具灵活性，且RVM更为稀疏，试验时间更短，更适合于在线检测。

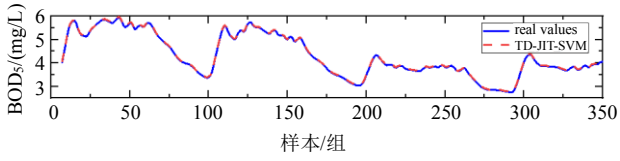
在4个基础模型中分别引入TD算法和JIT算法，图5　对比模型预测结果如图5所示。



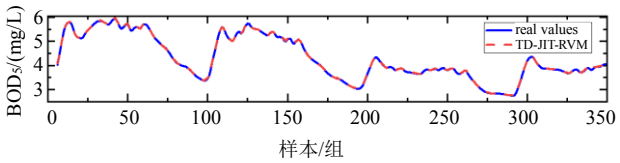
(a)　TD-JIT-ELM



(b) TD-JIT-PLS



(c) TD-JIT-SVM



(d) TD-JIT-RVM

图 5 对比模型预测结果

由图 5 可以看出，ELM、PLS、SVM、RVM 4 个基础模型的预测精度都得到了有效提升。因为这 4 个基础模型都是全局离线模型，在线应用时，模型参数没有得到及时更新，难以适应时变过程，但引入 TD-JIT 算法后，采用即时自适应微调策略，根据查询样本动态实时更新模型参数，可以选择最相关的标记样本来快速微调模型到最新的进程运行状态。模型预测结果对比如表 3 所示。

表 3 模型预测结果对比表

	模型名称	R_{RMSE}	r
基础模型	ELM	0.557 0	0.910 4
	PLS	0.152 4	0.988 9
	SVM	0.349 3	0.921 5
	RVM	0.203 7	0.974 3
在线软测量模型	TD-JIT-ELM	0.049 6	0.998 6
	TD-JIT-PLS	0.087 1	0.995 3
	TD-JIT-SVM	0.023 9	0.999 9
	TD-JIT-RVM	**0.011 0**	**0.999 9**

由表 3 可以看出，引入 TD 算法和 JIT 算法后的模型比 4 个基础模型的预测精度分别提升了 91%、423%、93%以及 94%，且 TD-JIT-RVM 模型的预测精度高于其他 7 个模型。这是因为 TD 算法通过比较当前时刻的数据与前一时刻的数据之差来提取变量之间的动态关系。JIT 算法通过局部建模，动态地更新模型参数，使得模型能够快速地适应实时数据的变化，两者结合，使基础模型不仅拥有更好的鲁棒性和自适应性，且更适用于污水处理过程中的复杂性和不稳定性。对比文献[27]，本文提出的 TD-JIT-RVM 模型优于对比模型，且具有较高的稳定性。

3.2 UCI 真实数据实验分析

为进一步测试 TD-JIT-RVM 模型的性能，采用 UCI 真实数据集进行实验分析[26]。该数据集从西班牙巴塞罗那市某个污水处理厂收集，为方便阐述，简称 UCI 数据集。该污水处理厂的污水处理工艺为活性污泥法，处理过程图如图 6 所示。其中包含来自不同时间段的污水水样数据和多个水质指标，如 pH 值、COD（化学需氧量）、BOD（生化需氧量）、溶解氧、氨氮等。数据采集的频率为每隔一天采集 1 次，采集的数据中包括 38 个与污水生化系统中有机物和微生物相关的指标变量。

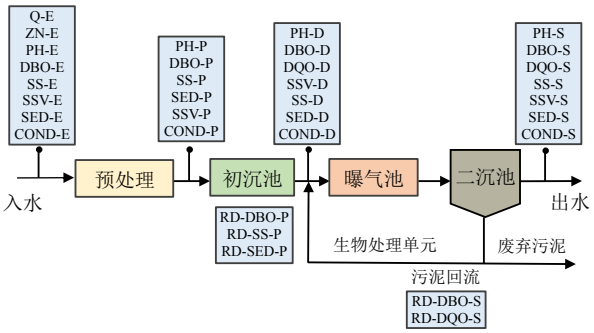
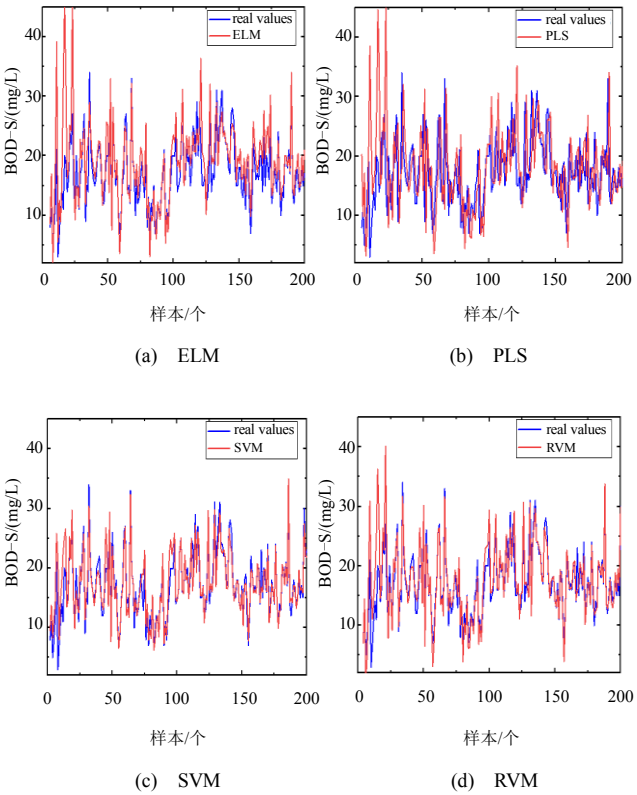


图 6 西班牙巴塞罗那某污水处理厂的污水处理过程图[26]

UCI 数据集经过前期预处理[32]共包括 400 个样本数据。其中，前 200 个样本数据训练模型，剩余的 200 个样本数据验证模型的性能。辅助变量的选择与文献[32]一致。先对样本数据进行均值归一化预处理，再利用预处理后的数据构建 BOD_5 软测量模型。

选用 PLS、ELM、SVM 3 个模型作为对比模型，分别引入 TD 算法和 JIT 算法进行软测量建模性能测试分析，模型参数与 BSM1 案例一致，相似因子 d = 14，预测结果如图 7 所示。



(a) ELM (b) PLS

(c) SVM (d) RVM

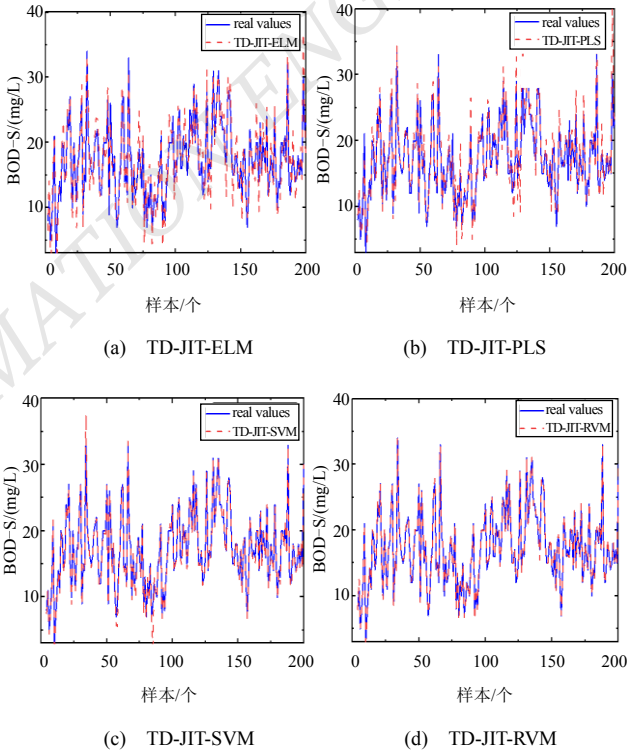
图 7 对比模型预测结果图

由图 7 可以看出，4 个基础模型的前期追踪效果较差，特别是 PLS 模型，这是由于 UCI 数据为强非线性数据，虽然 PLS 模型可以引入非线性因素来提高模型的拟合度，但拟合能力不足；ELM 采用的单层随机投影（single layer random projection, SLRP）神经网络结构无法充分拟合数据，且污水数据较少，无法保证其泛化能力，导致模型预测精度低，追踪效果差；相较于 PLS 和 ELM 模型，SVM 模型可以处理高维数据、非线性数据和小样本数据，且具有良好的泛化能力，但 SVM 模型需要选择的参数较多，如核函数类型、核函数参数、惩罚因子等，不同的参数组合会对模型性能产生不同的影响，对于比较复杂的污水强非线性数据集，参数选择的难度较大且需要大量的实验，费时费力；相较于 SVM 模型，RVM 模型可以使用更少的支持向量来拟合数据，从而简化模型并提高计算效率。与其他机器学习方法相比，RVM 的结果易解释，更容易应用于真实数据。

对比模型的预测结果如表 4 和图 8 所示。

表 4 模型预测结果对比表

	模型名称	R_{RMSE}	r
基础模型	ELM	4.441 0	0.808 0
	PLS	8.053 7	0.207 8
	SVM	2.541 0	0.898 8
	RVM	2.434 6	0.910 8
在线软测量模型	TD-JIT-ELM	2.698 2	0.911 9
	TD-JIT-PLS	2.224 3	0.947 0
	TD-JIT-SVM	0.8102	0.992 5
	TD-JIT-RVM	**0.431 8**	**0.997 3**



(a) TD-JIT-ELM (b) TD-JIT-PLS

(c) TD-JIT-SVM (d) TD-JIT-RVM

图 8 对比模型预测结果图

由表 4 可知，TD-JIT-ELM、TD-JIT-PLS 和 TD-JIT-SVM 模型的预测结果没有 TD-JIT-RVM 模型效果好。与 RVM 模型相比，TD-JIT-RVM 模型的 *RMSE* 提高了 82.26%，这是因为 TD 算法在分析强非线性数据时，采用一阶差分算法，不需要对状态空间或动作空间进行显式建模或者假设线性关系；相反，它能够根据前一个数据自动调整学习状态，且基于数据的特征和结构自适应地调整模型的参数和结构，RVM 结合 2 种算法的优势则可以更好地捕捉变量之

间的非线性关系。

4 结论

本文提出的 TD-JIT-RVM 模型，能够在实时数据处理中学习和优化模型参数，从而实现数据的动态建模和实时预测。通过 2 个案例分析得出，对比 RVM 模型，在加入 TD 算法和 JIT 算法后，BSM1 案例中的 *RMSE* 提高了 94.59%，UCI 案例中的 *RMSE* 提高了 82.26%，验证了该模型的有效性。然而，本文所提在线模型的参数更新时间较长，下一步工作将引入快速边际似然算法来加快模型参数的更新速度。

参考文献

[1] JIANG Y, YIN S, DONG J, et al. A review on soft sensors for monitoring, control, and optimization of industrial processes[J]. IEEE Sensors Journal, 2020,21(11):12868-12881.

[2] KADLEC P, GRBIĆ R, GABRYS B. Review of adaptation mechanisms for data-driven soft sensors[J]. Computers & Chemical Engineering, 2011,35(1):1-24.

[3] SUN Q, GE Z. A survey on deep learning for data-driven soft sensors[J]. IEEE Transactions on Industrial Informatics, 2021, 17(9):5853-5866.

[4] SOUZA F A A, ARAÚJO R, MENDES J. Review of soft sensor methods for regression applications[J]. Chemometrics and Intelligent Laboratory Systems, 2016,152:69-79.

[5] SEDGHI S, SADEGHIAN A, HUANG B. Mixture semisupervised probabilistic principal component regression model with missing inputs[J]. Computers & Chemical Engineering, 2017,103:176-187.

[6] DONG J, ZHANG K, HUANG Y, et al. Adaptive total PLS based quality-relevant process monitoring with application to the Tennessee Eastman process[J]. Neurocomputing, 2015,154: 77-85.

[7] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504-507.

[8] NI J, ZHANG C, YANG S X. An adaptive approach based on KPCA and SVM for real-time fault diagnosis of HVCBs[J]. IEEE Transactions on Power Delivery, 2011, 26(3): 1960-1971.

[9] VAPNIK V. The nature of statistical learning theory[M]. Springer Science & Business Media, 1999.

[10] SUYKENS J A K, VANDEWALLE J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9:293-300.

[11] GE Z, SONG Z. Nonlinear soft sensor development based on relevance vector machine[J]. Industrial & Engineering Chemistry Research, 2010,49(18):8685-8693.

[12] LYU Y, YANG T, LIU J. An adaptive least squares support vector machine model with a novel update for NOx emission prediction[J]. Chemometrics and Intelligent Laboratory Systems, 2015,145:103-113.

[13] TIPPING M E. Sparse Bayesian learning and the relevance vector machine[J]. Journal of machine learning research, 2001, 1(Jun):211-244.

[14] WANG J, QIU K, WANG R, et al. Development of soft sensor based on sequential kernel fuzzy partitioning and just-in-time relevance vector machine for multiphase batch processes[J]. IEEE Transactions on Instrumentation and Measurement, 2021,70:1-10.

[15] CHEN C, WANG Y, ZHANG Y, et al. Indoor positioning algorithm based on nonlinear PLS integrated with RVM[J]. IEEE Sensors Journal, 2017,18(2):660-668.

[16] 王春雷,赵琦,秦孝丽,等.基于改进相关向量机的锂电池寿命预测方法[J].北京航空航天大学学报,2018,44(9):1998-2003.

[17] 吴菁,刘乙奇,刘坚,等.基于动态多核相关向量机的软测量建模研究[J].化工学报, 2019,70(4):1472-1484.

[18] 谭承诚,于广平,邱志成.基于集成相关向量机的水质在线预测模型[J].计算机测量与控制,2018,26(3):224-227.

[19] 许玉格,刘莉,罗飞.基于自优化的多属性高斯核函数相关向量机方法[J].华南理工大学学报(自然科学版),2017, 45(1): 88-94.

[20] 乔俊飞,孙子健,汤健.面向工业过程软测量建模的概念漂移检测综述[J].控制理论与应用,2021,38(8):1159-1174.

[21] YAO L, GE Z. Online updating soft sensor modeling and industrial application based on selectively integrated moving window approach[J]. IEEE Transactions on Instrumentation and Measurement, 2017,66(8):1985-1993.

[22] KANEKO H, FUNATSU K. Discussion on time difference models and intervals of time difference for application of soft sensors[J]. Industrial & Engineering Chemistry Research, 2013,52(3):1322-1334.

[23] GUO F, XIE R, HUANG B. A deep learning just-in-time modeling approach for soft sensor based on variational autoencoder[J]. Chemometrics and Intelligent Laboratory Systems, 2020,197:103922.

[24] 吴菁.污水处理非稳态特性下核建模方法关键问题的研究[D].广州:华南理工大学,2020.
[25] QIN S J. Recursive PLS algorithms for adaptive data modeling[J]. Computers & Chemical Engineering, 1998, 22(4-5): 503-514.
[26] KANEKO H, ARAKAWA M, FUNATSU K. Development of a new soft sensor method using independent component analysis and partial least squares[J]. AIChE Journal, 2009, 55(1): 87-98.
[27] 袁凌玲.面向污水处理过程的软测量若干关键技术研究[D].广州:华南理工大学,2021.

作者简介：
金绍琴，女，1997 年生，研究生，主要研究方向：软测量建模及优化等。E-mail: airen_jsq@163.com
唐丽丽，女，1997 年生，研究生，主要研究方向：软测量建模及优化等。
吴菁（通信作者），女，1988 年生，博士，副教授，主要研究方向：软测量建模、深度学习等。E-mail: ipicq@gzmu.edu.cn
李明珠，女，1983 年生，硕士，副教授，主要研究方向：智能检测与控制、软测量建模等。
黄道平，男，1961 年生，博士，教授，主要研究方向：智能检测与控制、软测量技术等。

（上接第 13 页）

[19] XIONG J, LIANG J, ZHUANG Y, et al. Real-time localization and 3D semantic map reconstruction for unstructured citrus orchards[J]. Computers and Electronics in Agriculture, 2023, 213: 108217.
[20] 马鑫,梁新武,蔡纪源.基于点线特征的快速视觉 SLAM 方法[J].浙江大学学报(工学版),2021,55(2):402-409.
[21] 陈兴华,蔡云飞,唐印.一种基于点线不变量的视觉 SLAM 算法[J].机器人,2020,42(4):485-493.
[22] 贾松敏,丁明超,张国梁.RTM 框架下基于点线特征的视觉 SLAM 算法[J].机器人,2019,41(3):384-391.
[23] 李海丰,胡遵河,陈新伟.PLP-SLAM:基于点、线、面特征融合的视觉 SLAM 方法[J].机器人,2017,39(2):214-220;229.
[24] 林凯,梁新武,蔡纪源.基于重投影深度差累积图与静态概率的动态 RGB-D SLAM 算法[J].浙江大学学报(工学版),2022,56(6):1062-1070.
[25] 王浩,卢德玖,方宝富.动态环境下基于增强分割的 RGB-D SLAM 方法[J].机器人,2022,44(4):418-430.
[26] 谷晓琳,杨敏,张燚,等.一种基于半直接视觉里程计的 RGB-D SLAM 算法[J].机器人,2020,42(1):39-48.
[27] 徐陈,周怡君,罗晨.动态场景下基于光流和实例分割的视觉 SLAM 方法[J].光学学报,2022,42(14):147-159.
[28] 冯明驰,刘景林,李成南,等.一种多焦距动态立体视觉 SLAM[J].仪器仪表学报,2021,42(11):200-209.
[29] 魏彤,李绪.动态环境下基于动态区域剔除的双目视觉 SLAM 算法[J].机器人,2020,42(3):336-345.
[30] 王云峰,翁秀玲,吴炜,等.基于贪心策略的视觉 SLAM 闭环检测算法[J].天津大学学报(自然科学与工程技术版),2017,50(12):1262-1270.
[31] 史殿习,童哲航,杨绍武,等.面向场景变化的动态自适应同时定位与地图构建[J].中国科学:技术科学,2018,48(12):1373-1391.
[32] 安平,王国平,余佳东,等.一种高效准确的视觉 SLAM 闭环检测算法[J].北京航空航天大学学报,2021,47(1):24-30.
[33] 曹剑飞,余金城,潘尚杰,等.采用双视觉里程计的 SLAM 位姿图优化方法[J].计算机辅助设计与图形学学报,2021,33(8):1264-1272.

作者简介：
孟繁，男，1998 年生，硕士研究生，主要研究方向：农业电气化与自动化。E-mail: 3270530377@qq.com
周馨塑，女，1997 年生，博士，主要研究方向：智慧农业、机器视觉、农业电气化与自动化。E-mail: zxinzhao@126.com
吴烽云（通信作者），女，1988 年生，博士，主要研究方向：智慧农业、机器视觉。E-mail: fyseagull@163.com
邹天龙，男，1986 年生，大专，主要研究方向：测控系统集成应用。E-mail: 84174619@qq.com