

采用最小二乘支持向量机的青霉素发酵过程建模研究 Modelling a Penicillin Fed-batch Fermentation Using Least Squares Support Vector Machines

刘 毅, 王海清*

LIU Yi and WANG Hai-Qing*

工业控制技术国家重点实验室, 浙江大学工业控制技术研究所, 杭州 310027

National Laboratory of Industrial Control Technology, Institute of Industrial Process Control, Zhejiang University, Hangzhou 310027, China

摘 要 生化过程通常是严重非线性和时变的复杂动态系统,而且重要过程参数缺少在线测量仪表,对其建立机理模型往往非常耗时和困难。采用最小二乘支持向量机(LS-SVM)并以 Pensim 仿真平台为例对青霉素发酵这一典型生化过程进行建模研究。给出了 LS-SVM 参数的调整策略和分析结果,建立了青霉素产物浓度、菌体浓度和底物浓度等重要过程变量的在线预报模型。仿真结果表明用 LS-SVM 建立的在线预报模型拟合误差小,推广性能好,可以作为发酵过程的进一步控制和优化的参考依据。

关键词 生化建模, 支持向量机, 青霉素发酵过程, 在线预报模型

中图分类号 TP301.6, TQ465.1, TQ920.6 **文献标识码** A **文章编号** 1000-3061(2006)01-0144-06

Abstract The biochemical processes are usually characterized as seriously time varying and nonlinear dynamic systems. Building their first-principle models are very costly and difficult due to the absence of inherent mechanism and efficient on-line sensors. Furthermore, these detailed and complicated models do not necessary guarantee a good performance in practice. An approach via least squares support vector machines (LS-SVM) based on Pensim simulator is proposed for modelling the penicillin fed-batch fermentation process, and the adjustment strategy for parameters of LS-SVM is presented. Based on the proposed modelling method, the predictive models of penicillin concentration, biomass concentration and substrate concentration are obtained by using very limited on-line measurements. The results show that the models established are more accurate and efficient, and suffice for the requirements of control and optimization for biochemical processes.

Key words biochemical modelling, support vector machines, penicillin fed-batch fermentation, on-line predictive model

大多数的生化反应在通常都是极其复杂的,反应过程中的非线性、时变性和不确定性严重,而且目前缺少对重要过程参数的在线检测仪器,造成生化过程的自动化水平远低于其它工业生产过程。生化过程的建模是对其进一步实施控制和优化的前提,

因此研究满足工业应用要求的在线建模理论有着极其重要的理论和现实意义^[1]。

传统过程建模主要有机理分析(白箱模型)、系统辨识(黑箱模型)和混合建模(灰箱模型)三类方法。由于生化过程的特点,采用机理建模一般是很

Received: June 28, 2005; Accepted: October 11, 2005.

This work was supported by the grants from the National Natural Sciences Foundation of China(No.20206028, 20576116) and Alexander von Humboldt-Stiftung (Deutschland).

* Corresponding author. Tel:86-571-87951442-803; Email:hqwang@iipc.zju.edu.cn

国家自然科学基金(No. 20206028,20576116)和德国洪堡基金会资助项目。

困难的;在非机理建模方面,已有许多理论成果(比如统计回归^[2]、神经网络^[3]等),最近的研究表明,结合优化算法的神经网络^[4]应用到生化过程的建模中能得到不错的效果。但这些方法基本是在大样本假设下才有效,而且对过程时变特性的处理效果并不理想。因此得到的模型终究不能准确、及时地反映实际生化过程的变化,这对后续的基于模型的控制和优化必将造成一定的负面影响。

针对生化过程的特点,本文主要研究了在小样本情况下如何对数据进行处理,怎样利用少量的可测变量的数据进行快速建模。首先介绍了支持向量机(Support Vector Machines, SVM)的小样本建模理论,提出将最小二乘支持向量机(Least Squares Support Vector Machines, LS-SVM)这种基于SVM改进后的快速算法,应用到青霉素发酵这一典型的生化过程中来。基于Pensim仿真平台的仿真结果表明,该方法仅通过很少批次的样本数据的学习,就建立了青霉素产物浓度、菌体浓度和底物浓度等只能离线检测的重要过程变量的在线预报模型。

1 SVM与LS-SVM回归估计理论

1.1 SVM回归估计理论

SVM是上个世纪90年代初由Vapnik等人提出的一种新的学习方法^[5]。它的数学理论基础严密,与其它学习方法相比有更好的非线性处理能力和推广能力,特别是由于SVM使用了结构风险最小化(Structural Risk Minimization, SRM)原则避免了神经网络等学习方法可能出现的过学习和局部极小等问题,而且SVM的拓扑结构可由支持向量决定,避免了神经网络拓扑结构需要经验试凑的局限性,被认为是目前针对小样本分类和回归等问题的最佳理论。因此在模式识别、回归估计、概率密度函数估计等方面都有广泛应用并取得了很好的效果。但其在工业领域的应用研究还相对较少,在生化过程的应用更是少见。

首先考虑用线性函数 $f(x) = \langle w, x \rangle + b$ 来拟合样本数据集 $\{x_i, y_i\} (i = 1, 2, \dots, n; x_i \in R^d; y_i \in R)$,并采用线性不敏感损失函数如(1)式所示,其中 x 为自变量, y 为因变量, ε 代表用于控制拟合精度的不敏感参数。

$$\begin{aligned} L[y, f(x, a)] &= L[y - f(x, a)] \\ &= \begin{cases} 0 & |y - f(x, a)| \leq \varepsilon \\ |y - f(x, a)| - \varepsilon & \text{others} \end{cases} \quad (1) \end{aligned}$$

SVM就是在式(1)的约束条件下寻找最优拟合超平面,即最小化 $\frac{1}{2} \|w\|^2$,使所有训练数据都可以在精度 ε 下用线性函数拟合,其中 y 为标量, x_i 和 w 为 d 维向量。并引入(非负)松弛变量 ξ_i, ξ_i^* 和用于控制松弛变量在目标泛函中作用的(正的)惩罚系数 C, C 越大代表对超出误差 ε 的惩罚力度越强。SVM控制函数集复杂性的方法是使回归函数最平坦,因此回归建模将转化为式(2)所示的目标泛函在约束条件式(3)下的最小化问题。

$$\min \Phi(x, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (2)$$

$$s. t. \begin{cases} y_i - \langle w, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0, i = 1, 2, \dots, n \end{cases} \quad (3)$$

采用拉格朗日乘子法求解这个具有线性不等式约束的二次规划问题,即可得到对偶最优化问题如式(4)所示

$$\begin{aligned} \max W(\alpha, \alpha^*) &= -\varepsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) + \sum_{i=1}^n y_i (\alpha_i^* - \alpha_i) \\ &\quad - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) (\langle x_i, x_j \rangle) \\ s. t. &\begin{cases} \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0 \\ 0 \leq \alpha_i^* \leq C, 0 \leq \alpha_i \leq C, i = 1, 2, \dots, n \end{cases} \end{aligned} \quad (4)$$

求解式(4)的拉格朗日乘子 α_i^*, α_i 便可得到回归函数的参数 w 和 b 如下:

$$w = \sum_{i=1}^n (\alpha_i^* - \alpha_i) x_i \quad (5)$$

$$\begin{cases} b = y_i - \langle w, x_i \rangle - \varepsilon & \text{if } \alpha_i^* - \alpha_i < 0 \\ b = y_i - \langle w, x_i \rangle + \varepsilon & \text{if } \alpha_i^* - \alpha_i > 0 \end{cases} \quad (6)$$

由式(5)和式(6)即可得到如下的线性回归函数

$$f(x) = \langle w, x \rangle + b = \sum_{i=1}^n (\alpha_i^* - \alpha_i) (\langle x_i, x \rangle) + b \quad (7)$$

对于非线性问题,估计函数可表示为 $f(x) = \langle w, \phi(x) \rangle + b$,其中 $\phi(x)$ 是一非线性函数,它将训练数据集映射到高维线性特征空间,并在该特征空间里寻求最优回归函数。按照上面的方法同样可以得到回归函数

$$f(x) = \langle w, \phi(x) \rangle + b$$

$$= \sum_{i=1}^n (\alpha_i^* - \alpha_i) (\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle) + b \quad (8)$$

由于在式(7)和式(8)两式中均只涉及到内积计算 $\langle \mathbf{x}_i, \mathbf{x} \rangle$ 和 $\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle$, 根据 Hilbert-Schmidt 理论, 只要满足 Mercer 定理, 即可用核函数 $K(\mathbf{x}_i, \mathbf{x})$ 将高维特征空间的内积 $\langle \phi(\mathbf{x}_i), \phi(\mathbf{x}) \rangle$ 转化为原空间的内积函数进行计算, 而无须显式的计算该非线性函数 $\phi(\mathbf{x})$, 从而避免高维特征空间引起的维数灾难问题。即所得的回归函数式(7)和式(8)两式均可表示为式(9):

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b \quad (9)$$

常见的核函数有线性核函数、多项式核函数、径向基核函数(Radial Basis Function, RBF)、多层感知器核函数四种, 本文的核函数采用通用的 RBF 形式。

1.2 最小二乘支持向量机建模理论

众所周知, 最小二乘法(Least Squares, LS)是解决多元函数回归的经典方法。Suykens 等人^[6]在 SVM 的基础上引进了 LS, 将不等式约束改为等式约束, 避免了求解耗时的 QP 问题, 从而很大程度上降低了运算时间, 为在线估计创造了有利条件。而且相对于常用的线性不敏感损失函数, LS-SVM 不再需要指定逼近精度 ϵ , 只是 LS-SVM 的解不具有稀疏性。LS-SVM 和 SVM 的主要区别如下:

优化目标和约束条件分别变为式(10)和式(11):

$$\begin{aligned} \min J(\mathbf{w}, b, \xi) &= \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^n \xi_i^2 \quad (10) \\ \text{s.t. } y_i &= \langle \mathbf{w}, \Phi(\mathbf{x}_i) \rangle + b + \xi_i, i = 1, 2, \dots, n \end{aligned} \quad (11)$$

与 SVM 的求解方法相似, 可以得到以下方程组

$$\begin{bmatrix} 0 & \mathbf{e}^T \\ \mathbf{e} & Q + \frac{1}{\gamma} I \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \quad (12)$$

其中新出现参数的定义如式(13)所示, 其余参数与 SVM 中相同。

$$\begin{aligned} \mathbf{e} &= [1 \cdots 1]^T \in R^n, \quad \boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_n]^T \in R^n, \\ \mathbf{y} &= [y_1 \cdots y_n]^T \in R^n, \\ I &\in R^{n \times n}, \quad Q = [Q_{ij}]_{n \times n}, \quad q_{ij} = y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) \end{aligned} \quad (13)$$

求解式(12)最终得到建模函数

$$\begin{aligned} f(\mathbf{x}) &= \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b \\ &= \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b \end{aligned} \quad (14)$$

2 基于 LS-SVM 的青霉素发酵过程关键变量在线预报

青霉素发酵过程是较典型的生化过程, 由于发酵过程的强非线性、时变性和不确定性, 目前又缺少对关键过程变量如产物浓度、菌体浓度和底物浓度的在线检测仪器, 严重影响了对发酵过程进行有效的控制和优化, 因此建立这些关键过程变量的在线预报模型有着重要的现实意义。需要指出的是, 本文虽是以青霉素发酵这一典型的生化过程为例进行研究, 但下面所提出的方法完全适用于其他生化过程的建模问题。

2.1 基于 Pensim 仿真平台在线预报模型的建立

本文用于建模的数据来自 Pensim 仿真平台, 该软件的内核采用基于 Bajpai 机理模型改进的 Birol 模型, 在此平台上可以简易实现青霉素发酵过程的一系列仿真, 相关研究已表明该仿真平台的实用性与有效性^[7-9]。在对青霉素发酵过程的研究过程中, 已经有很多机理模型。典型的非结构模型如 Bajpai 模型等, Birol 模型主要是根据 Pirt 等^[10]的实验数据对 Bajpai 模型进行改进。在此模型中, 不仅考虑了 pH 值、温度、空气流量、搅拌功率、底物流加速率等控制变量对菌体及青霉素产量的影响, 而且把菌体生长、二氧化碳、青霉素产量, 底物消耗量、反应热等因素也全面包含到模型中去, 因此能比较全面地反映青霉素发酵过程。

影响青霉素发酵过程的可在线测量变量主要有空气流量、搅拌功率、温度、pH 值、溶解氧浓度(记为 DO)、溶解二氧化碳浓度(记为 P_{CO_2})等。其中, 温度和 pH 值一般都控制为常量故不予考虑。空气通入量和搅拌功率对 DO 和 P_{CO_2} 都会有影响, 进而会影响到菌体浓度等其它变量。 DO 是需氧发酵控制的最重要的参数之一, 上述几个变量都会对其产生影响, 它同时对菌体生长及产物的性质和产量也会产生很大的影响, 可以很好地反映整个发酵过程。此外, 排气二氧化碳浓度(E_{CO_2})也在一定程度上反映了菌体生长、产物的合成和代谢。 E_{CO_2} 是细胞代谢流, 是反应菌体生理特性的状态变量。由于 P_{CO_2} 和 DO 等并不是完全意义上的(独立)操作变量, 因此在建模中适当引入一些重要的状态变量将有助于模型精度的提高。但由于在本仿真模型中也没有输出该变量的测量值, 因此文中 E_{CO_2} 不考虑作为建模输

入变量。在 Pensim 仿真平台中,温度和 pH 值这两个被控变量均可采用闭环(PID 或开关)控制,其值几乎为恒定;而空气流量和搅拌功率这两个输入变量也控制为恒定值附近。此外, Pensim 仿真平台默认菌种质量、搅拌桨特性等因素都不变。

本文的出发点是研究小样本情况下如何对数据进行处理,怎样利用少量的数据进行快速建模。因此为保证所建模型的简单,只选择 DO 这个变量用 LS-SVM 方法建立产物浓度、菌体浓度和底物浓度的在线预报模型如式(15)所示,而在实际生产中,完全可以根据具体工况选择合适的可在线测量的变量。

$$\begin{aligned} P(t) &= f_1[DO(t)], X(t) = f_2[DO(t)], \\ S(t) &= f_3[DO(t)] \end{aligned} \quad (15)$$

其中, $DO(t)$ 、 $P(t)$ 、 $X(t)$ 、 $S(t)$ 分别代表 t 时刻的溶解氧浓度、产物浓度、菌体浓度和底物浓度。本文对 Pensim 仿真平台默认的初始条件进行了建模研究,仿真时对同一工况下不同批次的初始条件及仿真后的采样数据分别加入了一定的噪声以更加符合实际的检测情况。

以建立 $P(t) = f_1[DO(t)]$ 形式的在线预报模型为示例。样本数据可以表示为 $I \times J \times K$ 的三维结构,如图 1 所示。其中 I 代表反应批次,为更好地考验所提方法的小样本建模能力,这里仅取 5 批样本数据; J 代表过程变量,此处为 $DO(t)$ 和 $P(t)$; K 代表每一批次的反应时间,为简单起见,均取为 400h, $DO(t)$ 和 $P(t)$ 样本每隔 4h 采样一次,故 $K = 100$ 。若要对第 i 批次进行预测,则将前 $i - 1$ 批次的样本按批次进行切片,并按照批次顺序排成一矩

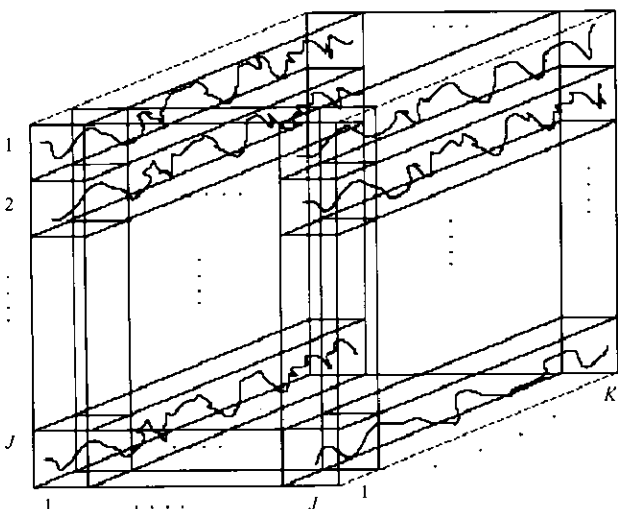


图 1 青霉素发酵过程数据的三维表现形式
Fig. 1 Data presentation in three-way array of penicillin fed-batch fermentation

阵,如图 2 所示,当成一批样本进行训练得到预测模型,并将其用以预测第 i 批次。

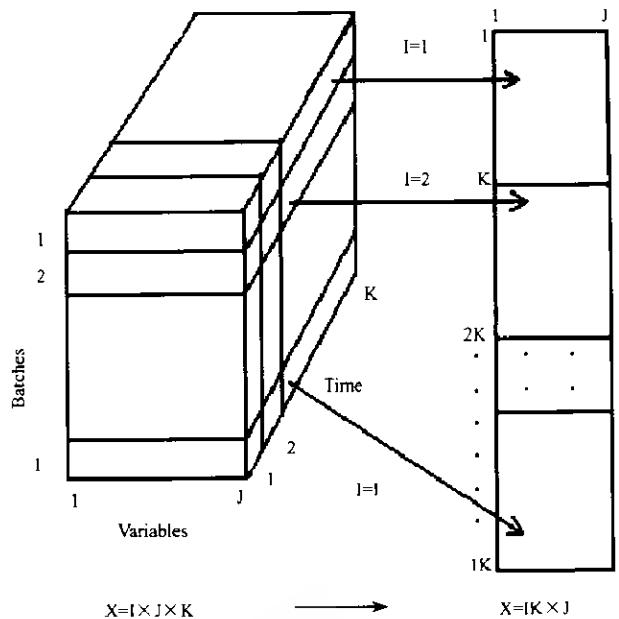


图 2 用切片法将数据按批次方向展开
Fig. 2 Unfolding of batch process data by preserving variable direction

2.2 LS-SVM 参数调整及分析

本文采用 RBF 形式的 LS-SVM,有两个重要的参数 γ 和 σ^2 需要调整。 γ 为正则化参数也即惩罚系数,用于控制函数的拟合误差, γ 值越大,拟合误差越小,相应的训练时间也就越多,但 γ 过大会导致过拟合; σ^2 是核函数参数,代表着 RBF 的带宽,随着 σ^2 的变小,拟合误差会变小,相应的训练时间也就变长,但 σ^2 过小会导致过拟合。目前关于对参数的选择还没有一套系统有效的方法,本文针对青霉素发酵这一具体过程,提出了一套确定上述参数的方法。

其基本思想是:首先将 σ^2 固定,让 γ 不断增大,因为随着 γ 的增大拟合误差和预测误差都会先变小。以预测均方误差 $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2}$ 为标准计算样本拟合误差和预测误差,其中 y_i^* 和 y_i 分别为 LS-SVM 建模的结果和样本的实际值,当预测误差首次出现极小值时即停止训练,相应的 γ 值作为选定值。然后将此 γ 值作为初值,用同样的方法训练得到 σ^2 ,于是 LS-SVM 的参数 (γ, σ^2) 即可确定。这样得到的 (γ, σ^2) 不仅保证了预测误差小即模型的推广能力好,而且训练时间短,可以适应实时在线预报的要求。

以 2 批次为例,拟合误差和预测误差及训练时

间随 γ 和 σ^2 的变化趋势分别如图 3、4 所示。从图中我们可以很清楚地看到仿真结果很好地反映了上述的理论分析。

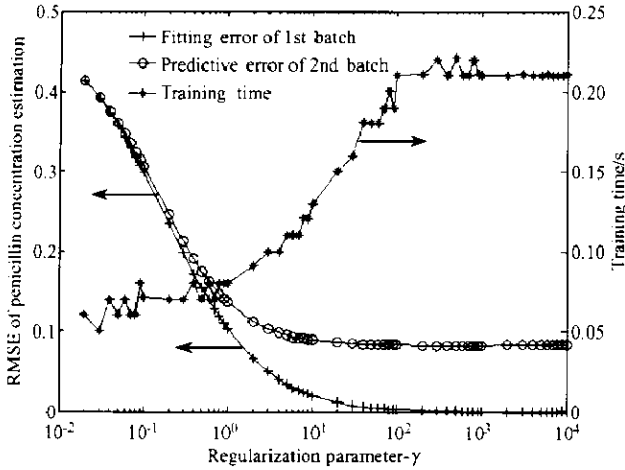


图 3 正则化参数 γ 对拟合误差, 预测误差和训练时间的影响
Fig. 3 Effect of regularization parameter γ on fitting error, prediction error and training time

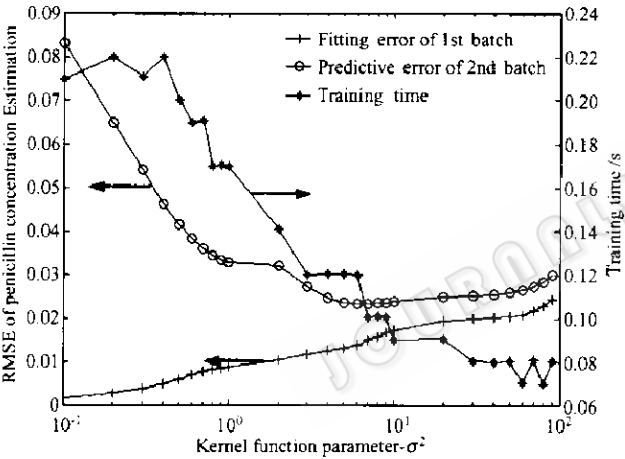


图 4 核函数参数 σ^2 对拟合误差, 预测误差和训练时间的影响
Fig. 4 Effect of kernel function parameter σ^2 on fitting error, prediction error and training time

2.3 青霉素发酵过程重要变量的在线预报模型

分别对 2~5 批次的样本分别建立在线预报模型 $P(t) = f_1[DO(t)]$ 。当一共有 i 批次的样本时, 取前 $i-1$ 批次进行训练并计算拟合误差, 第 i 批用来检验模型的推广能力并计算预测误差。表 1 列出了所建预报模型的拟合误差和预测误差。

从表中可以看出预报模型的精度很高, 而且通过批次的增加预测误差还将慢慢变小, 说明模型还有一定的自学习能力。为全面检验模型的推广能力, 再采用交叉验证法进行仿真, 即任取 5 批数据中的 4 批按上述方式训练一预报模型, 第 5 批次作为检验, 分别计算每次的拟合误差、预测误差和训练时

间(CPU2.4GHz, 内存 256MB, Matlab6.5)列于表 2。

表 1 青霉素产物浓度预报模型的拟合误差和预测误差
Table 1 Fitting error or predictive error of penicillin concentration

Batch data	RMSE; fitting error or predictive error of each batch				
	1st batch	2nd batch	3rd batch	4th batch	5th batch
2 batches	0.0139	0.0233			
3 batches	0.0193	0.0159	0.0384		
4 batches	0.0165	0.0213	0.0276	0.0250	
5 batches	0.0173	0.0176	0.0290	0.0182	0.0179

表 2 交叉验证法建立 5 批次的青霉素产物浓度在线预报模型的性能比较
Table 2 Performance comparison of predictive penicillin concentration using cross-validation of 5 batches

	RMSE: fitting error or predictive error of each batch					Training time/s
	1st batch	2nd batch	3rd batch	4th batch	5th batch	
1	0.0173	0.0176	0.0290	0.0182	0.0179	2.9040
2	0.0142	0.0162	0.0230	0.0268	0.0252	2.5340
3	0.0229	0.0216	0.0196	0.0184	0.0336	1.0720
4	0.0244	0.0219	0.0122	0.0179	0.0247	2.8040
5	0.0218	0.0276	0.0214	0.0163	0.0200	1.4220

从表 2 得知虽然训练样本改变, 但所建预报模型的预测误差变化并不大, 说明此方法有一定的鲁棒性, 而且训练时间都很短可以对 $P(t) = f_1[DO(t)], X(t) = f_2[DO(t)], S(t) = f_3[DO(t)]$ 的在线预报模型和 Pensim 仿真数据的比较结果。

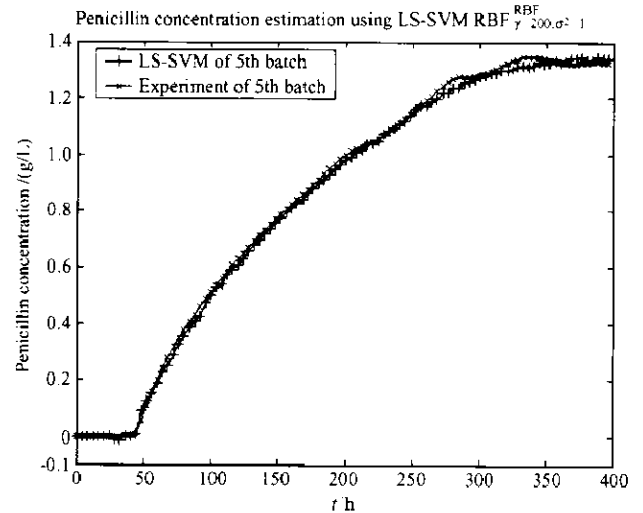


图 5 第 5 批次的产物浓度在线预报模型和仿真比较
Fig. 5 Comparison of predictive penicillin concentration using LS-SVM and experiment of 5th batch

由图 5~7 可知, 所建立的三种预报模型都有较高的精度, 这表明在 Pensim 仿真平台设定的操作条件下, 仅通过 $DO(t)$ 一个输入变量就可以实现对产物浓度、菌体浓度和底物浓度进行准确的在线预报。

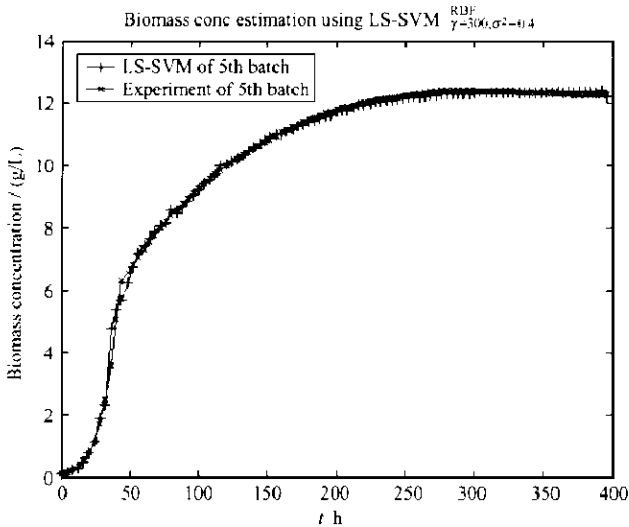


图6 第5批次的菌体浓度在线预报模型和仿真比较

Fig. 6 Comparison of predictive biomass concentration using LS-SVM and experiment of 5th batch

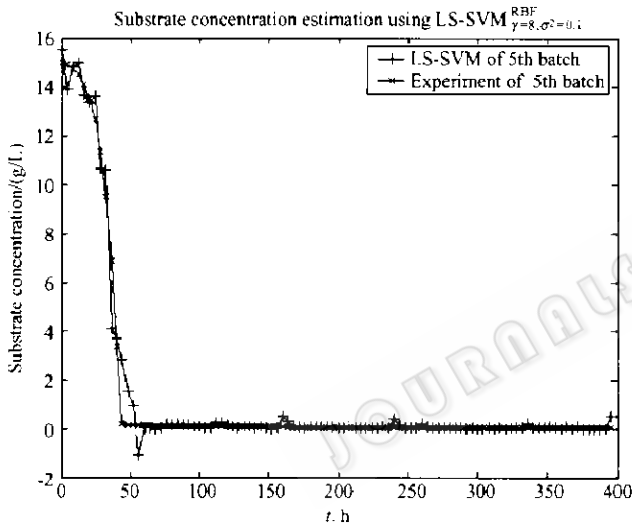


图7 第5批次的底物浓度在线预报模型和仿真比较

Fig. 7 Comparison of predictive substrate concentration using LS-SVM and experiment of 5th batch

需要指出的是,实际生产中仅仅根据 $DO(t)$ 一般是无法得到产物浓度等变量的准确估计的,必须考虑更多的过程因素。但本文所提出的方法仍旧适用,在实际应用中只需将输入变量(如本文的 $DO(t)$)改为与实际工况所要估计的变量相关的可测变量,而且可以选取多个可测变量作为 LS-SVM 的输入来实现对产物浓度等变量的估计。当然,非机理建模必须结合过程知识才能取得更好的效果,如在图5、图7中出现的少数浓度为负的离散点则由常理可以排除。

3 结论

青霉素发酵过程中重要过程变量无法在线测量

得到,严重影响了对其进行控制和优化操作。本文基于 Pensim 仿真平台对此进行了研究,利用可在线测量的 DO ,提出了基于 LS-SVM 的青霉素生产快速建模方法,获得了产物浓度、菌体浓度和底物浓度的在线预报模型。本方法仅利用 2~5 个批次的样本便能快速建立用于下一批次的在线预报模型,在对数据的要求方面远远低于神经网络和多元回归等传统方法,而且得到的预报精度很高。这表明 SVM 及其改进算法如 LS-SVM 完全适用于生化过程的建模。需要指出的是,本文所用数据是在 Pensim 仿真平台得到的(仅涉及某些设定的生产工况),这和实际生产还是有一定差别的。比如发酵过程中泡沫的形成, DO 将会显著上升;菌体形态的变化(即使菌浓没有变)引起 $K_L a$ 的变化,由此也会引起 DO 变化。在这些情况下,仅采用 DO 建立青霉素产物浓度等重要变量的准确在线预报模型就显然是不够的(仅采用 P_{CO_2} 在实际中也可能会有类似的问题)。因此必须再结合具体的工况进行研究,在建模中考虑更多的过程知识才能使所建模型精确、可靠,这也是我们下一步的研究工作。

REFERENCES (参考文献)

- [1] Wang SQ (王树青), Yuan YJ (元英进). Automatic Technology of Biochemical Process. Beijing: Chemical Industry Press (化学工业出版社), 1999
- [2] Nomikos P, MacGregor JF. Monitoring batch processes using multiway principal component analysis. *AIChE J*, 1994, 40(8): 1361-1375
- [3] Massimo CD, Montague GA, Willis MJ et al. Towards improved penicillin fermentation via artificial neural networks. *Computers and Chemical Engineering*, 1992, 16(4): 283-291
- [4] Xiong ZH, Zhang J. Neural network model-based on-line re-optimisation control of fed-batch processes using a modified iterative dynamic programming algorithm. *Chemical Engineering and Processing*, 2005, 44(4): 477-484
- [5] Vapnik VN. The Nature of Statistical Learning Theory. New York: Springer, 1995
- [6] Suykens JAK, Vandewalle J. Least squares support vector machine classifiers. *Neural Processing Letters*, 1999, 9(3): 293-300
- [7] Birol G, Undey C, Cinar A. A modular simulation package for fed-batch fermentation; penicillin production. *Computers and Chemical Engineering*, 2002, 26(11): 1553-1565
- [8] Bajpai, R, Reuss, M. A mechanistic model for penicillin production. *Journal of Chemical Technology and Biotechnology*, 1980, 30: 330-344
- [9] Undey C, Tataru E, Cinar A. Intelligent real-time performance monitoring and quality prediction for batch/fed-batch cultivations. *Journal of Biotechnology*, 2004, 108(1): 61-77
- [10] Pirt, S, Righetto, R. Effect of growth rate on the synthesis of penicillin by *Penicillium chrysogenum* in batch and chemostat cultures. *Applied Microbiology*, 1967, 15: 1284-1290