

基于阶段辨识的诺西肽发酵过程菌体浓度软测量

杨强大¹, 王福利^{2,3}, 常玉清^{2,3}

(1. 东北大学 材料与冶金学院, 辽宁 沈阳 110004; 2. 东北大学 信息科学与工程学院, 辽宁 沈阳 110004;

3. 东北大学 流程工业综合自动化教育部重点实验室, 辽宁 沈阳 110004)

摘要: 由于发酵过程中系统非线性特性与发酵阶段密切相关的实际特点, 针对诺西肽发酵过程菌体浓度的估计问题, 提出了一种基于阶段辨识的软测量方法. 首先以分阶段的诺西肽发酵过程非结构模型为基础, 根据隐函数存在定理进行辅助变量的合理选择; 然后利用经数学推导得到的指示变量“伪比生长率”完成发酵阶段的在线辨识, 并采用神经网络构建出对应于各阶段的局部软测量模型. 实际应用结果表明, 所提方法有效、预估精度较高.

关键词: 发酵; 软测量; 辅助变量选择; 阶段辨识; 神经网络

中图分类号: TP274 **文献标识码:** A

Phase-identifying-oriented soft sensor for biomass in Nosiheptide fermentation process

YANG Qiang-da¹, WANG Fu-li^{2,3}, CHANG Yu-qing^{2,3}

(1. School of Materials and Metallurgy, Northeastern University, Shenyang Liaoning 110004, China;

2. School of Information Science and Engineering, Northeastern University, Shenyang Liaoning 110004, China;

3. Key Laboratory of Integrated Automation of Process Industry, Northeastern University, Ministry of Education, Shenyang Liaoning 110004, China)

Abstract: Fermentation processes have different nonlinear characteristics in different fermentation phases. A Phase-identifying-oriented soft sensor approach is proposed for estimating the biomass in Nosiheptide fermentation process. Based on the segmented unstructured model of Nosiheptide fermentation process, the secondary variables are selected according to the implicit function existence theorem. The on-line identification of fermentation phases is accomplished by using an indicator variable which is gained by mathematical inference, and for each phase, a local soft sensor model is developed. The practical application results show the effectiveness and validity of the presented approach.

Key words: fermentation; soft sensor; selection of secondary variables; phase identification; neural networks

1 引言(Introduction)

发酵是现代流程工业中常见的一种生产方式, 被广泛应用于医药、食品和化工等领域^[1]. 在发酵过程中, 菌体浓度的实时获取, 对于过程的优化与控制具有十分重要的意义. 然而, 目前还没有实用的在线传感器能直接检测该参数, 工业生产中大都采用实验室取样分析的方法来获得. 取样分析具有较大时间滞后, 给菌体浓度的实时监测以及发酵过程的直接质量控制带来了很大的困难. 因此, 应用软测量技术, 实现菌体浓度的在线预估具有重要的实际意义.

本文针对诺西肽发酵过程, 并结合该过程所具有的多阶段特性, 提出了基于阶段辨识的菌体浓度软测量方法. 依据分阶段的诺西肽发酵过程非结构模

型和隐函数存在定理进行辅助变量的合理选择; 基于菌体生长动力学模型和摄氧率方程, 经数学推导得到阶段辨识所需的指示变量“伪比生长率”, 从而实现发酵阶段的在线辨识; 采用RBF神经网络构建出对应于各阶段的局部软测量模型. 最后, 基于诺西肽发酵实际生产过程验证了所提方法的有效性.

2 辅助变量的选择(Selection of secondary variables)

本文在分阶段的诺西肽发酵过程非结构模型的基础上, 根据隐函数存在定理进行辅助变量的合理选择, 其选择过程如下所述.

诺西肽批发酵过程从接种后开始, 相继经历延迟期(lag phase, LP)、指数生长期(exponential phase,

EP)和稳定期(stationary phase, SP)3个阶段. 根据诺西肽发酵实际情况和相关文献^[2~8], 分阶段的诺西肽发酵过程非结构模型可写成如下形式:

$$\begin{cases} \frac{dX}{dt} = \mu X, \\ \frac{dS}{dt} = -(m_S + \frac{\mu}{Y_{X/S}} + \frac{\beta}{Y_{P/S}})X, \\ \frac{dP}{dt} = \beta X - \lambda P, \\ \frac{dC_O}{dt} = g(N, Q) \cdot (C^* - C_O) - (m_O + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}})X, \\ CER = (m_{CO_2} + \frac{\mu}{Y_{X/CO_2}})X. \end{cases} \quad (1)$$

式(1)中,

$$\mu = \begin{cases} \frac{\mu_{\max} S}{k_S X + S} \cdot F_1(pH) \cdot A(T), & LP, \\ \mu_{\max}, & EP, \\ \frac{\mu_{\max} S}{k_S X + S} \cdot \frac{C_O}{k_O + C_O} \cdot F_2(pH), & SP. \end{cases} \quad (2)$$

式(1)(2)中: X, S, P 和 C_O 分别为菌体、基质、产物和溶解氧浓度(g/L); m_S, m_O 和 m_{CO_2} 分别为基质、溶解氧和 CO_2 维持因子((g/g)/h); $Y_{X/S}, Y_{X/O}$ 和 Y_{X/CO_2} 分别为基质、溶解氧和 CO_2 到菌体的得率常数(g/g); $Y_{P/S}$ 和 $Y_{P/O}$ 分别为基质和溶解氧到产物的得率常数(g/g); μ 为比生长率(h⁻¹); β 为产物合成常数((g/g)/h); λ 为水解速率常数, (h⁻¹); N 为搅拌转速, (r/min); Q 为进气流量, (m³/h); C^* 为氧饱和浓度(g/L); $\mu_{\max}$ 为最大比生长率(h⁻¹); k_S 为Contois半饱和常数; k_O 为饱和常数(g/L); T 为发酵液温度(°C); pH 为发酵液pH值; $g(N, Q)$ 为 N 和 Q 的函数, $g(N, Q) > 0$; $A(T)$ 为 T 的函数, $0 < A(T) \leq 1$; $F_1(pH)$ 和 $F_2(pH)$ 为 pH 的函数, $0 < F_1(pH), F_2(pH) \leq 1$.

构成如式(3)~(5)所示子系统.

1) 对于延迟期, 由式(1)中的第4, 5个子式并结合式(2)中的第1个子式构成如下子系统:

$$\begin{cases} \frac{dC_O}{dt} = g(N, Q) \cdot (C^* - C_O) - (m_O + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}})X \\ \quad + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}})X \\ \quad = f_{41}(X, S, C_O, pH, T, N, Q), \\ CER = (m_{CO_2} + \frac{\mu}{Y_{X/CO_2}})X = f_{51}(X, S, pH, T). \end{cases} \quad (3)$$

2) 对于指数生长期, 由式(1)中的第4个子式并结

合式(2)中的第2个子式构成如下子系统:

$$\begin{aligned} \frac{dC_O}{dt} &= g(N, Q) \cdot (C^* - C_O) - \\ &\quad (m_O + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}})X \\ &= f_{42}(X, C_O, N, Q). \end{aligned} \quad (4)$$

3) 对于稳定期, 由式(1)中的第4, 5个子式并结合式(2)中的第3个子式构成如下子系统:

$$\begin{cases} \frac{dC_O}{dt} = g(N, Q) \cdot (C^* - C_O) \\ \quad - (m_O + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}})X \\ \quad = f_{43}(X, S, C_O, pH, N, Q), \\ CER = (m_{CO_2} + \frac{\mu}{Y_{X/CO_2}})X = f_{53}(X, S, C_O, pH). \end{cases} \quad (5)$$

式(3)和式(5)所示子系统, 其雅克比矩阵

$$J_1 = \partial[f_{41} \ f_{51}]/\partial[X \ S]$$

和

$$J_3 = \partial[f_{43} \ f_{53}]/\partial[X \ S]$$

所对应的行列式都不为零; 而式(4)所示子系统, 其对应的偏导数 $\partial f_{42}/\partial X$ 也不为零(限于篇幅, 证明过程略), 则由隐函数存在定理^[9]可知, 式(3)~(5)所示子系统能够确定如下函数关系式.

$$X = \begin{cases} G_1(\dot{C}_O, C_O, CER, pH, T, N, Q), & LP, \\ G_2(\dot{C}_O, C_O, N, Q), & EP, \\ G_3(\dot{C}_O, C_O, CER, pH, N, Q), & SP. \end{cases} \quad (6)$$

由式(6)可知, 实现诺西肽发酵过程菌体浓度软测量需要选择如下辅助变量:

- 1) 延迟期为 $\dot{C}_O, C_O, CER, pH, T, N, Q$;
- 2) 指数生长期为 $\dot{C}_O, C_O, N, Q$;
- 3) 稳定期为 $\dot{C}_O, C_O, CER, pH, N, Q$.

由于非结构模型中存在未知函数表达式, 难以用解析式来描述辅助变量和菌体浓度之间的关系. 本文利用神经网络来逼近非线性函数关系 G_1, G_2 和 G_3 .

3 基于阶段辨识的菌体浓度软测量(Phase-identifying-oriented soft sensor for biomass)

由第2节的证明与分析可知, 在不同的诺西肽发酵阶段, 过程特性有所不同, 体现过程特性的辅助变量信息也不同. 为此, 本文提出了基于阶段辨识的诺西肽发酵过程菌体浓度软测量方法, 其模型结构如图1所示. 该软测量模型包括两部分: 第1部分利用经数学推导得到的指示变量“伪比生长率”完成发酵

阶段的在线辨识,并根据辨识结果确定出所需的局部软测量模型;第2部分包含3个局部神经网络软测量模型,每个模型对应一个阶段,该部分的作用是完成对菌体浓度的在线预估。

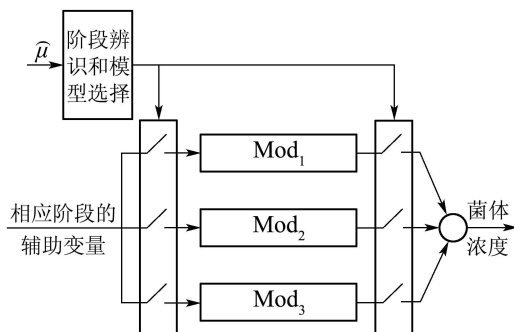


图1 基于阶段辨识的菌体浓度软测量模型

Fig. 1 Phase-identifying-oriented biomass soft sensor model

3.1 诺西肽发酵阶段的在线辨识(On-line identification of Nosiheptide fermentation phases)

发酵过程的不同阶段是由比生长率定义的. 在分批发酵过程中, 比生长率随时间变化的趋势如图2所示. 图2中, $0 \sim t_1$ 为延迟期, $t_1 \sim t_2$ 为指数生长期, 过了 t_2 点为稳定期. 对于某种确定的发酵过程, 指数生长期的比生长率为常数且最大^[8].

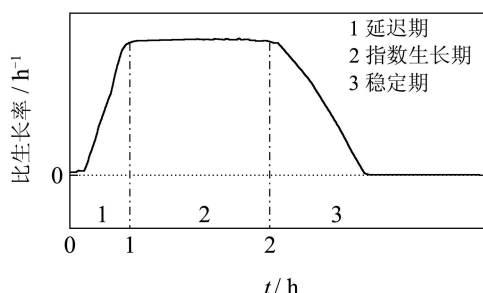


图2 比生长率变化趋势示意图

Fig. 2 Schematic diagram of specific growth rate change tendency

比生长率难于在线检测, 因此不能利用它进行发酵阶段的在线辨识. 本文基于菌体生长动力学模型和诺西肽发酵过程摄氧率(oxygen uptake rate, OUR)方程, 经数学推导得到指示变量“伪比生长率”, 进而完成诺西肽发酵阶段的在线辨识.

对于分批发酵过程, 菌体生长动力学模型为^[8]

$$\frac{dX}{dt} = \mu X. \quad (7)$$

依据文献[7]和诺西肽发酵实际情况, 诺西肽发酵过程摄氧率方程可表示为

$$r_{\text{OUR}} = [m_O + \frac{\mu}{Y_{X/O}} + \frac{\beta}{Y_{P/O}}]X. \quad (8)$$

式(8)中, r_{OUR} 为摄氧率, 它在线可测.

式(8)两边对时间求导, 并结合式(7)(8)可得

$$\dot{\mu} = \mu + a. \quad (9)$$

式(9)中

$$a = \frac{\dot{\mu}}{[\mu + Y_{X/O}m_O + \beta \frac{Y_{X/O}}{Y_{P/O}}]}; \hat{\mu} = \frac{r_{\text{OUR}}}{r_{\text{OUR}}}.$$

本文称 $\hat{\mu}$ 为“伪比生长率”, 它在线可测.

令 $\mu_i (i = 1, 2, 3)$ 分别表示诺西肽发酵过程延迟期、指数生长期和稳定期的比生长率; $\hat{\mu}_i (i = 1, 2, 3)$ 分别表示上述3个阶段的“伪比生长率”. 则由式(9)和分批发酵过程的特点可得出如下结论:

- 1) $\mu_1 < \mu_2, \mu_2 = \mu_{\max}, \mu_3 < \mu_2$;
- 2) $\hat{\mu}_1 > \mu_1$, 其原因是 $\dot{\mu}_1 > 0$, 即 $a > 0$;
- 3) $\hat{\mu}_2 = \mu_2$, 其原因是 $\dot{\mu}_2 = 0$, 即 $a = 0$;
- 4) $\hat{\mu}_3 \leq \mu_3$, 其原因是 $\dot{\mu}_3 \leq 0$, 即 $a \leq 0$.

由上述结论1)3)和4)可知, $\hat{\mu}_2 = \mu_{\max}$ 且 $\hat{\mu}_2 > \hat{\mu}_3$; 而由上述结论1)~3)可知, 存在 $\hat{\mu}_1 \geq \mu_{\max}$ 的可能. 但在实际应用中, $\hat{\mu}_1$ 表现出跟 μ_1 相似的变化趋势, 且并未出现 $\hat{\mu}_1(t) \geq \mu_{\max}$ 这一情况. 综上, $\hat{\mu}_1 < \hat{\mu}_2$, $\hat{\mu}_2 = \mu_{\max}$, $\hat{\mu}_3 < \hat{\mu}_2$, 即可以把 $\hat{\mu}$ 作为指示变量进行诺西肽发酵阶段的在线辨识.

3.2 局部软测量模型的建立(Establishment of local soft sensor model)

将建模数据按其所属阶段的不同分成3类后, 利用神经网络建模方法, 建立3个与不同发酵阶段相对应的局部软测量模型. 当所有局部软测量模型构建完成后, 一个如图1所示的分布式预估网络就确定下来了. 图1中, 子模型 Mod_1 , Mod_2 和 Mod_3 分别对应于诺西肽发酵过程的延迟期、指数生长期和稳定期.

RBF神经网络作为神经网络的一种, 其训练方法不存在局部最优问题, 且具有全局逼近的性质和最佳逼近性能^[10], 故本文采用RBF神经网络建立局部软测量模型. 其中, 隐含层节点数和各节点中心由改进的k-means聚类算法^[10]确定; 隐含层各节点的宽度由P近邻规则^[11]确定; 输出层连接权由递推最小二乘法确定.

3.3 菌体浓度的在线预估(On-line estimation of biomass)

菌体浓度的在线预估过程如下:

1) 首先利用子模型 Mod_1 对菌体浓度进行在线预估. 当出现 $|\hat{\mu}(k) - \mu_{\max}| \leq \sigma$, ($k = i, i+1, i+2$) ($\mu_{\max}$ 为常数, 可通过实验获得; σ 为待定正数)时, 认为发酵过程从第 $i+3$ 时刻起进入指数生长期.

2) 从第 $i+3$ 时刻起, 利用子模型 Mod_2 对菌体

浓度进行在线预估. 当出现 $|\hat{\mu}(k) - \mu_{\max}| > \sigma$, ($k = j, j+1, j+2$)时, 认为发酵过程从第 $j+3$ 时刻起进入稳定期.

3) 从第 $j+3$ 时刻起, 利用子模型Mod₃对菌体浓度进行在线预估.

4 实验研究(Experimental research)

为了验证本文提出的基于阶段辨识的诺西肽发酵过程菌体浓度软测量方法的有效性, 利用东北大学“211工程”生物发酵实验室中的诺西肽发酵过程, 对该软测量方法进行检验. 诺西肽分批发酵过程持续时间为96 h左右, 选取5批正常发酵过程所得样本数据建立软测量模型. 其中, 菌体浓度通过实验室取样分析获得, 其采样周期为3 h左右; 辅助变量通过DCS系统获得, 其采样周期为5 min.

将所建软测量模型应用于某批诺西肽发酵实际生产过程, 其应用结果如下所述.

1) 发酵阶段的在线辨识结果.

由本文提出的辨识方法得到的阶段辨识结果如图3所示, 其中参数 σ 取 0.003 h^{-1} . 由图3可知, 0~11.5 h左右为延迟期, 11.5~36.1 h左右为指数生长期, 过了36.1 h以后为稳定期. 该辨识结果与实际过程相符, 表明本文提出的诺西肽发酵阶段在线辨识方法是有效的.

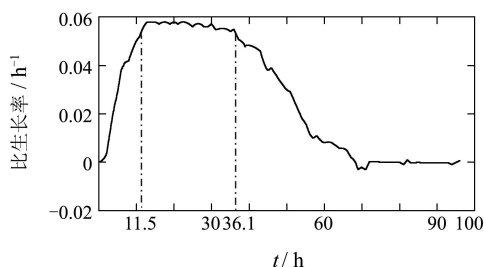


图3 发酵阶段辨识结果

Fig. 3 Identification result of fermentation phases

2) 菌体浓度的在线预估结果.

菌体浓度的在线预估结果如图4所示. 由图4可以看出, 基于阶段辨识的软测量方法能以较高精度对诺西肽发酵过程菌体浓度进行在线预估.

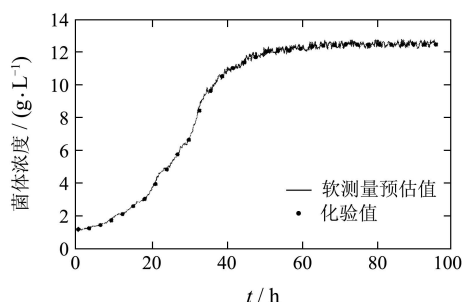


图4 菌体浓度预估结果

Fig. 4 Estimation result of biomass

另外, 将基于阶段辨识的软测量方法与全局软测量方法进行了比较. 全局软测量模型仍采用RBF神经网络进行构建, 其输出为菌体浓度, 输入为 DO , DO , CER , pH , T , N , Q . 利用第4节的5批建模数据对全局软测量模型进行训练, 其实际应用结果如表1所示. 由表1可以看出, 采用基于阶段辨识的软测量方法进行菌体浓度软测量其预估精度优于采用全局软测量方法的结果, 表明对菌体浓度进行分阶段软测量更符合诺西肽发酵过程所具有的多阶段特性.

表1 两种软测量方法泛化误差比较

Table 1 Comparison of generalization errors of the two soft sensor methods

软测量方法	分阶段软测量方法	全局软测量方法
平均相对误差	2.23%	4.96%
最大相对误差	5.01%	7.82%

5 结论(Conclusion)

诺西肽发酵过程具有明显的阶段性, 且在不同的发酵阶段, 菌体浓度软测量模型的输入也不同. 针对上述特点, 本文提出了基于阶段辨识的软测量方法. 实际应用结果表明, 该软测量方法具有较高预估精度, 且与全局软测量方法相比, 预估精度明显提高.

此外, 本文提出的根据隐函数存在定理选择辅助变量的方法, 为辅助变量的选择提供了新的思路.

参考文献(References):

- [1] SCHÜGERL K. Progress in monitoring, modeling and control of bioprocesses during the last 20 years[J]. *Journal of Biotechnology*, 2001, 85(2): 149 – 173.
- [2] KOUTINAS A A, WANG R, KOOKOS I K, et al. Kinetic parameters of *Aspergillus awamori* in submerged cultivations on whole wheat flour under oxygen limiting conditions[J]. *Biochemical Engineering Journal*, 2003, 16(1): 23 – 34.
- [3] BONA R, MOSER A. Modeling of the L-glutamic acid production with *Corynebacterium glutamicum* under biotin limitation[J]. *Bioprocess and Biosystems Engineering*, 1997, 17(3): 139 – 142.
- [4] NOOR S K, INDRA M M, SINGH R P. Modeling the growth of *Corynebacterium glutamicum* under product inhibition in L-glutamic acid fermentation[J]. *Biochemical Engineering Journal*, 2005, 25(2): 173 – 178.
- [5] RUCHI S, SUBHASH C, ASHOK K S. Batch kinetics and modeling of gibberellic acid production by *Gibberella fujikuroi*[J]. *Enzyme and Microbial Technology*, 2005, 36(4): 492 – 497.
- [6] CLAUDIO G, RICARDO P C, EDUARDO A. Modeling *Gibberella fujikuroi* growth and GA3 production in solid-state fermentation[J]. *Process Biochemistry*, 2002, 37(9): 1033 – 1042.
- [7] BIROL G, UNDEY C, CINAR A. A modular simulation package for fed-batch fermentation: penicillin production[J]. *Computers and Chemical Engineering*, 2002, 26(11): 1553 – 1565.

- [8] 戚以政, 汪叔雄. 生化反应动力学与反应器(第二版)[M]. 北京: 化学工业出版社, 1999.
(QI Yizheng, WANG Shuxiong. *Biochemical Reaction Kinetics and Reactor II*[M]. Beijing: Chemical Industry Press, 1999.)
- [9] 陆启韶. 现代数学基础[M]. 北京: 北京航空航天大学出版社, 1997.
(LU Qishao. *Fundamentals of Modern Mathematics*[M]. Beijing: Beijing University of Aeronautics and Astronautics Press, 1997.)
- [10] LI C, WANG G. A clustering method and its application to the modeling of RBF networks[C] // *Proceedings of the Sixth International Conference for Young Computer Scientist: in Computer Science and Technology in New Century*. Hangzhou: International Academic Publishers, 2001: 173 – 177.
- [11] MOODY J, DARKEN C J. Fast learning in networks of locally-tuned processing units[J]. *Neural Computation*, 1989, 1(2): 281 – 294.

作者简介:

杨强大 (1979—), 男, 讲师, 主要从事复杂工业过程建模与软测量, E-mail: daqiang9666@163.com;

王福利 (1957—), 男, 教授, 博士生导师, 校长助理, 主要从事复杂工业过程建模、故障诊断与智能控制等;

常玉清 (1973—), 女, 副教授, 主要从事复杂工业过程建模与软测量.

下 期 要 目

- 基于滑模观测器的车辆电子稳定性控制系统故障重构 于金泳, 刘志远, 陈 虹
- 考虑阻塞的集装箱堆场集卡路径控制策略 李 东, 汪定伟, 刘黎黎
- 自适应变尺度混沌免疫优化算法及其应用 杨海东, 鄂加强
- PS转炉造渣过程的动态优化控制 宋海鹰, 桂卫华, 阳春华, 王雅琳
- 求解PMU多目标优化配置问题的非劣排序微分进化算法 彭春华, 孙惠娟, 郭剑峰
- 动态不确定非线性系统直接自适应模糊Backstepping控制 贺向雷, 佟绍成
- 非平衡拓扑结构的多智能体网络系统一致性协议 谭拂晓, 关新平, 刘德荣
- 分数阶系统频域辨识算法 李 旺, 张国庆, 王 永
- 导弹动、静试验差异的参数化建模及补偿 谢美华
- 机翼主动颤振抑制输出微分调节器设计 张子健, 徐 敏, 陈士槽
- 基于高阶滑模和加速度反馈的机械谐振抑制方法 于 晶, 冯 勇, 郑剑飞
- 基于变折扣因子EWMA的混合产品批间控制方法 侯明阳, 郑 英, 艾 兵, 郑 燕, 张 杰
- 水下机器人广义预测控制算法及能耗问题研究 徐建安, 任立国, 杨立平, 张铭钧