

# 发酵过程生物量软测量技术的研究进展

王建林, 于涛

(北京化工大学信息科学与技术学院, 北京 100029)

**摘要:**生物量是发酵过程中的关键过程参数之一,它直接影响着发酵过程的优化和控制。综述了近年来发酵过程生物量软测量技术的研究现状,讨论了基于过程机理分析、回归分析、状态估计和神经网络等的软测量建模方法,对基于神经网络和改进的神经网络建模方法进行了分析。指出基于多尺度建立软测量混合模型,是实现发酵过程生物量在线测量的有效方法,并给出了建立混合模型需要解决的关键问题。

**关键词:**生物量;软测量;发酵

中图分类号:TP18;TP274

文献标识码:A

文章编号:0253-4320(2005)06-0022-04

## Research progress in soft-sensor techniques for on-line biomass estimation in fermentative process

WANG Jian-lin, YU Tao

(School of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)

**Abstract:** Biomass is one of the most important variables in the fermentative process, and it directly influences the automatic control performance of fermentation system as well as in system monitoring and on-line optimization. The researches of soft-sensor techniques for on-line biomass estimation in the fermentative process are reviewed in past years, and modeling methods based on the fermentative process mechanism analysis, regression analysis, state estimation, artificial neural network (ANN), are discussed, and furthermore the modeling methods based on ANN and improved ANN are analyzed. It is pointed out that the soft-sensor hybrid model is an effective approach to the on-line biomass estimation in the fermentative process, which tries to use the best available process mechanism model to describe the fermentative process dynamics, coupled with the so-called "black box" models such as ANN or support vector machine (SVM). The key problems involved in the hybrid modeling are presented.

**Key words:** biomass; soft-sensor; fermentation

发酵过程中的生物量是关键的过程参数,由于受生物传感技术发展水平的限制,一直没有很好地解决生物量的在线测量问题,导致先进优化控制算法和策略只能停留在理论探讨上,难以在工业上得到实际应用<sup>[1]</sup>。

生物量测量大多采用手工离线取样测量,如干重法、离心叠集细胞体积法、直接染色法、光密度法和细胞计数法等<sup>[2]</sup>。这些方法滞后时间长,测量误差大,且测量精度受死细胞影响,不能及时反映发酵状态,难以满足发酵过程实时控制的要求。生物量在线测量对发酵过程的优化和控制具有重要意义。在以硬件形式直接在线测量的方法中,代表性的方法有光浊度法、荧光法、介电常数法、超声波法等<sup>[3-4]</sup>。这些方法在使用上都有局限性,每种测量方法都有各自的特点和适用范围,且在线分析仪器价格昂贵,维护费用高,测量滞后大,难以适用于所有发酵过程的生物量在线测量。

近年来,发酵过程生物量软测量技术的研究已取得一些进展,其主要成果有:基于过程机理分析、回归分析、状态估计建立软测量模型;基于人工神经网络和改进的人工神经网络建立软测量模型。

### 1 基于过程机理分析的软测量

基于过程机理分析的软测量是从质能平衡、Monod型方程、Arrhenius型方程等出发,建立以过程动力学为基础的、反映生物量与可测变量之间关系的机理模型("白箱"模型),实现生物量在线估计。

Zhao对发酵过程的机理建模方法进行了分析<sup>[5]</sup>,并从质能平衡出发建立了机理模型,对发酵过程生物量进行在线估计和自适应控制。

机理建模是建立在对发酵过程机理有深刻认识的基础上,需要对发酵过程机理较为清楚。虽然机理模型中各参数的物理意义明确,能给出生物量和易测辅助变量的关系,然而,基于质能平衡、Monod

收稿日期:2004-12-29;修回日期:2005-04-18

基金项目:国家自然科学基金资助项目(20476007)

作者简介:王建林(1965-),男,博士,教授,博士生导师,研究方向为生化过程智能检测与控制,010-64433803, wangjl@mail.buct.edu.cn。

型方程等所建立的“白箱”软测量模型,因在发酵过程中菌体生长反应等理论处理上大为简化,导致模型不能反映微生物生长反应的真实性质,适应性差;而且目前对微生物发酵过程仍缺乏足够的了解,很多未知微生物反应根本无法直接建模,因此直接建立有效的发酵过程机理模型难度较大。

对于机理研究不充分、尚不完全清楚的发酵过程,难以建立合适的机理模型,有时可以与其他参数估计方法相结合建模。如 Pinchuk 等<sup>[6]</sup>在脱氮假单胞菌自循环发酵过程中,由质能平衡建立了机理模型,并采用基因算法实现系数匹配,对生物量进行估计,取得了较好效果。

## 2 基于回归分析的软测量

以部分最小二乘(PLS)为基础的一元和多元线性回归辨识建模方法比较成熟,广泛用于线性模型辨识。实际应用中,常采用主元回归(PCR)法和部分最小二乘法回归(PLSR)法建立简化的软测量模型,对于线性系统,PCR法和PLSR法均有较好的效果,对于非线性系统,PLSR法的效果较好。

Feng 等<sup>[7]</sup>在重组埃希氏杆菌(*Escherichia coli*)发酵过程中建立了基于PLSR的软测量模型,对菌体和蛋白浓度在线估计。Chae 等<sup>[8]</sup>在埃希氏杆菌培植过程的在线优化研究中,从重组埃希氏杆菌流加发酵的机理出发,采用物质平衡方程,用连续二次规划方法估计方程参数,推导在线估计和优化算法,取得一定效果。

Warnes 等<sup>[9]</sup>在重组埃希氏杆菌发酵过程中,分别采用多变量线性回归(MLR)、PCR、PLSR、自回归滑动平均(ARMAX)、非线性自回归滑动平均(NARMAX)等建模方法对生物量和重组蛋白浓度进行了在线估计,并对每种建模方法进行了比较。从比较结果看,虽然MLR、PCR、PLSR、ARMAX等方法方便、直观,但对非线性过程的跟踪特性差、预测误差大,不稳定;有非线性项的NARMAX法得到的估计性能较好。

由于发酵过程具有高度的非线性,采用单一回归分析方法难以建立合适的软测量模型。实际应用中,通过对发酵过程机理分析,给出模型各变量组合的大致框架,将回归分析和其他辨识方法结合起来获得生物量软测量模型。如张泉灵等<sup>[10]</sup>在链霉素发酵过程中,采用主元回归分析法,结合模糊集理论,构建了模糊分段线性模型,实现产物浓度的在线估计,该模型与线性多元回归模型相比有较高的预

测精度。Lennox 等<sup>[11]</sup>在流加发酵装置的状态监控系统,采用了非线性多路最小二乘法(MPLS)和人工神经网络方法来建立生物量软测量模型。McGovern 等<sup>[12]</sup>在赤霉素生产过程中,采用PLSR和人工神经网络相结合,建立了赤霉素浓度的估计模型,用PLSR来降低系统维数,用人工神经网络来估计目标变量,取得了较好的预测精度。

回归分析法算法简单,但它需要较多的数据样本,且对测量误差比较敏感。

## 3 基于状态估计的软测量

基于状态估计的软测量是从发酵过程的状态空间模型出发,将对于所选辅助变量是完全可观测的生物量作为系统的状态变量,结合构造自适应观测器、卡尔曼滤波器(Kalman filter, KF)或扩展卡尔曼滤波器(extended Kalman filter, EKF)实现在线估计。

Lubenova<sup>[13]</sup>在好氧发酵过程中通过一般模型构造自适应观测器,导出生物量自适应观测算法,并用仿真方法对算法进行了验证。Yuan 等<sup>[14]</sup>采用基本的氨合成过程模型,研究了生物量最大生长率的在线估计,给出了滤波估计算法。Takiguchi 等<sup>[15]</sup>在赖氨酸发酵过程中,采用EKF估计算法预测了谷氨酸棒杆菌(*Corynebacterium glutamicum* AJ-3462)浓度。

基于状态估计的软测量模型能反映生物量与辅助变量之间的动态关系,有利于处理各变量间动态特性的差异和系统滞后,但该软测量模型是在发酵过程状态空间模型的基础上建立的,发酵过程的复杂性增加了建模难度,虽然可通过简化系统描述来降低建模难度,但会导致在线估计误差增大,且持续缓慢不可测的扰动也会影响该模型的估计精度。

## 4 基于神经网络的软测量

### 4.1 基于标准神经网络的软测量

基于人工神经网络的软测量可在不具备对象先验知识的条件下,根据对象的输入输出数据直接建模,模型的在线校正能力强,适合解决高度非线性和严重不确定性系统的软测量建模问题。

随着神经网络技术的发展,1992年英国纽卡斯尔大学(Newcastle University)的Di Massimo和Montague等<sup>[16]</sup>在青霉素发酵过程中首先应用人工神经网络技术构造了软传感器,实现生物量在线估计,取得较好的估计效果。随后一些研究者开始用人工神经网络来描述发酵过程和在线预测生物量<sup>[17]</sup>。

Warnes 等<sup>[9]</sup>通过对几种基于回归分析的建模方

法与基于标准人工神经网络建模方法的比较,认为在高度非线性的发酵过程中,采用人工神经网络建立生物量“黑箱”模型,估计结果较好,但较难选取其中涉及的非线性多项式的阶数及相互作用项。Glassey 等<sup>[18]</sup>采用前馈网络和标准反向传播(BP)算法建立“黑箱”模型,研究了青霉素 G 的发酵过程中生物量的估计问题。Bhowmik 等<sup>[19]</sup>以普遍的微生物发酵过程生物量在线估计为研究对象,采用标准人工神经网络建立了生物量“黑箱”模型,并通过主元分析(PCA)法降低数据维数,得到了较好的估计结果,但文中没有研究人工神经网络的结构和学习算法的有效性。

采用标准 BP 算法对网络进行训练,网络学习的收敛速度慢,需要较多的实验数据才能达到要求的精度,对于一些无法提供大量数据的发酵过程,标准 BP 算法不适用。当输入数据维数较大,实验数据量又较少时,需采用结构不同、性能更好的网络结构,以及改进的网络学习算法来提高网络的收敛速度,通过数据预处理减小输入数据维数。

人工神经网络能以任意精度逼近任意连续非线性函数,对复杂不确定性问题具有自适应和自学习能力,其信息处理具有并行性和冗余性,可用于求解大规模问题,有很好的容错性能。然而,人工神经网络对超过训练数据覆盖的数据区域以外的输入误差无法控制,其性能对实际环境的变化非常敏感,通常需要较多的历史数据进行学习,并可能会出现训练过程中网络参数收敛速度较慢的问题。因此,采用人工神经网络方法建模必须考虑网络结构和学习算法。网络结构的选择与实际解决的问题类型有关,而所采用的学习方法将直接影响网络的收敛、收敛速度以及最后的精确度。

#### 4.2 基于改进的神经网络的软测量

改进人工神经网络的训练方法、网络拓扑结构,能大大提高生物量软测量模型的性能和估计精度。

Bachinger 等<sup>[20]</sup>在重组埃希氏杆菌发酵过程中,应用人工神经网络建立了生物量和特征生长率的模型,采用改进的列文伯格-玛尔特(Levenberg-Marquardt, LM)网络训练方法,该方法可以调整每次计算的步长,能获得较好的收敛速度,与使用标准 BP 训练方法相比,其减少实验数据量的效果非常明显。Cimander 等<sup>[21]</sup>在酸奶发酵过程中对乳糖、半乳糖、乳酸盐进行在线估计,使用改进的拟牛顿(quasi-Newton)训练方法,以加快人工神经网络学习的收敛速度。James 等<sup>[22]</sup>在有机聚合物的流加发酵过程中

采用径向基函数(RBF)神经网络在线估计生物量,给出了多神经网络结构的建模方法,使模型具有更好的逼近性。黄明志等<sup>[23]</sup>在红霉素发酵过程状态预估中,也采用 RBF 神经网络结构,建立了菌体浓度和葡萄糖浓度的软测量模型,模型具有相当好的学习能力,收敛速度快。Ronen 和 Shabtai<sup>[24]</sup>在酵母发酵生物量和普鲁兰多糖的在线估计中,将模糊集理论和人工神经网络相结合,提出模糊人工神经网络建模方法,有效地减少了输入变量的维数,估计效果较好。方千山和王永初<sup>[25]</sup>在木糖醇发酵过程中,对人工神经网络结构和训练方法进行改进,将均匀设计和遗传算法与人工神经网络结合起来,建立了木糖醇浓度在线软测量模型,实现了发酵过程的实时控制,提高了木糖醇得率。

人工神经网络在网络结构形式上的改进主要有采用 RBF 网络结构,以及与模糊理论相结合构建模糊人工神经网络。RBF 网络具有广泛的非线性适应能力,使用的学习算法有 Mooky 和 Darken 算法、局部训练算法、正交优选算法、聚类和 Givens 变换联合算法。RBF 网络的学习算法速度快,能避免局部极小值问题,但在通过数据来确定 RBF 基函数的中心时,可能需要较多实验数据。

标准 BP 算法是多层前向人工神经网络结构中常用的学习算法,是一种基于梯度下降的最优化算法。该算法充分利用了多层前向网络的结构优势,在正反向传播过程中每一层的计算都是并行的,算法在理论上比较成熟;然而该算法存在着收敛速度慢,目标函数存在局部极小值点的问题。对标准 BP 算法的改进主要有加入动量项,采用共轭梯度法、牛顿法、Levenberg-Marquart 算法等。Levenberg-Marquart 算法结合了梯度下降法和牛顿法的优点,具有更好的学习质量。

基于人工神经网络的“黑箱”软测量模型,虽无须深入了解发酵过程机理,能实现逼近性能良好的非线性映射,但所选网络结构和学习训练算法对模型的精度以及泛化能力影响较大,仅仅采用人工神经网络建立“黑箱”模型适用范围有限,也难以推广。

近年来,将先验知识、机理分析和人工神经网络相结合建立软测量“灰箱”模型的方法也开始出现,该方法既能利用已有的先验知识和已知过程机理,又能降低建模难度。如 Shene 等<sup>[26]</sup>建立了单孢菌发酵过程中生物量和产物浓度在线估计的“灰箱”模型,该模型采用了机理模型,使用人工神经网络对模型中的非线性系数进行估计,简化了建模难度。

Zorzetto 等<sup>[27]</sup>在啤酒发酵过程中,从质能平衡方程出发,采用 Monod 模型,通过人工神经网络辨识 Monod 模型中的 3 个系数,建立了“灰箱”模型,对发酵罐中主要底物浓度进行了在线估计,并对“黑箱”模型和“灰箱”模型的效果进行了比较,从比较结果可知“灰箱”模型的效果较好。隋青美和王正欧<sup>[28]</sup>对于一般的发酵过程,从质能平衡方程出发给出机理参数模型,由人工神经网络估计模型中的不可测参数,建立了生物量在线估计的“灰箱”模型,得到较好的仿真结果。Hagedorn 等<sup>[29]</sup>也采用了“灰箱”模型研究了真氧产碱菌(*Alcaligenes eutrophus*)流加发酵过程中葡萄糖和生物量的预测,模型预测精度高。

“灰箱”模型采用机理参数模型,由人工神经网络辨识机理模型中的参数,因此,机理模型的简化和未能进行机理分析的过程部分仍影响“灰箱”模型估计的准确性,且模型的针对性强。

## 5 结语

发酵过程是时变、非线性、带约束的多输入多输出复杂系统,其机理十分复杂,由于对发酵过程缺乏足够的了解,很难建立准确的“白箱”机理模型。人工神经网络建模受网络拓扑结构和类型、学习数据数量和质量的影响较大,用单一人工神经网络建立的模型有效域较窄,适用范围有限。“灰箱”模型实质上是机理参数模型,缺乏对生物量和辅助变量之间关系的全面描述。

发酵过程生物量软测量建模实质上是一个多尺度建模问题,应从先验知识和过程机理分析、统计学习等尺度上出发,充分考虑发酵过程微生物生长反应特性,利用发酵工程技术研究的最新成果,应用人工智能理论和技术来建立由发酵过程机理模型和辨识模型构造的软测量混合模型,全面反映生物量和各变量之间的关系,才是解决生物量在线估计的有效方法。目前,建立软测量混合模型,实现发酵过程生物量在线估计还有一些关键问题需要解决,如数据的预处理方法,不同尺度上模型的分析 and 划分方法,有效辅助变量的选取原则以及软测量混合模型的在线校正方法等。

## 参考文献

- [1] De Assis A J, Filho R M. [J]. Computers and Chemical Engineering, 2000, 24(2-7): 1099-1103.
- [2] Olsson L, Nielson J. [J]. Trends in Biotechnology, 1997, 15(11): 517-522.
- [3] 王贻俊, 樊育. [J]. 生物化学与生物物理进展, 2000, 27(4): 387-390.
- [4] 王武, 杨海麟. [J]. 无锡轻工大学学报, 2002, 21(3): 322-326.
- [5] Zhao Y. Studies on modeling and control of continuous biotechnical processes [D]. Norway: Norwegian University of Science and Technology, 1996.
- [6] Pinchuk R J, Brown W A, Hughes S M, et al. [J]. Biotechnology and Bioengineering, 2000, 67(1): 19-24.
- [7] Feng M, Glassey J. [J]. Biotechnology and Bioengineering, 2000, 69(5): 495-503.
- [8] Chae Hee Jeong, Delisa M P, Cha Hyung Joon, et al. [J]. Biotechnology and Bioengineering, 2000, 69(3): 275-285.
- [9] Warnes M R, Glassey J, Montague G A. [J]. Process Biochemistry, 1996, 31(2): 147-155.
- [10] 张泉灵, 金晓明, 王树青, 等. [J]. 无锡轻工大学学报, 2000, 19(5): 447-451.
- [11] Lennox B, Montague G A, Hiden H G, et al. [J]. Biotechnology and Bioengineering, 2001, 74(2): 125-135.
- [12] McGovern A C, Broadhurst D, Taylor J, et al. [J]. Biotechnology and Bioengineering, 2002, 78(5): 527-538.
- [13] Lubenova V. [J]. Journal of Biotechnology, 1996, 46(3): 197-207.
- [14] Yuan Z, Bogaert H, Devisscher M, et al. [J]. Biotechnology and Bioengineering, 1999, 65(3): 265-273.
- [15] Takiguchi N, Shimizu H, Shioya S. [J]. Biotechnology and Bioengineering, 1997, 55(1): 170-181.
- [16] Di Massimo C, Montague G A, Willis M J, et al. [J]. Computers and Chemical Engineering, 1992, 16(4): 283-291.
- [17] Dacosta P, Kordich C, Williams D, et al. [J]. Artificial Intelligence in Engineering, 1997, 11(4): 383-392.
- [18] Glassey J, Ignova M, Ward A C, et al. [J]. Journal of Biotechnology, 1997, 52(3): 201-205.
- [19] Bhowmik U K, Saha G, Barua A, et al. [J]. Chemical Engineering and Technology, 2000, 23(6): 543-549.
- [20] Bachinger Th, Mårtensson P, Mandenius C-F. [J]. Journal of Biotechnology, 1998, 60(1-2): 55-66.
- [21] Cimander C, Carlsson M, Mandenius C-F. [J]. Journal of Biotechnology, 2002, 99(3): 237-248.
- [22] James S, Legge R, Budman H. [J]. Journal of Process Control, 2002, 12(1): 113-121.
- [23] 黄明志, 杭海峰, 储炬, 等. [J]. 华东理工大学学报, 2000, 26(2): 162-165.
- [24] Ronen M, Shabtai Y. [J]. Biotechnology and Bioengineering, 2002, 77(4): 420-429.
- [25] 方千山, 王永初. [J]. 厦门大学学报(自然科学版), 2001, 40(6): 1324-1327.
- [26] Shene C, Diez C, Bravo S. [J]. Computers and Chemical Engineering, 1999, 23(8): 1097-1108.
- [27] Zorzetto L F M, Filho R M, Wolf-Macieli M R. [J]. Computers and Chemical Engineering, 2000, 24(2-7): 1355-1360.
- [28] 隋青美, 王正欧. [J]. 系统仿真学报, 2002, 14(4): 415-417.
- [29] Hagedorn A, Legge R L, Budman H. [J]. Biotechnology and Bioengineering, 2003, 83(1): 104-111. ■