

生物发酵过程的建模、 优化与故障诊断

高学金 齐咏生 王 普 著



科学出版社

生物发酵过程的建模、 优化与故障诊断

高学金 齐咏生 王 普 著

科学出版社

北 京

内 容 简 介

本书结合具体的发酵过程实例,分别对发酵过程的建模、控制优化,特别是在线动态建模、在线状态预测以及在线过程监控与故障诊断技术和方法方面进行了比较系统的介绍,并引入了模糊逻辑推理、支持向量机、主元分析、偏最小二乘回归、核熵成分分析等新型的控制、优化、状态预测以及故障诊断等方法和技术,具有很强的理论与实践指导意义,读者能学以致用。

本书内容深入浅出,系统性强,注重理论联系实际,可作为控制科学、发酵控制工程、智能控制等专业高年级本科生、研究生和教师的参考书,也可供工业控制、生物制药等领域的工程师和科研人员阅读和参考。

图书在版编目(CIP)数据

生物发酵过程的建模、优化与故障诊断/高学金,齐咏生,王普著. —北京:科学出版社,2016

ISBN 978-7-03-048867-1

I. ①生… II. ①高…②齐…③王… III. ①微生物-发酵-过程控制-研究 IV. ①TQ920.6

中国版本图书馆 CIP 数据核字(2016)第 136624 号

责任编辑:姚庆爽 高慧元 / 责任校对:蒋 萍

责任印制:张 伟 / 封面设计:迷底书装

科学出版社 出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京教图印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 6 月第 一 版 开本:720×1000 B5

2016 年 6 月第一次印刷 印张:18

字数:358 000

定价:108.00 元

(如有印装质量问题,我社负责调换)

前 言

生物发酵产业是生物领域最具潜力的分支,其在现代食品、医药等高附加值工业中所发挥的作用越来越大。《国民经济和社会发展第十二个五年规划纲要》明确指出,把生物产业作为重点培育的战略性新兴产业之一。经过“十二五”的建设与发展,生物发酵产业正成为我国国民经济发展的主要支柱之一。其中发酵过程建模、优化控制与故障诊断技术,既关系到能否发挥菌种的最大生产能力,又影响到整个发酵过程的安全性和重复性,在整个发酵过程中是一项承上启下的关键技术。

本书将生物发酵过程的建模、优化控制与故障诊断技术融合在一起进行介绍,缘于这些技术与方法在发酵工程中变得越来越关键。另外,由于系统的复杂性越来越高,将三种技术结合在一起,可构成一个完整的控制整体,集中各自的优势进行发酵过程指导,对于许多其他工业过程也是非常有益的。目前,能够完整介绍发酵过程控制理论的相关书籍还比较少,本书的出版会为上述问题提供有效的总结和参考,读者亦可以了解相关理论和技术的新进展。

本书共八章。第1章是绪论,介绍了课题研究的背景意义、国内外的研究现状及研究思路;第2章针对生物发酵过程,介绍了生物发酵过程的概念、特点和发酵工业现状等,并概述了一般生物发酵过程的控制方式以及不同的参数对于发酵过程的影响;第3章对发酵过程的建模方法进行了详细论述,重点给出了基于数据驱动的发醇过程建模方法与应用实例;第4章在第3章基础上,介绍了发酵过程在线动态建模思想与方法;第5章对生物发酵过程的优化控制问题进行了深入的阐述和总结,并给出了3种不同类型的典型优化控制方法在发酵过程中的应用;第6章深入介绍了生物发酵过程状态预测模型思想和体系结构,基于此思想和结构给出了基于人工神经网络及PLS技术建立预测模型的方法和案例;第7、8章为发酵过程故障诊断篇,这两章重点阐述了发酵过程状态监测与故障诊断技术的发展与目前的最新方法,并给出相应的应用案例,同时也表明了该方向研究的重要性。

其中,高学金副教授撰写了第3~5章,齐咏生副教授撰写了第6~8章,王普教授撰写了第1、2章。

本书在撰写过程中得到了国家自然科学基金(60704036,61174109,61364009,21466026)、高等学校博士学科点专项科研基金(20101103110009)、北京市教委科技计划面上项目(NO.KM201210005002)、内蒙古自治区自然科学基金(2015MS0615)等的资助,得到了数字社区教育部工程研究中心、城市轨道交通北

京实验室、计算智能与智能系统北京市重点实验室的支持。在本书的撰写过程中，周东华教授、孙富春教授、易建强研究员、侯忠生教授、赵春晖教授对本书进行了审阅，并对本书提出了宝贵的意见，在此，对他们表示最诚挚的感谢。本书的出版得到了科学出版社的大力支持和帮助，在此一并致以诚挚的感谢。研究生王锡昌、张亚潮、胡永兵、薛攀娜、张海利、解亚萍等完成了文字的录入编辑工作，感谢他们的辛勤工作。

由于作者水平有限，书中难免存在不妥之处，恳请读者批评指正。

作 者

2015年9月于北京

目 录

前言

第 1 章 绪论	1
1.1 引言	1
1.2 研究背景	1
1.2.1 发酵技术的历史	1
1.2.2 生物发酵工艺	2
1.2.3 生物发酵特点	4
1.3 研究意义	5
1.4 发酵过程的建模、优化控制与故障诊断发展现状	6
1.4.1 数据驱动方法在发酵过程中的推广与应用	6
1.4.2 发酵过程数学建模方法的研究进展	7
1.4.3 发酵过程优化控制技术的研究进展	10
1.4.4 发酵过程故障诊断技术的研究进展	13
1.5 结束语	23
参考文献	23
第 2 章 生物发酵过程介绍	29
2.1 引言	29
2.2 生物发酵过程概述	29
2.2.1 生物发酵的概念和特点	29
2.2.2 发酵技术的发展	31
2.2.3 现代生物技术在发酵工业中的应用	35
2.2.4 发酵方式的分类	36
2.2.5 发酵过程的一般流程	39
2.2.6 生物发酵现状	40
2.3 生物发酵过程的参数影响及其控制	42
2.3.1 发酵过程的参数监测	43
2.3.2 温度对发酵过程的影响及其控制	44
2.3.3 pH 对发酵过程的影响及其控制	47
2.3.4 溶解氧浓度对发酵过程的影响及其控制	49

2.3.5	比生长速率对发酵过程的影响及其控制	51
2.3.6	补料方式与控制	52
2.4	结束语	53
	参考文献	54
第3章	生物发酵过程建模	55
3.1	引言	55
3.2	生物发酵过程建模问题描述	55
3.2.1	基于白箱模型的发酵过程建模	55
3.2.2	基于黑箱模型的发酵过程建模	57
3.2.3	基于灰箱模型的发酵过程建模	58
3.3	基于ANN的生物发酵过程黑箱建模	61
3.3.1	人工神经网络	61
3.3.2	人工神经网络模型	62
3.3.3	基于BP神经网络的发酵过程建模	64
3.3.4	基于RBF神经网络的发酵过程建模	70
3.4	基于支持向量机的发酵过程黑箱建模	74
3.4.1	统计学习理论	74
3.4.2	支持向量机	77
3.4.3	基于支持向量机的发酵过程建模方法	79
3.4.4	实例验证——工业青霉素生产过程的SVM预估模型	80
3.5	结束语	85
	参考文献	85
第4章	生物发酵过程在线动态建模	87
4.1	引言	87
4.2	问题描述	87
4.3	在线动态建模方法	88
4.3.1	在线全局建模	88
4.3.2	在线局部建模	92
4.4	基于DTW的发酵过程在线动态建模	94
4.4.1	动态时间规整理论	94
4.4.2	基于DTW的时间序列相似性搜索算法	96
4.4.3	基于DTW的LSSVM在线局部建模方法	99
4.4.4	基于DTW的相似样本构建方法及其改进	105
4.4.5	基于亲和度的加权欧氏距离	107

4.5	基于 SLPP 的发酵过程在线动态建模	114
4.5.1	基于 SLPP 的相似样本构建方法	114
4.5.2	案例研究	119
4.6	结束语	123
	参考文献	123
第 5 章	生物发酵过程优化控制	125
5.1	引言	125
5.2	问题描述	126
5.3	优化控制的研究内容、表述和方法	129
5.3.1	优化控制的研究内容及表述	129
5.3.2	优化控制的方法概述	130
5.4	基于 GMC 的比生长速率优化控制	132
5.4.1	通用模型控制	132
5.4.2	GMC 控制算法在比生长速率控制中的应用	133
5.4.3	扩展卡尔曼滤波器用于过程状态的估计	136
5.4.4	大肠杆菌重组蛋白生产过程应用实例	138
5.5	基于遗传算法的发酵过程最优控制	140
5.5.1	遗传算法	140
5.5.2	多目标遗传算法	142
5.5.3	遗传算法在发酵过程优化控制中的应用	143
5.5.4	应用实例	147
5.6	基于 SA-SQP 的发酵过程优化控制	150
5.6.1	序列二次规划算法	150
5.6.2	模拟退火算法	154
5.6.3	SA-SQP 混合优化算法	158
5.6.4	基于 SA-SQP 的发酵过程最优化控制	160
5.6.5	SA-SQP 混合优化算法的数值例验证	163
5.7	结束语	165
	参考文献	165
第 6 章	生物发酵过程状态预测	168
6.1	引言	168
6.2	基于神经网络的发酵过程状态预测	169
6.2.1	基于神经网络的发酵过程状态预测模型	169
6.2.2	神经网络预测器	170

6.2.3	多 BP 神经网络非线性并行预测	172
6.2.4	基于径向基神经网络的预测	172
6.2.5	预估器性能的评价	174
6.3	基于偏最小二乘回归的发酵过程状态预测	175
6.3.1	偏最小二乘	175
6.3.2	多向偏最小二乘	178
6.3.3	核偏最小二乘回归	182
6.3.4	仿真实例	186
6.4	引入效益函数的发酵过程质量监测与预报	188
6.4.1	效益函数的概念	188
6.4.2	效益函数应用举例	188
6.4.3	效益函数的预报技术	190
6.5	结束语	191
	参考文献	191
第 7 章	生物发酵过程状态监测	193
7.1	引言	193
7.2	多阶段 KPCA-PCA 监测算法在发酵过程中的应用	195
7.2.1	多阶段 KPCA-PCA 故障监测策略	196
7.2.2	仿真验证	202
7.2.3	应用实例	207
7.3	基于 GMM-DPCA 的非高斯动态过程故障监测	211
7.3.1	高斯混合模型理论	212
7.3.2	基于 GMM-DPCA 的故障监测策略	213
7.3.3	基于 GMM-DPCA 策略的离线建模和新批次监测	216
7.3.4	应用研究	217
7.4	基于核熵成分分析的发酵过程状态监测	224
7.4.1	核熵成分分析	225
7.4.2	基于多向 KECA(MKECA)的发酵过程状态监测	232
7.4.3	算法仿真验证	233
7.4.4	MKECA 监测算法在发酵过程中的应用	237
7.5	结束语	239
	参考文献	239
第 8 章	生物发酵过程故障诊断与预测	243
8.1	引言	243

8.2 工业过程故障预测技术简述	244
8.3 基于数据驱动的发酵过程故障诊断	245
8.3.1 基于贡献图的故障诊断方法	245
8.3.2 基于非线性重构的故障诊断方法	254
8.3.3 基于模糊自适应神经网络的故障诊断方法	260
8.4 基于数据驱动的发酵过程故障预测方法	270
8.4.1 基于 ARMA 模型的故障预测	271
8.4.2 基于 Fisher 特征投影差异度的故障预测	273
8.5 结束语	275
参考文献	275

第 1 章 绪 论

1.1 引 言

人类利用微生物的代谢产物作为食品和医药,已有几千年的历史了。发酵是微生物在一定培养环境中生长并形成代谢产物的过程。发酵工程是化学工程与生物技术相结合的产物,是生物技术的重要分支,是生物加工与生物制造实现产业化的核心技术之一。现代发酵工程是指利用微生物的特定性状,通过现代工程技术,在发酵罐中生产有用物质的一种技术系统,是生物技术的重要组成部分,是生物技术产业化的重要技术基础。

《国民经济和社会发展的第十二个五年规划纲要》明确指出:“把生物产业作为重点培育的战略性新兴产业之一。”生物发酵产业是生物领域最具潜力的分支,其在现代食品、医药等高附加值工业中所发挥的作用越来越大。经过“十二五”的建设与发展,生物发酵产业正受到政府的高度重视,在未来几年内生物发酵产业将成为我国国民经济发展的主要支柱之一。我国生物发酵产业是全球最大的生物发酵板块,已形成一个门类齐全的现代生产体系,谷氨酸、赖氨酸、柠檬酸、结晶葡萄糖、麦芽糖浆、葡萄糖酸钠、木糖醇、低聚异麦芽糖、黄原胶等产品的产能和单一企业的规模已达世界最大。在雄厚的技术实力背景下涉及医药、食品等关乎经济发展和人民生活水平的生物发酵产业得到了长足发展,尤其是基因工程技术在发酵产业中的应用,使得制药工业等生命科学工程的生产方式发生了重大变革^[1]。

然而,微生物发酵过程,不同于一般的工业过程,它涉及生命体的生长繁殖过程,机理十分复杂,生长环境的微小变化都可能造成显著的个体差异,而且至今仍有许多发酵过程信息,如菌体浓度、产物浓度、基质浓度等,无法在线测量,这些都给发酵过程的建模、监测、优化与控制带来了极大的困难^[2]。因此,如何更好地对工业发酵过程进行建模、在线监测、控制与优化将是本书介绍的重点。

1.2 研究背景

1.2.1 发酵技术的历史

几千年来,人类一直是依靠经验法采用微生物发酵技术来生产各种酒类、醋、

泡菜以及各种奶制品等,而对其中蕴涵的原理知之甚少。直到近代发酵技术才被逐渐认知,快速发展起来。大体说来,发酵技术的发展可以分为以下四个阶段^[3]。

第一阶段:经验发酵技术时期。此时期由于不了解发酵的本质,人们只能依靠经验对发酵过程加以干预,属于天然发酵时期。

第二阶段:纯培养发酵技术时期。19世纪末,柯赫发明了固体培养基,得到了微生物的纯培养物,由此建立了单种微生物的分离与纯培养技术。19世纪末到20世纪30年代,发酵产品主要是一些厌氧发酵和表面固体发酵产生的初级代谢产物。

第三阶段:深层培养发酵技术时期。二战爆发后,对抗感染药物的极大需求促进了青霉素制备技术的发展。由于青霉素生产是需氧发酵,易受到杂菌污染,因此建立了深层通气培养法和一整套培养技术,使培养过程中的温度、pH、通气量和营养物的供给都得到了控制。

第四阶段:定向育种发酵技术时期。20世纪60年代末,开发微生物菌体作为饲料蛋白的研究促使发酵工业取得了重大进展。同时,生产规模得到扩大,氧传递速率高的加压喷射发酵罐和加压循环发酵罐开发出来,取代了传氧效率低的机械搅拌罐。为了进一步提高经济效益,连续培养技术也被开发出来。

1.2.2 生物发酵工艺

微生物发酵过程作为复杂系统中的一类,是一个非线性、大滞后、多变量、多输入多输出、强耦合的生化反应过程^[4]。在整个发酵过程中,无论微生物生长阶段还是代谢产物生成阶段发酵过程所需的操作条件(即发酵参数)均是不同的,甚至在相同阶段的不同时刻也不相同。如果按单一给定值控制整个发酵过程,就容易导致发酵单位低而且波动性大。目前,国内外微生物发酵工厂大多采用单一给定值控制生产,远远不能获得应有的投入产出效益。随着复杂工业系统的出现以及生化工程技术的发展,应进一步重视发酵条件。发酵条件的控制是从微生物生长和产物生成动力学原理出发,区分不同的发酵阶段,寻找与菌体生长、基质消耗、产物生成三者相关联的一系列发酵参数的最优控制轨线,以此为操作条件,实现发酵过程的最优控制,获得最高的发酵产物量。

在发酵生产设备方面,目前国内一些生产厂仍然采用“第一代生物反应器”,如图1-1所示。根据混合方式可分为机械搅拌式、循环泵式和空气鼓泡式(气升式)三大类发酵设备,它们都只是一般性地控制微生物的营养条件和环境条件,生产各种有用的发酵产品。这些发酵设备在设计上都基本达到了合理而实用的要求。但是,任何设备的应用都有其局限性。多年来一直使用的标准式发酵罐(图1-2),存在能耗大、结构复杂、易污染、容积不易放大、搅拌产生的过强剪切力影响培养物固有生理特性的表现,以及在黏度较大的培养液中气液接触不良等

缺点。因此,近二十年又开发了气升式、自吸式、喷射式、筛板塔式等一些新型发酵罐,如图 1-3 和图 1-4 所示。从机械搅拌过渡到气流搅拌,更加有利于改善气液接触和节约能源。目前,发酵设备正处于基于计算机控制系统的“第二代生物反应器”全面更新时期。

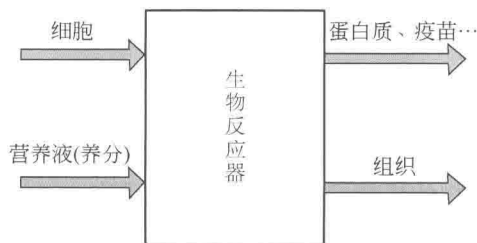


图 1-1 生物反应器示意图

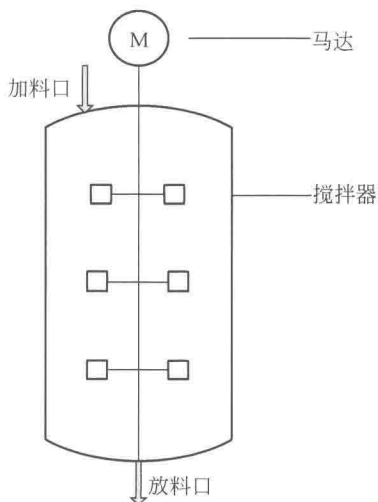


图 1-2 搅拌式生物反应器

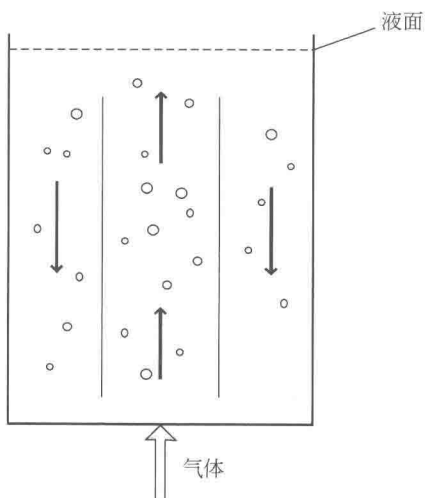


图 1-3 气升式反应器示意图

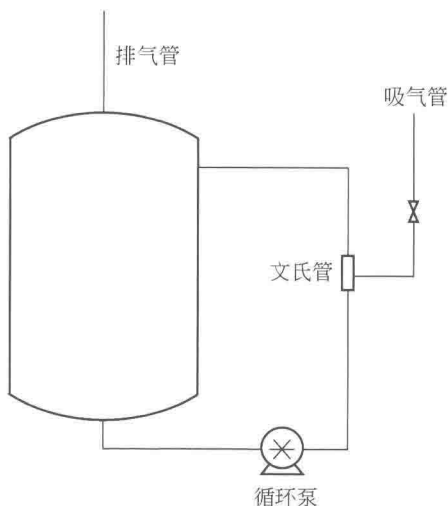


图 1-4 自吸式反应器示意图

在微生物发酵生产过程的自动控制方面,同样也经历了不同的发展阶段。起

初,技术人员凭借生产操作经验和实验室结果,手工设定发酵操作条件,这些操作条件在整个发酵过程中通常是一成不变的,这大大影响了生产效率。更为重要的是,操作条件的不准确以及控制精度差,致使出现产量不稳定、发酵单位低、能耗大、成本高等问题。随着生物工程技术的迅速发展,发酵工业的生产规模不断扩大,对自动控制的要求也越来越迫切。一些发酵工厂采用常规的控制方法,如定值控制、伺服控制和PID控制等。然而,由于微生物发酵过程不同于一般的工业过程,它是一个时变、非线性、大滞后、不确定的多变量输入输出系统,涉及生命体的生长繁殖过程,机理十分复杂,因此现有的控制方法只是稳定操作条件,远不能对控制变量进行优化控制。

今后,针对“第二代生物反应器”利用计算机实现优化控制,重点工作是研制能在线反映生物反应器内重要生物参数的传感器,以及完善计算机控制系统的硬件和软件,实现最佳控制和自适应控制,这也是作者撰写本书的初衷所在。

1.2.3 生物发酵特点

微生物发酵,在生物学上是指微生物无氧呼吸和有氧呼吸以外的一种生物氧化作用。在工业生产中,发酵是利用微生物的机能获得工业产品的泛称。通过微生物的培养,使某种特定的代谢产物得以大量积累,如各种有机酸发酵、维生素发酵、酶制剂发酵、氨基酸发酵、核酸发酵以及抗生素发酵等。

发酵工程可以认为是直接利用微生物的机能将物料加工为所需产品的过程。发酵过程的微生物学属性,使得其不同于一般的化学反应过程,其特点主要表现在如下几个方面^[5]。

(1)微生物发酵过程是一个生命过程,不可逆转,一般按照菌体生长和代谢规律可以分为四个时期^[6]:迟滞期(调整期)、对数生长期(生长旺盛期)、稳定期(平衡期)和死亡期(衰退期)。

(2)微生物以“纯种”状态在发酵罐中进行发酵,培养基和发酵罐、管道等设备都必须在接种前进行灭菌,使整个发酵过程保持无杂菌污染状态。

(3)培养基成分的微小差别会使发酵结果迥然不同,这就是发酵过程的混沌现象。

(4)发酵过程中,初级代谢和次级代谢两种不同的代谢途径交织在一起。发酵过程有生长期和生产期两个截然不同的生理阶段。次级代谢产物的形成与微生物的生长不同步,当微生物生长速度减慢或停止生长后,次级代谢产物才开始合成。

(5)对发酵过程中的相当一部分生物化学变化(即从培养基转化为菌种和产物的过程)还不太清楚,难以定量计算。此外,影响这些变化的参数(温度、pH等)相互耦联,给建模带来了很大的困难。

(6)菌体生长速率和产物生成速率与培养基所提供的条件呈非线性关系,菌体

量与产物量之间不具有明显的对应关系。发酵产物理论产量与实际产量往往相差甚远,即使同一菌种,在同一厂家,也会因生产设备、原料等差别,使菌种的生产能力大不相同。

(7)温度、pH 等环境变量常常具有非线性、大滞后等特点,难于精确控制。

(8)菌体生长阶段和产物生成阶段所需要的最优控制条件往往是不同的。

(9)菌体生长速率和产物生成速率随时间和环境条件的变化而变化。

(10)发酵过程中,温度、压力、溶解氧等物理和化学参数能在线测量,而一些重要的生物变量,如菌体浓度,至今还很难在线测量^[7,8],这也给建模带来了相当大的困难。

综上所述,微生物发酵过程是一个非线性、时变性、大滞后、阶段性、多变量输入输出、强耦合的生化反应过程。

1.3 研究意义

随着生物工程的不断发展和人民生活水平的不断提高,生物发酵呈现出广阔的发展前景。近年来,生物发酵技术和发酵装置作为生物技术产业化的基础,已经引起人们的广泛关注。然而,发酵过程本身比较复杂,因此控制难度较大;批次间的复现性很差,造成产量与质量波动较大。此外,随着产量需求的增加,环保意识的增强,对发酵过程控制的精度要求也越来越高。目前,对生物发酵过程的简单、常规控制已不能满足其快速发展的需求。因此,对发酵过程进行建模、监控及优化控制也变得越来越紧迫。

发酵过程的全方位自动化就是通过建立预报、监控与故障诊断模型,用于在线监视与优化控制,及时发现并消除异常状况,实现过程的高效、安全、稳定运行,最终达到提高产品质量一致性和企业经济效益的目的。为实现这一目标,就必须对发酵过程进行深入而细致的研究,开发出适用于工业级的生物发酵自动控制系统,为企业的生存与发展提供可靠保障。

此外,现代生物技术不仅在制药、饲料生产和新型食品生产中发挥重要作用,还能改进传统生物技术或一般化学方法难以制备特殊化学品的过程,使其既经济又清洁地生产,而这些过程都需要对发酵过程实现精确的优化与控制。以工业微生物为例,选育或构建一株优良菌株仅仅是开始,要使菌株的潜力充分发挥出来就必须实现对发酵过程的优化与控制,以此来获得较高的产物浓度,较高的底物转化率和较高的生产强度。因此,可以说,发酵过程优化与控制在解决人类面临的健康、环境、粮食、安全、人口等重大问题上也将起到积极的作用,这正是发酵过程优化控制的另一重大意义。

1.4 发酵过程的建模、优化控制与故障诊断发展现状

生物反应过程与化学反应过程不同,它不但可在常温常压下进行,而且操作和反应条件温和,对环境的污染相对较小。但是,一方面生物过程的反应速度较慢,目的产物浓度、生产强度、反应物质向目的产物的转化率较低。要解决上述问题,最终途径还是要依靠分子生物学和微生物学的方法来选育和改良优质生产菌种。另一方面,生物反应过程的环境因子,也就是通常所说的操作条件,诸如温度、压力、pH、培养基浓度等,也是影响生物过程生产水平的重要因素。利用过程控制和优化的方法,将生物过程准确地控制在最优的环境或操作条件下,是提高整体生产水平的一个捷径或者说是一种更简单易行的方法,其重要性绝不亚于菌种改良的方法。

与物理过程和化学过程不同,生物发酵过程建模往往要涉及成百上千个物理过程和化学反应。因此,与上述过程相比,生物过程有着以下迥然不同的特征:①动力学模型呈高度的非线性;②随着发酵或生物反应的进行,或随着发酵批次的不同,过程的动力学模型参数常常变化不定,呈现强烈的时变特性,对于某些生物过程甚至无法用数学模型来对动力学特性进行定量的描述;③由于测量噪声、稳定性、苛刻的操作维护条件、传感器价格的制约,除了某些简单的物理和化学状态变量,如温度、pH、压力、气体分压、溶氧浓度外,绝大多数生物状态变量是很难在线测量的;④最后,由于涉及诸多物理过程和化学反应,其相互间的作用和影响必然导致生物过程的响应速率慢、在线测量带有大幅时间滞后的特征。生物过程的上述特性,使得基于线性动力学模型的传统控制与优化理论难以适应和满足生物过程控制与优化的要求。

1.4.1 数据驱动方法在发酵过程中的推广与应用

数据驱动方法(或大数据分析技术)以采集的海量过程数据为基础,通过各种数据处理与分析方法(如多元统计方法、聚类分析、频谱分析、小波分析等)挖掘出数据中隐含的信息与特征,从而指导生产,提高产品质量和故障诊断能力,且符合流程工业的特点,通用性强。

现代流程工业过程,无论整个工厂还是单独一台大型设备,都是大系统、大数据量。21世纪是大数据的时代,以大数据为依托的数据驱动方法是建立在比较严格的统计研究基础上,能把最重要的信息捕获到较低维的空间上的方法。过程数据是空间相关的,这是因为从过程中测得的大量传感器读数以及过程变量的变化数据通常被限制在较低的维度上;同样,过程数据也是序列相关的,是因为采样间隔相对较小,标准的过程控制器不能消除惯性部件在系统的影响,如水箱、反应

器和再循环流,而多元统计方法本身就是解决变量间相关性的一种数学方法。正因为如此,数据驱动技术非常适合于大型流程工业系统的过程建模、优化控制与故障诊断。

虽然现代流程工业有大量的智能仪器仪表,可以获得大量的过程信息,但是要从观测的数据中实现对过程运行情况的评估,已超出了操作员或工程师的能力范围,发酵过程尤其如此。而数据驱动技术的优势在于它们能够将大量高维的数据转换成低维,并从海量数据中获取重要的信息,这些信息不仅能让工程师实时掌握有关过程的运行状态,提高控制的准确性和快速性,还可以利用这些数据提前预测出系统的运行情况或将要出现的故障,将有利于提高生产效率、有效排除故障,减小企业不必要的损失。除此之外,基于数据驱动的方法只需要过程数据就可以实现过程建模、优化与故障诊断。这一特点在实际应用中有着非常重要的意义,因为在某些工业过程中,可能唯一能利用的信息就是过程数据。因此,该方法很容易结合操作经验、工艺知识、历史故障记录等信息,且这些信息的正确利用往往会起到事半功倍的效果。上述特点都是基于数学模型方法所不具备的。

综上,流程工业的迅速发展,加之集散控制系统(DCS)、数据库和智能仪表的广泛使用,使得工业过程采集并记录了大量的工业数据,考虑到数据驱动方法分析大数据的优势,这对流程工业的建模、优化和故障诊断提供了一条简单有效的途径。数据驱动的方法不仅可以提高流程工业中产品的质量和产量、减少能耗、节约资源,还可以排除安全隐患、保障生命和财产安全,这对企业的兴衰、国家经济发展、社会安定都有着重要的意义^[9]。

有鉴于此,作者近年来一直从事基于数据驱动方法的发酵过程建模、优化与故障诊断研究工作。因此,在接下来的研究中,我们也重点阐述基于数据驱动和大数据分析方法的国内外最新研究成果。

1.4.2 发酵过程数学建模方法的研究进展

发酵过程最优控制需要根据过程模型对温度、pH、溶解氧浓度、补料速率等发酵参数进行寻优计算,以确定这些参数的最优控制轨线。而数学建模是对重要的过程输入(菌种、培养基、补料、环境条件等)和过程输出(生物量、产物、pH、温度、溶解氧(DO)、尾气成分等)的关联,揭示了发酵状态变量的特性,有助于改进过程的控制。输入变量和输出变量根据建立数学模型的目的来选择。

代谢产物的生成是大量细胞反应的结果,细胞活性及功能的非均一性进一步增加了过程的复杂性。限定模型的结构(或指定模型的复杂度)是影响发酵过程模型的一个极为重要的因素,一个通用的规则是:越简单越好,只要模型中包括基本的机理及模型结构即可。因此,最优控制一般采用非结构模型。非结构模型是指