

研究论文

基于机理模型的 2-KGA 混菌发酵过程 关键状态变量预报

丰强强¹, 潘玉霖¹, 成宝琨¹, 孙君伟², 袁景淇¹

(¹上海交通大学自动化系, 系统控制与信息处理教育部重点实验室, 上海 200240;

² 河北维尔康制药有限公司, 河北 石家庄 050031)

摘要: 维生素 C 生产的前体 2-KGA 是巨大芽孢杆菌和普通生酮古龙酸菌混合发酵的产物。利用先前建立的 2-KGA 混菌发酵动力学模型对背景厂 80 个批次的实测数据进行分析, 结果表明该模型能够很好地符合工业生产的实际情况。在模型参数灵敏度分析的基础上固定了部分模型参数, 并选取具有代表性的 3 个罐批 (劣等、普通、优势), 利用移动数据窗口技术和滚动参数辨识方法成功地进行了 2-KGA 浓度和底物浓度的超前 4 h 和 8 h 拟在线预报, 预报误差均在 5% 以内。同时还比较了固定长度时间窗口和变长度时间窗口的预报结果, 并根据现场实际数据特点分析了二者的优劣。

关键词: 2-KGA; 动力学模型; 移动数据窗口; 滚动参数辨识; 产物浓度预报

DOI: 10.3969/j.issn.0438-1157.2013.08.030

中图分类号: TP 273

文献标志码: A

文章编号: 0438-1157 (2013) 08-2913-05

Key variables prediction based on mechanism models of 2-KGA mixed culture fermentation

FENG Qiangqiang¹, PAN Yulin¹, CHENG Baokun¹, SUN Junwei², YUAN Jingqi¹

(¹Key Laboratory of System Control and Information Processing, Ministry of Education, Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China; ²Hebei Welcome Pharmaceutical Co., Ltd., Shijiazhuang 050031, Hebei, China)

Abstract: 2-Keto-L-gulonic acid (2-KGA), the precursor for vitamin C synthesis, is produced by the mixed culture of *Ketogulonigenium vulgare* and *Bacillus megaterium*. In this paper, the previously established kinetic model for 2-KGA mixed culture was firstly tested with the data of 80 industrial batches. Based on sensitivity analysis, it was found that some insensitive parameters might be assigned fixed values to minimize computing time. Then, the model was used to predict the most important state variables, *i. e.*, substrate and product concentrations. Moving data window technique and rolling parameter identification approach were used in the prediction process. 4 h and 8 h ahead prediction errors for 2-KGA concentration were less than 5%.

Key words: 2-keto-L-gulonic acid; dynamic model; moving data window; rolling parameter identification; prediction of product concentration

2012-12-09 收到初稿, 2013-01-30 收到修改稿。

联系人: 袁景淇。第一作者: 丰强强 (1987—), 男, 硕士研究生。

基金项目: 博士学科点专项科研基金项目 (20110073110018); 国家自然科学基金项目 (60974068)。

Received date: 2012-12-09.

Corresponding author: Prof. YUAN Jingqi, jqyuan@sjtu.edu.cn

Foundation item: supported by the Research Fund for the Doctoral Program of Higher Education (20110073110018) and the National Natural Science Foundation of China (60974068).

引 言

两步发酵法生产维生素 C 前体 2-KGA 的第二步是利用酮古龙酸菌 (*Ketogulonigenium vulgare*, 俗称小菌) 和巨大芽孢杆菌 (*Bacillus megaterium*, 俗称大菌) 混菌发酵由 L-山梨糖转化而来^[1]。混菌发酵由于不同菌种生长条件的差异和复杂的相互作用方式, 使得该过程难以预测与控制^[2]。

发酵过程关键状态变量预报可以采用数据驱动的预报方法^[3-4], 如主元回归^[5]、偏最小二乘回归^[6]、人工神经网络^[7]、关键核网络^[8]等。对 2-KGA 混菌发酵, Cui 等^[9]提出了基于支持向量机的产量预报方法, 但基于机理模型的状态预报研究报道较少^[1,10]。本文利用先前构建的动力学模型, 对实际工业罐批进行拟在线仿真, 实现了较高精度的 2-KGA 产物浓度和底物浓度超前预报。

1 2-KGA 混菌发酵的机理模型

通过前人的实验研究了解到: 2-KGA 发酵体系中小菌是产酸菌, 大菌是伴生菌^[1]; 小菌产酸是一种典型的酶促反应过程, 小菌具有 2-KGA 生产的全部酶系^[1]; 大菌并不能利用山梨糖, 仅能利用玉米浆中的碳源生长^[11]; 大菌活菌代谢物、芽胞生成和自溶释放物质都对小菌的生长有益^[12-13]。

经过深入研究, 可将大、小菌的耦合关系归纳如下: ①大菌通过群体感应^[14]感知小菌存在, 调整自身生长; ②大菌活菌及自溶释放物质弥补小菌代谢缺陷, 促进小菌生长; ③小菌释放溶菌酶^[15]促使大菌自溶以利于自身生长; ④大菌自溶释放蛋白酶类物质^[16]提高小菌的 SDH^[17]比活性。

本文不考虑发酵体系的分布特性, 根据物料守恒原理、生化反应动力学和大菌、小菌之间的耦合机理, 构造了 2-KGA 混菌发酵动力学模型 [包括: 大菌生长模型 (1), 大菌死亡自溶模型 (2), 小菌生长模型 (3), 产物生成模型 (4), 底物消耗模型 (5), 物料衡算方程 (6)] 如下

$$\frac{d(X_1 V)}{dt} = \mu_{1\max} \frac{X_2}{K_q + X_2} X_1 V - U_a \frac{X_2 X_1}{K_a + X_1} V - F_O X_1 \quad (1)$$

$$\frac{d(X_{1a} V)}{dt} = U_a \frac{X_2 X_1}{K_a + X_1} V - F_O X_{1a} \quad (2)$$

$$\frac{d(X_2 V)}{dt} = \mu_{20} \left(1 + \alpha X_1 + \beta \frac{dX_a}{dt} \right) \left(1 - \frac{X_2}{X_{2m}} \right) X_2 V - F_O X_2 \quad (3)$$

$$\frac{d(PV)}{dt} = \frac{U_P (1 + K_f X_a) X_2 S}{K_P + S} V - F_O P \quad (4)$$

$$\frac{d(SV)}{dt} = F_S S_R - m_2 X_2 V - \frac{1}{Y_{X_2/S}} \frac{d(X_2 V)}{dt} - \frac{1}{Y_{P/S}} \frac{d(PV)}{dt} - F_O S \quad (5)$$

$$\frac{dV}{dt} = F_S + F_B - F_O - rV \quad (6)$$

该模型为微分方程组, 只需确定其模型参数和初态, 即可用数值计算的方法来求解^[18-19]。

2 模型验证及参数敏感度分析

大、小菌经过若干次放大培养后接种到发酵罐中混合发酵。背景厂利用 200 m³ 气升式生物反应器补料-分批发酵生产 2-KGA。为避免底物抑制初始糖浓一般较低, 在发酵进行到 20 h 左右会进行一次补糖。

本文对从现场获得 2011~2012 年间 80 个批次的离、在线数据进行仿真, 其中 20 个批次为效益较差的罐批, 40 个为普通罐批, 另外 20 个为较好的罐批。在线数据包括: 发酵罐液位高度, 累计补料体积; 离线数据包括: 底物 (L-山梨糖) 浓度和产物 (2-KGA) 浓度, 其中底物浓度只测开始和终点的几个数据, 当残糖低于 1% 开始停罐。令蒸发系数 γ 为 0.001 L · h⁻¹, 利用以上模型对 80 个批次的数据进行参数辨识, 各参数的均值和相对方差见表 1。图 1 展示了其中一个批次的拟合情况。

14 个模型参数可分为敏感参数和非敏感参数两类: 敏感参数可设为定值, 非敏感参数必须在线辨识。如此不仅可以简化模型结构、加快计算速度, 也可以提高辨识结果的稳定性和准确性。本文

表 1 参数辨识结果

Table 1 Identified model parameters

Item	$\mu_{1\max}$ /h ⁻¹	K_q /g · L ⁻¹	U_a /h ⁻¹	K_a /g · L ⁻¹	μ_{20} h ⁻¹	α	β	U_P /L · (g · h) ⁻¹	K_P /g · L ⁻¹	X_{2m} /g · L ⁻¹	m_2 /h ⁻¹	$Y_{X_2/S}$ /g · g ⁻¹	$Y_{P/S}$ /g · g ⁻¹	K_f /L · g ⁻¹
mean	0.30	0.054	173.7	41.43	0.059	39.29	22.44	19.20	15.67	1.71	0.033	1.01	1.007	33.55
variance	0.13	0.171	0.727	0.149	0.474	0.421	0.225	0.716	0.172	0.03	0.381	0.06	0.002	0.315

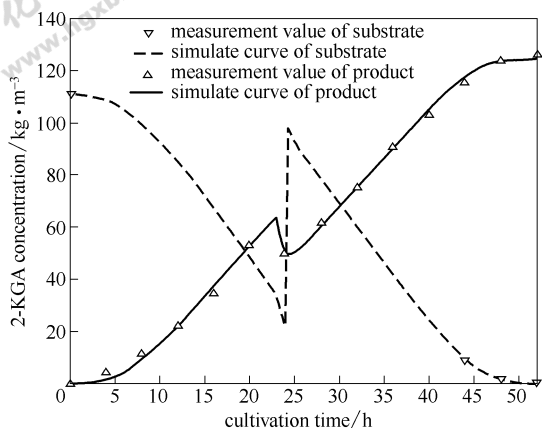


图1 模型仿真结果与测量值的比较

Fig. 1 Comparison between model simulation and off-line assays

中将 $\mu_{1\max}$ 、 U_a 、 X_{2m} 、 $Y_{X_2/S}$ 、 $Y_{P/S}$ 、 m_2 作为非敏感参数，令各自取其均值，如此辨识参数可减少近一半。

3 产物浓度预报方法

发酵过程关键变量（如产物浓度、底物浓度等）反映了过程所处状态，是进行补料、供氧等动态控制和优化操作的重要依据^[9]。但在实际生产中这些变量通常难以在线测量，离线分析也难以满足发酵过程的实时优化和控制需要。在线软测量与软预报技术是一种可行的方案。

假设 u 为发酵过程的控制输入， d 为可测扰动， y 为发酵过程可测输出变量， x 为发酵过程的状态变量集， x^* 为可能有的离线分析变量或者在线测量变量，关键状态变量 x 的在线预测就是根据可测数据得到变量 x 的最优估计 $\hat{x}$ ，即

$$\hat{x} = f(y, u, d, x^*, t) \quad (7)$$

在预估模型中还采用了移动数据窗口技术和滚动参数辨识技术^[20]。移动数据窗口技术指通过时间段“数据窗口”的移动，来跟随系统特性的变化。滚动参数辨识技术指每次移动数据窗口时都要对模型参数进行辨识，以跟随系统的时变特性。移动数据窗口技术和滚动参数辨识技术的结合可以很好地进行关键变量的预报。

实际计算机控制中应用到的过程数据都是时间的序列，数据窗口是指过程中有效的输入、输出数据序列段的时间范围。移动数据窗口技术通过输入时间窗口来获取输入数据序列，再通过输出窗口预

测未来时刻过程变量的取值，两个窗口按一定的步调沿时间轴移动，形成相应的输入输出向量对便能跟随过程对象特性随时间的变化。

由于发酵初始数据积累有限，而且实际工艺需求是在发酵的中后期要判断是否要进行补料，同时补料的时间段和多少是优化的目标对象，因此更关注中后期关键变量的预报。预报软件的流程如下：待发酵进行了 20 h 时，准备好数据窗口并确定初态，然后启动参数辨识模块。根据参数辨识结果进行模型计算，获得超前 4 h 和 8 h 的状态预报结果同时更新数据窗口初态。每 4 h 更新数据窗口，如此滚动进行直到发酵结束。

4 仿真结果

本文选取 3 个典型（劣势批次、普通批次、优势批次）批次进行拟在线产物浓度预报，这 3 个批次的批号分别为 399，489 和 504。

由于离线分析数据（间隔为 4 h），因此本文设计了两种数据窗口进行预报：一种为固定窗口长度的移动数据窗口（左侧图），一种为将既有数据尽限利用的变长度时间窗口（右侧图）。两种数据窗口对 3 个罐批的预报结果对比见图 2~图 4。

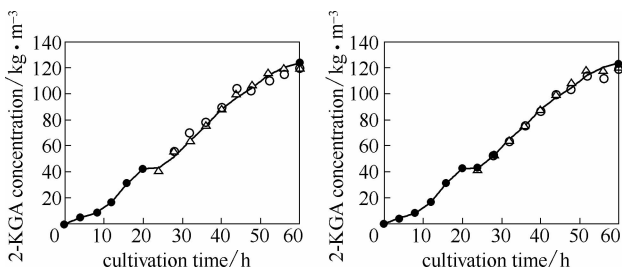


图2 第399批次（劣等）超前4 h与8 h产物浓度预报
Fig. 2 4 h and 8 h ahead prediction of product concentration for batch 399

● measurement value; △ 4 h predict value; ○ 8 h predict value

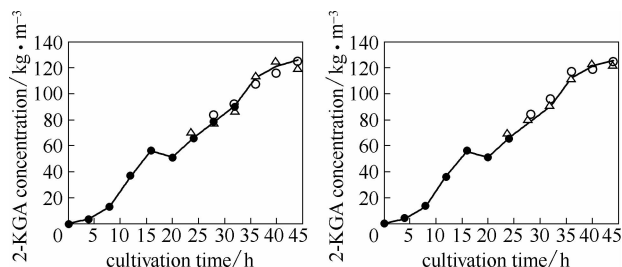


图3 第489批次（普通）超前4 h与8 h产物浓度预报
Fig. 3 4 h and 8 h ahead prediction of product concentration for batch 489

● measurement value; △ 4 h predict value; ○ 8 h predict value

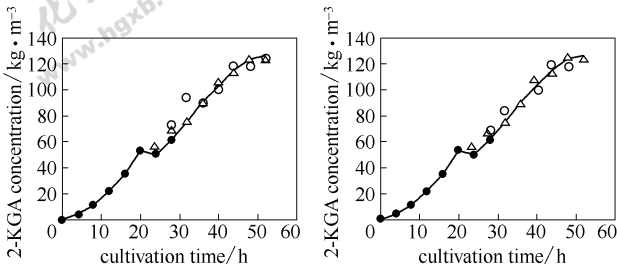


图 4 第 504 批次（优势）超前 4 h 与 8 h 产物浓度预报
Fig. 4 4 h and 8 h ahead prediction of product concentration for batch 504

● measurement value; △ 4 h predict value; ○ 8 h predict value

为比较两种方案的性能，可定义预报误差 (RMSE) 如下

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^q e(T_k + T_p)^2}{q}} \quad (8)$$

$$e(T_k + T_p) = \frac{P_{\text{real}}(T_k + T_p) - P_{\text{model}}(T_k + T_p)}{P_{\text{real}}(T_k + T_p)} \quad (9)$$

式中 $P_{\text{real}}(T_k + T_p)$ 为 $(T_k + T_p)$ 时刻的实际产物浓度值， $P_{\text{model}}(T_k + T_p)$ 为模型预报值。

对两种方案的预报结果进行误差分析，结果见表 2。

表 2 两种方案的预报误差

Table 2 Comparisons of two forecast scenarios/%				
Batch	Fixed data window		Variable data window	
	4 h ahead	8 h ahead	4 h ahead	8 h ahead
batch 399	2.13	4.06	1.67	1.83
batch 489	2.44	3.29	1.79	2.89
batch 504	3.34	4.83	1.61	1.78
av. err.	2.64	4.06	1.69	2.17

从两种方案对 3 个典型批次的预报结果看，变长数据窗口的性能要优于定长数据窗口。

5 结 论

上述仿真结果说明了 2-KGA 混菌发酵动力学模型的可靠性和可应用性。两种不同数据窗口下的预报结果表明变长数据窗口的预报性能更好。虽然定长数据窗口能更好地反映过程的时变特性，但由于一些关键变量并不实时测量，而且测量间隔过长，20 h 内的实测数据有限，从而导致其性能反而不如后者。变长数据窗口将已知数据都囊括进来，能够反映发酵开始以来的总体参数水平，其预报值则体现了这一水平的延续趋势。

文献 [9] 采用支持向量机方法对 2-KGA 浓

度进行的提前 8 h 预报精度可达 1.71%~2.85%。本文基于机理模型的预报精度与之相当，但预报方法结构简单，在遇到工艺调整时（如：补料时间、补料量和浓度等）仍可以获得较好的预报结果。产物浓度预报可以用于异常批次的早发现以及指导补料策略（补料时机的选取、补料量以及补料方式等），从而达到提高产率、降低单耗的目的。

符 号 说 明

- F_B ——碱液流加速率， $\text{L} \cdot \text{h}^{-1}$
- F_O ——采样放料速率， $\text{L} \cdot \text{h}^{-1}$
- F_S ——L-山梨糖流加速率， $\text{L} \cdot \text{h}^{-1}$
- K_a ——大菌自溶的米氏常数， $\text{g} \cdot \text{L}^{-1}$
- K_f ——辅酶对 SDH 的促进系数， $\text{L} \cdot \text{g}^{-1}$
- K_P ——产物形成的米氏常数， $\text{g} \cdot \text{L}^{-1}$
- K_q ——群体感应阈值浓度， $\text{g} \cdot \text{L}^{-1}$
- m_2 ——小菌的维持系数， h^{-1}
- P ——产物浓度， $\text{g} \cdot \text{L}^{-1}$
- S ——发酵液中底物浓度， $\text{g} \cdot \text{L}^{-1}$
- S_R ——补料料液中底物浓度， $\text{g} \cdot \text{L}^{-1}$
- U_a ——大菌自溶速度系数， h^{-1}
- U_P ——产物形成速度系数， $\text{L} \cdot (\text{g} \cdot \text{h})^{-1}$
- V ——发酵液体积， L
- X_a ——大菌自溶累积浓度， $\text{g} \cdot \text{L}^{-1}$
- X_1 ——大菌浓度， $\text{g} \cdot \text{L}^{-1}$
- X_2 ——小菌浓度， $\text{g} \cdot \text{L}^{-1}$
- X_{2m} ——小菌的最大浓度， $\text{g} \cdot \text{L}^{-1}$
- $Y_{P/S}$ ——产物对基质的得率系数， $\text{g} \cdot \text{g}^{-1}$
- $Y_{X_2/S}$ ——小菌对基质的得率系数， $\text{g} \cdot \text{g}^{-1}$
- α, β ——大菌及其自溶物对小菌的促进作用
- γ ——发酵液蒸发系数， $\text{L} \cdot (\text{L} \cdot \text{h})^{-1}$
- $\mu_{1\max}$ ——大菌最大比生长速率， h^{-1}
- μ_{20} ——大菌单独培养的比生长速率， h^{-1}

References

[1] Wei Dongzhi (魏东芝), Yuan Weikang (袁渭康), Yin Guanglin (尹光琳), Yuan Zhongyi (袁中一), Chen Minheng (陈敏恒). Study on kinetic model of vitamin C two-step fermentation process [J]. *Chinese Journal of Biotechnology* (生物工程学报), 1992, **8**: 277-282

[2] Xu Yingxuan (徐颖宣), Xu Erni (徐尔尼), Feng Naixian (冯乃宪), Xing Feng (幸峰). Research progress on application of microbial mixed fermentation [J]. *China Brewing* (中国酿造), 2008, **4**: 1-4

[3] De A, Adilson J, Rubens M. Soft sensors development for on-line bioreactor state estimation [J]. *Computers and Chemical Engineering*, 2000, **24** (2): 1099-1103

[4] Kadlec P, Gabrys B, Strandt S. Data-driven soft sensors in

- the process industry [J]. *Computers and Chemical Engineering*, 2009, **33**: 759-814
- [5] Zhao Y H. A soft sensor based on orthogonal nonlinear principal component analysis [J]. *Developments in Chemical Engineering and Mineral Processing*, 2005, **13**: 233-242
- [6] Zamprogna E, Barolo M, Seborg D E. Estimating product composition profiles in batch distillation via partial least squares regression [J]. *Control Engineering Practice*, 2004, **12** (7): 917-929
- [7] Su H B, Fan L T, Schlup J R. Monitoring the process of curing of epoxy/graphite fiber composites with a recurrent neural network as a soft sensor [J]. *Engineering Applications of Artificial Intelligence*, 1998, **11** (2): 293-306
- [8] Liu Yi (刘毅), Wang Haiqing (王海清), Li Ping (李平). Key kernel network and its application to online modeling for fermentation process [J]. *Journal of Chemical Industry and Engineering (China)* (化工学报), 2008, **59** (5): 1194-1199
- [9] Cui Lei, Xie Ping, Sun Junwei, Yu Tong, Yuan Jingqi. Data-driven prediction of the product formation in industrial 2-keto-L-gulonic acid fermentation [J]. *Computers and Chemical Engineering*, 2012, **36**: 386-391
- [10] Li Qiang (李强), Cao Zhu'an (曹竹安). Kinetics study of producing 2-keto-L-gulonic acid in mixed fermentation (II): Analyze fermentation process by kinetic model// Proceedings of the 9th National Conference of Biochemical Engineering [C]. Beijing: Chemical Industry Press, 2000: 167-171
- [11] Chen Kejie (陈克杰), Zhou Jingwen (周景文), Liu Liming (刘立明), Liu Jie (刘杰), Du Guocheng (堵国成), Chen Jian (陈坚). Enhancing 2-keto-L-gulonic acid production under hyperosmotic stress by adding sucrose [J]. *Chinese Journal of Biotechnology* (生物工程学报), 2010, **26**: 1507-1513
- [12] Lü Shuxia (吕淑霞), Feng Shu (冯树), Zhang Zhongze (张忠泽), Liu Yang (刘阳), Xie Zhanwu (谢占武), An Haiying (安海英). The effect of *Bacillus megaterium* in vitamin C two-step fermentation [J]. *Microbiology China* (微生物学通报), 2001, **5**: 10-13
- [13] Gamer M, Stammen S. *Bacillus megaterium*—an alternative expression system [J]. *Journal of Biotechnology*, 2007, **131**: 220
- [14] Waters C M, Bassler B L. Quorum sensing: cell to cell communication in bacteria [J]. *Annual Review of Cell and Developmental Biology*, 2005, **21**: 319-346
- [15] Zhang Jing, Liu Jie, Shi Zhongping, Liu Liming, Chen Jian. Manipulation of *B. megaterium* growth for efficient 2-KLG production by *K. vulgare* [J]. *Process Biochemistry*, 2010, **45**: 602-606
- [16] Lu Shujuan, Wang Jun, Yao Jianming, Yu Zengliang. Study on the effect of mutated *Bacillus megaterium* in two-stage fermentation of vitamin C [J]. *Plasma Science and Technology*, 2003, **5**: 2011-2016
- [17] Hoshino T, Sugisawa T, Tazoe M, Shinjoh M, Fujiwara A. Metabolic pathway for 2-keto-L-gulonic acid formation in *Gluconobacter melanogenus* IFO3293 [J]. *Agricultural and Biological Chemistry*, 1990, **54**: 1211-1218
- [18] Zhou Yuping (周玉平), Zhou Jie (周杰). A review on models of microorganism continuous fermentation and its application [J]. *Microbiology China* (微生物学通报), 2010, **37** (2): 269-273
- [19] Andres M B, Parag G, Elnashaie SSEH. Non-linear characteristics of a membrane fermentor for ethanol production and their implications [J]. *Nonlinear Analysis*, 2006, **7** (3): 432-457
- [20] Yuan Jingqi, Vanrolleghem P A. Rolling learning-prediction of product formation in bioprocesses [J]. *Journal of Biotechnology*, 1999, **69**: 47-62