

增广KALMAN滤波器在酵母菌发酵中的应用

施源 李友荣 陈敏恒 袁渭康

(华东化工学院, 上海)

在微生物发酵系统中存在着各种随机干扰,从而使实验测定值带有不可避免的误差。为消除与估算菌体浓度有关的三个实验测定变量;糖消耗速率(SUR),氧消耗速率(OUR),二氧化碳释放速率(CPR)中干扰误差成分,应用增广Kalman滤波方法于菌体浓度的估算,有效地抑制了在菌体生长动力学的递推运算中造成的偏差积累,使估算结果能相应正确地模拟实际真值,从而较好地完善发酵实验数据处理方法。在运算过程中同时还得到菌体收率系数和维持能耗系数随发酵的变化规律。

关键词 增广Kalman滤波器;菌体浓度的估算;菌体收率和维持能耗计算;酵母菌发酵

微生物培养对于它们生长环境的变化是非常敏感的。因此大多数的发酵过程都建立了较严格的操作条件。尽管在发酵实验中对系统的温度、搅拌转速、发酵pH以及通气流量都加以控制,但在酵母菌发酵操作中^[1],培养基基础液和糖蜜水解和灭菌处理的差异,菌种贮存和菌种本身性质的差异,糖蜜的阶跃流加,起沫和消沫的波动,取样对系统的扰动以及在线测定排气中氧浓度和二氧化碳浓度的仪器对信号响应延迟、漂移等不但造成微生物在生长和代谢时的变化波动,也造成在过程的监测和控制方面出现了难以预料和描述的波动。为此我们把这些波动归结为发酵系统受到随机干扰和噪声的侵袭。图1显示

了酵母菌发酵操作中溶解氧含量(DO%)、排气中氧浓度(O₂%)和二氧化碳浓度(CO₂%)随动态发酵过程的变化情况。

由于菌体生长的改变相对于氧或二氧化碳的变化要来的缓慢,所以在氧和二氧化碳的波动信号中既体现了菌体的生长也夹带了其他环境的和操作条件引起的干扰信息和一些不甚了解的随机噪声。它们总是既存在于过程中,又存在于测量分析中。用这些实验测定变量值来关联菌体生长和代谢的动力学方程^[1]

$$\frac{dn_1}{dt} = -4.0 \left(3.0 + \frac{dn_1}{dt} + \frac{dn_2}{dt} + 1.5 \frac{dn_7}{dt} \right) \cdot \frac{1}{12.703 - 0.074t} \quad (1)$$

$$SUR = - \frac{dn_1}{dt} \quad (2)$$

$$OUR = - \frac{dn_2}{dt} \quad (3)$$

$$CPR = \frac{dn_7}{dt} \quad (4)$$

势必会引起估算菌体浓度的方差随时间不

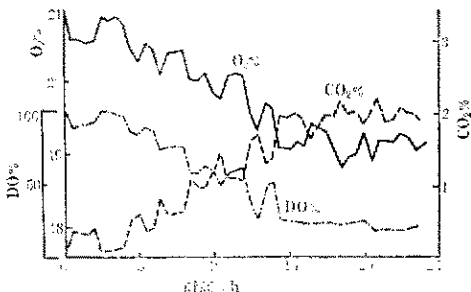


图1 酵母菌发酵的代谢曲线

Fig.1 Metabolism curves of yeast fermentation

本文于1985年12月14日收到,

断增加, 以至使估算结果偏离真值。分析 积累型干扰误差, 如果方程 (1) 为:

$$\frac{dn_4}{dt} = f[n_1(t), n_2(t), n_7(t), t] + W_4(t) \quad (5)$$

假设 $W_4(t)$ 为简化了的零均值干扰噪声, 是: 方差为 σ_4^2 。那么 $n_4(t)$ 的期望和方差分别

$$E[n_4(t)] = n_4(t_0) + \int_{t_0}^t f[n_1(\tau), n_2(\tau), n_7(\tau), \tau] d\tau \quad (6)$$

$$\text{Var}[n_4(t)] = \text{Var}[n_4(t_0)] + \sigma_4^2 t \quad (7)$$

除非测量值不带任何干扰, 否则随时间的延续, 方差总会增加。为了有效地处理实验数据, 必须以控制方差为目标。另外, 由于发酵的速率表达方程数和变量维数的一致^[1], 为求得菌体浓度的变化过程, 就得输入实验观察的变量数据, 从而导致估算过程为递推的形式。因此, 采用动态增广 Kalman 滤波器用以估算噪声^[2,3], 分析 Q 和 R 矩阵, 最终得到噪声近似最小条件下的菌体浓度随发酵时间的变化规律以及其他一些发酵参数。

理论分析

(一) 增广 Kalman 滤波器

考察微生物发酵过程, 其特点为周期都比较长。发酵状态量的变化相对时间是缓慢的响应, 采样频率低, 发酵中各种随机干扰的出现不可预料, 但可以认为它们服从某种统计规律。

假设干扰是随机出现, 相互之间相关性较弱, 随机的干扰有的对过程起促进作用, 有的则相反。在整个发酵周期内用统计概念归纳上述特点和假设, 认为发酵中各种干扰都偏向于零均值的平稳随机过程, 称此为白噪声序列。

离散的非线性动态系统可表示为:

$$X(k) = \Psi[X(k-1) | k-1] + W(k-1) \quad (8)$$

通过观察得到和状态向量 $X(k)$ 有关的输

出为:

$$Z(k) = h[X(k) | k] + V(k) \quad (9)$$

离散时间的白噪声干扰序列相关函数为:

$$\text{Cov}[X(k)X(j)] = c\delta(k|j) \quad (10)$$

$$\delta(k|j) = \begin{cases} 1 & k=j \\ 0 & k \neq j \end{cases} \quad (11)$$

向量白噪声干扰序列相关函数为:

$$\begin{aligned} \text{Cov}[U(k)U^T(j)] &= E[U(k)U^T(j)] \\ &= S(k)\delta(k|j) \end{aligned} \quad (12)$$

此, 白噪声序干扰列对不同时刻 ($k > 0$) 变量是不相关的。方程 (8)、(9) 中的干扰项 $W(k-1)$, $V(k)$ 有下列关系:

$$E[W(k)] = 0 \quad (13)$$

$$\text{Cov}[W(k)W(j)] = Q(k)\delta(k|j) \quad (14)$$

$$E[V(k)] = 0 \quad (15)$$

$$\text{Cov}[V(k)V(j)] = R(k)\delta(k|j) \quad (16)$$

$$\text{Cov}[W(k)V(k)] = 0 \quad (17)$$

系统初态统计特征为:

$$E[X(0)] = \bar{X}(0) \quad (18)$$

$$\text{Var}[X(0)] = E\{[X(0) - \bar{X}(0)][X(0) - \bar{X}(0)]^T\} = P(0) \quad (19)$$

$$\text{Cov}[X(0)W(k)] = 0 \quad (20)$$

$$\text{Cov}[X(0)V(k)] = 0 \quad (21)$$

根据以上方程 (13)~(21) 的条件可得到离散时间动态增广 Kalman 滤波器递推公式^[4]:

$$\hat{X}(0) = \bar{X}(0) \quad (22)$$

$$P(0) = \text{Var}[X(0)] \quad (23)$$

步预报方程

$$\hat{X}(k|k-1) = \Psi[\hat{X}(k-1) | k-1] \quad (24)$$

离散滤波Riccati方程

$$P(k|k-1) = \phi(k|k-1)P(k-1) \cdot \phi^T(k|k-1) + Q(k-1) \quad (25)$$

滤波协方差方程

$$P(k) = [I - L(k)H(k)] \cdot P(k|k-1) \quad (28)$$

滤波增益方程

$$L(k) = P(k|k-1)H^T(k)[H(k) \cdot P(k|k-1)H^T(k) + R(k)]^{-1} \quad (26)$$

$k = 1, 2, 3, \dots$

(二) 发酵状态方程

分析酵母发酵过程, 其状态变量可定为糖消耗、氧消耗、二氧化碳释放和菌的生长代谢^[1]。根据发酵中氧的消耗机理和二氧化碳释放机理, 导出下列状态方程组:

滤波估算方程

$$\hat{X}(k) = \hat{X}(k|k-1) + L(k)[Z(k) - H(k)\hat{X}(k|k-1)] \quad (27)$$

$$-\frac{dn_1}{dt} = C_{IN}F_{IN} - \frac{\Delta S}{\Delta t} + W_1(t)$$

$$-\frac{dn_2}{dt} = Y_{O_2} \frac{dn_4}{dt} + m_{O_2}n_4 + W_2(t)$$

$$\frac{dn_7}{dt} = Y_{CO_2} \frac{dn_4}{dt} + (m_{CO_2} + Y_{ETH})n_4 + W_3(t)$$

$$\frac{dn_4}{dt} = -\frac{4.0}{12.703 - 0.074t} \left(3.0 \frac{dn_1}{dt} + \frac{dn_2}{dt} + 1.5 \frac{dn_7}{dt} \right) + W_4(t)$$

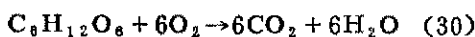
$$\frac{d}{dt}(Y_{O_2}) = 0 + W_5(t)$$

$$\frac{d}{dt}(m_{O_2}) = 0 + W_6(t)$$

$$\frac{d}{dt}(Y_{CO_2}) = 0 + W_7(t)$$

$$\frac{d}{dt}(m_{CO_2} + Y_{ETH}) = 0 + W_8(t) \quad (29)$$

方程组(29)中 Y_{O_2} 和 Y_{CO_2} 分别为单位菌体呼吸时消耗的氧量和释放的二氧化碳量, 一般可认为它们和时间没有很强的联系^[5]。 m_{O_2} 为菌体生长和代谢维持能耗。指菌体对于细胞养分活性输送过程中的能量消耗以及维持所有结构组分和蛋白质运动方面的能量需要。对应于



把1克分子葡萄糖通过氧化, 全部转化为 CO_2 和 H_2O 的同时得到38克分子ATP高能

物质, 用以推动生物合成过程中不可避免的能量消耗。可认为:

$$m_{O_2} = m_{CO_2} \quad (31)$$

由于酵母发酵过程中还有分解代谢产物乙醇生成, 它的形成同样要释放出二氧化碳, 用 Y_{ETH} 表示。在正常发酵时, 由于 Y_{O_2} 、 Y_{CO_2} 、 m_{O_2} 、 Y_{ETH} 的变化缓慢, 近似作为时不变参数, 得它们的微分为零。但实际过程它们受随机干扰, 认为它们的波动属白噪声序列, 在定态中叠加动态影

响，作为动态参数。实验观察方程组为：

$$- 1)T + W(k - 1) \quad (35)$$

$$Z_1(t) = \int_{t_0}^t SUR dt + V_1(t)$$

$$Z(k) = H \tilde{X}(k) + V(k) \quad (36)$$

其中

$$Z_2(t) = \int_{t_0}^t OUR dt + V_2(t)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (37)$$

$$Z_3(t) = \int_{t_0}^t CPR dt + V_3(t) \quad (32)$$

(三) 修正状态转移矩阵

把方程组 (29) 和 (32) 写成向量形式：

$$\dot{X}(t) = f[X(t)] + W(t) \quad (33)$$

$$Z(t) = \tilde{X}(t) + V(t) \quad (34)$$

线性、离散化处理

$$X(k) = X(k - 1) + f[X(k$$

$$\Phi(k|k - 1) = I + J(k - 1)T + [\dot{J}(k - 1) + J^2(k - 1)]T^2/2 + [\ddot{J}(k - 1) + 2\dot{J}(k - 1)J(k - 1) + J(k - 1)\dot{J}(k - 1) + J^3(k - 1)]T^3/6 \quad (38)$$

用修正和未加修正的状态转移矩阵模拟发酵中流加糖蜜过程，得到图 2 所示的曲

线。从中可见，采用修正后的状态转移矩阵计算结果更接近实际过程。

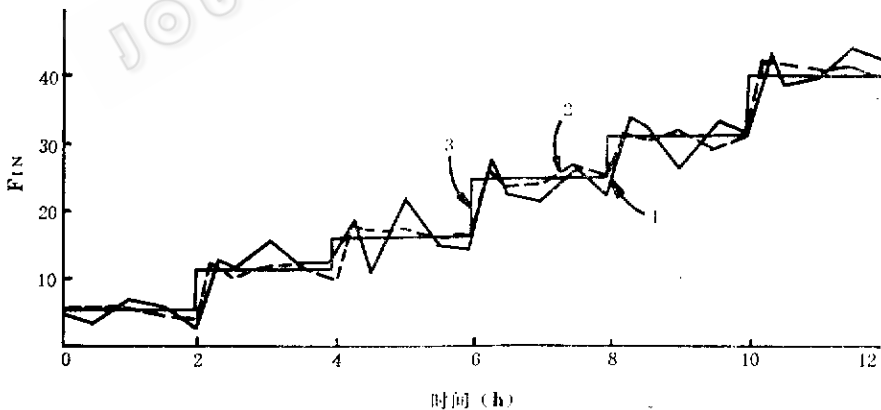


图 2 糖蜜流速实际值和估算值之间的比较

Fig.2 Comparison of estimated values with measured values of various molasses feeding rates

- 1. 采用未修正的状态转移矩阵所估算结果
Estimated values of feeding rates using unmodified state transfer matrices
- 2. 采用方程 (38) 后估算结果
Estimated values of feeding rates using modified state transfer matrices
- 3. 实际糖蜜流速
Measured values of molasses feeding rates

(四) 噪声矩阵Q和R

从方程 (14) 和 (16) 中得到:

$$R(k) = \begin{pmatrix} \sigma_1^2(k) & 0 & 0 \\ 0 & \sigma_2^2(k) & 0 \\ 0 & 0 & \sigma_3^2(k) \end{pmatrix} \quad (39)$$

$$R \equiv R(k) = \begin{pmatrix} \sigma_1^2(k) & 0 & 0 \\ 0 & \sigma_2^2(k) & 0 \\ 0 & 0 & \sigma_3^2(k) \end{pmatrix} = \begin{pmatrix} 0.074 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0.25 \end{pmatrix} \quad (41)$$

$$Q(k-1) = IB(k-1) \quad (42)$$

$$B(k-1) = [\sigma_{11}^2(k-1), \sigma_{22}^2(k-1), \dots, \sigma_{88}^2(k-1)]^T \quad (43)$$

$$\sigma_{ij}^2(k-1) = E[W_{ij}(k-1)W_{ij}^T(k-1)] \quad (44)$$

$j, ij = 1, 2, \dots, 8$

Q 矩阵不仅包括信号的影响, 而且当模型化简时还包括了模型误差的影响。要确切地描述 Q 矩阵每一主对角线上方差的变化, 就必须掌握过程中各种干扰噪声的统计特性。但是对于动态系统, 这难以做到。

考察 Q 矩阵中 σ_{ij}^2 , 它表示环境干扰

$$\sigma_{ij}^2(k) = \begin{cases} 0 & \lambda_i \leq 0, \lambda_i \geq 0.4 \\ \lambda_i = \{ [Z_j(k-1) - \hat{X}_j(k-1)]^2 - P_{ij}^2(k-1) - \sigma_{ij}^2 \} \theta_i & 0 < \lambda_i < 0.4 \end{cases} \quad (45)$$

$j, ij = 1, 2, 3$

$$0.87 < \sigma_i \leq 0.91 \quad (46)$$

θ_i 作为计算调节参数:

$$\sigma_{44}^2(k) = \begin{cases} 0 & \lambda_4 \leq 0, \lambda_4 \geq 0.4 \\ \lambda_4 = \{ [\hat{X}_4(k-1) - \hat{X}_4(k-1|k-2)]^2 - P_{44}^2(k-1) - 0.04 \} 0.93 & 0 < \lambda_4 < 0.4 \end{cases} \quad (47)$$

对于状态参数, 近似为:

$$\sigma_{ij}^2(k) = P_{ij}^2(k-1) \quad (48)$$

$jj = 5, 6, 7, 8$

结果和讨论

图 3 表示用方程 (1) 和用增广 Kalman 滤波方法得到的计算结果与实验测定值之间的对照。可以看到随递推计算的逐步延续, 积累误差使得前者偏离实测结果。

$$\sigma_i^2(k) = E[V_i(k)V_i^T(k)] \quad (40)$$

$i = 1, 2, 3$

由于实验测得的三个独立变量中 SUR 为手工测定, OUR 和 CPR 分别由仪器分析测定, 现给出它们的最大可能偏差, 并简化为:

或模型噪声的强弱程度。如果定性地认为测量值带有观察过程的不确定性, 而滤波估算值中带有系统在预报过程中的不确定性和系统滤波值偏离真值的程度。采用滤波协方差矩阵 P 以及 R 来近似描述上面的分析, 通过计算, 在一定限制范围内, 整理得出表达关系:

对于菌体浓度的变化

分析图 4, 当乙醇含量增高时, 菌体的收率就有所下降。这表明有一部分碳源用于副产物的生成, 菌体生长受到阻遏, 收率逐步下降低于 0.5。随着发酵的进行, 不利于菌体生长的有害物质积累。另外, 在不改变通气条件的情况下, 随着菌体浓度的提高, 氧的需求量增加, 氧传递过程逐步成为控制步骤, 相应呼吸减慢 (Y_{O_2} 下降), 乙醇代谢提高 (Y_{CO_2} 上升)。菌量增加使维持能耗相应增加 (m_{O_2} 上升), 用于菌生长部分的碳能源的利用下降

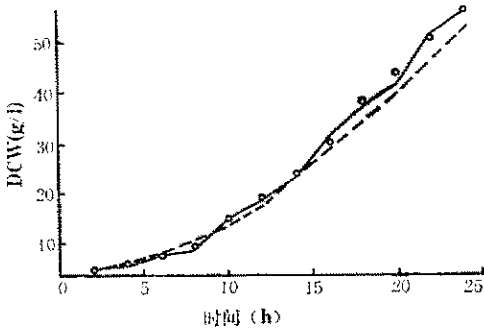


图 3 菌体浓度估算值和实验值的比较

Fig.3 Comparison of estimated cell concentration with measured cell concentration

—经增广Kalman滤波方法估算的DCW值
 DCW estimates using extended Kalman filter
 ...未经增广Kalman滤波方法估算的DCW值
 DCW estimates without using extended Kalman filter
 •实验测定值 DCW experimental data

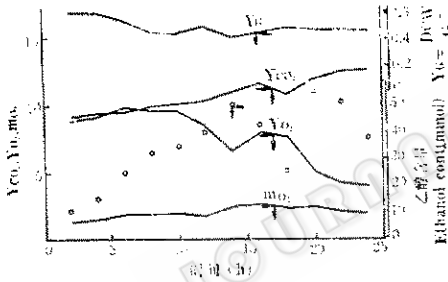


图 4 收率系数、维持系数以及乙醇含量随发酵过程的变化

Fig.4 Change of yields, maintenance and ethanol content in fermentation

(Y_G 下降)。发酵后期菌体不再适应于此生长环境，其活力下降，比生长率下跌， Y_{O_2} 急剧下降。以上分析证明，随发酵操作过程进行的整体计算很好地模拟了菌体在发酵进行中的外在表现和内部的变化，这和描述菌体生长的一般理论是相应的⁶¹。

经增广 Kalman 滤波方法计算菌体浓度和实验测定值的平均相对误差为 3.39%，满足实验要求。不经增广 Kalman 滤波方法使得计算菌体浓度和实验测定值的平均相对误差达 7.45%，是前者的 2.2

倍，这也证明发酵操作中确实不可低估各种干扰的影响。

为检验菌体生长动力学模型的精度和检验增广 Kalman 滤波器计算方法的适宜范围，并在实验室试验装置¹¹寻找较好的工艺条件，于众多的监测、控制参数中选择 pH、温度、搅拌转速、通气量以及流加糖方案为五个独立实验因素（不考察它们的交互作用），组成 $L_{16}(4^5)$ 正交实验。以平均菌体收率为优化目标。通过全部实验数据的计算表明，菌体生长动力学模型和增广 Kalman 滤波器计算方法能适应各种操作条件。所有实验结果和计算值之间的平均相对误差都小于 5.0%，可认为此处理方法能提供较宽的范围，在去除提供的信息中不确定成分后，使菌体生长动力学模型能真实地反映出酵母菌在发酵过程中的生长和代谢规律。

经正交试验统计分析，搅拌转速对实验影响最大，随着菌量增加，对溶解氧的摄取增加。因此，搅拌转速通过氧传递的改变影响菌体正常的代谢活动。从实验室小发酵罐装置中得正交试验较优工艺条件为 rpm = 800, pH = 5.0, T = 33℃, 通气量 1.5L/L/min, 总流加糖蜜量 650ml/2.2L (发酵液), 可使收率达 $Y_G = 0.48-0.50$ 。

符号说明

- $B(k-1)$ 状态方程随机干扰的方差向量
- c 常数
- C_{IN} 流加糖蜜中葡萄糖浓度 mol/L
- CO_V 协方差符号
- CPR 二氧化碳释放速率 mmol/h
- E 期望符号
- $f(X(t))$ or $f(X(k-1))$ 向量 $X(t)$ or $X(k-1)$ 的函数关系
- F_{IN} 糖蜜流加速率 ml/h
- $H(k)$ 系统在时刻 $t(k)$ 的 $m \times n = 3 \times 8$ 维观察矩阵

$h(X(k) k)$	系统在时刻 $t(k)$ 的 $m=3$ 维非线性向量值函数	$X(0)$	状态向量初值的期望值
I	$n \times n = 8 \times 8$ 单位矩阵	$\hat{X}(k)$	滤波估算状态向量值
$J(X(k-1))$	向量 $X(k-1)$ 的雅可比矩阵	$\tilde{X}(k)$	观察向量的数值积分值
k	计算的递推次数	$Z(t)$ or $Z(k)$	时刻 $t(k)$ 获得的 $m=3$ 维观察向量
m_{CO_2}	维持系数 $mmol CO_2/DCW/h/L$	Y_{CO_2}	收率系数 $mmol CO_2/DCW/L$
m_{O_2}	维持系数 $mmol O_2/DCW/h/L$	Y_{O_2}	收率系数 $mmol O_2/DCW/L$
n_i	发酵基质、产物的量 $mmol i=1,2,4,7$	希腊文符号	
OUR	氧耗速率 $mmol/h$	$\sigma_i^2(k)$	第 i 个观察值的随机干扰方差 $i=1,2,3$
$P(0)$	计算的初值方差	$\sigma_{ij}^2(k-1)$	在时刻 $t(k-1)$ 第 j 个状态方程随机干扰的方差
$Q(k)$	系统在 $t(k)$ 时的误差协方差矩阵, 非负定	$j, ij=1,2,\dots,8$	
$R(k)$	系统在 $t(k)$ 时的观察误差协方差矩阵, 正定	$\Psi(X(k-1) k-1)$	系统在时刻 $t(k-1)$ 的 $n=8$ 维非线性向量值函数
$S(k)$	正定矩阵	$\phi(k k-1)$	系统在时刻 $t(k-1)$ 到时刻 $t(k)$ 的 $n \times n = 8 \times 8$ 维可逆状态转移矩阵
SUR	糖耗速率 $mmol/h$	λ_i	状态方程随机干扰的方差运算符 $i=1,2,3$
T	计算的递推步长	θ_i	调节参数 $i=1,2,3$
$U(k)$	向量白噪声干扰序列	$\delta(k j)$	Dirac- δ 函数
$V(t)$ or $V(k)$	相应于时刻 $t(k)$ 的 $m=3$ 维随机向量		
$W(t)$ or $W(k)$	相应于时刻 $t(k)$ 的 $n=8$ 维随机向量		
Var	方差符号		
$X(t)$ or $X(k)$	系统在时刻 $t(k)$ 的 $n=8$ 维状态向量		

参 考 文 献

- (1) 施源等: 生物工程学报, 2(4):61-66, 1986.
- (2) Jazwinski, A.H.: "Stochastic Processes and Filtering Theory", Academic Press., New York, 1970.
- (3) Stephanopoulos, G.: "State Estimation for Computer Control of Biochemical Reactors", Proc. Vth Int. Ferm. Symp., London, Canada, 1:399, 1981.
- (4) 刘豹: 现代控制理论, 机械工业出版社, 1983.
- (5) Humphrey, A.E.: Process Biochemistry, 12:19, 1977.
- (6) 李友荣: "发酵生理学" (上册), 华东化工学院, 1984

APPLICATION OF EXTENDED KALMAN FILTER IN YEAST FERMENTATION PROCESS

Shi Yuan Li Yourong Chen Minheng Yuan Weikang
(*East China Institute of Chemical Technology, Shanghai*)

Due to existence of white noise, measurement error are inevitable in fermentation systems. Extended Kalman filtering theory is employed to eliminate noise in measurement of the fermentation state variables, such as substrate uptake rate, oxygen uptake rate, carbon dioxide produced rate. This paper proposes a concept for restricting error accumulation using recurrence calculation of cell growth model, leading to an accurate, noise-free estimation of the cell concentration.

Key words

Extended Kalman filter; estimate of biomass concentration; calculation of cellular yield and maintenance energy coefficients; yeast fermentation.