

# 基于神经网络的滤波器<sup>1)</sup>

胡泽新

(华东理工大学自动化所 上海 200237)

## 摘 要

利用函数连接型网络理论,提出了一种新的基于神经网络的非线性滤波器(NNNF),证明了NNNF的无偏性和最小方差性.将NNNF用于谷氨酸和红霉素发酵过程的状态估计,结果表明,NNNF滤波估计值与实验结果吻合得相当好,对噪声特性无特殊要求,对初始状态估值具有一定的鲁棒性,NNNF可利用有限的状态量测信息在线推算其它不可测量的状态变化,为非线性生化过程的在线优化奠定了基础.

**关键词:** 非线性随机系统,状态估计,神经网络,生化过程.

## 1 引言

在生化过程中,由于缺乏可靠的传感器用于菌体浓度、基质浓度和产物浓度的在线测量,难以获取关键发酵变量的数据和实现优化控制.因此,有必要研究复杂系统状态估计的有效方法.近年来,“软传感器”的研究受到了极大重视<sup>[1]</sup>.目前,估计算法可以分为两大类:增广卡尔曼滤波器算法(EKF)<sup>[2]</sup>和专门用于生化反应器状态估计的特殊算法<sup>[3]</sup>.特殊算法尽管对具体问题效果较好,但不具备通用性,故不能推广应用于任意的非线性系统.EKF技术虽已成功地解决了一些生化过程的非线性滤波问题,但是在其估计值上尚存在偏差,这个偏差与状态方程及测量方程的非线性程度有关,当非线性程度较严重、模型不准确或不合适时,就会出现较大的估计偏差,甚至导致EKF发散.

本文提出了一种新的神经网络非线性滤波器(NNNF),将其用于生化反应器的状态估计,仿真结果证实了NNNF的有效性.

## 2 自适应神经元及学习策略

一个基本的自适应神经元由 $q$ 个输入量 $I_i(t)(i=1, 2, \dots, q)$ 和一个输出量 $O(t)$ 组成,其输入输出关系为

$$O(t) = F \left( \sum_{i=1}^q W_i(t) I_i(t) \right), \quad (1)$$

1) 本文得到国家自然科学基金资助.  
本文于1993年11月23日收到

式中  $W_i(t)$  是相应于第  $i$  个输入  $I_i(t)$  的加权值, 由某种学习算法确定,  $F(\cdot)$  为神经元传递函数.

通常假定在学习期间, 神经元的关联加权值  $W_i(t)$  正比于学习信号  $r_i(t)$ , 同时它又慢衰减, 则神经元的学习规则可表示为<sup>[4]</sup>

$$W_i(t+1) = W_i(t) + d \cdot r_i(t), \quad (2)$$

或

$$\frac{dW_i}{dt} = d \cdot r_i(t), \quad (3)$$

式中  $d$  为大于零的学习步长.  $W_i(t)$  的监督学习策略为

$$r_i(t) = Z(t) \cdot I_i(t). \quad (4)$$

表示对一个动态特性未知的环境, 自适应神经元在教师信号  $Z(t)$  的指导下进行强迫学习, 从而对外界作出反映和作用.

对于上述递推式学习, 运用随机逼近理论, 可以证明, 当学习速率  $d$  充分小时, 权值  $W_i(t)$  将收敛到某一稳态值  $W_i^*$ , 使  $W_i^*$  与期望值的偏差达到允许的范围.

### 3 基于神经网络的滤波器 (NNNF)

#### 3.1 随机非线性系统的描述

$$\mathbf{X}(t) = \mathbf{f}(\mathbf{X}(t-1)) + \boldsymbol{\xi}(t-1), \quad (5)$$

$$\mathbf{y}(t) = \mathbf{H}(t)\mathbf{X}(t) + \mathbf{v}(t), \quad (6)$$

式中  $\mathbf{X}(t)$  表示  $n$  维状态向量,  $\mathbf{f}$  为  $n$  维非线性向量函数,  $\mathbf{y}(t)$  表示  $m$  维检测向量,  $\mathbf{H}$  为  $m \times n$  维矩阵,  $\boldsymbol{\xi}(t)$  为  $n$  维动态噪声序列,  $\mathbf{v}(t)$  为  $m$  维观测噪声. 通常  $\boldsymbol{\xi}(t)$  与  $\mathbf{v}(t)$  均是互不相关的零均值正态白噪声序列, 即对所有  $t, j$ , 有如下的噪声统计特性:

$$E\xi_i = 0, \text{cov}(\xi_i, \xi_j) = E\xi_i \xi_j^T = Q_i \delta_{ij}, \quad (7)$$

$$E v_i = 0, \text{cov}(v_i, v_j) = E v_i v_j^T = R_i \delta_{ij}, \quad (8)$$

$$\text{cov}(\xi_i, v_j) = E \xi_i v_j^T = 0. \quad (9)$$

当然,  $\boldsymbol{\xi}(t)$  和  $\mathbf{v}(t)$  亦可是均值不为零的有色噪声.

#### 3.2 非线性滤波器

根据新息原理, 有如下的非线性滤波器:

$$\hat{\mathbf{X}}(t) = \hat{\mathbf{X}}_{t/(t-1)} + g(t)[y(t) - H(t)\hat{\mathbf{X}}_{t/(t-1)}], \quad (10)$$

$$\hat{\mathbf{X}}_{t/(t-1)} = \mathbf{f}(\hat{\mathbf{X}}(t-1)), \quad (11)$$

式中  $\wedge$  符号代表估计值,  $g(t)$  为滤波器增益矩阵,  $g(t) \in R^{n \times m}$ .

非线性滤波器的设计, 就是使某一性能指标函数最小求取增益矩阵  $g(t)$ , 并保证状态估计值  $\hat{\mathbf{X}}(t)$  最终收敛到其真值  $\mathbf{X}(t)$ . 将式 (10) 看作是一个函数连接型网络<sup>[5]</sup>, 根据前面的自适应神经元监督学习策略, 提出基于神经网络滤波器的增益矩阵求取方法如下:

$$g_{ij}(t+1) = g_{ij}(t) + \eta_i E \{ X_i(t) - \hat{X}_i(t) \} \left[ y_j(t) - \sum_{k=1}^n h_{jk} f_k(\hat{\mathbf{X}}(t-1)) \right], \quad (12)$$

或

$$\frac{dg_{ij}}{dt} = \eta_i E \{ X_i(t) - \hat{X}_i(t) \} \left[ y_j(t) - \sum_{k=1}^n h_{jk} f_k(\hat{X}(t-1)) \right], \quad (13)$$

其中  $\eta_i$  为学习步长。

$$g(t) = \begin{bmatrix} g_{11}(t) & g_{12}(t) \cdots g_{1m}(t) \\ \vdots \\ g_{n1}(t) & g_{n2}(t) \cdots g_{nm}(t) \end{bmatrix}, \quad H(t) = \begin{bmatrix} h_{11}(t) & h_{12}(t) \cdots h_{1n}(t) \\ \vdots \\ h_{m1}(t) & h_{m2}(t) \cdots h_{mn}(t) \end{bmatrix}.$$

### 3.3 无偏性和最小方差性

**定义 1.** 神经网络滤波器的估计误差定义为

$$\tilde{X}(t) = X(t) - \hat{X}(t). \quad (14)$$

**定义 2.** 神经网络滤波器增益矩阵学习误差定义为

$$\begin{aligned} J &= \frac{1}{2} E \{ (\tilde{X}(t) - E\{\tilde{X}(t)\})^T (\tilde{X}(t) - E\{\tilde{X}(t)\}) \} = \frac{1}{2} E \{ \tilde{X}(t)^T \tilde{X}(t) \} \\ &= \frac{1}{2} E \left\{ \sum_{i=1}^n (X_i(t) - \hat{X}_i(t))^2 \right\}. \end{aligned} \quad (15)$$

**定理 1.** 对于式 (10)–(13) 构成的神经网络非线性滤波器, 有

1) 无偏性  $E\{\tilde{X}(t)\} = 0$ .

2) 最小方差性即有  $J \geq 0$ ,  $\frac{dJ}{dt} \leq 0$ , 当且仅当  $\frac{dg_{ij}}{dt} = 0$  时, 才有  $\frac{dJ}{dt} = 0$ .

证明. 根据 Lebesgue 控制收敛定理<sup>[5]</sup>, 即

$$\frac{\partial}{\partial t} E\{f(t)\} = E \left\{ \frac{\partial f}{\partial t} \right\}, \text{ if } \left| \frac{\partial f}{\partial t} \right| \leq M. \quad (16)$$

其中  $E\{\}$  为取均值运算,  $M$  为一有界的常数. 由式 (16), 有

$$\frac{dJ}{dt} = -E \left\{ \sum_{i=1}^n (X_i(t) - \hat{X}_i(t)) \sum_{j=1}^m \frac{\partial \hat{X}_j(t)}{\partial g_{ij}(t)} \cdot \frac{dg_{ij}}{dt} \right\}, \quad (17)$$

将式 (9), (10) 代入式 (17), 有

$$\frac{dJ}{dt} = -E \{ X_i(t) - \hat{X}_i(t) \} \sum_{i=1}^n \left( y_j(t) - \sum_{k=1}^n h_{jk} f_k(\hat{X}(t-1)) \right) \frac{dg_{ij}}{dt}, \quad (18)$$

1) 最小方差性

由式 (13), (18), 可推得

$$\frac{dJ}{dt} = - \sum_{i=1}^n \sum_{j=1}^m \frac{1}{\eta_i} \left( \frac{dg_{ij}}{dt} \right)^2. \quad (19)$$

由于学习步长  $\eta_i > 0$ , 故有  $\frac{dJ}{dt} \leq 0$ , 当且仅当  $\frac{dg_{ij}}{dt} = 0$  时, 才有  $\frac{dJ}{dt} = 0$ .

2) 无偏性

由最小方差性可知, 当  $t \rightarrow \infty$  时, 有

$$\frac{dg_{ij}}{dt} = 0, \quad (20)$$



即

$$\eta_i E\{\tilde{X}_i(t)\} \left[ y_i(t) - \sum_{k=1}^n h_{ik} f_k(\hat{X}(t-1)) \right] = 0, \quad (21)$$

由式(21)可推得

$$E\{\tilde{X}_i(t)\} = 0. \quad (22)$$

## 4 在生化反应器状态估计中的应用

### 4.1 谷氨酸发酵过程的状态估计

仿真用谷氨酸发酵过程动力学模型为

$$\frac{dX_1}{dt} = \mu_m X_1 \left( 1 - \frac{X_1}{X_m} \right) + \xi_1, \quad (23)$$

$$\frac{dX_2}{dt} = b X_1 \frac{X_3}{K_s + X_3} + \xi_2, \quad (24)$$

$$\frac{dX_3}{dt} = -\frac{1}{Y_G} \mu_m X_1 \left( 1 - \frac{X_1}{X_m} \right) - \frac{1}{Y_P} b X_1 \frac{X_3}{K_s + X_3} - m X_1 + \xi_3, \quad (25)$$

式中  $X_1(t), X_2(t), X_3(t)$  分别为菌体浓度, 谷氨酸浓度和葡萄糖浓度 (g/l),  $\xi(t) = [\xi_1(t), \xi_2(t), \xi_3(t)]^T$  是动态系统随机噪声向量,  $\mu_m$  为菌体最大比生长速率 ( $0.767\text{h}^{-1}$ ),  $X_m$  为菌体的最大合成量 ( $6.43\text{g/l}$ ),  $b$  为最大产酸比速率 ( $0.358\text{h}^{-1}$ ),  $K_s$  对葡萄糖的饱和常数 ( $12.04\text{g/l}$ ),  $Y_G$  为葡萄糖的菌体得率系数 ( $0.436$ ),  $Y_P$  为葡萄糖的谷氨酸得率系数 ( $0.645$ ),  $m$  是菌体维持系数 ( $0.105\text{h}^{-1}$ ).

量测方程为

$$\mathbf{y}(t) = H\mathbf{X}(t) + \mathbf{v}(t), \quad (26)$$

式中  $\mathbf{y}(t) = [y_1(t), y_2(t), y_3(t)]^T$  为检测向量,  $\mathbf{v} = [v_1, v_2, v_3]^T$  是零均值正态白噪声向量,  $H = I_{3 \times 3}$  是 3 阶单位检测系数矩阵.

在谷氨酸发酸实验中, 对菌体、谷氨酸和葡萄糖三个状态变量进行了离线测量, 结果如图 1—3 所示, 采样间隔  $T_s = 5\text{min}$ , 发酵周期  $T_F = 26\text{h}$ , 滤波器增益变量初值为  $g_{1j}(0) = 0.1$ ,  $g_{2j}(0) = 10^{-2}$ ,  $g_{3j}(0) = 10^{-2} (j = 1, 2, 3)$ , 学习步长  $\eta = 10^{-4}$ . 实验过

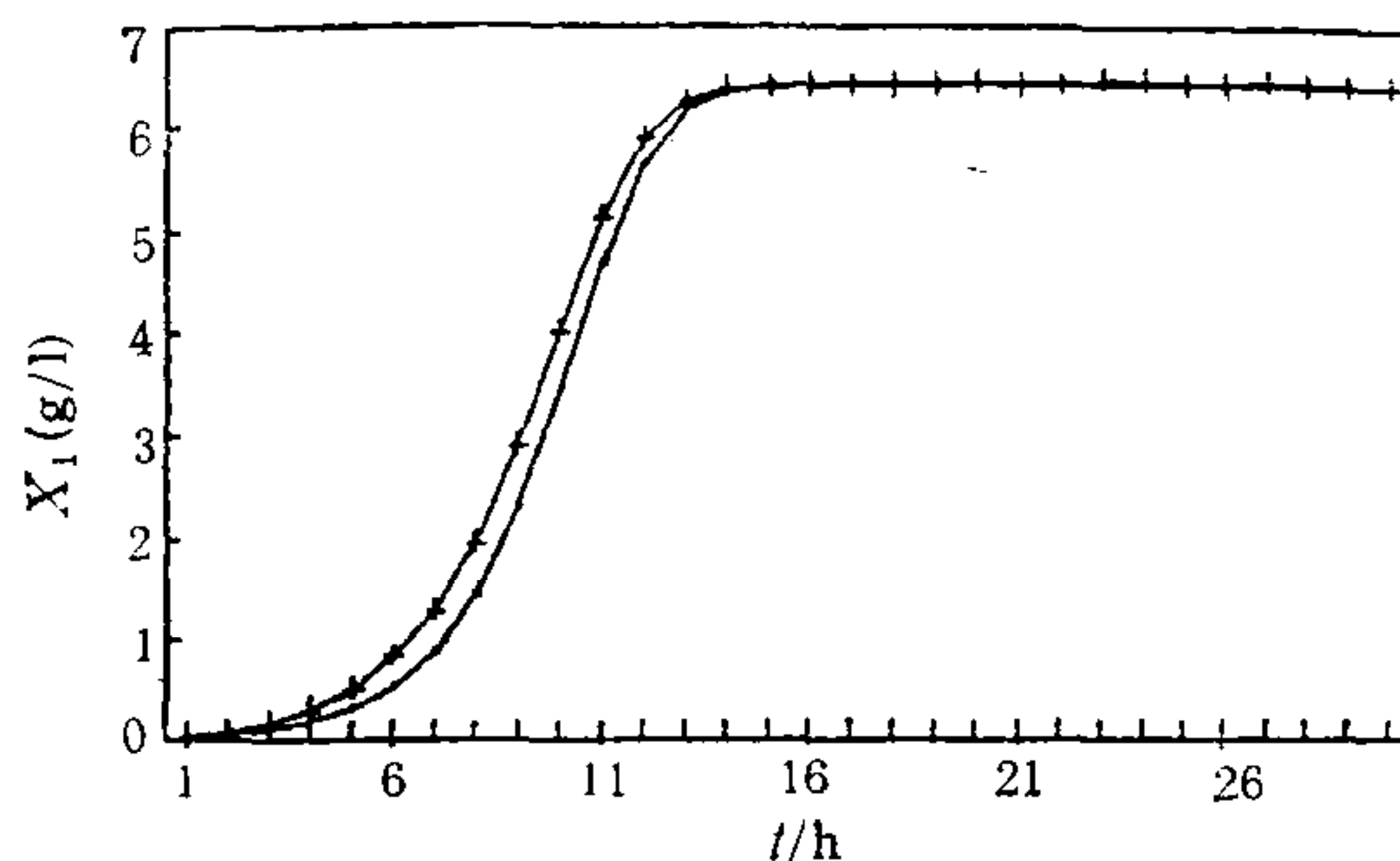


图 1 菌体浓度滤波值与实测值比较  
+ 滤波值    • 实测值

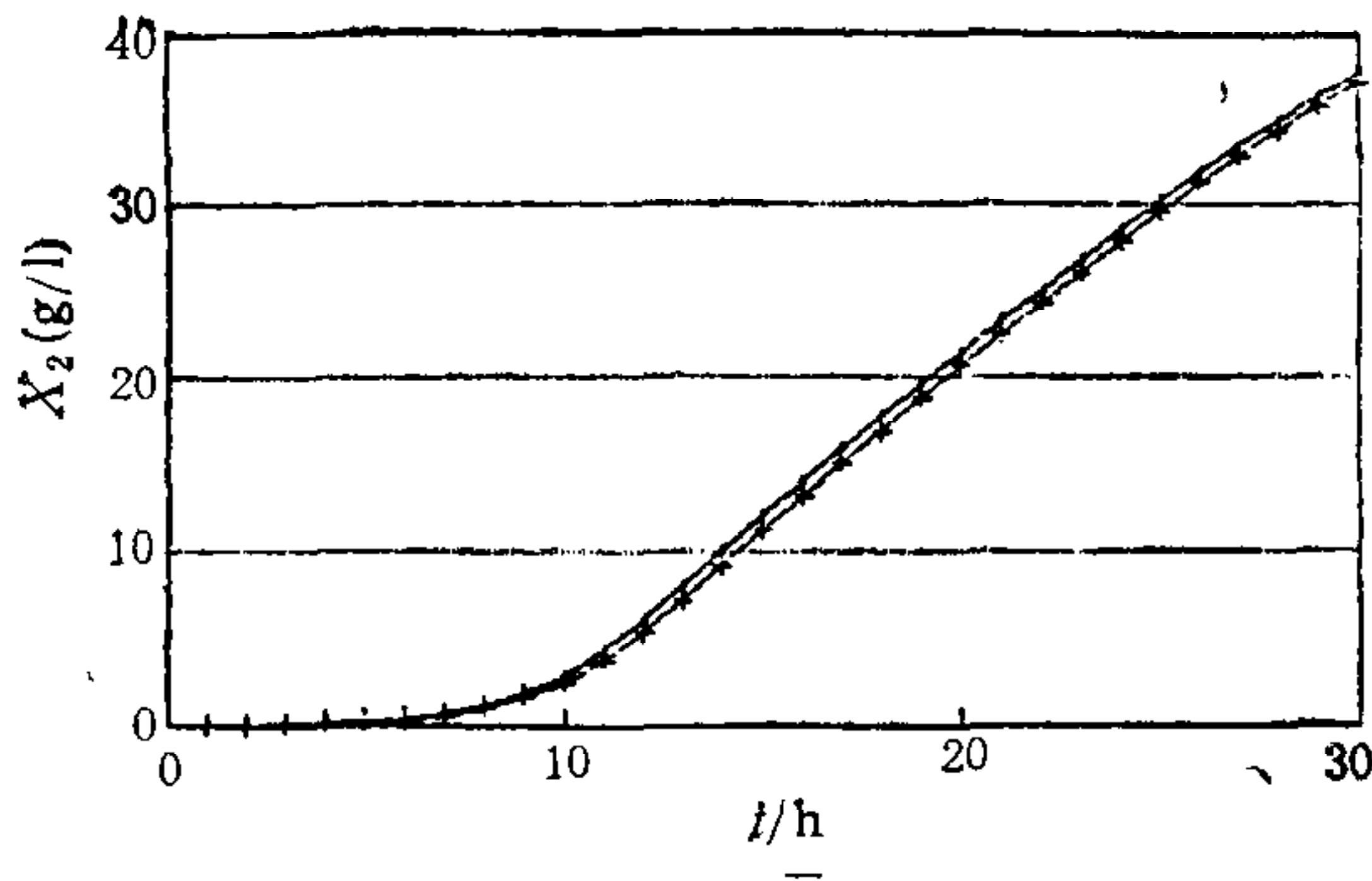


图2 谷氨酸浓度滤波值与实测值比较  
+滤波值 ●实测值

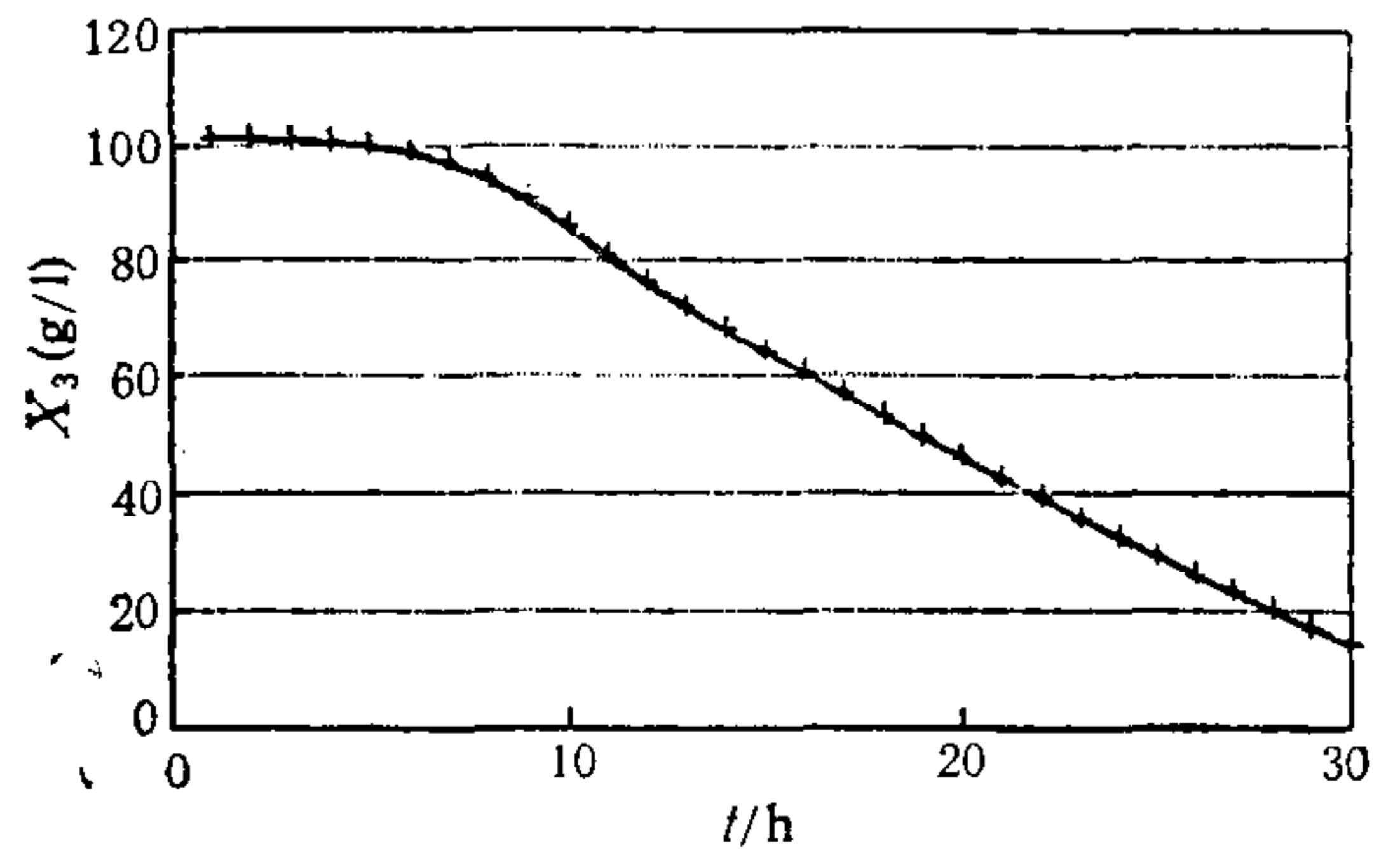


图3 葡萄糖浓度滤波值与实测值比较  
+滤波值 ●实测值

程中的动态噪声方差和观测噪声方差分别为

$$Q = \begin{bmatrix} 0.01 & & \\ & 0.003375 & \\ & & 0.04 \end{bmatrix}, \quad R = \begin{bmatrix} 0.04 & & \\ & 0.0225 & \\ & & 0.04 \end{bmatrix}.$$

在保证与实验过程中的随机噪声相一致的前提下,进行仿真研究.菌体、谷氨酸和葡萄糖的初始状态估计值以发酵初始时刻的实测值代入,即

$$\hat{X}_1(0) = 0.031 \text{ (g/l)}, \hat{X}_2(0) = 0 \text{ (g/l)}, \hat{X}_3(0) = 101.2 \text{ (g/l)}.$$

NNNF 产生的三个状态估计值如图 1—3 所示. NNNF 估计值与实测值呈现了较好的一致性.

## 4.2 红霉素生物合成过程的状态估计

描述微生物链丝菌生长和红霉素生物合成过程的数学模型为

$$\frac{dX_1}{dt} = \mu(\theta, \text{pH})X_1 \left[ 1 - \left( \frac{X_1}{X_{1f}(\theta, \text{pH})} \right)^a \right] + \xi_1, \quad (27)$$

$$\frac{dX_2}{dt} = b_1(\theta, \text{pH})X_1 - b_2(\theta, \text{pH})X_2 + \xi_2, \quad (28)$$

$$y = HX + v, \quad (29)$$

其中  $X_1$  和  $X_2$  分别为菌体浓度和红霉素浓度,  $\xi = [\xi_1, \xi_2]^T$  为系统噪声,  $v = [v_1, v_2]^T$  为检测噪声,  $H = I_{2 \times 2}$  为二阶单位观测矩阵.  $\theta$  发酵介质温度的约束范围为  $28^\circ\text{C} \leq \theta \leq 40^\circ\text{C}$ ; pH 发酵介质的 pH 值的约束范围为  $6.4 \leq \text{pH} \leq 7.2$ ;  $\mu$  为生长速率;  $X_{1f}$  为微生物最终浓度;  $a, b_1$  和  $b_2$  均为模型参数. 其中

$$\mu = 0.075 - 0.065\text{pH} + 0.05\text{pH}^2 + 0.0519\text{pH} \cdot \theta, \quad (30)$$

$$X_{1f} = 9.26 + 26.85\theta - 4.19\text{pH}, \quad (31)$$

$$b_1 = 0.08 - 0.15\theta - 0.02\theta^2 - 0.008\text{pH}^2 + 0.03\text{pH} \cdot \theta, \quad (32)$$

$$b_2 = 0.013 - 0.025\theta - 0.04\theta^2 - 0.01\text{pH}^2 + 0.054\text{pH} \cdot \theta, \quad (33)$$

$$a = 2.0, \quad (34)$$

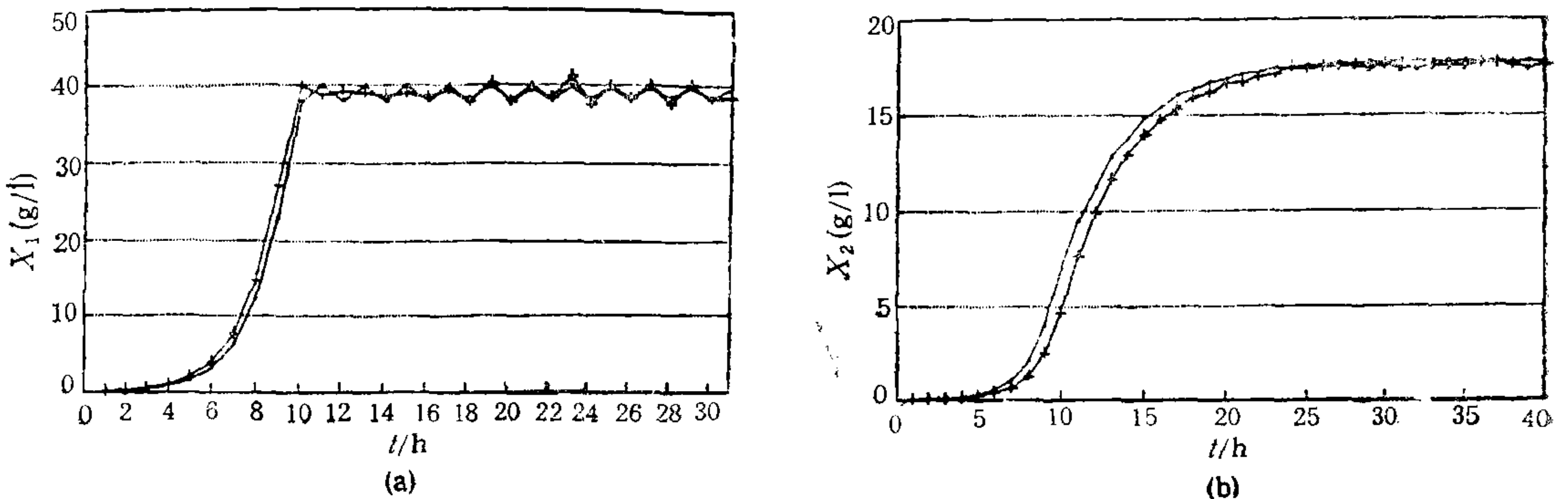
式中  $\theta, \text{pH}$  均作无因次化处理,取  $\theta = 37^\circ\text{C}$ ,  $\text{pH} = 7.2$ , 无因次化后  $\theta = 0.75$ ,  $\text{pH} = 1$ , 代入式 (30)—(33) 中, 计算得  $\mu = 0.09893$ ,  $X_{1f} = 39.02$ ,  $b_1 = 0.010125$ ,  $b_2 =$

0.0225.

1) 对过程施加较弱的随机干扰

$$Q = \text{diag}\{0.04, 0.01\} \quad R = \text{diag}\{9, 4\}.$$

初始状态估值  $\hat{X}(0) = [0.1, 0]^T$ , 采样周期  $T_s = 10\text{min}$ , 滤波器增益初始值  $g_{1j}(0) = g_{2j}(0) = 10^{-2} (j = 1, 2)$ , 学习步长  $\eta = 10^{-4}$ . 仿真结果如图 4 所示. NNNF 估计值与实测值吻合得很好.



(a) 菌体浓度

(b) 红霉素浓度

图 4 弱随机干扰时滤波值与实测值比较

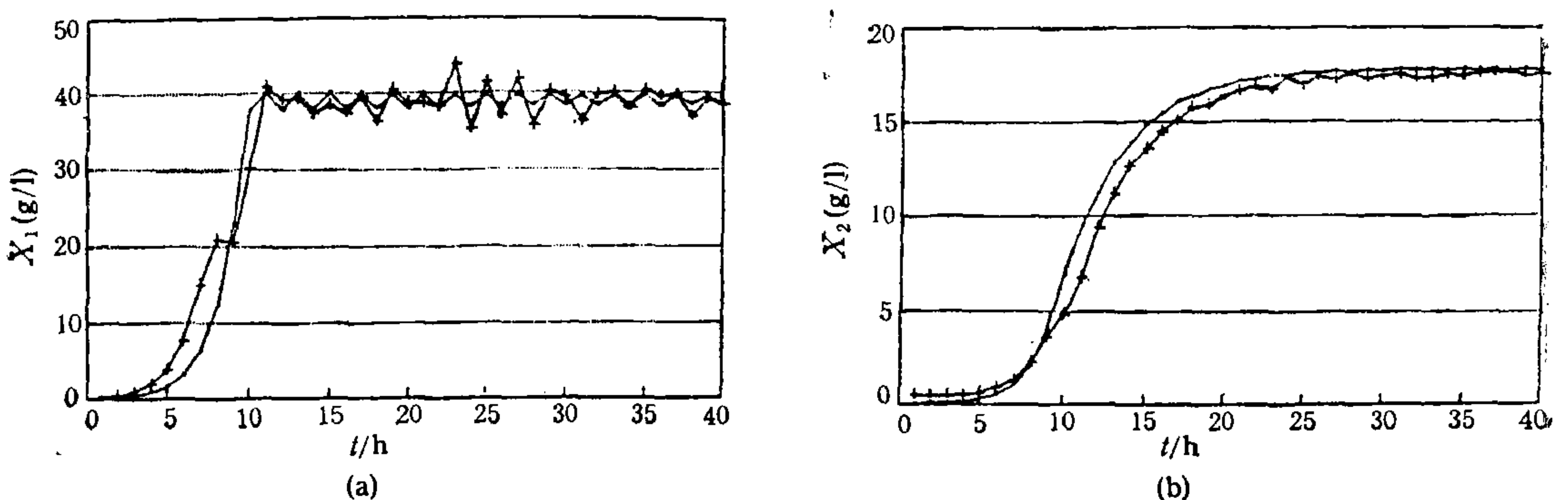
+ 滤波值 ● 实测值

2) 加大随机噪声强度,同时只有部分状态可量测,另外还改变初始状态估计值,来考察 NNNF 的滤波能力.

所施加的动态系统随机噪声幅度超出过程正常响应值的 10% 左右, 检测噪声幅度超出过程正常响应值的 15%.  $Q = \text{diag}\{9, 2.25\}$ ,  $R = 25$ , 初始状态估值由原来的  $\hat{X}(0) = [0.1, 0]^T$  变为  $\hat{X}(0) = [0.2, 0.5]^T$ , 采样周期  $T_s = 10\text{min}$ , 滤波器增益  $g_{1j}(0) = g_{2j}(0) = 10^{-2} (j = 1, 2)$ , 学习步长  $\eta = 10^{-4}$ , 这时量测方程为

$$y_1(t) = X_1(t) + v_1(t). \tag{35}$$

仿真结果如图 5 所示,可以看出,尽管只对  $X_1$  分量作了离线测量,且处在大的随机干



(a) 菌体浓度

(b) 红霉素浓度

图 5 强随机干扰时滤波值与实测值比较

+ 滤波值 ● 实测值



扰下,同时改变初始状态估值,但经过 NNNF 滤波后,依然能够得到与系统真实状态拟合良好的状态估计结果,尤其对  $X_2$  分量颇有实用价值。因为在没有对  $X_2$  作量的前提,通过滤波却能真实地反映出  $X_2$  分量的实际响应情况,NNNF 确实起到了在线观测器的作用。

### 参 考 文 献

- [1] McAvoy T J. contemplative stance for chemical process control—an IFAC report. *Automatica*. 1992, 28(2): 441—442.
- [2] Bastin G and Dochain D. On-line estimation and adaptive control of bioreactors. Amsterdam, Netherlands: Elsevier Science Publishers, 1990.
- [3] Stephanopoulos G N and San K Y. Studies On On-line Bioreactor Identification, I, Theory. *Biotechnology and Bioengineering*. 1984, 26: 1176—1188.
- [4] Hunt K J. Neural networks for control systems—a survey. *Automatica*. 1992, 28(6): 1083—1112.
- [5] Torchinsky A. Real Variable. New York: Addison-Wesley Publishing Company, Inc. 1987. 120.

## FILTER BASED ON NEURAL NETWORKS

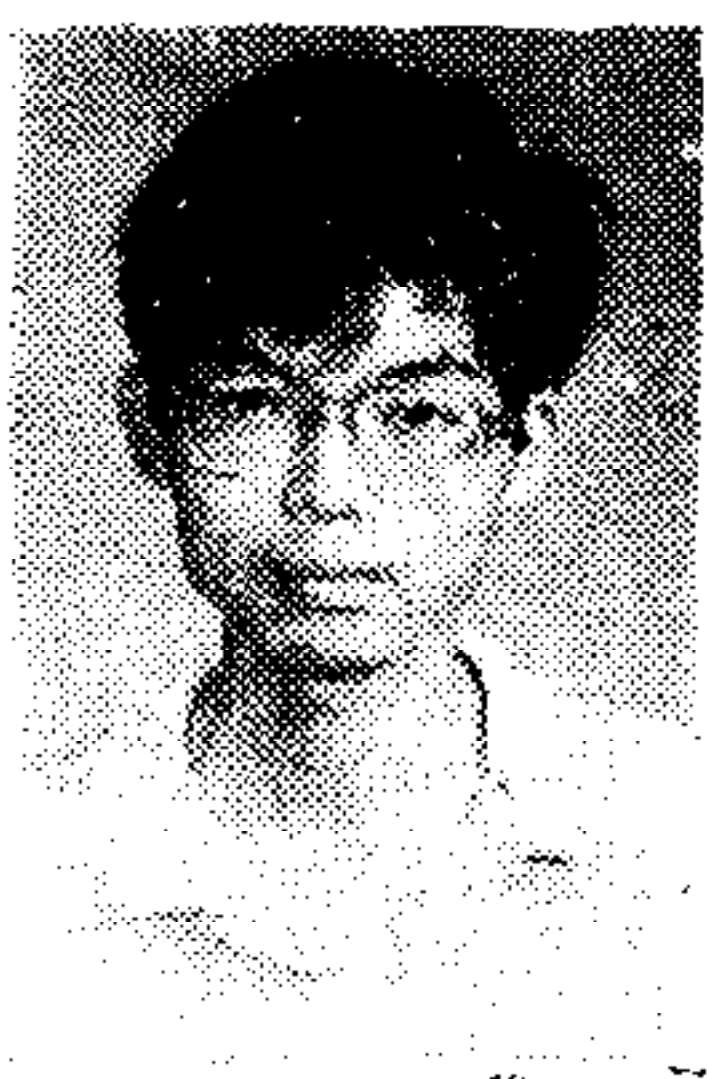
HU ZEXIN

(Res. Inst. of Automatic Control, East China Univ. of  
Sci. and Tech. Shanghai 200237)

### ABSTRACT

Introducing the functional-like network theory, nonlinear filter based on neural networks(NNNF) for state estimation of stochastic nonlinear systems is proposed by the using of data of on-line measurements or available off-line measurements data. The performance of biaslessness and minimal error variances of NNNF are proven. NNNF is applied to the state estimations of glutamic acid fermentation and erythromycin fermentation processes disturbed by noise, and the estimation by experimental errors. The estimated results and experimental online state estimation coincide very well. NNNF is insensitive to noise distributions and initial state estimation, NNNF can be used in on-line measurements of biomass, substrate, and product concentrations.

**Key words:** Nonlinear stochastic systems, state estimation, neural networks, biotechnical processes.



**胡泽新** 1990年在华东理工大学获得工学博士学位,现为该校自动化研究所副教授,从事工业过程模型化与控制的教学和科研工作。主要研究方向为工业过程优化控制与智能控制,生物反应器的建模和控制,目前感兴趣的领域为计算机网络的管理与控制。