



一种基于时序预报神经网络的故障 预报方法及其应用¹⁾

冯恩波 肖德云 方崇智

(清华大学自动化系 北京 100084)

摘 要

提出一种基于时序预报神经网络的工业过程故障预报方法,同时给出了描述神经网络预报和外推能力的表达方式,并以氯碱电解工艺的现场数据验证了这种故障预报方法的有效性。实验结果表明,该方法可成功地用以实现氯中含氢的24小时预报。

关键词: 故障预报,神经网络,氯碱工业。

1 引言

利用神经网络的非线性映射关系进行故障诊断已有不少结果^[1-5],但这些方法的最大局限在于难以处理具有动态特性的早期预报问题。这里论述一种基于时序预报神经网络的工业过程故障预报方法。此方法将时序数据引入神经网络的训练过程,在训练样本中引进遗忘因子和期望因子,并形成相应的神经网络结构。同时详细研究了神经网络的外推性,提出了一种外推性或预报能力的刻划方法。

2 时序预报神经网络的基本结构

预报模型的结构是由输入数据加至神经网络输入节点的组织方式决定的。用 $x(t)$ 和 $y(t)$ 分别表示对象的输入和输出。记当前时刻为 t_0 ,移动数据窗口中则包含了具有成对关系的数据 $\{[x(j), y(j)]: j = t_0 - d, t_0 - d + 1, \dots, t_0\}$ 和输出变量数据 $\{y(j): j = t_0 + 1, \dots, t_0 + k\}$ 。为了使网络预报能力得以实现,对神经网络施以训练样本

$$\{x(j), y(j): j = t_0 - d, t_0 - d + 1, \dots, t_0 | y(j): j = t_0 + 1, \dots, t_0 + k\}.$$

随着 t_0 的右移,数据窗生成具有时序关系的训练样本集。

为了能体现对远期效应的遗忘过程,在时序输入样本集中引入遗忘因子 γ 。与此相仿,实际应用中往往希望近期预报能更准确一些,而长期预报可适当放宽要求。为此,在输出样本集中又引入期望因子 β 。故训练样本集成为如下形式:

本文于1992年8月28日收到。

1) 国家“八五”重点科技攻关项目资助课题。

$$\{\gamma^{t_0-i}x(j), \gamma^{t_0-i}y(j): j = t_0 - d, \dots, t_0 | \beta^{i-t_0}y(j): j = t_0 + 1, \dots, t_0 + k\}, \quad (1)$$

式中因子取值范围为 $0 \leq \gamma \leq 1, 0 \leq \beta \leq 1$.

图 1 所示为故障预报方法采用的模型基本结构。

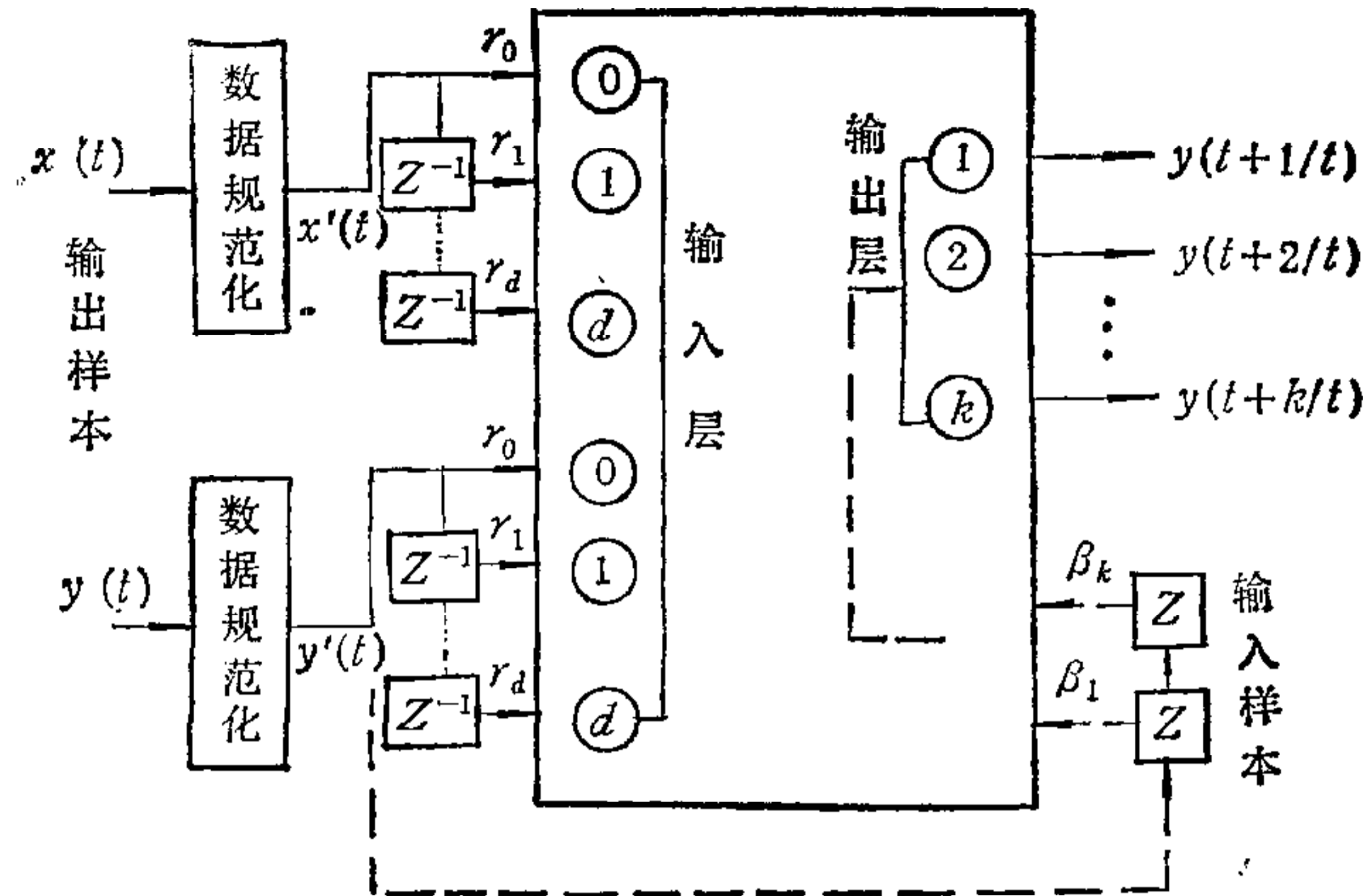


图 1 预报神经网络结构

其网络表达式为

$$s_0(t) = N[w, s_i(t)]. \quad (2)$$

其中 $s_0(t)$ 是网络输出, $s_i(t)$ 是网络输入, w 是网络内部神经元可训练参数。用向量表示, 则有

$$s_0(t) = \begin{bmatrix} y(t+1|t) \\ y(t+2|t) \\ \vdots \\ y(t+k|t) \end{bmatrix}; \quad s_i(t) = \begin{bmatrix} \gamma^{t-i}x(j) (j = t-d, \dots, t) \\ \vdots \\ \beta^{j-t}y(j) \\ (j = t+1, \dots, t+k) \end{bmatrix}; \quad (3)$$

$$w = [w_1, w_2, \dots, w_n]^T.$$

3 神经网络外推性的度量

时序预报神经网络要求有较强的外推性, 即在训练完成后, 不仅能复现原训练样本, 还要求在输入新样本时, 能外推出正确结果。事实上, 外推的效果与训练样本的完备性直接相关。为了定量地描述这一点, 下面均假定网络的所有训练数据是有限的, 或者可用有限组数据表示, 即一个完备的训练样本可写成如下形式:

$$\{s_i^{(j)} | s_0^{(j)}: j = 1, 2, \dots, p\}.$$

其中 $s_i^{(j)}, s_0^{(j)}$ 分别是训练样本的输入和输出集, p 是最大组数。

当用 p 组样本训练网络之后, 每组样本引入的平均误差可用二次型表示为

$$E_p(w^*, p) = \frac{1}{p} \sum_{j=1}^p [s_0^{(j)} - s_0^{(j)}(w^*)]^T [s_0^{(j)} - s_0^{(j)}(w^*)]. \quad (4)$$

式中 $s_0^{(j)}(w^*)$ 是计算输出, w^* 是网络训练过程收敛时的网络参数, 即 w^* 满足:

$$\begin{cases} \min_w J(w) = \frac{1}{2} \sum_{j=1}^p [s_0^{(j)} - \tilde{s}_0^{(j)}(w)]^T [s_0^{(j)} - \tilde{s}_0^{(j)}(w)], \\ \tilde{s}_0^{(j)}(w) = N[w, s_i^{(j)}], j = 1, 2, \dots, p. \end{cases} \quad (5)$$

当样本不完备时,也可设法使训练过程达到收敛,即使网络参数 \tilde{w}^* 满足:

$$\begin{cases} \min_w J(w) = \frac{1}{2} \sum_{j=1}^q [s_0^{(j)} - \tilde{s}_0^{(j)}(w)]^T [s_0^{(j)} - \tilde{s}_0^{(j)}(w)], \\ \tilde{s}_0^{(j)}(w) = N[w, s_i^{(j)}], j = 1, 2, \dots, q. \end{cases} \quad (6)$$

其中 $q < p$ 是训练样本实际组数,此时相对整个训练样本集的样本平均引入误差为

$$E_q(\tilde{w}^*, p) = \frac{1}{p} \sum_{j=1}^p [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)]^T [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)]. \quad (7)$$

将上式分解为两部分,可得

$$\begin{aligned} E_q(\tilde{w}^*, p) &= \frac{1}{p} \sum_{j=1}^q [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)]^T [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)] \\ &\quad + \frac{1}{p} \sum_{j=q+1}^p [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)]^T [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)] \\ &\triangleq \frac{q}{p} E_q(\tilde{w}^*, q) + \frac{1}{p} \tilde{E}_{p-q}(\tilde{w}^*). \end{aligned} \quad (8)$$

式中

$$\tilde{E}_{p-q}(\tilde{w}^*) = \sum_{j=q+1}^p [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)]^T [s_0^{(j)} - \hat{s}_0^{(j)}(\tilde{w}^*)], \quad (9)$$

它表示训练样本的不完备性所带来的残差。

定义预报能力指数为

$$K = \frac{E_p(w^*, p)}{E_q(\tilde{w}^*, p)}. \quad (10)$$

此指数的物理意义反映了新息引入残差能力。显然, $0 \leq K \leq 1$, 且 K 值越大,表明新息影响越小,也就是网络外推性越好。

为了进一步得出所需样本组数与预报指数的关系,利用以上诸式,易得

$$K = \frac{qE_p(w^*, q) + \tilde{E}_{p-q}(w^*)}{qE_q(\tilde{w}^*, q) + \tilde{E}_{p-q}(\tilde{w}^*)}. \quad (11)$$

这里

$$\tilde{E}_{p-q}(w^*) = \sum_{j=q+1}^p [s_0^{(j)} - \hat{s}_0^{(j)}(w^*)]^T [s_0^{(j)} - \hat{s}_0^{(j)}(w^*)]. \quad (12)$$

为简化起见,假设训练过程(完备样本时和非完备样本时)收敛时,各样本引入误差是等概率的,并且其大小均为相同数量级,近似有

$$E_p(w^*, q) \approx E_p(\tilde{w}^*, q), \quad (13)$$

$$\frac{\tilde{E}_{p-q}(w^*)}{qE_p(w^*, q)} \approx \frac{p-q}{\frac{q}{p}q} = \frac{p(p-q)}{q^2}. \quad (14)$$

将式(13),(14)代入式(11)中,并记

$$\alpha = \frac{\frac{1}{p-q} \tilde{E}_{p-q}(\tilde{\omega}^*)}{E_p(\tilde{\omega}^*, q)} \text{ 和 } \delta = p/q, \quad (15)$$

则整理后得

$$K = \frac{1 + \delta(\delta - 1)}{1 + (\delta - 1)\alpha}. \quad (16)$$

事实上,系数 α 反映了新息引入样本平均残差和训练样本平均残差的相对大小; δ 表示样本相对数量。由式(16)看出,在要求新息引入残差小于原训练样本平均残差的条件下,训练样本愈多,则网络模型的预报能力愈强。

例如,若要网络每步预报最大误差不大于样本平均引入残差的两倍,即 $\alpha \leq 2.0$, 样本总数 $p = 20$, 实际训练样本数 $q = 16$ 。通过计算可求得 $K \geq 87.5\%$, 这意味着新输入数据将引入12.5%的新残差。

4 氯碱工业电解过程的氯中含氢预报

氯中含氢的预报是氯碱工业一个极为迫切而难度甚大的研究课题。我们对某大型化工厂的分析数据进行了分类整理,采用神经网络实现氯中含氢的预报。训练算法采用 BP 算法,预报长度为 24 小时。图 2 所示分别是两个月的预报值和实际值。

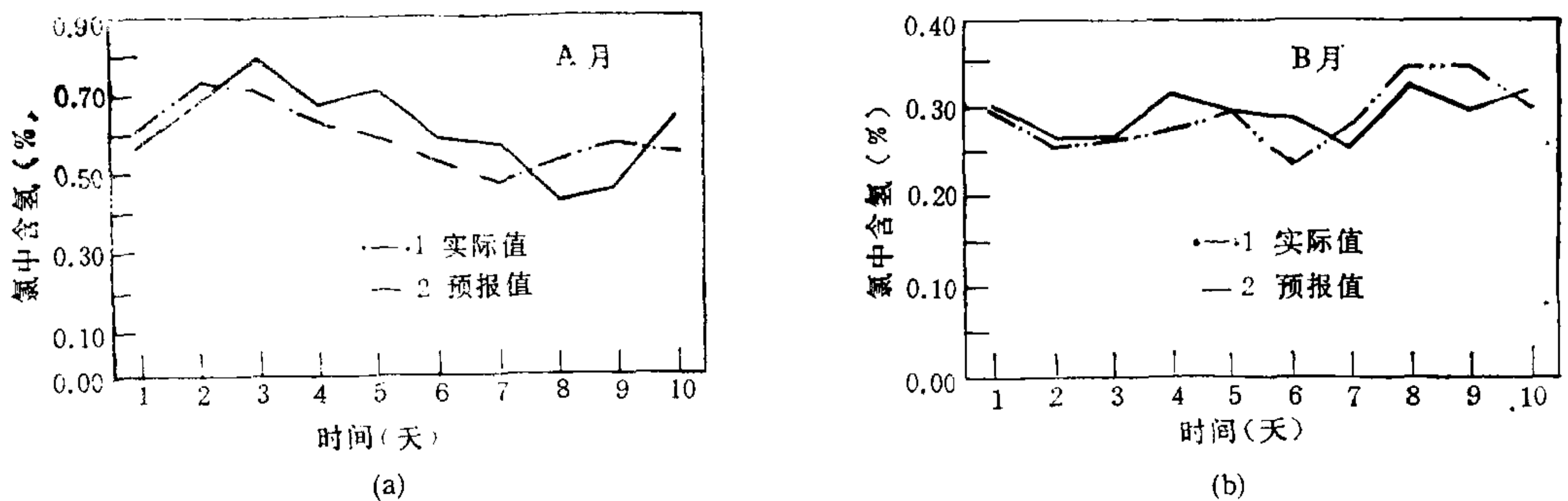


图 2 预报结果与实际值比较

训练和预报是滚动进行的,因为前 20 天的数据只用于启动过程的训练,而不用于预报,所以只给出了第 21 天后的滚动预报结果。

5 结束语

本文提出的预报神经网络结构,以时序训练方式使其具有了动态特性,其预报能力可以通过样本的挑选加以控制,因而可以解决故障早期预报问题。

参 考 文 献

- [1] Myers W. Expert systems and neural networks can work together interview with E. rich. *IEEE Expert*, 1990, 5(5): 5—7.
- [2] 冯恩波等. 基于神经网络技术的容错控制策略. 工业过程模型化及控制, 1992, 4: 302—307.
- [3] Hoskins J C et al. Fault diagnosis in complex chemical plants using artificial neural networks. *AIChEJ*, 1991, 37(1): 137—141.
- [4] Venkatasubramanian V et al. Process fault detection and diagnosis using neural networks-I. steady-state processes. *Computers Chem Engng*, 1990, 14(7): 699—712.
- [5] Muller B, Reinhardt J. Neural networks—An introduction. Springer-Verlag, New York: 1990.

A FAULT PREDICTION METHOD VIA TIME SERIES PREDICTIVE NEURAL NETS AND ITS APPLICATION

FENG ENBO XIAO DEYUN FANG CHONGZHI

(Dept. of Automation, Tsinghua University Beijing 100084)

ABSTRACT

A novel method of fault prediction in industrial systems which is based on the predictive neural network trained using time series data, is proposed in this paper. Then descriptions of the ability of prediction or generalization of time series predictive neural networks are discussed. Also we apply the new approach to the prediction of hydrogen-in-chlorine in electrolytic soda industry to prevent an explosion caused by the mixture of them. The results show that the new method is effective for 24 hour prediction.

Key words: Fault prediction, neural network, electrolytic soda process.