

控制与决策

Control and Decision

带自相关约束的NARX动态软测量模型

熊伟丽, 孙文心, 马君霞

引用本文:

熊伟丽, 孙文心, 马君霞. 带自相关约束的NARX动态软测量模型[J]. *控制与决策*, 2020, 35(4): 816–822.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2018.0586>

您可能感兴趣的其他文章

Articles you may be interested in

需求损失下两产品联合生产动态批量决策及预测时阈

Dynamic lot-size model and forecast horizon study for two products under joint production and lost sales

控制与决策. 2019, 34(2): 429–436 <https://doi.org/10.13195/j.kzyjc.2017.1104>

基于偏最小二乘的质量相关多模态故障检测技术

Quality-related multimodal fault detection technique based on partial least squares

控制与决策. 2019, 34(12): 2547–2557 <https://doi.org/10.13195/j.kzyjc.2018.0282>

偏最小二乘线性模型及其非线性动态扩展模型综述

Review of partial least squares linear models and their nonlinear dynamic expansion models

控制与决策. 2018, 33(9): 1537–1548 <https://doi.org/10.13195/j.kzyjc.2017.1306>

基于雅可比预测的机器人无模型视觉伺服定位控制

Jacobian prediction-based for robot model-free visual servoing positioning control

控制与决策. 2018, 33(10): 1813–1819 <https://doi.org/10.13195/j.kzyjc.2017.0717>

一种多变量预测控制的分程控制策略实现方法

An implementation of split control strategy for multi-variable predictive control

控制与决策. 2017, 32(4): 746–750 <https://doi.org/10.13195/j.kzyjc.2016.0040>

具有交互效应的多变量GM(1,N)模型

Multivariable GM(1, N) model with interaction effects

控制与决策. 2017, 32(3): 515–520 <https://doi.org/10.13195/j.kzyjc.2016.0138>

具有未建模动态和输出约束系统的自适应输出反馈控制

Adaptive output feedback control of systems with unmodeled dynamics and output constraint

控制与决策. 2017, 32(1): 55–62 <https://doi.org/10.13195/j.kzyjc.2015.1613>

正交信号校正的自回归模型及其在动态过程监测中的应用

Orthogonal signal correction based auto-regression model with application to dynamic process monitoring

控制与决策. 2016, 31(8): 1505–1508 <https://doi.org/10.13195/j.kzyjc.2015.0750>

带自相关约束的NARX动态软测量模型

熊伟丽^{1,2†}, 孙文心², 马君霞²

(1. 江南大学 轻工过程先进控制教育部重点实验室, 江苏 无锡 214122;

2. 江南大学 物联网工程学院, 江苏 无锡 214122)

摘 要: 非线性带外输入自回归模型(NARX)在进行预测估计时依赖于主导变量的实时测量,因此在实际工业过程中存在一定的实施难度.针对该问题,利用神经网络构造一种新型NARX动态软测量模型,当工业过程无法及时提供上时刻主导变量测量值时,能通过多步预测方法来确保主导变量的实时预测,通过设计模型结构来降低预测序列的自相关性,从而抑制由多步估计造成的累积误差,以适当降低单步预测精度为代价,使模型在主导变量检测时间长、采样周期长、测量存在噪声的工业场合下得到更好的预测效果.通过数学分析和脱丁烷塔数据仿真实验验证了所构建模型的有效性.

关键词: 动态软测量建模; 神经网络; NARX模型; 多步预测

中图分类号: TP319.9

文献标志码: A

Autocorrelation constrained NARX dynamic soft sensing model

XIONG Wei-li^{1,2†}, SUN Wen-xin², MA Jun-xia²

(1. China Key Laboratory of Advanced Process Control for Light Industry Ministry of Education, Jiangnan University, Wuxi 214122, China; 2. School of the Internet of Things Engineering, Jiangnan University, Wuxi 214122, China)

Abstract: Due to the excessive dependence on the real time measurement of dominant variables during prediction, it is difficult to apply the nonlinear autoregressive with exogenous input (NARX) model in real industrial processes. In order to solve the difficulty, a new NARX model structure is built by combining two network models. When the industrial environment cannot provide the history information of dominant variables in real time, the real-time prediction of dominant variables can be guaranteed by using the multi-step prediction method. By reducing the autocorrelation of prediction sequence, the cumulative error caused by multi-step estimation is suppressed. At the cost of reducing the accuracy of single-step prediction properly, the model can obtain better prediction effect in the much rough industrial environment. Simulation results verify the effectiveness of the proposed model.

Keywords: dynamic softsensor modeling; neural networks; NARX model; multi-step-ahead prediction

0 引 言

软测量建模能够利用工业过程中易于测量的辅助变量对不可直接测量或测量代价大的主导变量进行实时估计,已在环保^[1-2]、冶金^[3-4]、化工^[5]等领域得到了广泛应用.通常,软测量模型可以分为静态模型和动态模型两类,基于稳态假设^[6]所建立的模型称作静态模型,反之称为动态模型.为了能在非稳态工况下正常工作,动态模型在做预测时不仅要考虑当前时刻的过程状态,还需要考虑过程的历史信息.相比静态模型,动态模型对于时间序列信息具有更强的描述能力,能更好地适应工业过程中的非稳态情况^[7],由于更贴近实际过程,动态建模方法得到了工业界的广

泛关注,现已取得许多研究成果^[7-9].

带输入自回归模型(ARX)是最常见的一种线性动态模型,具有模型结构简单、易于实现的优点,已在许多时间序列预测中得到了应用.文献[10]介绍了该模型的非线性拓展形式,即NARX模型,在ARX的基础上结合神经网络(NN)的非线性逼近特性,使得模型能够更好地适应具有强非线性特性的工业过程.在此基础上,NARX模型已得到进一步研究,文献[11]研究了一种利用容积卡尔曼滤波来优化模型参数的算法,具有计算简单、训练高效等优点;文献[12]利用BPSO算法解决了模型结构的选择问题;文献[13]讨论了NARX模型在预测过程中存在的“大

收稿日期: 2018-05-05; 修回日期: 2018-10-22.

基金项目: 国家自然科学基金项目(61773182, 61702228); 国家重点研发计划子课题(2018YFC1603705-03).

责任编委: 刘向杰.

[†]通讯作者. E-mail: greenpre@163.com.

误差点”问题,并提出了误差修正方法,使得模型预测精度有了很大提升.

一方面,在许多工业过程中,主导变量的历史值与其当前值有很强的相关性,NARX模型在许多应用中表现出较高的预测精度^[12-13].另一方面,传统自回归模型依赖于主导变量的实时采集,而大部分工业过程中主导变量的测量存在采样周期长、测量有延迟以及代价大等问题,无法实时获得主导变量的测量值,这种情况下NARX模型的可靠性受到了极大的限制^[10].目前,基于NARX的软测量建模研究大多建立在主导变量的数据能被实时采集的假设之下,而摒弃这一假设来构建NARX软测量模型的研究仍然较少.

为解决以上问题,本文利用两个神经网络模型构造一种带自相关约束的NARX模型,通过约束神经网络的输出范围减小模型对于主导变量实时测量值的依赖,在保证建模精度的情况下,能够抑制由主导变量历史信息缺失带来的误差累积.

1 NARX模型介绍与符号说明

定义 \mathbf{X}_t 、 y_t 为过程在 t 时刻辅助和主导变量的取值. 辅助变量的上标表示变量中的具体特征,特征指过程中易测的物理变量,如温度、压力等,而辅助变量是由特征组成的列向量,设其维度为 m . 本文中主导变量为标量,有

$$\mathbf{X}_t = [x_t^1 \ x_t^2 \ \dots \ x_t^m]^T.$$

定义建模对象的真实模型为 f , 所建立的NARX模型为 \hat{f} , t 时刻主导变量的预测值为 \hat{y}_t , 有

$$\begin{aligned}
y_t &= f([\varphi^{n_a}(y_{t-1}) \ \varphi^{n_b}(\mathbf{X}_t)]), \\
\hat{y}_t &= \hat{f}([\varphi^{n_a}(y_{t-1}) \ \varphi^{n_b}(\mathbf{X}_t)], \theta), \\
\varphi^{n_a}(y_{t-1}) &= [y_{t-1} \ y_{t-2} \ \dots \ y_{t-n_a}], \\
\varphi^{n_b}(\mathbf{X}_t) &= [\mathbf{X}_t \ \mathbf{X}_{t-1} \ \dots \ \mathbf{X}_{t+1-n_b}]. \quad (1)
\end{aligned}$$

其中: θ 为模型的参数向量, n_a 和 n_b 为模型的阶次,模型 \hat{f} 可用图1表示.

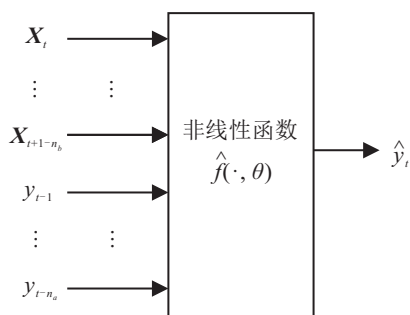


图1 NARX模型示意图

通常,利用神经网络作为非线性函数 $\hat{f}(\cdot, \theta)$ 描述 \hat{y}_t 与模型输入变量之间的关系,并利用已有的训练样

本对模型参数 θ 进行优化,使所定义的神经网络成为理想的预测模型. 虽然NARX仅需少量的模型参数便能建立精确的模型^[9],但该模型依赖于主导变量的历史信息,这使得NARX模型在主导变量采样缓慢或无法实时采样的工业场合下无法正常使用.

为应对上述工业环境,在主导变量历史信息存在缺失时,一般利用多步预测方法,即用 $\varphi^{n_a}(\hat{y}_{t-1})$ 代替 $\varphi^{n_a}(y_{t-1})$ 来完成对 y_t 的实时估计^[9]. 然而,这种方式会产生较大的预测误差,原因在于预测序列 $\{\hat{y}_0 \ \hat{y}_1 \ \dots \ \hat{y}_{t-1} \ \hat{y}_t\}$ 存在较大的自相关性,若某一时刻的预测存在误差,则该误差会对下一时刻的预测值产生影响,随时间的累计传统NARX模型将无法正常工作. 因此,本文构造了一种新的NARX模型,通过模型结构限制了预测结果 \hat{y}_t 与上时刻主导变量测量值 y_{t-1} (或 \hat{y}_{t-1}) 的相关性,并以此来避免上述问题的发生.

2 带自相关约束的NARX模型

2.1 模型的建立

在NARX模型中,辅助变量和主导变量的历史信息(即 $\varphi^{n_b}(\mathbf{X}_t)$ 和 $\varphi^{n_a}(y_{t-1})$) 都是决定模型预测结果 \hat{y}_t 的关键因素. 本文构造一种新的NARX模型,该模型将这两个关键因素对预测结果 \hat{y}_t 的作用分开表示,通过模型结构适当降低 $\varphi^{n_a}(y_{t-1})$ 对 \hat{y}_t 的作用强度. 一方面,考虑到大部分工业过程是稳定的,可以用一阶模型做近似描述;另一方面,考虑一阶模型更易于分析和实施,因此本文设定所建立的模型阶次 n_a 等于1.

利用两个输入变量为 $\varphi^{n_b}(\mathbf{X}_t)$ 的网络构建预测模型,分别为A网络模型和B网络模型,模型结构如图2所示.

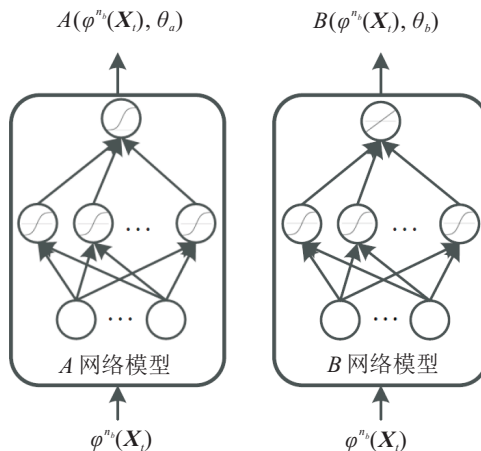


图2 网络结构示意图

A、B网络模型都是普通的单隐层、多输入单输出神经网络. 其中, A网络模型的输出层激活函数

为Sigmoid函数,其输出被限制在0~1之间;B网络模型的输出层为线性层,其输出可以是实数范围内的任意数。 θ_a 和 θ_b 分别表示A、B网络模型的模型参数。

利用上述两个神经网络构造新的NARX模型

$$\hat{y}_t = g([y_{t-1}, \varphi^{n_b}(\mathbf{X}_t)], \theta) = B(\varphi^{n_b}(\mathbf{X}_t), \theta_b) + cA(\varphi^{n_b}(\mathbf{X}_t), \theta_a)y_{t-1}. \quad (2)$$

其中: $\theta = \{\theta_a, \theta_b\}$; $c \in (0, 1)$,为自定义常量; θ 为整个模型的参数,包含A网络模型和B网络模型的模型参数。当主导变量的自相关性较大时可以适当增大 c 的取值。容易看出,当 c 取零时,模型(2)退化为一个标准的静态模型或NMA^[9]预测模型(非线性滑动平均模型)。模型结构的设计理由如下:首先,考虑到存在过度过程,令模型预测 \hat{y}_t 与 y_{t-1} 相关以提高预测精度;另外,考虑到大多数过程是稳定的,会逐渐趋于稳态,且过度依赖历史信息 y_{t-1} 会导致预测误差随时间累积,因此加入 c 项并限制A网络的输出范围来适当降低预测序列的自相关性。

由于模型(2)的复杂度低于式(1)模型 \hat{f} ,在单步预测方面其精度会略低于普通NARX模型,但在多步预测方面,模型(2)从结构上抑制了预测序列的自相关性,某一时刻的预测误差不会对后续预测结果造成很大影响,相比NARX模型,模型(2)能够更好地抑制多步预测时的误差累计。

2.2 代价函数的构造

利用已有的训练样本序列构造代价函数,以此评价预测模型 $g(\cdot, \theta)$ 的拟合精度。假设训练样本个数为 N ,构造代价函数为

$$J(\theta) = \sum_{i=n_b-1}^{N-1} (g([y_{i-1}, \varphi^{n_b}(\mathbf{X}_i)], \theta) - y_i)^2 + \lambda \sum_{i=n_b-1}^{N-1} A(\varphi^{n_b}(\mathbf{X}_i), \theta_a). \quad (3)$$

该代价函数由误差平方和与一个正则项构成, θ 为正则项系数,通常设为很小的值,本文默认设置为0.01。采用Adam^[13]算法对模型参数 θ 进行调节,使代价函数尽可能小。为了提高参数调整的效率,主导变量与辅助变量一般先经过标准化处理后再输入网络模型^[14]。

当建模对象的动态特性明显时,正则项基本不起作用;当仅利用辅助变量 $\varphi^{n_b}(\mathbf{X}_t)$ 便能精确预测当前主导变量 y_t 时,正则项会使A网络模型的输出惩罚至零。模型参数 θ 经过调整后成为预测模型的一部分并不再更改,为了表示简便,后文在表示训练完成的模型时略写其参数项。

2.3 多步预测的实现

通常,工业现场难以实时提供上一时刻主导变量的测量值。式(2)在对主导变量 y_t 进行预测时依赖 y_{t-1} ,无法直接使用,该情况下可以用其估计值 \hat{y}_{t-1} 替代。因此,本文方法的预测结果可分为两类:基于测量值 y_{t-1} 对 y_t 作出的预测和基于估计值 \hat{y}_{t-1} 对 y_t 作出的预测。为了便于区分,定义预测量 $\hat{y}_{t_2}^{t_1}$ 为

$$\hat{y}_{t_2}^{t_1} = \begin{cases} g([y_{t_2-1}, \varphi^{n_b}(\mathbf{X}_{t_2})]), & t_2 = t_1 + 1; \\ g([\hat{y}_{t_2-1}^{t_1}, \varphi^{n_b}(\mathbf{X}_{t_2})]), & t_2 > t_1 + 1. \end{cases} \quad (4)$$

其中: $\hat{y}_{t_2}^{t_1}$ 的上标 t_1 表示该预测值是基于测量值 y_{t_1} 利用多步或单步预测所得到的;下标 t_2 表示预测量 $\hat{y}_{t_2}^{t_1}$ 是模型对 y_{t_2} 作出的预测。图3对此作了进一步说明。

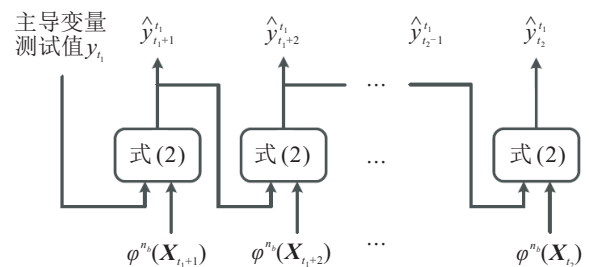


图3 多步预测结果说明

根据式(2)和(4)可以得到

$$\begin{aligned} \hat{y}_{t_2}^{t_1} &= B(\varphi^{n_b}(\mathbf{X}_{t_2})) + cA(\varphi^{n_b}(\mathbf{X}_{t_2}))\hat{y}_{t_2-1}^{t_1} = \\ & \sum_{i=0}^{t_2-t_1-1} c^i A(\varphi^{n_b}(\mathbf{X}_{t_2-i}))B(\varphi^{n_b}(\mathbf{X}_{t_2-i-1})) + \\ & y_{t_1} c^{t_2-t_1} \prod_{i=t_1}^{t_2} A(\varphi^{n_b}(\mathbf{X}_i)). \end{aligned}$$

由图3可见, t_1 越接近 t_2 , $\hat{y}_{t_2}^{t_1}$ 所含的过程信息越新。因此希望建立一种递推关系,若 $t(t_1 < t < t_2)$ 时刻所采用的主导变量 y_t 在 t_2 时刻被测量或化验得到,则能够利用该递推关系更新当前估计量 $\hat{y}_{t_2}^{t_1}$ 并得到 $\hat{y}_{t_2}^t$ 。进一步推导可以得到

$$\begin{cases} \hat{y}_{t_2}^t = y_t c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{n_b}(\mathbf{X}_i)) + M, \\ \hat{y}_{t_2}^{t_1} = \hat{y}_t^{t_1} c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{n_b}(\mathbf{X}_i)) + M, \\ M = \sum_{i=0}^{t_2-t-1} c^i A(\varphi^{n_b}(\mathbf{X}_{t_2-i}))B(\varphi^{n_b}(\mathbf{X}_{t_2-i-1})). \end{cases}$$

最后,利用递推关系更新原估计 $\hat{y}_{t_2}^{t_1}$,从而得到新的估计量 $\hat{y}_{t_2}^t$,有

$$\begin{aligned} \hat{y}_{t_2}^t &= y_t c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{n_b}(\mathbf{X}_i)) + M = \\ & \hat{y}_t^{t_1} c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{n_b}(\mathbf{X}_i)) + M + \end{aligned}$$

$$(y_t - \hat{y}_t^{t_1})c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{nb}(\mathbf{X}_i)) = \hat{y}_{t_2}^{t_1} + (y_t - \hat{y}_t^{t_1})c^{t_2-t} \prod_{i=t}^{t_2} A(\varphi^{nb}(\mathbf{X}_i)). \quad (5)$$

本文所提出方法在作预测时需要提供过程初始时刻的主导变量测量值 y_0 . 预测主要过程如下: 利用式(4)预测当前时刻的主导变量值, 然后判断是否得到更新的主导变量测量值, 若得到则利用式(5)对当前主导变量估计值作更新, 否则利用多步预测方法(即式(4))预测当前主导变量. 具体实施步骤如图4所示.

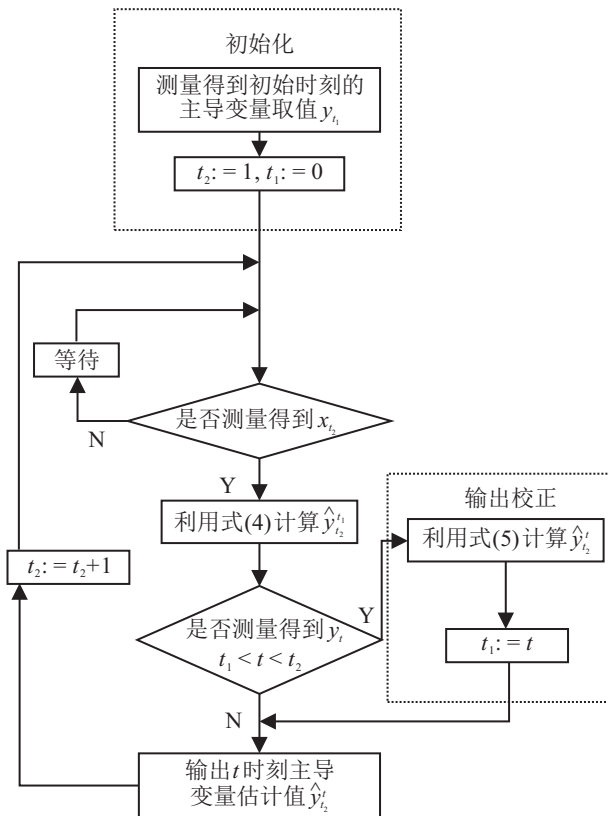


图4 多步预测方法流程

由图4可见, 在预测主导变量时本文方法不会严格依赖于上一时刻主导变量的测量值, 因此适用于主导变量采样缓慢、存在较大时滞或无法采样的场合.

2.4 模型特性的分析

当本文模型能够精确描述整个工业过程时, 上述预测方法能够很好地保证精度, 即使在无法提供主导变量测量信息(除了初始时刻)的情况下, 也能保证预测误差不会随时间发散. 为了方便分析, 定义以下集合表示预测模型 g 能够准确描述真实对象 f 的范围.

定义1 将能够使以下不等式成立的向量 $[y, \varphi^{nb}(\mathbf{X})]$ 全体称为预测模型 g 相对于真实对象 f 的 ω 准确集合 ($\omega > 0$), 并记作 $D_\omega(g, f)$:

$$|g([y, \varphi^{nb}(\mathbf{X})]) - f([y, \varphi^{nb}(\mathbf{X})])| \leq \omega.$$

引理1 假设已知初始时刻主导变量的测量值 y_0 和过程运行的某一时刻 t_0 , 若能保证对于任意 $t \in [0, t_0]$, 都能使得所属关系 $[y_{t-1}, \varphi^{nb}(\mathbf{X}_t)] \in D_\omega(g, f)$ 成立, 则能保证预测误差不至于发散, 即

$$|\hat{y}_t^0 - y_t| < \frac{\omega}{1-c}, \quad 0 < t \leq t_0. \quad (6)$$

证明 利用数学归纳法, 先假设存在 t_1 使得以下不等式成立:

$$|\hat{y}_{t_1}^0 - y_{t_1}| < \frac{\omega}{1-c}, \quad t_1 > 0.$$

有

$$\begin{aligned} \hat{y}_{t_1+1}^0 &= g([\hat{y}_{t_1}^0, \varphi^{nb}(\mathbf{X}_{t_1+1})]) = \\ &= cA(\varphi^{nb}(\mathbf{X}_{t_1+1}))\hat{y}_{t_1}^0 + B(\varphi^{nb}(\mathbf{X}_{t_1+1})) = \\ &= cA(\varphi^{nb}(\mathbf{X}_{t_1+1}))(\hat{y}_{t_1}^0 - y_{t_1}) + \\ &= g([y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})]). \end{aligned}$$

由于 $[y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})] \in D_\omega(g, f)$, 由模型结构可知 $A(\mathbf{X}_{t_1+1}) < 1$, 有

$$\begin{aligned} |\hat{y}_{t_1+1}^0 - y_{t_1+1}| &= \\ &= |cA(\varphi^{nb}(\mathbf{X}_{t_1+1}))(\hat{y}_{t_1}^0 - y_{t_1}) + \\ &= |g([y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})]) - f([y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})])| \leq \\ &= |cA(\varphi^{nb}(\mathbf{X}_{t_1+1}))(\hat{y}_{t_1}^0 - y_{t_1})| + \\ &= |g([y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})]) - f([y_{t_1}, \varphi^{nb}(\mathbf{X}_{t_1+1})])| < \\ &= c|\hat{y}_{t_1}^0 - y_{t_1}| + \omega < \frac{c\omega}{1-c} + \omega = \frac{\omega}{1-c}. \end{aligned}$$

又由于

$$|\hat{y}_1^0 - y_1| \leq \omega < \frac{\omega}{1-c},$$

使假设成立的 t_1 存在. \square

为了进一步分析测量误差对预测过程造成的影响, 假设初始时刻主导变量在测量上存在噪声污染, 定义为 \tilde{y}_0 , 根据该测量值得到的主导变量预测值记为 \tilde{y}_t^0 , 其关系描述为

$$\begin{aligned} \tilde{y}_0 &= y_0 + \varepsilon; \\ \tilde{y}_{t+1}^0 &= \begin{cases} g([\tilde{y}_0, \varphi^{nb}(\mathbf{X}_{t+1})]), & t = 0; \\ g([y_t^0, \varphi^{nb}(\mathbf{X}_{t+1})]), & t > 0. \end{cases} \end{aligned} \quad (7)$$

其中 ε 为未知的测量噪声. 根据迭代关系(5)可得到以下结论.

引理2 设初始时刻主导变量真实值为 y_0 , 测量值为 \tilde{y}_0 , 本文模型可以保证测量噪声对模型预测的影响随预测步数的增大而减小, 即

$$|\tilde{y}_t^0 - \hat{y}_t^0| < \varepsilon c^t, \quad t > 0. \quad (8)$$

证明 首先假设如下 t_1 时刻不等式成立:

$$|\tilde{y}_{t_1}^0 - \hat{y}_{t_1}^0| < \varepsilon. \quad (9)$$

可以推知

$$|\hat{y}_{t+1}^0 - \hat{y}_{t+1}^0| = cA(\varphi^{n_b}(\mathbf{X}_{t+1}))|\hat{y}_{t_1}^0 - \hat{y}_{t_1}^0| < c\varepsilon. \quad (10)$$

由式(4)和(7)可以得到

$$|\hat{y}_1^0 - \hat{y}_1^0| = cA(\varphi^{n_b}(\mathbf{X}_1))|\tilde{y}_0 - y_0| < c\varepsilon. \quad (11)$$

在不等式(11)的基础上利用(10)作递推,可得到式(8)成立. □

综上,当本文模型能够准确描述整个工业过程时,该预测方法具有以下两个特点:1)减少了对主导变量测量值的依赖,可适用于主导变量采样缓慢、存在较大时滞或无法采样的场合;2)对主导变量的测量噪声不敏感.

3 仿真实验

脱丁烷塔是脱硫石脑油分离装置的一部分,其主要用途是分离石油脑中的成分,分离出的碳三、碳四(丁烷)从塔顶输出,剩下碳五从塔底输出到下一个工业流程,过程工艺如图5所示.

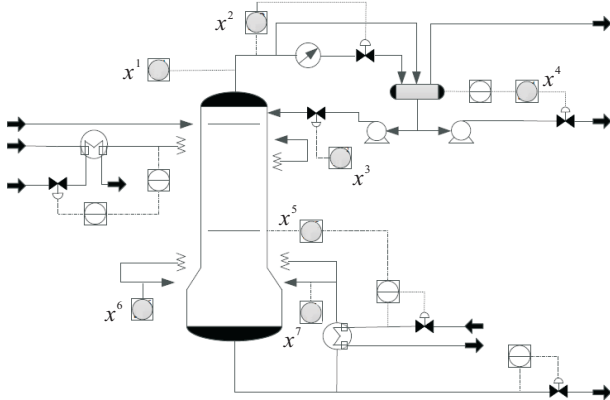


图5 脱丁烷塔过程示意图

为了使分离尽可能彻底,需要对塔底丁烷浓度进行实时测量.由于检测装置安装位置的原因,主导变量的测量存在大量延迟^[15],需要用软测量对塔底丁烷浓度进行实时估计.所选辅助变量如表1所示.

表1 建立模型所用变量

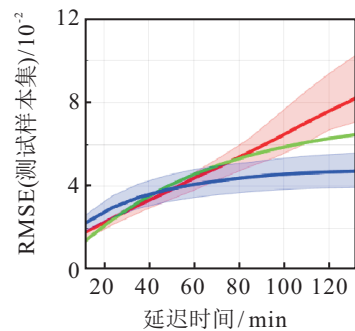
符号	变量名称	符号	变量名称
x^1	塔顶温度	x^4	塔顶出料量
x^2	塔顶压强	x^5	塔板6温度
x^3	塔顶回流量	x^6, x^7	塔底温度

以6 min作为采样间隔共采集2394组样本,取一半作为训练样本,另一半作为测试样本.分别利用静态神经网络模型、NMA模型、ARX模型、NARX模型和本文模型对数据进行建模和比较,仿真中设置NMA模型的阶次 n_b 为5,其余模型的阶次为1,式(2)中的 c 值默认设置为0.9.由于前两者在测试中不依

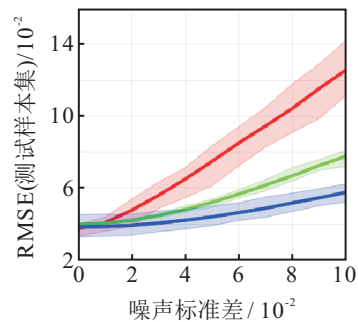
赖于主导变量的测量值,将仿真分为两组.首先,考虑主导变量采样存在噪声、时延和采样频率低的情况,将本文方法与ARX模型、NARX模型进行对比;然后,考虑无法对主导变量采样的情况,将本文方法与静态神经网络模型、NMA模型对比,从而验证本文方法的有效性.

3.1 仿真实验1

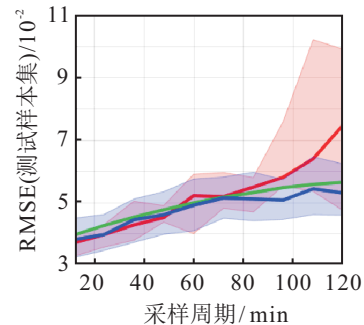
神经网络建模存在随机性,由于初始模型参数随机给定,每次建模结果不一定相同,为了减小随机性对结论造成的影响,对各方法进行100次建模实验.统计每次实验中各方法的均方根误差(RMSE),结果如图6所示.



(a) 存在时延



(b) 存在测量噪声



(c) 测量周期长



图6 不同情况下建模误差比较

由于过程中丁烷检测的时延普遍为45 min^[16],在模型抗噪声实验和低采样频率实验(图6(b)和图6(c))均对主导变量的测量值引入相应的时延.可以看出,当丁烷浓度能够及时测量时,NARX模型和ARX模型都有非常准确的预测效果,但由于这些模型过度依赖主导变量的测量值,在主导变量测量存在困难的情况下(如图6的3种情况)预测精度会受到影响.本文方法利用模型结构减小了对主导变量测量值的依赖性,这种做法一定程度上降低了模型的单步预测精度,但能够使模型更好地抑制由多步预测带来的误差积累,因此能够适用于主导变量采样缓慢、存在较大时滞或测量存在噪声的场合.

3.2 仿真实验2

建立一个普通NARX模型,并在两种不同的情况下对测试样本进行预测,如图7所示,理想情况表示前一时间刻主导变量能被实时测量获得的情况.可以看出,在该情况下预测模型对主导变量的预测值与实际测量值几乎完全贴合,表明NARX模型对该过程的动态具有非常好的描述能力.然而,即使在能够准确描述整个过程的情况下,传统NARX模型仍然无法保证预测误差在序列中收敛,在主导变量无法测量的工业场合下,误差容易随时间而累积形成较大的偏移.

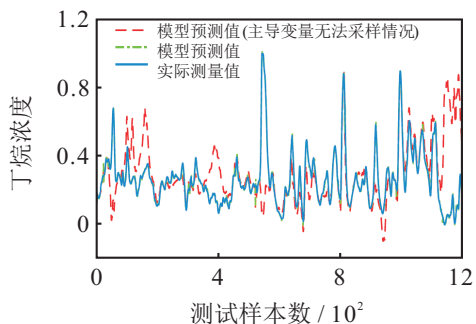
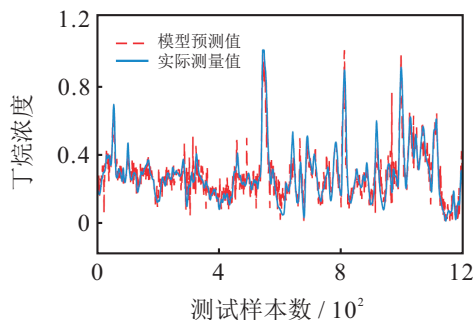


图7 NARX模型预测结果

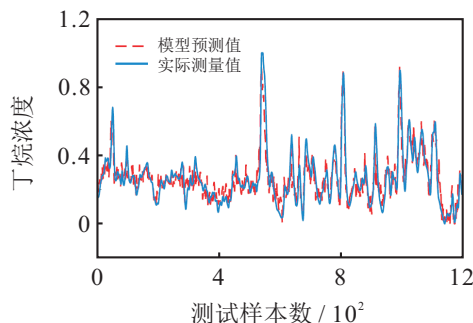
静态神经网络模型和NMA模型由于完全不考虑主导变量历史信息,能够在该情况下正常工作,图8为各方法的预测效果对比.仿真中仅提供初始时刻主导变量测量值,以此来模拟主导变量无法采样的场合.可以看出,NMA模型由于较好地体现了过程的动态特性,预测精度高于普通的静态模型,但相比之下不及理想情况下的NARX模型.本文方法结合了NARX模型精度高、参数少的优点,并通过设计模型结构缓解了普通NARX模型对主导变量测量值的过度依赖性,因此在测试中本文方法很好地抑制了误差的累计并表现出很好的预测效果.

最后,同样为了减小建模随机性对结论造成的影

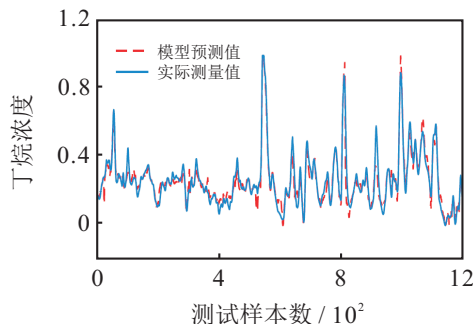
响,在模拟主导变量无法采样的情况下进行100次实验,结果如表2所示.可以看出,本文方法在100次实验中均未发生误差发散的情况.



(a) 静态神经网络预测效果



(b) NMA模型预测效果



(c) 本文模型预测效果

图8 仿真结果对比

表2 建立模型所用变量

模型名称	最大RMSE	平均RMSE	最小RMSE
静态模型	0.080	0.073	0.068
NMA模型	0.074	0.065	0.060
本文方法	0.054	0.047	0.039

4 结论

传统NARX对主导变量的历史信息有较强的依赖,当主导变量的测量存在困难时,预测精度会随着预测步数的增加而下降.本文利用两个神经网络构造了一种带自相关约束的NARX模型,通过模型结构有效约束了模型对于主导变量测量的依赖性,虽然在一定程度上会降低单步预测精度,但使得模型更好地适应了主导变量测量存在困难的场合.

本文方法还存在一些需要完善的方面,如将本文方法拓展到更高阶的系统上,在保证误差收敛的条件下如何更准确地描述系统的动态特性.以后的工作会在本文基础上作进一步讨论.

参考文献(References)

- [1] 韩红桂, 陈治远, 乔俊飞, 等. 基于区间二型模糊神经网络的出水氨氮软测量[J]. 化工学报, 2017, 68(3): 1032-1040.
(Han H G, Chen Z Y, Qiao J F, et al. Soft-sensor method for effluent ammonia nitrogen based on interval type-2 fuzzyneural networks [J]. *CIESC Journal*, 2017, 68(3): 1032-1040.)
- [2] 刘林, 马邕文, 万金泉, 等. 基于pso-SVM的废水厌氧处理过程软测量模型[J]. 环境科学学报, 2017, 37(6): 2122-2129.
(Liu L, Ma Y W, Wan J Q, et al. 2017. An accuracy soft-sensing model for the estimation of anaerobic digestion process based on pso-SVM model[J]. *Acta Scientiae Circumstantiae*, 2017, 37(6): 2122-2129.)
- [3] 于清, 贺建军. 模压时效炉锻件温度软测量方法[J]. 信息与控制, 2017, 46(3): 328-334.
(Yu Q, He J J. ITemperatureSoftSensorMethod for Molded Aging Ovens[J]. *Information and Control*, 2017, 46(3): 328-334.)
- [4] 蒋朝辉, 董梦林, 桂卫华, 等. 基于Bootstrap的高炉铁水硅含量二维预报[J]. 自动化学报, 2016, 42(5): 715-723.
(Jiang Z H, Dong M L, Gui W H, et al. Two-dimensional prediction for silicon content of hot metal of blast furnace based on Bootstrap[J]. *Acta Automatica Sinica*, 2016, 42(5): 715-723.)
- [5] 李妍君, 熊伟丽, 徐保国. 一种带过程变量时滞估计的在线软测量建模方法[J]. 信息与控制, 2016, 45(6): 641-646.
(Li Y J, Xiong W L, Xu B G. An onlines soft-sensor modeling method including process variable time-delay estimation[J]. *Information and Control*, 2015, 45(6): 641-646.)
- [6] Yong M A, Huang D, Jin Y. Discussion about dynamic soft-sensing modeling[J]. *Journal of Chemical Industry & Engineering*, 2005, 56(8): 1516-1519.
- [7] 杜文莉, 官振强, 钱锋. 一种基于时序误差补偿的动态软测量建模方法[J]. 化工学报, 2010, 61(2): 439-443.
(Du W L, Guan Z Q, Qian F. Dynamic soft sensor modeling based On time series error compensation[J]. *CIESC Journal*, 2010, 61(2): 439-443.)
- [8] Wu Q Y, Chan L L T, Chen J. Active learning dynamic soft sensor with forward-update scheme[C]. *Advanced Control of Industrial Processes*. Taipei: IEEE, 2017: 306-311.
- [9] Wang Z, Luo X. Modeling study of nonlinear dynamic soft sensors and robust parameter identification using swarm intelligent optimization CS-NLJ[J]. *Journal of Process Control*, 2017, 58: 33-45.
- [10] Fortuna L, Graziani S, Rizzo A, et al. Soft sensors for monitoring and control of industrial processes[M]. Berlin: Springer-Verlage Press, 2007: 116.
- [11] 胡振涛, 袁光耀, 胡玉梅, 等. 基于容积卡尔曼滤波的神经网络训练算法[J]. 控制与决策, 2016, 31(2): 355-360.
(Hu Z T, Yuan G Y, Hu Y M, et al. Training method of neural network based on cubature Kalman filter[J]. *Control and Decision*, 2016, 31(2): 355-360.)
- [12] Yassin I M, Taib M N, Adnan R, et al. Effect of swarm size parameter on binary particle swarm optimization-based NARX structure selection[C]. *Industrial Electronics and Applications*. Bandung: IEEE, 2012: 219-223.
- [13] 钱欣瑞, 史彬, 鄢烈祥. 基于NARX神经网络的油品性质软测量建模[J]. 计算机与应用化学, 2017, 34(1): 64-68.
(Qian X R, Shi B, Yan L X. Soft-sensing modeling for crude products properties based on NARX neural network[J]. *Computers and Applied Chemistry*, 2017, 34(1): 64-68.)
- [14] Goodfellow I, Bengio Y, Courville A, et al. Deep learning[M]. Cambridge: MIT Press, 2016: 508.
- [15] LeCun Y, Bottou L, Orr G B, et al. Efficient backprop[M]. Berlin: Springer Heidelberg, 1998: 9-50.
- [16] Shao W, Tian X. Semi-supervised selective ensemble learning based on distance to model for nonlinear soft sensor development[J]. *Neurocomputing*, 2017, 222: 91-104.

作者简介

熊伟丽(1978—),女,教授,博士生导师,从事复杂工业过程建模及其优化等研究, E-mail: greenpre@163.com;

孙文心(1993—),男,硕士生,从事复杂工业过程建模的研究, E-mail: 13373647858@163.com;

马君霞(1982—),女,讲师,博士,从事系统辨识、参数估计和过程控制等研究, E-mail: junxia.20@163.com.

(责任编辑: 郑晓蕾)