

烟气含氧量软测量新方法研究

卢 勇, 徐向东

(清华大学 热能工程系, 北京 100084)

摘 要: 针对目前大量工业现场使用的氧量分析仪成本昂贵、维护费用高且易损坏等问题, 在几种常见方法对比讨论的基础上, 提出了利用基于统计分析和神经网络技术的 NNPLS 方法建立烟气含氧量软测量模型的方法。该方法综合了 PLS 和神经网络技术的优点, 能够利用过程历史数据辨识对象模型; 利用现场实际数据对该方法进行了仿真验证, 并将仿真结果与传统的线性 PLSR 方法和直接神经网络建模方法作了比较, 结果显示 NNPLS 方法所建立的软测量模型具有更强的泛化能力。文中还对静态模型向动态模型进行了扩展。

关 键 词: 软测量; 偏最小二乘; 神经网络; 交叉验证; 泛化能力

中图分类号: O241; TP183

文献标识码: A

1 引 言

维护锅炉燃烧过程的最佳状态和经济性是锅炉燃烧过程自动控制的重要任务。通常工程上通过测量与空气过剩系数密切相关的烟气含氧量来间接的反映燃料、风配比的好坏。大多数动力锅炉目前所安装使用的氧量计, 由于测量原理复杂, 探头工作环境恶劣, 因此日常维护工作量相当繁重, 经常出现氧量计正常工作几个月后即失效的情况。

软测量技术是利用一些较易在线测量的辅助变量(与被测变量密切联系), 通过在线分析, 去估计不可测或难测变量的方法。其核心内容是发现和建立辅助变量与主要变量之间的模型关系。主要方法可以分为三类: 基于工艺机理的软测量模型(first principle model), 基于统计分析的软测量模型(如回归分析及相关分析), 以及基于人工智能的软测量模型。工艺机理模型由于人们对大多数过程机理的认识不足, 难以实际应用; 而利用统计分析和人工智能方法建立的经验模型(empirical model)应用范围较广, 前者依靠对现场收集到的试验和历史数据进行统计分析, 发现数据间潜在的规律和关系, 建立预测模型; 而后者则是通过机器学习等方法, 将来自实践中的数据和经验整理成一种“知识性”模型用以监测和预

报, 其中基于神经网络的软测量技术是研究最多、发展最快和应用范围最广的一种软测量技术。

正是基于上述原因, 采用在对象过程机理分析的基础上, 结合统计分析和神经网络非线性系统建模两种方法的优点, 通过构造 NNPLS 网络结构和算法, 建立烟气含氧量的稳态和动态模型。利用该软测量模型, 不仅可以取代现有的氧量监测设备, 同时还可以与燃烧自动控制系统结合组成推理控制器(inference controller), 彻底消除由于氧量计失效所导致的燃烧误调整问题。

2 偏最小二乘回归方法

许多系统分析人员为了更完备的描述系统, 往往倾向于较周到的选取有关的指标作为自变量, 这样选取的变量之间通常存在较严重的多重相关性; 同时回归建模时由于样本点数量较少所带来的问题也是导致常规方法失效的原因。偏最小二乘方法的出现有效的解决了上述问题, 它集多元线性回归分析、典型相关分析和主成分分析的基本功能于一体, 在 PLS 算法统一框架下, 可以同时实现回归建模、数据结构简化以及两组变量间的相关分析。下面简要的介绍一下 PLS 方法的基本原理: 设有 q 个因变量和 p 个自变量, 由观测得到的 n 个样本点组成了两个数据表 $X = [x_1, \dots, x_p]_{n \times p}$ 和 $Y = [y_1 \dots y_q]_{n \times q}$ 。PLS 方法分别在两个数据表中提取成分 t_h 和 u_h , 两个成分应尽可能的携带各自数据表中的变异信息, 同时两个成分间的相关程度要最大。其原理可归纳为外部关系式和内部关系式(如下所示):

PLS 外部关系: $X = TP^T + E = \sum_{h=1}^a t_h p_h^T + E$
(E 为成分提取后的残余信息)。

$Y = UQ^T + F = \sum_{h=1}^a u_h q_h^T + F$
(F 为成分提取后的残余信息)。

PLS 内部关系: $\hat{u}_h = b_h t_h$

PLS 通常采用 NIPALS 递推算法计算, 步骤简述如下(假设数据矩阵已经过标准化处理(aut oscaling)):

(1) 取 $E_0 = X, F_0 = Y, h = 1$; (1)

(2) 取 F_{h-1} 的某一列作为其主元得分向量 u_h ;

(3) 将 u_h 对 X 回归计算输入权向量 w , 并将其规范化:

$$w_h^T = u_h^T E_{h-1} / (u_h^T u_h); w_h^T = w_h^T / \|w_h^T\|$$

(4) 计算输入得分向量 t_h : $t_h = E_{h-1} w_h$;

(5) 计算输出负荷向量 q_h , 并将其规范化:

$$q_h^T = t_h^T F_{h-1} / (t_h^T t_h); q_h^T = q_h^T / \|q_h^T\|;$$

(6) 计算新的输出得分向量 u_h : $u_h = F_{h-1} q_h$;

(7) 检查得分向量 u_h 的收敛性; 若不收敛, 返回步骤 3 重新计算, 否则继续;

(8) 计算输入负荷向量 p_h : $p_h^T = t_h^T E_{h-1} / (t_h^T t_h)$;

(9) 计算内部关系回归模型系数 b_h :

$$b_h = u_h^T t_h / (t_h^T t_h);$$

(10) 计算输入输出残余信息矩阵 E_h 和 F_h :

$$E_h = E_{h-1} - t_h p_h^T; F_h = F_{h-1} - b_h t_h q_h^T.$$

(11) 如果精度未达到要求, $h = h + 1$ 返回步骤 2, 重复上述计算过程; 否则退出。

3 基于 FFNN 的非线性系统建模

PLS 方法是一种线性系统的回归建模方法, 当应用在非线性系统中时, 较次要的一些 PLS 成分并不完全代表噪声, 相反它们中包含了许多反映系统非线性本质的信息。因此简单的舍弃这些成分将会导致模型的预测精度降低。在非线形系统建模方法中, 前向神经网络(FFNN 或 MLP)不仅理论相对比较完善, 而且在实际中也得到了广泛的应用验证。在 FFNN 的训练建模过程中, 一个很突出的问题是过拟合问题, 它通常发生在以下几种情况中: (1)选择的输入变量数目较多且存在严重相关性, 由此导致网络结构复杂(需要较多的隐层单元数); (2)相对于复杂的网络结构, 训练样本数较少, 导致网络参数不能唯一确定。过拟合将导致 FFNN 网络模型的泛化能力大大降低, 最终影响模型的预测精度。提高模型泛化能力(generalization)的方法很多, 比较常见的有 Regularization 规则化方法和 early stopping 的训练方法, 但实用效果并不理想。

4 基于 FFNN 的非线性 PLS 方法

基于上述考虑, 结合 PLS 方法和 FFNN 方法各

自的优势, 美国学者 S. J. Qin 和 T. J. McAvoy 提出了基于神经网络的 PLS 方法, 它除具有神经网络优秀的非线性连续系统建模能力之外, 还保留了线性 PLS 回归方法稳健的特性。另外, 在 NNPLS 的计算过程中, 一个多输入多输出(MIMO)系统的建模问题被分解成两部分, 分别对应线性的外部关系求解过程和若干 SISO 的单变量非线性内部关系求解过程(若干 SISO 前向神经网络的训练过程), 这种分解不仅带来了问题求解过程的简化, 同时由于网络结构由多入多出变为单入单出, 隐层单元数的减少也在一定程度上克服了单纯神经网络建模中的“过拟合”问题。

改进后非线性 PLS 方法的结构如图 1 所示, 其算法结构与算法(1)基本相似, 主要是有以下两点区别:

(1) 步骤 9 中的内部关系求解由原来的线性回归, 变为内部单入单出神经网络的训练过程, 即训练网络使下述误差函数最小: $J_h = \|u_h - N(t_h)\|^2$;

(2) 步骤 10 中的因变量残余信息矩阵计算式变为: $F_h = F_{h-1} - N(t_h) q_h^T$, 其中 $N(t_h)$ 代表 SISO 神经网络的函数映射关系。

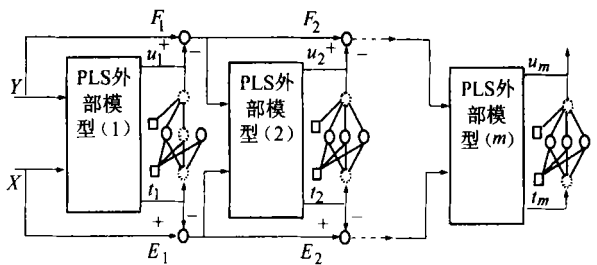


图 1 NNPLS 算法结构框图

5 基于 NNPLS 的氧量软测量模型设计实例

用于 NNPLS 软测量模型训练的数据, 主要由辽宁开原热电有限公司的 75 t/h 燃煤锅炉 DCS 系统采集而来。依赖这些历史运行数据, 分别建立了烟气含量的静态和动态软测量模型。

5.1 静态模型建立

主要步骤如下:

5.1.1 模型输入变量选择及分析

输入变量的选择主要依据对过程机理的分析, 同时为保证模型的预测精度, 尽量全面的选取了相关的信息变量。用于建立氧量软测量模型的输入变量共有 15 个, 分别是 8 个给粉机转速(反映各燃烧器的燃料量, 但各燃烧器燃烧出力不尽相同)、送风

挡板开度、引风挡板开度、排烟温度、炉膛温度、以及反映锅炉负荷情况的主蒸汽流量、压力、温度；除去前几个操作变量(manipulate variable)外,其余变量反

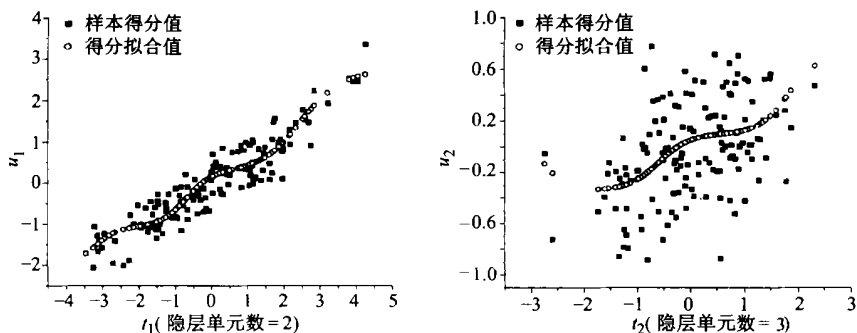


图2 前2组PLS成份间的非线性函数映射关系

映了锅炉的运行工况信息(state variable)。上述输入变量间存在明显的多重相关性；首先无论在自动和手动状态下,8个给粉机转速必然存在较强的相关性；另外一些状态变量与操作变量间也存在较强的相关关系；例如通过对训练样本集的统计分析,给粉机转速与主蒸汽流量、炉膛温度之间的相关系数分别为0.8273和0.7014。

5.1.2 稳态运行数据准备及数据预处理

对锅炉历史运行数据经过稳态判据过滤后,得到用于模型训练用的数据,经过简单的统计分析,去除数据中的奇异点,所用到的训练数据基本覆盖了该锅炉系统日常运行的所有工况,将提取的输入输出变量整理成数据矩阵的形式,并对其实施标准化处理(Normalize)。

5.1.3 模型训练过程

将经过预处理的数据样本共210个分为两组,分别为模型训练集(140个)和验证集(70个);另外验证数据集还用来通过Cross Validation的方法决定最终模型中的PLS成分提取个数;PLS成分间的FFNN训练(SISO对象),其首要问题是如何确定网络隐层单元个数,本文中仍然采用了Cross Validation的方法,在一定范围内优选网络结构;SISO网络的训练采用的是批处理的方式,算法选用的是共扼梯度法;为提高网络的预测能力,训练中采用了early stopping的方法,一旦网络的验证均方误差(Validation MSE)随着训练误差的减小反而增大时,停止训练过程,并返回对应验证均方误差最小的网络结构。

5.1.4 结果分析

由交叉验证方法最终获得PLSR和NNPLS方法提取的成分数分别为4和6;表1中列出了线性

PLSR方法各主要成分对因变量的解释情况;表2列出了三种方法各自的训练误差、验证预测误差;图2为前2组PLS成分间的非线性函数映射关系;图3为三种方法的预测性能比较;由表2和图3所示数据可以看出,PLSR方法对存在明显非线性特征的系统预测性能不很理想,而直接神经网络方法造成网络结构复杂,训练参数过多,虽然训练中采用了交叉验证的方法提高网络泛化能力,但预测精度仍不理想(尽管训练精度较其他方法为优);NNPLS方法正如前文讨论的一样,很好的结合了前两种方法各自的优势,具有最佳

的预测性能。

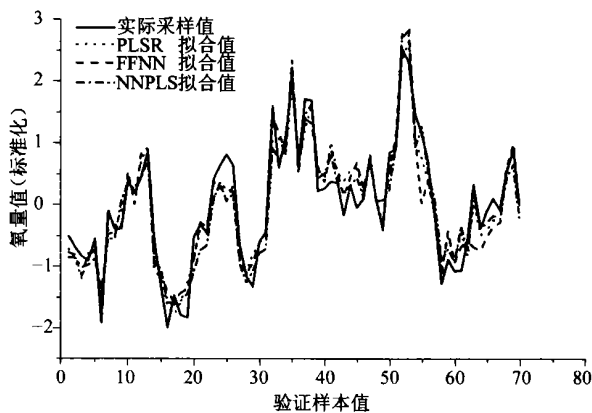


图3 三种方法预测精度比较

表1 PLS成份对因变量的解释情况 $Rd(t_h, y) = r^2(t_h, y)$

PLS 成分	解释能力 $Rd(t_h, y)$	累积解释能力 $\sum Rd_h$
t_1	0.704 0	0.704 0
t_2	0.138 0	0.842 0
t_3	0.075 2	0.917 2

表2 三种方法的训练误差、验证预测误差比较

	训练MSE	验证预测MSE
PLSR	0.085 3	0.107 2
FFNN	0.043 3	0.123 2
NNPLS	0.065 1	0.093 2

5.2 动态模型建立

5.2.1 动态模型与稳态模型比较

含氧量信息除可以帮助操作人员及时了解炉内燃烧情况外,锅炉送风控制回路中也都加入了氧量校正环节,用以调整风煤配比至最佳燃烧状态。若

要将推理模型 (inference model) 替代传统的信号变送器, 与自控系统结合组成推理控制器, 则要求模型能够反映变量动态变化信息 (动态系统建模)。

5.2.2 动态非线性 PLS 建模方法介绍

目前比较常见的动态系统扩展方法主要有三

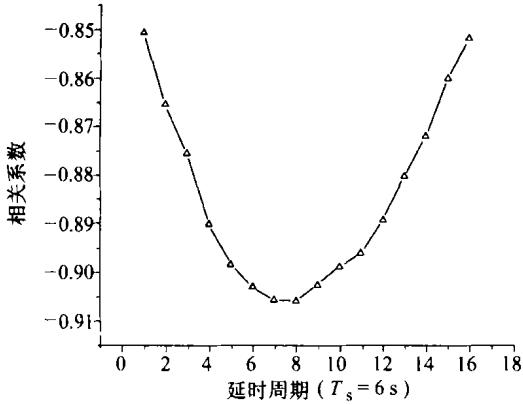


图 4 纯延时相关分析
(给粉机转速—氧量 $n_k=7$)

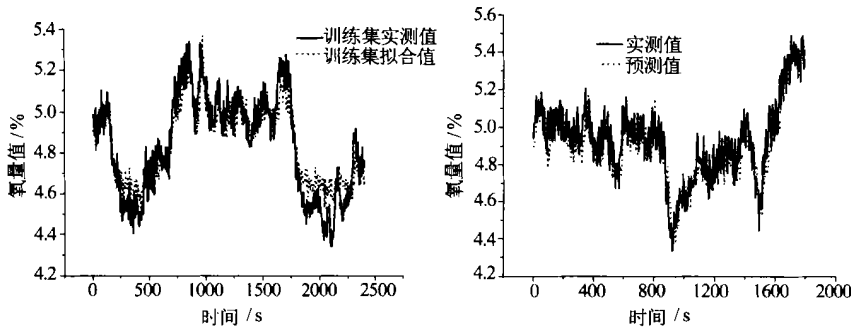


图 5 动态 NNPLS 方法氧量软测量模型训练(左)、验证(右)结果

种: (1) 直接在输入数据矩阵中加入输入输出变量的过去时刻值, 前述的建模方法和步骤仍然适用; 将输入数据经过滤波处理后 (滤去动态成分) 再利用常规 PLS 方法建模 (Kaspar, 1992); (3) 在输入输出数据的得分向量间, 用动态模型拟合取代原来的静态映射, 将一个 MIMO 系统的动态建模问题分割成若干 SISO 系统的动态模型辨识问题。根据研究对象的特点, 本文采用方法一作为氧量动态建模的主要方法。

5.2.3 模型输入变量选取

上一节中的稳态模型偏重于描述辅助变量 (锅炉稳态运行工况信息) 与主变量 (烟气含氧量) 之间的函数映射关系; 而动态模型的应用背景决定了其主要目的是反映操作变量 (manipulate variable)、外部扰动 (disturbance) 和被控变量 (quality variable) 之间

的动态关系, 因此输入变量的选取有所不同。本模型中用到的输入变量主要包括给粉机转速、送风门挡板和引风门挡板开度, 以及各自变量和因变量的时间延迟项。

5.2.4 训练过程

操作变量和扰动和延迟时间辨识采用的是相关分析的方法 (如图 4 所示)。而模型的阶次确定则采用了经验分析的方法 (3 阶系统), 因为对于非线性的 MISO 系统, 通常的模型定阶准则 (如 AIC 准则) 计算量大, 难于直接应用。训练集数据共 402 个样本 (40 min), 验证测试集数据样本共 302 个 (30 min), 采样时间间隔为 6 s, 单个网络的训练过程仍采用 early stopping 的技术以提高泛化能力。

5.2.5 结果分析

经过多次训练 (消除网络权系数初始值对训练结果的影响), 最终得到的氧量 NNPLS 动态模型具有良好的预测精度, 如图 5 所示。而利用常规的前向神经网络所建立的动态预测模型, 由于模型阶次较高造成的过参数化问题依然导致模型泛化能力不强、预测精度偏低。由于文章篇幅关系, 这里没有作详细的对比分析。

6 结论

由以上的方法介绍和实例对比分析可以看出, NNPLS 建模方法确实有效的综合了线性 PLSR 方法和非线性神经网络建模方法的优势, 所建模型无论静态或是动态都有良好的泛化能力; 将该方法应用于锅炉烟气含氧量的软测量, 具有预测精度高、适应范围广、计算方法简单等优点, 有很高的应用推广价值。

参考文献:

- [1] 陈德钊. 多元数据处理 [M]. 北京: 化学工业出版社, 1998.
- [2] 王惠文. 偏最小二乘回归方法及其应用 [M]. 北京: 国防工业出版社, 2000.
- [3] QIN S J, MCAVOY T J. Nonlinear PLS modeling using neural networks [J]. *Computer and Chemical Engineering*, 1992, 16: 379—391.
- [4] LAKSHMINA RAYANAN S. Modeling and control of multivariable processes: dynamic PLS approach [J]. *AIChE Journal* 1997, 43 (9): 2307—2322.

(辉 编辑)

various channels as well as the structural parameters of the heat exchanger. As a result, obtained were the fluid temperature distribution of various channels and the fluid temperature difference of neighboring channels. Furthermore, analyzed was the effect of the variation of fluid parameters, flow modes and structural parameters on the fluid temperature crossover of the neighboring channels. **Key words:** multi-stream plate-fin heat exchanger, temperature crossover, fin bypass, flow mode

烟气含氧量软测量新方法研究 = **The Study of a New Method Incorporating the Soft Sensing of Oxygen-content in Flue Gases** [刊, 汉] / LU Yong, XU Xiang-dong (Department of Thermal Engineering, Tsinghua University, Beijing, China, Post Code: 100084) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(6). — 614 ~ 617

In view of the high first cost of conventional oxygen-content analyzers for industrial applications, their high maintenance expenses and low durability the authors have on the basis of comparing several commonly used methods come up with a new method for measuring oxygen content in flue gases. The proposed method involves an oxygen-content soft sensing model set up through the use of a NNPLS (neural network partial least square) approach based on statistical analyses and neural network technology. It enjoys both the merits of PLSR (partial least square regression) and neural network technology, making it possible to identify a target model by utilizing historical process data. A simulation verification of the method has been conducted by using on-site industrial data. In addition, the simulation results are compared with traditional linear PLSR method and the direct neural network-based modeling method. The results of comparison indicate that the soft sensing model based on the NNPLS approach features a more effective generalizing ability. Furthermore, an extension of a static model to a dynamic one was also performed. **Key words:** soft sensing, partial two least squares, neural network, cross validation, generalizing ability

用预报误差校正的锅炉燃烧系统预测控制研究 = **A Study of the Predictive Control of a Boiler Combustion System through the Correction of a Forecast Error** [刊, 汉] / ZHU Xue-li, QI Wei-gui, LI Li-yan (School of Electric Engineering and Automation under the Harbin Institute of Technology, Harbin, China, Post Code: 150001) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(6). — 618 ~ 621

To improve the performance of a boiler-combustion control system, a dynamic matrix control (DMC) - based algorithm with the correction of a forecast error has been put forward to fulfill relevant control functions. After a brief description of the DMC composition and an internal-model control structure a model error is predicted based on a time sequence analysis, parameter estimation and an optimum forecast theory. Furthermore, by forecasting the model error and using the model forecast error to replace the model error the rolling optimization of a predictive control can be duly corrected. Finally, through the simulation tests of the predictive control for the boiler control system it is shown that the correction of the forecast error can result in a marked improvement in such characteristics as tracking ability, anti-interference and robustness when compared with an error correction algorithm in general. **Key words:** combustion system, predictive control, time sequence, forecast error

劳伦斯法在热工对象动态特性辨识中的应用 = **The Application of Lawrence Algorithm in the Identification of Dynamic Behavior of Thermodynamic Objects** [刊, 汉] / XU Hou-qian, JIANG Gui-zhen (Power Engineering College under the Nanjing University of Science & Technology, Nanjing, China, Post Code: 210094) // Journal of Engineering for Thermal Energy & Power. — 2002, 17(6). — 622 ~ 624

Described is the process of a transfer-function fitting performed through the use of Lawrence algorithm by way of frequency domain data. Furthermore, the above method was employed to conduct the fitting of transfer functions for a split-shaft gas turbine under three operating conditions. Under various operating conditions the results of fitting agree relatively well with those of experiments, testifying to the credibility of the Lawrence algorithm. In addition, by using a method, which com-