测试与故障诊断

文章编号:1671-4598(2020)08-0062-05 DOI:10.16526/j.cnki.11-4762/tp.2020.08.013

.013 **中图分类号:**TM62; TP274 **文献标识码:**A

基于 SVR 的燃煤机组 NOx 含量的软测量模型

李雅晶,辛妍丽

(华南理工大学电力学院,广州 510640)

摘要:作为火电厂燃煤脱硝工作的基础,选择性催化还原(SCR)脱硝反应器入口氮氧化物(NO_x)含量的测量至关重要; 针对难以准确现场实时测量 NO_x含量的问题,提出了基于回归型支持向量机(SVR)的软测量模型;首先对 SCR 反应器生成 NO_x的过程进行机理分析,并结合相关性分析、主成分分析等数据处理方法选取辅助变量,然后基于回归型支持向量机算法建 立模型,最后运用 BP 神经网络对模型进行检验,解决了 SCR 反应器入口 NO_x 的含量的难以准确预测问题;为 SCR 反应器入口 NO_x含量的实时、准确测量打下基础。

关键词:NOx 含量; 主成分分析; 回归型支持向量机; BP 神经网络; 软测量

Soft Sensor Model Based on SVR for the Measurement of NO_x Concentration of Coal—fired Unit

Li Yajing, Xin Yanli

(School of Electric Power Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: As a basis for denitration of coal—fired power plant, the measurement of nitrogen oxide (NO_X) content at the inlet of the Selective Catalytic Reduction (SCR) reactor is critical. In order to solve the problem that the NO_X content cannot be accurately measured in real time, a soft sensing model based on regression support vector machine (SVR) is proposed. Firstly, the process of generating NO_X at the inlet of SCR reactor is analyzed. Then auxiliary variables are selected by correlation analysis and principal component analysis, and the mathematics model based on support vector machine for regression algorithm is built. Finally, the model is tested by the method of BP artificial neural network. The proposed model lays the foundation for the real—time and accurate measurement of the NO_X concentration at the inlet of the SCR reactor.

Keywords: NO_x; PCA; SVR; BP neural network; soft sensor

0 引言

煤炭燃烧是造成我国大气污染的主要原因,也是出现 雾霾天气的主要原因^[1],影响人体健康^[2]。燃煤电厂是我国 煤炭主要消耗源^[3],因此对燃煤电厂进行除尘、脱硫、脱 硝等控污方法很重要。火电厂脱硝主要用选择性催化还原 (SCR)方式,即在催化剂催化下,用氨或其他还原剂将氮 氧化物 (NO_x)还原为氮气和水。现场存在着因 NO_x 测量 不准确而导致还原剂用量不正确,造成脱硝不完全或还原 剂浪费。需要的喷氨量根据反应器入口 NO_x含量与反应器 出口 NO_x含量差值计算出,因此喷氨量大小需要及时并准 确地测量。但是常规分析仪测量的 NO_x需要 40~60 s 的时 间,并不能满足电厂脱硝的要求^[4]。为解决上述问题,可 以通过寻找测量量之间的相关关系,采用间接测量的方法,

收稿日期:2020-01-07; 修回日期:2020-02-13。

基金项目:中央高校基本科研业务费专项资金(x2dl-D2181850)。

作者简介:李雅晶(1996-),女,河北邯郸人,硕士研究生,主要 从事软测量技术、电力系统调度方向的研究。

通讯作者:辛妍丽(1988-),女,河南南阳人,博士,助理研究员, 主要从事海上风电场暂态建模、内部过电压产生机理及防护措施方 向的研究。 建立 NO_x 排放量的预测模型。此方法为软测量技术的应用。

Brosillow 在 1978 年首次提出了由估计器和控制器组成 的推断控制思想^[5]。而软测量技术就源于推断估计器^[6]:即 通过分析变量间数学关系来建立目标函数值的预测模型, 实现对目标函数值的准确预测,进而有效控制系统。近些 年随着软测量技术的迅速发展,在工业领域的应用逐渐广 泛。软测量问题的解决过程主要分为3个步骤^[7]:采集及处 理数据、选取辅助变量、建立和校正模型。软测量建模是 软测量的重要步骤,近几年主要的建模方法有机理建 模^[8-10]、回归分析建模^[11]、状态分析建模^[12]、神经网络建 模^[13-15]、支持向量机建模^[16-17]等。

本文以火电厂燃煤锅炉 NO_x 含量作为研究对象,建立 了基于回归支持向量机(SVR)的软测量模型。不使用硬 件检测,预测 NO_x 排放浓度。首先通过对 SCR 反应器生成 NO_x 的过程机理分析,并结合相关性分析、主成分分析等 数据处理方法选取辅助变量,然后基于 SVR 算法建立软测 量模型,最后通过电厂锅炉不同工况下的运行数据建模测 试,运用 BP 神经网络对模型效果进行对比检验。

1 辅助变量的选择

1.1 机理分析

本模型的主导变量为脱硝 SCR 反应器入口的 NO_x 含

量,要确定影响此主导变量的辅助变量,首先分析电厂锅 炉 NOx 生成过程的影响因素:

1) 燃料特性: 燃料中氮的存在形式不同, NO_x 生成量 也相应变化; 煤挥发成分中的各种元素比也是一个重要影 响因素。

2) 过量空气系数:当空气分级时,可降低 NO_x 排放 量,随着一次风量减少,二次风量增加,氮被氧化的速度 降低,NO_x 排放量也随之下降。

3) 燃烧温度: 炉内燃烧温度越高, NO_x 排放量越大。

4) 一次风率:为了有效控制 NO_x 的含量,减弱 NO_x 生成环境,二次风送入点上部应维持富氧区,下部应维持 富燃料区。

5)负荷率:负荷率越大,给煤量越大,燃烧室及尾部 受热面处的烟温也相应提高,挥发分氮生成的 NO_x 相应 增加^[18]。

6) 风煤比:即总风量与总煤量的比值,反映了送粉的 效率。

7) A、B、C、D、E、F 磨的单台磨风煤比:即给煤机 的瞬时流量和磨煤机的入口一次风流量,反映了单台磨的 给煤效率。

8) 燃尽风门开度与所有风门开度总和比值:反映了空 气在炉膛内的流动情况。

结合电厂现场的实际测量数据,挑选出相关变量作为 备选的辅助变量。

1.2 数据的采集和处理

本文所用数据均采自某火电厂现场运行7日的数据。 由于数据在火电厂分布式控制系统(DSC)中已经进行了 滤波处理,因此在建模前不需要对其进行滤波。接下来, 对数据进行标准化,在建模过程中,避免因数据大小对样 本造成的影响,用以下公式对其进行标准化处理:

$$x'_{i} = \frac{x_{i} - \overline{x}}{\delta} \tag{1}$$

其中:

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{N} x_i \tag{2}$$

$$\delta = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2}$$
(3)

式中, N 为每组数据的总个数, x_i 为需处理的原始数据, x' 为标准化后的数据, \overline{x} 为每组数据的均值, δ 为每组数据的标准差。

1.3 相关分析

首先对备选的辅助变量原始变量和主导变量进行相关 分析。相关系数公式如下:

$$R_{xy} = \frac{Cov(x, y)}{\sqrt{D(x)}\sqrt{D(y)}} \tag{4}$$

式中, *x* 是备选辅助变量的测量值, *y* 是该辅助变量对应的 主导变量, 即 NOx 含量的值。*Cov(x*, *y*) 是两者的协方差。 *D(x)*,*D(y)* 分别是两者的方差。 选取相关系数大于 0.4 的变量作为原始辅助变量,结果如表 1 所示。

表 1 原始辅助变量				
序号	名称	序号	名称	
1	总风量	7	二次风率	
2	总煤量	8	B磨风煤比	
3	机组负荷	9	C磨风煤比	
4	风煤比	10	D磨风煤比	
5	空预器二次风温	11	锅炉氧量	
6	一次风率	12	燃尽风门开度与所有风门开度总 和比值	

由表1可看出,NO_x的生成影响因素复杂,影响主导 变量的原始辅助变量较多。

1.4 主成分分析

由于表1中筛选出的12个原始辅助变量之间的相关性 很强,而且变量的个数较多。如果直接采用这些原始辅助 变量作为软测量模型的辅助变量,不但会增加训练的成本, 而且会造成信息的重叠,影响预测结果的客观性。为了将 上述的相关性很高的变量转化成彼此互相独立或不相关的 变量,并且减少变量个数降低训练成本,考虑对原始辅助 变量进行主成分分析。

通过 Matlab 软件对 12 个原始辅助变量进行主成分分 析,分析其相关系数矩阵,前几个特征根及其贡献率见 表 2。

表 2 原始辅助变量主成分分析结果

序号	特征根	贡献率	累计贡献率
1	8.542 8	71.190 0	71.190 0
2	1.918 1	15.984 1	87.174 1
3	0.653 4	5.444 9	92.619 0
4	0.519 8	4.3317	96.950 7
5	0.105 6	0.879 6	97.830 3
6	0.084 0	0.699 3	98.529 6
7	0.071 3	0.594 3	99.123 9
8	0.059 7	0.497 5	99.621 4
9	0.039 4	0.328 4	99.949 8
10	0.003 9	0.032 8	99.982 6

从表 2 可以看出,前 3 个特征根的累积贡献率超过 90%,主成分分析的效果明显。前 6 个主成分的主成分达 到 98%以上,因此选取前 6 个主成分作为本文软测量模型 的辅助变量。

1.5 辅助变量的确定

为具体说明选取的 6 个辅助变量与原始辅助变量的关系,表 3 列出了标准化变量的前 6 个特征根对应的特征向量。其中, x_1 , x_2 ,..., x_{12} 表示表 1 中的 12 个原始变量标准化后的数值, p_1 , p_2 ,..., p_6 表示选取的辅助变量,即前 6 个主成分。

表 3 前 6 个主成分(辅助变量)对应的特征向量

	p1	р 2	р3	p 4	р 5	р б
x 1	-0.311 3	-0.311 2	-0.3061	0.322 8	0.301 5	-0.295 8
\mathbf{x}^2	0.270 2	0.215 3	0.309 3	0.151 1	0.133 0	0.304 2
x 3	-0.154 9	-0.1776	-0.1221	-0.0801	-0.236 3	-0.1215
x 4	-6.0271	-0.307 2	0.016 2	0.218 7	-0.538 4	0.132 5
x 5	0.010 9	0.0125 3	-0.034 8	-0.036 5	-0.048 3	0.417 9
x 6	0.300 4	0.382 8	0.111 7	-0.3586	-0.100 6	-0.3550
x 7	-0.147 6	-0.088 7	-0.0464	-0.2254	-0.081 8	-0.325 9
x 8	0.056 0	0.062 8	0.054 2	-0.178 7	-0.061 6	-0.491 3
x 9	0.390 1	0.024 4	0.228 9	0.715 2	-0.059 4	-0.3678
x 10	-0.262 2	-0.4035	0.835 6	-0.1553	0.083 3	-0.056 0
x11	-0.472 1	0.438 2	0.112 8	0.180 9	0.438 3	-0.022 6
x 12	0.496 2	-0.465 7	-0.115 5	-0.194 1	0.020 4	0.565 9

每个选取作为辅助变量的主成分 pi 是 12 个原始的辅助 变量 x_i 的线性组合,具体系数为对应的特征向量,如表 3 所示。可以看出, 第1主成分主要反映了总风量、总煤量、 燃尽风门开度与所有风门开度总和的比值和锅炉氧量,第2 主成分主要反映了风煤比和燃净风门开度与所有风门开度 总和的比值以及锅炉氧量,第3主成分主要反映了D磨风 煤比, 第4 主成分主要反映了 C 磨风煤比, 第5 主成分主 要反映了空预器二次风温、B磨风煤比以及一次风率和二 次风率,第6主成分主要反映了机组负荷、风煤比、燃净 风门开度与所有风门开度总和的比值以及锅炉氧量。

由此可得,选取的前6个主成分作为软测量模型的辅助 变量,既全面的反映了原始的12个辅助变量,又消除了原始 变量之间的相关性,而且减少了变量的个数节约了训练的成 本。因此,本文 NO_x 软测量模型辅助变量的选择恰当。

回归型支持向量机(SVR)的原理 2

SVR 是在支持向量机分类的基础上,引入 ε 线性不敏 感损失函数,其基本思想是寻找一个最优分面,使所有训 练样本离这个分类面的误差最小。如图1所示。





通过非线性映射函数 $\varphi(\bullet)$ 将训练集样本 { $(x_i, y_i), i =$ $1,2,\dots,l$ } 的输入列向量 x 映射到高维特征空间 R^d , 建立 SVR 的线性回归函数:

$$f(x) = \omega \varphi(x) + b \tag{5}$$

其中: ω 为法向量,b是位移项,f(x)即为回归函数返回 的预测值。

定义 ε 线性不敏感损失函数 $L(f(x), y, \varepsilon)$,如图 2 所 示。表示如果预测值 f(x) 与真实值 y 的差不大于 ϵ ,则损

失为0。

寻找最优的分类面转化为寻找最优的ω和6,引入松弛 变量 ξ_i 、 ξ_i^* ,可把优化的目标函数具体表示为:

$$\min \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^{l} (\boldsymbol{\xi}_i + \boldsymbol{\xi}_i^*)$$
 (6)

其中:C是惩罚因子,表示训练误差超过误差要求 ε 的 惩罚系数; ω^2 是置信区间,与模型的复杂度有关; l 是训练 样本个数。



图 2 ε线性不敏感损失函数

可引入拉格朗日函数并转化为对偶形式求解,最终得 到最优的线性回归函数即为 SVR 预测函数:

因此,得到的回归函数即软测量模型为:

$$f(x) = \omega^* \bar{\omega}(x) + b^* = \sum_{i=1}^{l} (\alpha_i - \alpha_i^*) K(x_i, x_j) + b^*$$
(7)

式中,只有部分参数 ($\alpha_i - \alpha_i^*$) 不等于零,其对应的样本 x_i 即为本模型中的支持向量。其中核函数 $K(x_i, x_i) =$ $\varphi(x_i)\varphi(x_i)$ 应满足 Mercer 条件, 核函数不同, 构造的支持 向量机不同。常用核函数有多项式核函数、RBF 核函数、 线性核函数、Sigmoid 核函数。

此外,由式(7)可以看出,SVR最终的函数形式的结 构如图3所示。

在图 3 中,每个支持向量对应一个中间节点,中间节 点的线性组合即为输出。

3 基于 SVR 的软测量模型

3.1 产生训练集/测试集

SCR 反应器入口 NO_x 含量的训练数据来源是火电厂现 场 DCS 采集到每秒的数据,共 201 600 组。为减少软测量 模型的训练成本和时间,以20s为长度等间隔取值,构成 10080个新数组,使数据的变化更明显。然后对标准化后 原始的辅助变量进行主成分分析,取前六个主成分作为最 终的辅助变量,为不失一般性,选取前6500组数据作为样 本产生训练集,剩余的3580组数据作为产生测试集的样 本,来对模型的测试集进行评价。

3.2 构建/训练 SVR 模型

首先对于作为辅助变量的前 6 个主成分数据集,进行 再次归一化。不同的核函数和模型参数对 SVR 模型的性能 影响很大。因此,合适的核函数及参数组合的选择在 SVR 模型建立中极为重要。本模型需要选取的参数有两个,分 别是惩罚因子 C 和 RBF 核函数中的方差 g 。惩罚因子 C 和 RBF 核函数中的方差 g 的取值对 SVR 软测量模型的性能有 着重要影响,惩罚因子 C 用来控制样本偏差和机器泛化能 力之间的关系, RBF 核函数中的方差 g 太大或太小会造成 对样本数据的过学习或欠学习。

对于核函数的选择,本文采用 RBF 核函数。对于选择 最优的参数 g (即 RBF 核函数中的方差),和参数 C (即惩罚 因子),本文采用交叉验证的方法。特别说明的是,因为惩 罚因子参数 C 越大,最终得到的支持向量越多,计算量也 就随之增大。因此,当模型的性能相同时,为了节约运算 时间,优先选择惩罚因子参数 C 较小的参数组合。





3.3 仿真测试

本文选取的工具是 Matlab 中 libsvm 软件包,对 SVR 进行仿真测试,得到测试集和训练集的预测值,并且计算 测试集的均方误差 *E* 和决定系数 *R*²,具体计算的公式如式 (8)、(9) 所示:

$$E = \frac{1}{l} \sum_{i=1}^{l} (y_i^* - y_i)^2$$
(8)

$$R^{2} = \frac{\left(l\sum_{i=1}^{l} y_{i}^{*} y_{i} - \sum_{i=1}^{l} y_{i}^{*} \sum_{i=1}^{l} y_{i}\right)^{2}}{\left(l\sum_{i=1}^{l} y_{i}^{*} - \left(\sum_{i=1}^{l} y_{i}^{*}\right)^{2}\right)\left(l\sum_{i=1}^{l} y_{i}^{2} - \left(\sum_{i=1}^{l} y_{i}\right)^{2}\right)}$$
(9)

式中,*l* 是测试集的样本个数,*y_i* 表示第*i* 个样本预测值, *y_i** 表示第*i* 个样本真实值

均方误差 E 表示测量的精密度, E 越接近零测量的精度 越高; 决定系数 R² 表示模型的泛化能力, R² 越接近 1, 表 明模型对样本的依赖性越低,泛化能力越强。

此外,还要特别说明的是,在仿真测试时,对于直接 得到的训练集和样本集的预测值,即主导变量 NO_x 的预测 量,因为之前对数据已经做过了归一化处理,所以这里最 后的预测结果还要进行反归一化处理。

4 模型验证与检验

4.1 软测量模型的求解

因为随机产生训练集和测试集,所以每次运行的结果 会略有差异。由于 SVR 算法最终转化为一个二次规划问题,理论上可以避免陷入局部最优,得到全局最优解。因此,每次运行的结果虽略有不同,但结果基本稳定。某一 次运行的训练集、测试集结果分别如图 4~图 5 所示。



训练集均方误差的计算结果是 0.009 75,结合图 4 可 知,训练过程中的精密度很高;训练集决定系数的计算结 果是 0.889 56,表明训练的结果对训练样本的依赖性很小, 泛化能力很好。

测试集预测值和真实值具体对比结果如图 5 所示,测 试集的均方误差是 0.013 40,表示该软测量预测模型的精 密度高;训练集的决定系数是 0.857 95,表明模型的预测 的结果对样本的依赖性很小,该模型的泛化能力理想。

4.2 与 BP 神经网络模型的性能对比

软测量模型的建模方法中,神经网络也是做回归拟合 预测问题的选择之一,下面用 BP 神经网络的建模方法对该 SVR 软测量模型结果进行检验,并进行两种方法的性能 对比。

建立的 BP 神经网络模型对测试集预测的结果如图 6 所示,与4.1 节对应的结果为同一次运行所得。

经计算本次 BP 神经网络模型运行的均方根误差为 0.025 72,结合图 6 可以看出精度较高,说明此次的运行结 果没有陷入局部最优,结果有效。

对比图 6 和图 5 发现, SVR 的均方根误差小于 BP 神经 网络,说明在 NO_x 软测量的问题中 SVR 的精度较高; SVR 的决定系数大于 BP 神经网络的 0.816 81,说明对于此问题,SVR 的泛化能力优于 BP 神经网络。



4.3 不同核函数对模型性能的影响

从 SVR 的原理可知,不同核函数性能对软测量模型性 能的影响不同。本文之前的工作中,选取的是 RBF 核函数, 为了检验核函数选择的准确性,这里用某次随机产生的测 试机和训练集进行对比试验,具体结果如表 4 所示。

拉函粉米刊	训练	东集	测试集		
核函数关型	E	R^2	Ε	R^2	
线性	0.015 6	0.8598	0.0437	0.844 3	
多项式	0.013 2	0.879 4	0.027 4	0.783 5	
RBF	0.009 8	0.889 6	0.013 4	0.858 0	
Sigmoid	0.022 8	0.857 4	0.036 9	0.826 8	

表 4 不同核函数对软测量模型的性能影响

如表 4 所示, RBF 核函数对应的模型泛化能力最好, 和 Sigmoid 及线性核函数相比,虽然多项式对应的模型训练 性能较好,但泛化能力较差。因此,本文选取的 RBF 核函 数较为合适。

5 结束语

本文利用相关分析、机理分析和主成分分析选出影响 电站锅炉 NO_x 含量的辅助变量,使得辅助变量的选取更加 具有统计意义。由于现场检测 NO_x 含量分析仪具有一定的 延时特性,在数据处理时增加时序变换,使建立的模型更 加符合实际情况。仿真结果显示利用回归型支持向量机 (SVR)建立的预测模型具有很好的精度和泛化能力。本文 提出的软测量模型能够及时反映并预测 NO_x 含量,为解决 火电厂 SCR 反应器入口 NO_x 含量难以实时在线测量的问题 提供了参考,而且可以拓展应用到火电厂烟气含氧量、锅 炉炉膛温度、排放烟气含湿量等问题,甚至可以应用到工 业上,尤其是小样本数据的,很多难以实时在线测量数据 的问题。

参考文献:

- [1] Zhao Y, Zhang J, Nielsen C P. The effects of energy paths and emission controls and standards on future trends in China" s emissions of primary air pollutants [J]. Atmospheric Chemistry and Physics, 2014, 14 (17): 8849-8868.
- [2] Ashraf A, Butt A, Khalid I, et al. Smog analysis and its effect on reported ocular surface diseases: a case study of 2016 smog event of Lahore [J]. Atmospheric Environment, 2019, 198: 257-264.
- [3] 国网能源研究院. 2014 中国发电能源供需与电源发展分析报告 [R]. 北京:中国电力出版社, 2014.
- [4] Korpela T, Kumpulainen P, Majanne Y, et al. Indirect NO_x emission monitoring in natural gas fired boilers [J]. Control Eng. Pract., 2017, 65: 11-25.
- [5] Brosilow C, Tong M. Inferential control of processes: Part II. The structure and dynamics of inferential control systems [J]. Aiche Journal, 1978, 24 (3): 485 - 509.
- [6] 朱学峰. 软测量技术及其应用 [J]. 华南理工大学学报 (自然 科学版), 2012 (11): 61-67.
- [7] 俞金寿. 软测量技术及其应用 [J]. 自动化仪表, 2008 (1): 1 -7.
- [8] 孙丽娜,黄永红,蒋星红,等. 基于 KPCA-DFNN 海洋微生物发酵过程软测量建模 [J]. 计算机测量与控制,2018,26
 (7):41-43,98.
- [9] 王美琪, 陈恩利, 刘鹏飞, 等. 融合机理与数据的篦冷机温度 软测量模型 [J]. 仪器仪表学报, 2018, 39 (6): 182-188.
- [10] 李翔宇,高宪文,侯延彬. 基于示功图的抽油井动液面软测量 机理建模 [J]. 控制工程,2018,25 (3):464-471.
- [11] 赵 帅, 史旭东, 熊伟丽. 一种分层集成的高斯过程回归软测 量建模方法 [J]. 系统仿真报, 2019, 31 (10): 2042-2051.
- [12] 俞 华,王天正,芦竹茂,等.基于软测量的图谱变化检测技术在变电站状态智能分析中的研究与应用[J].现代电子技术,2017,40(13):171-174.
- [13] Ro K, Kim S, Park W, et al. Development of a wearable soft sensor system for measuring finger motions [A]. 2016 16th International Conference on Control, Automation and Systems (ICCAS) [C]. IEEE, 2016.
- [14] 朱湘临,宋 彦,王 博,等.基于改进布谷鸟算法-BP神 经网络的松茸发酵过程软测量建模[J].计算机测量与控制, 2019,27 (5):39-43.
- [15] Souza F, Member S, et al. Online mixture of univariate linear regression models for adaptive soft sensors [J]. IEEE Transactions on Industrial Informatics, 2013, 10 (2): 937-945.
- [16] 丁续达,金秀章,张 扬. 基于最小二乘支持向量机的改进型在 线 NO_x 预测模型 [J]. 热力发电, 2019,48 (1):61-67.
- [17] 姜哲宇,刘元清,朱湘临,等. 基于 KFCM 和 AMDE-LSSVM 的软测量建模方法 [J]. 计算机测量与控制,2018,26(8):46-50,92.
- [18] 任建兴, 翟晓敏, 傅坚刚, 等. 火电厂氮氧化物的生成和控制 [J]. 上海电力学院学报, 2002 (3): 19-23.