

doi: 10.3969/j.issn.1005-3026.2017.05.002

# 基于 M-SVR 与 RVFLNs 的高炉十字测温中心温度估计

周平, 尤磊, 刘记平, 张兴

(东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819)

**摘要:** 由于高炉中心温度较高, 十字测温中心位置传感器极易损坏, 并且更换周期长, 因而导致无法及时判断炉顶煤气流分布. 采用多输出支持向量回归(M-SVR)和随机神经网络(RVFLNs)两种数据驱动智能建模方法建立高炉十字测温中心带温度估计模型, 并基于实际工业数据对建立的模型进行验证和比较分析. 结果表明, 在样本数量较小时, M-SVR模型和RVFLNs模型都具有较好的温度估计效果, 但当样本数量充足时, M-SVR模型的泛化性能和估计精度更优于RVFLNs模型.

**关键词:** 高炉炼铁; 十字测温; 温度估计; 多输出支持向量回归机; 随机神经网络

**中图分类号:** TP 182 **文献标志码:** A **文章编号:** 1005-3026(2017)05-0614-06

## Centre Temperature Estimation of Blast Furnace Cross Temperature Measuring Based on M-SVR and RVFLNs

ZHOU Ping, YOU Lei, LIU Ji-ping, ZHANG Xing

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China. Corresponding author: ZHOU Ping, E-mail: zhouping@mail.neu.edu.cn)

**Abstract:** Due to the high temperature in the middle of blast furnace, the central position sensor of the cross temperature measuring is very easy to be damaged, and the replacement period is always long, resulting in the gas flow distribution not being observed in time. To this end, two kinds of data-based intelligent modeling methods of multi-output support vector regression machine (M-SVR) and random vector functional-link networks (RVFLNs) were used to establish the temperature estimation model of cross temperature measuring center of blast furnace. Finally, the temperature estimation model based on industrial data was verified and compared. The results show that both M-SVR model and RVFLNs model have good temperature estimation effect when the sample size is small. However, when the sample size is large enough, the generalization performance and estimation accuracy of M-SVR model is better than those of the RVFLNs model.

**Key words:** blast furnace ironmaking; cross temperature measuring; temperature estimation; multi-output support vector regression (M-SVR); random vector functional-link networks (RVFLNs)

高炉炼铁作为钢铁生产的主要上游环节,其生产的铁水质量和产量决定着钢铁的产量和生产成本. 目前,对高炉炼铁的主要研究工作仍是保证高炉的稳定、顺行、高产,具体来说分为对高炉的上部调剂研究和下部调剂研究. 下部调剂主

要根据富氧顶压、喷煤、风温等高炉参数来调节铁水质量,这部分研究多集中在铁水温度、硅、磷、硫含量等的预报上<sup>[1-2]</sup>. 上部调剂主要是调整布料制度来控制炉料的分布,以此来调节煤气流的分布和煤气的利用率. 现今对高炉下部调剂研究较

收稿日期: 2015-12-25

基金项目: 国家自然科学基金资助项目(61473064, 61290323, 61333007, 61290321, 61621004); 中央高校基本科研业务费专项资金资助项目(N160805001, N160801001); 辽宁省教育厅科技项目(L20150186).

作者简介: 周平(1980-),男,湖南邵阳人,东北大学副教授,博士.

多,而对上部调剂研究较少.

### 1 十字测温

十字测温作为广泛使用的高炉炉顶温度监测装置,可以反映高炉炉顶煤气流的分布<sup>[3]</sup>.高炉操作人员根据煤气流的分布来调整布料制度,使炉料合理分布.如图 1 所示,十字测温装置通常由 4 个测温臂、21 个测温点组成,用来测量炉喉主要点位置的温度.准确地监测十字测温温度,估计

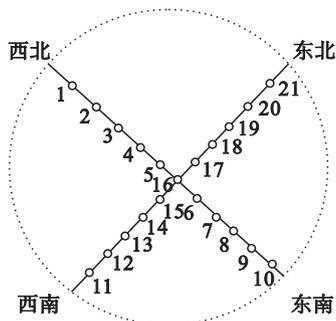


图 1 十字测温装置图

Fig. 1 Cross temperature measuring device

炉内煤气流分布,及时调整布料制度,能够辅助高炉稳定顺行,达到高产高效的效果.然而,由于高炉中心带温度过高,导致十字测温装置的传感器寿命较短,特别是高炉中心的 5 个测温点传感器极易损坏.图 2 为某高炉十字测温点 6 和 15 的热电偶传感器损坏现状.由于十字测温传感器更换周期较长,更换不及时会严重影响高炉操作人员对高炉内部煤气流分布的判断,导致无法及时地调整布料矩阵,降低了高炉的生产效率.本文针对上述问题对十字测温装置靠近中心的 5 个测温点(测温点 5,6,16,15,17)的温度进行估计,保证在高炉中心温度点传感器有损坏时,操作人员可以根据温度的估计值来判断炉内煤气流分布,及时调整布料矩阵.

高炉是一个多变量、强耦合、大滞后的系统,十字测温点温度受多种因素的影响,准确的机理

模型难以建立.本文基于数据驱动建模技术,采用多输出支持向量回归(M - SVR)和随机神经网络(RVFLNs)算法,建立高炉十字测温中心 5 点温度的估计模型.

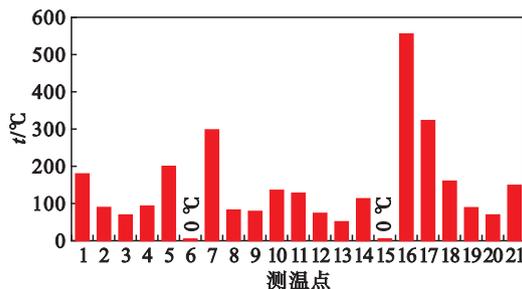


图 2 带有损坏点的十字测温实时温度

Fig. 2 Real time cross temperature with damage point

### 2 基于多元统计分析的建模输入变量选取

影响高炉十字测温中心 5 点温度的因素很多,主要包括十字测温其他温度点温度、顶温、雷达料线、顶压等 30 个变量.若 30 个变量同时作为输入变量则会造成信息冗余和模型计算量过大,严重影响模型的计算效率和估计精度.本文选取 5 000 组高炉实际生产数据利用多元统计分析对高炉采集到的实际数据进行因子分析和相关性分析<sup>[2,4]</sup>.对 5 个输出变量进行因子分析得到一个可以代表其他 5 个输出的主因子,再将得到的主因子与其余 30 个变量进行相关性分析,如表 1 所示.

根据以上对输入变量的分析并结合工程实际,选取相关性在 0.5 以上的 9 个因子作为输入变量:十字测温点 3,4,8,10,19,顶温东南、顶温西北、顶温西南、顶温西北.设 9 个输入变量为  $X_1(t), \dots, X_9(t)$ , 设 5 个输出变量为  $Y_1(t), \dots, Y_5(t)$ , 考虑到高炉炼铁过程的动态特性以及输入输出之间的时序和滞后关系,将上一时刻的输入和输出数据也作为输入变量<sup>[5]</sup>,从而构成 23 输入 5 输出的估计模型.

表 1 输出变量主因子相关性分析

Table 1 Correlation analysis of main factors of output variables

输入变量	T1	T2	T3	T4	T7	T8	T9	T10	T11	T12
相关性	0.377**	0.434**	0.549**	0.598**	0.344**	0.544**	0.087**	0.556**	0.448**	0.233**
输入变量	T13	T14	T18	T19	T20	T21	雷达料线	冷风流量	风温电偶	风温红外
相关性	0.466**	0.361**	0.401**	0.538**	0.270**	0.323**	0.463**	-0.141**	-0.031*	-0.045**
输入变量	热风压力	顶压	压差	透气性	鼓风湿度	富氧流量	顶温东南	顶温西北	顶温东北	顶温西南
相关性	-0.055**	-0.079**	-0.033*	-0.035*	-0.129**	0.041**	0.802**	0.834**	0.782**	0.907**

注: \*\* 表示在 0.01 水平显著相关; \* 表示在 0.05 水平显著相关; T1, T2, ..., T21 分别表示测温点 1, 2, ..., 21.

### 3 十字测温中心温度估计建模算法

#### 3.1 基于 M-SVR 的十字测温中心温度估计算法

SVR 是一种基于统计学习理论的机器学习方法,由于其在小样本情况下仍具有较好的建模和泛化能力,因而被广泛用于工业过程的数据驱动建模. M-SVR 是 SVR 的发展和改进,通过在超球空间重新定义  $\varepsilon$  不敏感损失函数,将多个输出同时纳入风险评估,最终采用迭代变权最小二乘求解对偶问题,从而获得多输入多输出的估计模型<sup>[5-6]</sup>. M-SVR 基本算法如下:

仿照单输出支持向量机的方法构造目标函数,求取  $w, b$  寻找最大分类面,最小化目标函数:

$$L_p(w, b) = \frac{1}{2} \sum_{j=1}^N \|w^j\|^2 + C \sum_{i=1}^L L(u_i). \quad (1)$$

$\varepsilon$  不敏感损失函数可扩展成多维形式,用输出变量误差的  $L_2$  范数代替原单维误差,定义:

$$L(u) = \begin{cases} 0, & u < \varepsilon; \\ u^2 - 2u\varepsilon + \varepsilon^2, & u \geq \varepsilon. \end{cases} \quad (2)$$

式中:  $u_i = \|e_i\| = \sqrt{e_i^T e_i}$ ;  $e_i = y_i - \tilde{y}_i = y_i - \varphi(x_i) \times w_i - b_i$ ;  $\varepsilon$  为允许偏差;  $C$  为惩罚因子;  $\varphi(x_i)$  为数据  $x_i$  通过 Sigmoid 核函数映射到高维 Hilbert 空间后的数据. 由于 M-SVR 不能直接转化为二次规划对偶问题求解,所以把目标函数展开为离散泰勒形式:

$$L'_p(w, b) = \frac{1}{2} \sum_{j=1}^N \|w^j\|^2 + C \left( \sum_{i=1}^L L(u_i^k) + \frac{dL(u)}{du} \Big|_{u_i^k} \frac{(e_i^k)^T}{u_i^k} [e_i - e_i^k] \right). \quad (3)$$

式中:  $e_i^k = y_i - \tilde{y}_i^k = y_i - \varphi(x_i) w_i^k - b_i^k$ ;  $u_i^k = \|e_i^k\| = \sqrt{(e_i^k)^T e_i^k}$ ;  $k$  为迭代次数. 随迭代次数增加,  $e_i^k \rightarrow e_i$ , 得到一个近似的二次规划形式:

$$L'_p(w, b) \approx \frac{1}{2} \sum_{j=1}^N \|w^j\|^2 + C \left( \sum_{i=1}^L L(u_i^k) + \frac{dL(u)}{du} \Big|_{u_i^k} \frac{u_i^2 - (u_i^k)^2}{2u_i^k} \right) = \frac{1}{2} \sum_{j=1}^N \|w^j\|^2 + \frac{1}{2} \sum_{i=1}^L a_i u_i^2 + CT. \quad (4)$$

式中:  $a_i = \frac{C dL(u)}{u_i^k du} \Big|_{u_i^k} = \begin{cases} 0, & u_i^k < \varepsilon; \\ \frac{2C(u_i^k - \varepsilon)}{u_i^k}, & u_i^k \geq \varepsilon. \end{cases}$

$T = \sum_{i=1}^L L(u_i^k) - \frac{dL(u)}{du} \Big|_{u_i^k} \frac{(u_i^k)^2}{2u_i^k}$  为常量. 由于

$L_p(w, b)$  与  $L'_p(w, b)$  具有相同极值, 可通过对  $L'_p(w, b)$  求导数得到, 运用线性搜索方法即可获取  $L_p(w, b)$  的最小值, 确定分类面, 得到最优值对应的  $w, b$ . 再依据式  $\tilde{y}_i = \varphi(x_i) w + b$  得到估计值.

#### 3.2 基于 RVFLNs 的十字测温中心温度估计算法

RVFLNs 由 Pao 于 1992 年在前馈神经网络强大的辨识能力基础上提出的<sup>[7]</sup>. 不同于传统的基于梯度下降的神经网络需要通过误差的反向传播迭代寻优, RVFLNs<sup>[8]</sup> 可以保证逼近任意连续函数的前提下, 随机给定隐含层的权值  $w$  和阈值  $b$ , 通过正则化原则计算输出权值建立模型. RVFLNs 算法如下:

给定  $N$  组数据样本  $(x_i, y_i)$ ,  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in \mathbf{R}^n$ ,  $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T \in \mathbf{R}^m$ ,  $L$  个隐含层节点, 激活函数为  $g(x)$  (本文选择 Sigmoid 函数), 依据 Schmidt<sup>[9]</sup> 指出的  $[-1, 1]$  区间内随机选取隐含层节点参数输入权向量  $w_i$  和阈值  $b_i$ , 得隐含层的输出矩阵为  $A = g(\langle w, x \rangle + b)$ . 从而 RVFLNs 的输出为

$$Y = \sum_{i=1}^L \beta_i g(\langle w_i, x_j \rangle + b_i). \quad (5)$$

式中:  $i = 1, 2, \dots, L$ ;  $j = 1, 2, \dots, N$ .

式(5)写成矩阵简化形式为

$$Y = A\beta. \quad (6)$$

式中:  $Y$  为期望输出;  $\beta$  为输出权值. RVFLNs 训练过程就是得到合适的  $\hat{w}_i, \hat{\beta}_i$  和  $\hat{b}_i$ , 使得

$$\|A(\hat{w}_i, \hat{b}_i)\hat{\beta}_i - Y\| = \min_{w_i, \beta_i, b_i} \|A(w_i, b_i)\beta_i - Y\|. \quad (7)$$

如果在网络学习过程中, 隐含层节点数  $L$  和学习样本数  $N$  相同时, 输出矩阵  $A$  是方阵且可逆, 此时单隐层随机神经网络可以以 0 误差逼近训练样本. 然而, 通常情况下隐含层节点数会远远小于训练样本数, 即  $L < N$ , 此时  $A$  不是方阵, 这时可能得不到  $\beta, w, b$  使  $A\beta = Y$  成立. 那么就需要采用最小二乘方法求解上述系统  $\hat{\beta} = A^\dagger Y = (A^T A)^{-1} A^T Y$ ,  $A^\dagger$  是根据 Moore-Penrose 产生的  $A$  的伪逆矩阵.

### 4 工业试验

为了验证建立温度估计模型的有效性, 选取某炼铁厂 2 号高炉 2014 年 9 月实际生产数据进行验证. 在正常生产情况下, 选取训练样本 700 组 (采样周期 10 s), 测试样本 400 组. 根据以上对输

入变量的分析,选取顶温等 9 个变量作为输入变量.考虑到测量装置的随机噪声和数据量纲对模型的影响,因此对数据进行离群点的剔除和归一化处理.同时,为了验证 M-SVR 和 RVFLNs 在高炉十字测温建模效果的精度,将其建模效果分别与常见线性回归建模方法偏最小二乘回归<sup>[10]</sup>(PLS)进行对比.PLS 在建模过程中集成主成分分析、典型相关分析和线性回归分析的特点,特别适用于当输入输出变量个数较多,且都存在较强的相关性的建模.

#### 4.1 M-SVR 模型估计效果

根据以上描述的 M-SVR 算法,需要对 M-SVR 的两个参数,即伸缩变量  $\sigma$  和误差惩罚因子  $C$  进行设定.通常采取试凑法来选取两个参数,本文通过遗传算法<sup>[11-12]</sup>对两个参数的选取进行优化,经过遗传计算得到当伸缩变量  $\sigma = 0.6$ ,误差惩罚因子  $C = 0.5$ ,达到最佳估计效果.应用 PLS 与 M-SVR 进行比较,图 3 为 M-SVR 模型对测试样本的测试效果,通过仿真结果明显可以看出, M-SVR 模型具有良好的估计精度,估计值与实际值已非常吻合,同时可以看出, M-SVR 的建模和泛化性能都明显优于线性建模方法 PLS.

#### 4.2 RVFLNs 模型估计效果

选择同样的训练样本和测试样本.通过多次实验,确定当模型的隐含层节点个数选择为 50 时取得最好的建模效果.最终构成一个 23-50-5 的随机神经网络结构,为了体现 RVFLNs 在十字测温温度估计的建模优势,仍采用常用 PLS 模型进行对比.图 4 为 RVFLNs 模型的测试效果.可以看出, RVFLNs 模型也具有良好的估计效果,模型估计值与实际值变化趋势一致且误差较小,

估计精度也明显优于 PLS.

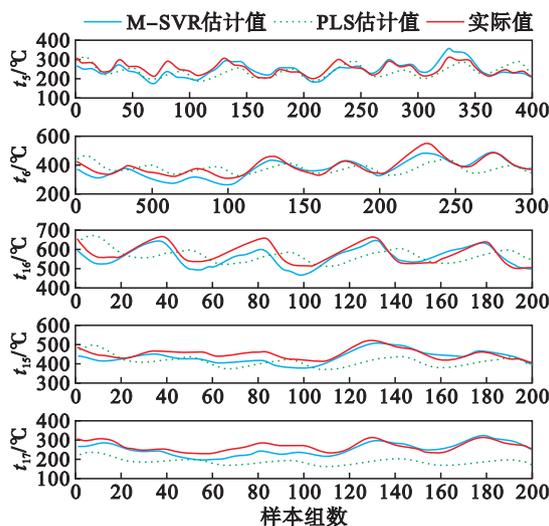


图 3 基于 M-SVR 的十字测温测试效果

Fig. 3 Test results of cross temperature measuring based on M-SVR

#### 4.3 M-SVR 模型与 RVFLNs 模型估计效果比较

从以上仿真实验可以看出, M-SVR 模型与 RVFLNs 模型都具有较好的估计能力和泛化性能,均可用于十字测温的在线估计应用.以下将对 M-SVR 模型和 RVFLNs 模型的十字测温泛化能力进行统计比较.表 2 给出了不同模型对十字测温中心 5 点温度 (T5, T6, T16, T15, T17) 估计的均方根误差 (RMSE) 比较,其中  $N$  表示所需估计的样本点个数,  $y_i$  是实际值,  $\tilde{y}_i$  是模型的估计值.表 3 给出了不同模型在十字测温中心 5 点温度估计误差的平均值和标准差.可以看出 M-SVR 模型和 RVFLNs 模型在建模精度方面相差不大,都具有较好的建模效果.但从模型测试误差指标比较可以看出 M-SVR 模型在不同测温点

表 2 基于 M-SVR 与 RVFLNs 的十字测温估计均方根误差

Table 2 RMSE of cross temperature measuring based on M-SVR and RVFLNs

算法	模型训练 RMSE = $\sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{y}_i - y_i)^2}$					测试 RMSE = $\sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{y}_i - y_i)^2}$				
	T5	T6	T16	T15	T17	T5	T6	T16	T15	T17
M-SVR	7.662 6	8.869 8	9.997 2	6.223 0	6.136 3	15.286 6	19.159 3	17.143 4	17.309 6	20.831 3
RVFLNs	<b>3.917 8</b>	<b>11.607 4</b>	<b>9.367 3</b>	<b>7.364 0</b>	<b>6.482 5</b>	<b>11.165 4</b>	<b>36.476 6</b>	<b>29.055 7</b>	<b>31.463 8</b>	<b>34.754 6</b>

表 3 基于 M-SVR 与 RVFLNs 的十字测温估计误差的平均值和标准差

Table 3 Mean and standard deviation of cross temperature measuring estimation error based on M-SVR and RVFLNs

算法	误差平均值 $\Delta \bar{y} = \frac{1}{N} \sum_{i=1}^N  \tilde{y}_i - y_i $					标准差 $\delta = \sqrt{\frac{1}{N} \sum_{i=1}^N ( \tilde{y}_i - y_i  - \Delta \bar{y})^2}$				
	T5	T6	T16	T15	T17	T5	T6	T16	T15	T17
M-SVR	11.527 0	15.190 4	14.801 3	15.102 8	17.686 1	10.040 3	11.676 1	8.649 8	8.457 4	11.006 7
RVFLNs	<b>9.313 1</b>	<b>30.270 0</b>	<b>26.675 5</b>	<b>26.941 8</b>	<b>31.510 0</b>	<b>6.159 0</b>	<b>20.353 5</b>	<b>11.517 4</b>	<b>16.251 4</b>	<b>14.662 8</b>

的精度更高,泛化能力更强.

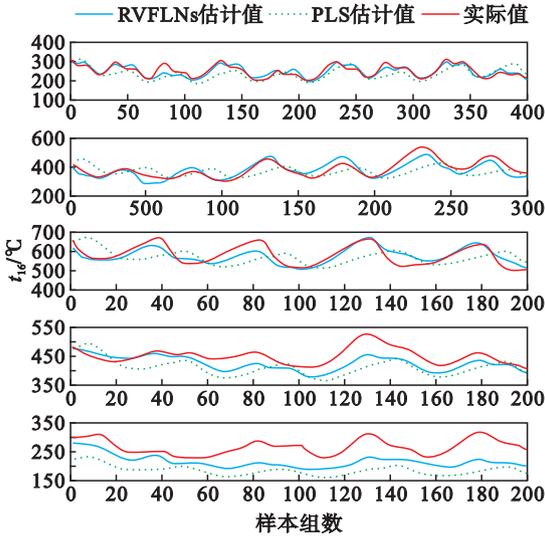


图 4 基于 RVFLNs 的十字测温测试效果  
 Fig. 4 Test results of cross temperature measuring based on RVFLNs

以上分析是根据固定样本数量进行分析的,为了得到更为一般的结论,从训练样本规模变化的角度,进一步比较两种算法建模和泛化能力. 本文通过求取从输入样本数量 100, 200, ..., 2 500, 输出样本数量为 400 的情况下,从建模和测试误差平均值  $\Delta\bar{y}$  角度比较两种方法的建模和泛化能力. 从图5可以看出,在不同训练样本下, M -

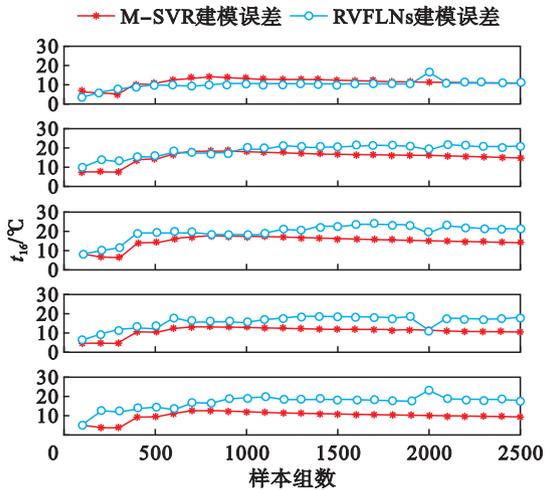


图 5 不同样本下 M - SVR 模型与 RVFLNs 模型建模误差比较  
 Fig. 5 Comparison of modeling error between M-SVR model and RVFLNs model under different samples

SVR 模型的建模误差整体要小于 RVFLNs 模型. 从图 6 可以看出,当训练样本数充足时, M - SVR 模型和 RVFLNs 模型的测试误差趋于下降,但 M - SVR 模型测试误差要略小于 RVFLNs 模型,

说明在高炉十字测温温度估计中, M - SVR 模型的估计效果相比于 RVFLNs 模型要好,因此 M - SVR 更适合用于高炉十字测温温度估计.

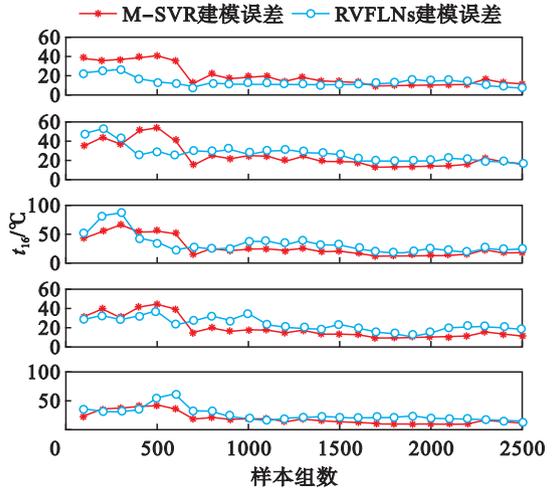


图 6 不同样本下 M - SVR 模型与 RVFLNs 模型测试误差比较  
 Fig. 6 Comparison of testing error between M-SVR model and RVFLNs model under different samples

### 5 软件实验

以上比较说明 M - SVR 模型估计效果和泛化性能要优于 RVFLNs 模型. 为此,将 M - SVR 模型编制成软件用于工业实际,以便于操作人员对炉顶温度进行观察. 十字测温估计软件采用通用 C/S (Client/Server) 结构,以专用网络为传输环境,服务器提供链接和数据交换服务,数据库采用 SQL 2012,编程环境为 .NET,编程语言为 C#, 界面编写为窗体形式,编程工具使用 Visual Studio 2010. 图 7 为工业现场的十字测温中心温度估计系统. 应用表明,该软件可根据实时数据对十字测温中心 5 点温度及其趋势进行精确估计,帮助现场高炉操作人员对布料操作进行决策.

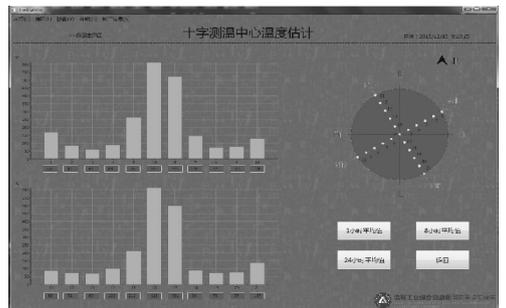


图 7 十字测温中心温度估计软件  
 Fig. 7 Centre temperature estimation software of cross temperature measuring

## 6 结 语

本文分别应用M-SVR和RVFLNs两种智能算法建立了高炉十字测温中心5点温度估计模型.与其他算法相比M-SVR具有运算速度快、精度高并且具有较好的泛化能力,RVFLNs具有运算速度快,建模效果好,估计精准等优势.并且本文将两种算法的建模和泛化能力进行了对比,两种算法都具有较好的建模和泛化性能,但整体而言当样本数量充足时,M-SVR的泛化性能优于RVFLNs.在工业现场的应用中基于M-SVR模型的软件估计准确,能够很好地辅助高炉操作人员对炉顶温度进行观察.因此,M-SVR更适用于高炉十字测温温度的估计.

### 参考文献:

- [1] Jian L, Gao C H, Li L, et al. Application of least squares support vector machines to predict the silicon content in blast furnace hot metal [J]. *ISIJ International*, 2008, 48 (11): 1659 - 1661.
- [2] Yuan M, Zhou P. Intelligent multivariable modeling of blast furnace molten iron quality based on dynamic AGA-ANN and PCA [J]. *Journal of Iron and Steel Research, International*, 2015, 22(6): 487 - 495.
- [3] Zhao J, Liu Y, Zhang X P, et al. A MKL based on-line prediction for gasholder level in steel industry [J]. *Control*

- Engineering Practice*, 2012, 20: 629 - 641.
- [4] Zhou P, Yuan M, Wang H, et al. Multivariable dynamic modeling for molten iron quality using online sequential random vector functional-link networks with self-feedback connections [J]. *Information Sciences*, 2015, 325 (12): 237 - 255.
- [5] Tuia D, Verrelst J, Alonso L, et al. Multioutput support vector regression for remote sensing biophysical parameter estimation [J]. *IEEE Geoscience and Remote Sensing Letters*, 2011, 8(4): 804 - 808.
- [6] Sanchez F, Cumplido D P, Arenas G, et al. SVM multiregression for non linear channel estimation in multiple-input multiple-output systems [J]. *IEEE Transactions on Signal Processing*, 2004, 52(8): 2298 - 2307.
- [7] Pao Y H, Takefuji Y. Functional-link net computing: theory, system architecture, and functionalities [J]. *Computer*, 1992, 25(5): 76 - 79.
- [8] Han H G, Wang L D, Qiao J F. Hierarchical extreme learning machine for feedforward neural network [J]. *Neurocomputing*, 2014, 128(5): 128 - 135.
- [9] Schmidt W F, Kraaijveld M, Duin R P W. Feedforward neural networks with random weights [C]//1992 Pattern Recognition Conference B: Pattern Recognition Methodology and Systems. Sydney, 1992: 1 - 4.
- [10] Martens H, Martens M. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression [J]. *Food Quality & Preference*, 2000, 99 (11): 5 - 16.
- [11] Saracoglu I, Topaloglu S, Keskinurk T. A genetic algorithm approach for multi-product multi-period continuous review inventory models [J]. *Expert Systems with Applications*, 2014, 41(18): 8189 - 8202.
- [12] Gao X X, Yang H X, Lin L, et al. Wind turbine layout optimization using multi-population genetic algorithm and a case study in Hong Kong offshore [J]. *Journal of Wind Engineering and Industrial Aerodynamics*, 2015, 139: 89 - 99.

(上接第613页)

取表征过程内在特征的少数主要变量,因此PCA和ICA广泛应用于线性工业流程的过程监测.基于核函数的KPCA和KICA通过非线性映射,巧妙地将一个非线性问题转化为线性问题,因此被广泛应用于非线性过程监测中.

2) 上述算法均有自身的适用条件,如果在监测过程中不考虑工业过程的数据特性,盲目选择监测算法,则可能给出错误结论.

3) 本文提出的基于数据特性分析的方法,对于先验知识未知的过程,检验过程数据的相关特性,并根据检验结果及各监测算法的适用条件自动选择合适的监测算法,从而有效解决了各方法对过程数据的限制问题,提高了监测性能.

4) 仿真验证了该方法选择的监测算法效果良好,进一步证明了该方法可行有效.

### 参考文献:

- [1] Zhao E D, Tan L S. A PCA based optimization approach for IP traffic matrix estimation [J]. *Journal of Network and Computer Applications*, 2015, 57: 12 - 20.
- [2] Ajami A, Daneshvar M. Data driven approach for fault detection and diagnosis of turbine in thermal power plant

- using independent component analysis (ICA) [J]. *Electrical Power and Energy Systems*, 2012, 43: 728 - 735.
- [3] Lee J M, Yoo C, Choi S W, et al. Nonlinear process monitoring using kernel principal component analysis [J]. *Chemical Engineering Science*, 2004, 59: 223 - 234.
- [4] Lee J M, Qin S J, Lee I B. Fault detection of non-linear processes using kernel independent component analysis [J]. *Canadian Journal of Chemical Engineering*, 2007, 85 (4): 526 - 536.
- [5] 刘桂珍.判断向量组线性相关性的常用方法[J].凯里学院学报, 2007, 25(3): 3 - 6.  
(Liu Gui-zhen. General kinds of methods for judging the related linearity of vectors group [J]. *Journal of Kaili University*, 2007, 25(3): 3 - 6.)
- [6] Shen Z Y, Wang Q, Shen Y. A new non-linear correlation measure [J]. *Information, Computing and Telecommunication*, 2009, 9: 11 - 14.
- [7] Zhang S M, Wang F L, Zhao L P, et al. A novel strategy of data characteristic test for selecting process monitoring method automatically [J]. *Industrial & Engineering Chemistry Research*, 2016, 55 (6): 1642 - 1654.
- [8] Hyvärinen A, Oja E. A fast fixed-point algorithm for independent component analysis [J]. *Neural Computation*, 1997, 9(7): 1483 - 1492.
- [9] Jiang Q C, Yan X F. Nonlinear plant-wide process monitoring using MI-spectral clustering and Bayesian inference-based multiblock KPCA [J]. *Journal of Process Control*, 2015, 32: 38 - 50.
- [10] Zhang S M, Wang F L, Wang S, et al. Multivariate process monitoring based on the distribution test of the data [C]// Proceedings of the 11th World Congress on Intelligent Control and Automation. New York: Institute of Electrical and Electronics Engineers Inc, 2015: 2723 - 2728.