# 高炉煤气流分布过程的多算法融合预测模型

吴晓阳,张 森<sup>†</sup>,陈先中,尹怡欣

(北京科技大学自动化学院,北京100083;北京科技大学工业过程知识自动化教育部重点实验室,北京100083)

摘要: 在高炉生产中, 准确的预测高炉煤气流分布状况将有助于保证高炉的稳定顺行. 针对传统高炉煤气流预测 模型的缺陷, 本文提出了一种将带遗传因子的自回归移动平均模型(FF-ARMAX)和基于限定记忆的正则化极限学 习机(RFMLS-RELM)相结合的高炉煤气流多步预测模型. 在数据预处理方面, 建立FF-ARMAX模型消除原始数 据中的测量误差, 同时采用傅里叶变换法消除数据中叠加的环境噪声. 最后采用RFMLS-RELM算法进行多步预 测, 对比试验表明, 该算法在应用于煤气流预测时, 预测精度更高, 适用于对煤气流分布状况的多步预测. 多步预测 实验结果表明, 该模型虽然仍旧无法完全解决预测误差随预测步数的增加而不断叠加的问题, 但相较于其他传统预 测模型能够实现更好的预测效果和更高的预测精度, 为高炉操作人员分析炉况提供了有效的帮助和支持.

关键词: 高炉; 煤气流分布; 自回归移动平均模型; 限定记忆; 极限学习机; 预测

引用格式:吴晓阳,张森,陈先中,等.高炉煤气流分布过程的多算法融合预测模型.控制理论与应用,2020,37(6): 1241-1252

DOI: 10.7641/CTA.2019.90146

# Multi-algorithm fusion prediction model for the blast furnace gas flow distribution process

WU Xiao-yang, ZHANG Sen<sup>†</sup>, CHEN Xian-zhong, YIN Yi-xin

(School of Automation and Electrical, University of Science and Technology Beijing, Beijing 100083, China; Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, University of Science and Technology Beijing, Beijing 100083, China)

**Abstract:** In blast furnace production, the prediction of gas flow will help to ensure the stability of the blast furnace. Aiming at the shortcomings of the traditional prediction model of blast furnace gas flow, this paper proposes a multistep prediction model of gas flow based on auto-regressive moving average model based on recursive least squares with forgetting factor (FF–ARMAX) and regularized extreme learning machine based on recursive finite memory least squares (RFMLS–RELM). In terms of data correction, construct FF–ARMAX model to eliminate the measurement error in the data. At the same time, the Fourier transform method is used to denoising. The RFMLS–RELM algorithm is used for multi-step prediction. The comparison experiments show that the proposed algorithm is faster in modeling and higher in prediction accuracy when applied to blast furnace gas flow prediction. It is suitable for the multi-step prediction of gas flow distribution of blast furnace. The experimental results show that the model can accurately predict the distribution of blast furnace gas flow, which provides effective help and support for blast furnace operators to analyze furnace conditions.

**Key words:** blast furnace; gas flow distribution; auto-regressive moving average model; finite memory; extreme learning machine; prediction

**Citation:** WU Xiaoyang, ZHANG Sen, CHEN Xianzhong, et al. Multi-algorithm fusion prediction model for the blast furnace gas flow distribution process. *Control Theory & Applications*, 2020, 37(6): 1241 – 1252

# 1 引言

钢铁工业是国民经济的支柱性产业.高炉是钢铁 生产过程中的核心设备,关系到钢铁行业的工业产量 和能源消耗.维持高炉的稳定顺行是整个钢铁行业共 同追求的重要目标<sup>[1]</sup>. 在高炉生产中, 炉顶的煤气流 分布直接反映了煤气热能的利用情况, 因而影响高炉 能耗水平和生产成本, 所以准确的预测高炉煤气流分 布状况将有助于保证高炉的稳定顺行<sup>[2]</sup>. 然而, 高炉

收稿日期: 2019-03-19; 录用日期: 2019-11-15.

<sup>&</sup>lt;sup>†</sup>通信作者. E-mail: zhangsen@ustb.edu.cn.

本文责任编委:阳春华.

国家自然科学基金项目(61673056, 61673055, 61671054), 北京市自然科学基金项目(4182039)资助.

Supported by the National Natural Science Foundation of China (61673056, 61673055, 61671054) and the National Natural Science Foundation of Beijing (4182039).

在运行过程中处于"黑箱"状态,使得高炉操作人员 无法直接观测煤气流的分布及其状态变化.因此关于 高炉煤气流分布状况的预测一直是学术界和工业界 的热门研究方向[3]. 高炉煤气流分布的预测有利于研 究高炉料面的变化规律,从而指导布料操作.随着检 测技术的发展,包括红外成像、十字测温等先进技术 较为成熟的运用在高炉煤气流的实时检测中,使得煤 气流的预测成为了可能. Li等人<sup>[4]</sup>基于小波概率神经 网络建立一种煤气流分布辨识模型. Zhang 等人<sup>[5]</sup>建 立了一种基于低频特征提取技术的支持向量回归模 型预测高炉煤气流分布状况. Shi等人<sup>[6]</sup>提出了一种基 于红外图像处理的高炉煤气流中心分布特征识别方 法,采用模糊C均值聚类和统计方法对高炉煤气流中 心的分布特征和相应的气体利用率进行了分类. An等 人[7]提出一种基于炉顶摄像检测的高炉炉喉煤气流形 态三维模型重建方法. Zhou等人<sup>[8]</sup>基于时间序列思 想,采用多输出自回归移动平均模型、主因子分析法、 Pearson相关性分析、赤池信息准则与模型拟合优度 等技术,提出一种结构简单、精度高且易于实现的十 字测温中心温度在线估计方法.虽然上述模型均可实 现对高炉煤气流分布状况的研究和预测,但是仍然存 在两个问题使模型的准确性无法令人完全信服.

问题1 在高炉运行过程中,基于十字测温装置 和红外图像处理技术所得到的温度数据总是不同程 度的带有异常数据.异常数据产生的原因,既有因高 炉内部粉尘与高温煤气密布等环境因素产生的环境 噪声,也有因长期处于高温高压环境下导致的测量仪 器损坏所产生的测量误差,而后者的影响更为恶劣. 使用滤波等数据预处理手段虽然可以有效去除环境 噪声但是却无法完全排除测量误差.

问题2 传统高炉煤气流温度预测模型主要集中 在优化模型,提高预测精度上,而忽视了在工程中的 实用性.大多数预测模型属于单步预测,虽然具有很 高的精度,但是只能为高炉操作人员提供一分钟甚至 几十秒的预测结果,而这对实际操作的帮助十分有限.

针对以上所提到的问题,本文基于真实的高炉十 字测温数据,将带遗忘因子的递推最小二乘法的自回 归移动平均模型(auto-regressive moving average model based on recursive least squares with forgetting factor, FF-ARMAX)和限定记忆的正则化极限学 习机 (regularized extreme learning machine based on recursive finite memory least squares, RFMLS-RELM) 两种数据驱动算法融合建立高炉煤气流多步预测模 型.首先,利用时间序列分析能够准确反映动态依存 关系的特性和带遗忘因子的递推最小二乘法能够降 低历史数据影响力的特性建立FF-ARMAX模型,通 过该模型对十字测温装置产生的测量误差进行修正, 并且运用傅里叶变换法去除内部环境产生的环境噪 声,以减少原始数据与真实情况之间的差距,提高预 测结果的可靠性.然后分别选取十字测温装置中的 17个测温点作为输出变量,如图1所示,使用经过 数据预处理的高炉数据,结合正则化极限学习 机(regularized extreme learning machine, RELM)算 法,建立高炉煤气流多步预测模型.考虑到传统的 RELM算法会随着训练样本的不断增加,而逐渐失去 学习能力,也就是新的样本无法对模型产生影响, 即"数据饱和"问题,本文运用限定记忆递推最小 二乘法(recursive finite memory least squares, RFMLS) 优化RELM,从而保证模型的预测精度和可信度.实 验结果表明,采用改进的RELM算法建立的高炉煤气 流多步预测模型可以准确预测高炉煤气流分布状况, 为高炉操作人员分析炉况提供了有效的帮助和支持.



Fig. 1 The crossing temperature device

# 2 煤气流分布影响因素分析及数据预处理

传统高炉煤气流分布预测模型中,模型输入变量的确定大多根据经验主观确定,容易造成模型特征冗余或关键特征丢失,影响预测精度和建模效率.因此,在建立高炉煤气流多步预测模型之前,首先需要确定影响高炉煤气流分布的相关影响因素作为预测模型的输入变量.同时,由于温度数据中存在由采集装置导致的测量误差和生产环境导致的高频噪声.因此需要对从高炉生产现场采集到的原始数据进行数据预处理,具体而言就是分别进行测量误差处理和数据去噪处理,使数据能够更加真实的反映高炉生产信息,从而提升预测模型的性能,提高预测精度.

### 2.1 煤气流分布影响因素分析

本文除采用十字测温装置的测温点作为影响因子外,还选择高炉炉况信息等其他影响高炉煤气流分布的相关因素作为影响因子,进行煤气流分布影响因素分析.由于本文需要对测温点进行预测,故选择十字测温装置的各测温点作为预测模型的输出变量.本文选取风量(FL)、风温(FW)、风压(FY)、顶压(DY)、富氧量(O2)、风速(FS)、压差(YC)和17个测温点,共计24个变量作为影响因子.选取1000组高炉现场数据,相关影响因子的详细数据信息如表1所示.由于篇幅

有限,此处以十字测温装置中的T25测温点为例进行 影响因素分析.分析采用Pearson相关性分析,分别计 算这23个影响因子与输出变量T25之间的相关程度. Pearson相关系数的计算公式如下:

$$r = \frac{\sum_{i=1}^{n} (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n} (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^{n} (b_i - \bar{b})^2}},$$
 (1)

式中:  $a_i$ ,  $b_i$ 表示第i时刻两个变量的采样值;  $\overline{a}$ ,  $\overline{b}$ 表示 变量均值; n为采样总量.

表	1 高	炉数据详	细信	息
Table 1	Blas	st furnace	e data	details

参数	平均值	最小值	最大值	标准差	量纲
FL	4183.9	4143.9	4242.1	12.52	$m^3/\min$
FW	1180.2	1173.7	1193.5	2.233	°C
FY	348.90	321.47	362.21	5.24	kPa
DY	190.19	178.56	220.68	2.61	kPa
O2	63.75	57.28	72.03	5.06	%
FS	254.24	246.96	272.84	3.23	m/s
YC	158.71	130.40	172.39	5.49	kPa
T39	284.38	129.33	476.62	61.25	°C
T38	141.04	83.86	301.45	35.77	°C
T37	119.46	71.35	264.83	32.54	°C
T36	264.96	78.67	638.06	108.17	°C
T25	<b>418.40</b>	-12.50	775.39	223.91	C
T29	65.28	-12.50	159.55	29.63	C
T28	72.23	51.51	154.66	16.70	°C
$T28 \\ T27$	72.23 101.98	51.51 71.66	154.66 204.41	16.70 17.87	°C °C
$T28 \\ T27 \\ T26$	72.23 101.98 323.58	51.51 71.66 178.16	154.66 204.41 694.82	16.70 17.87 81.65	ິC ິC ິC
$T28 \\ T27 \\ T26 \\ T21$	72.23 101.98 323.58 263.50	51.51 71.66 178.16 133.30	154.66 204.41 694.82 494.93	16.70 17.87 81.65 60.95	С С С С
$T28 \\ T27 \\ T26 \\ T21 \\ T22$	72.23 101.98 323.58 263.50 87.22	51.51 71.66 178.16 133.30 54.26	154.66 204.41 694.82 494.93 200.44	16.70 17.87 81.65 60.95 19.45	С С С С
$T28 \\ T27 \\ T26 \\ T21 \\ T22 \\ T23$	72.23 101.98 323.58 263.50 87.22 100.33	51.51 71.66 178.16 133.30 54.26 50.90	154.66 204.41 694.82 494.93 200.44 253.85	16.70 17.87 81.65 60.95 19.45 33.35	ິ ວິ ວິ ວິ ວິ
$T28 \\ T27 \\ T26 \\ T21 \\ T22 \\ T23 \\ T24$	72.23 101.98 323.58 263.50 87.22 100.33 239.91	51.51 71.66 178.16 133.30 54.26 50.90 81.42	154.66 204.41 694.82 494.93 200.44 253.85 606.02	16.70 17.87 81.65 60.95 19.45 33.35 114.42	つ つ つ つ つ
$T28 \\ T27 \\ T26 \\ T21 \\ T22 \\ T23 \\ T24 \\ T33$	72.23 101.98 323.58 263.50 87.22 100.33 239.91 319.15	51.51 71.66 178.16 133.30 54.26 50.90 81.42 82.03	154.66 204.41 694.82 494.93 200.44 253.85 606.02 606.32	16.70 17.87 81.65 60.95 19.45 33.35 114.42 106.70	С С С С С С С С
$\begin{array}{c} T28 \\ T27 \\ T26 \\ T21 \\ T22 \\ T23 \\ T24 \\ T33 \\ T32 \end{array}$	72.23 101.98 323.58 263.50 87.22 100.33 239.91 319.15 162.74	51.51 71.66 178.16 133.30 54.26 50.90 81.42 82.03 82.64	154.66 204.41 694.82 494.93 200.44 253.85 606.02 606.32 317.32	16.70 17.87 81.65 60.95 19.45 33.35 114.42 106.70 35.50	С С С С С С С С С
$\begin{array}{c} T28 \\ T27 \\ T26 \\ T21 \\ T22 \\ T23 \\ T24 \\ T33 \\ T32 \\ T31 \end{array}$	72.23 101.98 323.58 263.50 87.22 100.33 239.91 319.15 162.74 136.45	51.51 71.66 178.16 133.30 54.26 50.90 81.42 82.03 82.64 96.38	154.66 204.41 694.82 494.93 200.44 253.85 606.02 606.32 317.32 282.53	16.70 17.87 81.65 60.95 19.45 33.35 114.42 106.70 35.50 24.24	С С С С С С С С С С С

根据式(1)计算得到每个影响因子与输出变量*T*25 的相关系数,结果如图2所示.



Fig. 2 The correlation coefficient between the influence factor and T25

为降低模型复杂度, 需要进行数据降维, 因此选取 相关性大于 $r_0 = 0.5$ 的影响因子作为模型输入. 如图 2中所示, 灰色虚线标识的地方即为 $r_0 = 0.5$ 处, 本文 选取Pearson相关系数大于0.5的影响因子作为输入变 量. 由此得到高炉煤气流多步预测模型以T25为输出 变量所对应的输入变量为T23, T22, T21, T24, T32, T37, T33和T36.

#### 2.2 数据预处理

由于信息采集装置的老化、失效等问题和高炉内 部环境复杂多变,高炉生产现场采集到的数据不可避 免存在着各种误差.有因信息采集装置所处的恶劣工 作环境导致的仪器老化、损坏等问题而产生的测量误 差,也有因高炉内部本身复杂多变的环境而产生的环 境噪声.并且,使用存在着误差的高炉数据建立预测 模型会使预测效果变差,所以需要先对高炉生产数据 进行预处理.接下来数据预处理将会分为两个部分进 行,一个建立FF-ARMAX模型修正测量误差,一个 是通过傅里叶变换消除环境噪声.

# 2.2.1 基于FF-ARMAX的测量误差处理

十字测温装置的热电偶因为自身所处的恶劣环境, 极易损坏,导致测量数据出现严重的测量误差,本文 针对温度数据在时序上存在测量误差的问题,提出通 过建立FF-ARMAX模型对测量误差进行修正的方 法.时间序列分析的基本思想是根据观察到的历史数 据,建立能够比较精确地反映时间序列中所包含动态 依存关系的数学模型<sup>[9]</sup>,所以可以通过分析从高炉现 场获取的输入输出时间序列数据之间与输出时间序 列之间的相互关系来建立十字测温装置测量误差修 正模型.时间序列分析中ARMAX模型的一般结构为

$$A(z^{-1})y(k) = B(z^{-1})u(k) + \varepsilon(k),$$
 (2)

式中:

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{n_p} z^{-n_p}, \\ B(z^{-1}) = 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_q} z^{-n_q}; \end{cases}$$
(3)

 $z^{-1}$ 为延迟算子, 即 $z^{-1}y(k) = y(k-1)$ ;  $\varepsilon(k)$ 为零均 值高斯白噪声;  $n_{\rm p}$ ,  $n_{\rm q}$ 为滞后的阶次, k为采样时刻.

1) 变量选取.

建立FF-ARMAX模型是为了消除数据中存在的 测量误差,因此需要先确定哪些变量数据中存在测量 误差.由表1可知,变量T25和T29(字体加粗标明)的 最小值均为-12.5℃,显然高炉中不可能出现零下温 度,因此这两个变量必然存在测量误差,需要通过 FF-ARMAX模型修正.此处以T25为例.选取T25作 为输出变量,输入变量的选取采用第2.1节中的相关性 分析方法,经过分析最终选取T23,T22,T21,T24, T32,T37,T33和T36作为输入变量.

 $\phi(k -$ 

# 2) 模型定阶.

建模过程中增加模型阶数可以提高模型精度,但 是模型复杂度会增大,提高了建模难度,并易于导致 过拟合问题.为此,采用信息准则作为确定模型阶数 的依据,信息准则通过加入模型复杂度的惩罚项来避 免过拟合的问题.最常用的信息准则为赤池信息准则 (Akaike's information, AIC)<sup>[10]</sup>:

$$AIC = \log(v) + \frac{2d}{N},$$
(4)

式中: d为模型参数个数; v为损失函数; N为训练样本 容量.本文采用MATLAB中的armax(·)函数和aic(·)函 数进行模型定阶,确定式(2)-(3)中滞后阶次 $n_p$ ,  $n_q$ .滞后阶次的选取范围为 $n_p \in [1,5]$ ,  $n_q \in [1,5]$ , N = 1000, 实验结果如图3所示.



图 3(a) 模型AIC与不同输入输出阶次的三维立体图 Fig. 3(a) The AIC value corresponding to different order (three-dimensional map)



图 3(b) 模型AIC与不同输入输出阶次的俯视图 Fig. 3(b) The AIC value corresponding to different order (top view)

由图3(a)和图3(b)可以看出, 当 $n_p = 3$ ,  $n_q = 3$ 时, AIC函数值最高, 因此模型阶次确定为 $n_p = 3$ ,  $n_q = 3$ .

3) 模型参数辨识.

由式(2)-(3)可知, 需要建立的FF-ARMAX模型 表达式为

$$y(k) = A(z^{-1})y(k) + B_i(z^{-1})u_i(k) + \varepsilon(k),$$
 (5)  
式中:

$$\begin{cases} A(z^{-1}) = -a_1 z^{-1} - a_2 z^{-2} - \dots - a_{n_p} z^{-n_p}, \\ B(z^{-1}) = 1 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_{n_q} z^{-n_q}. \end{cases}$$
(6)

定义数据向量和参数向量:

$$-y(k-1) \cdots y(k-n_{\rm p}) \ u_i(k) \cdots u_i(k-n_{\rm q})],$$
(7)

$$\theta^{\rm T} = [a_1 \ a_2 \ \cdots \ a_{n_{\rm p}} \ b_1 \ b_2 \ \cdots \ b_{n_{\rm q}}], \tag{8}$$

其中 $i = 1, 2, \dots, N_{\text{R}}, N_{\text{R}}$ 为输入变量个数. 将式(7)和式(8)代入式(5),可得

$$y(k) = \phi(k-1)\theta.$$
(9)

为了解决"数据饱和"问题,本文采用带遗忘因子的递推最小二乘算法(the least square method with forgetting factor),其算法公式为

$$K(k) = \frac{P(k-1)\phi(k-1)}{\mu(k-1) + \phi^{\mathrm{T}}(k-1)P(k-1)\phi(k-1)},$$
 (10)

$$\theta(k-1) + K(k)(y(k) - \phi^{\mathrm{T}}(k-1)\theta(k-1)),$$
(11)

$$P(k) = \frac{1}{\mu(k-1)} (I + K(k)\phi^{\mathrm{T}}(k-1))P(k-1), \quad (12)$$

$$\mu(k) = \mu_0 \mu(k-1) + (1-\mu_0), \tag{13}$$

式中K(k),  $\theta(k)$ , P(k),  $\mu(k)$ 分别为增益矩阵、参数 向量、逆矩阵和遗忘因子. 注意到上述算法需要事先 给定初值 $\theta(0) = 0$ ,  $P(0) = \alpha I$ ,  $\mu_0 = 0.99$ ,  $\mu(0) = 0.95$ , 其中 $\alpha$ 为足够大的正数, 本文选用10<sup>4</sup>.

4) 实验结果.

 $\theta(k) =$ 

参数收敛曲线如图4所示. 图中曲线表示参数向 量θ中参数值的变化趋势, 当参数值随着递推次数的 增加趋于稳定时, 认为参数向量达到稳定, FF-ARMAX模型将通过稳定的参数向量构建. 从图4中 可以看出, 参数向量θ在经过400次迭代后达到稳定. 因此, 选取迭代400次时的参数向量θ代入式(9)中构建 FF-ARMAX模型.





#### 2.2.2 基于傅里叶变换的环境噪声处理

本文中采用的数据均是高炉生产数据,这些数据

中叠加的环境噪声会对建模与预测结果产生严重的 影响,因此建模之前需要剥离出数据中的高频环境噪 声<sup>[5]</sup>.为了验证信号的正确性,需要对数据进行频谱 分析,也就是对数据进行傅里叶变换,将数据由时间 域转换到频率域,再根据数据的频率特征剥离数 据中的噪声.三角函数形式的傅里叶变换(Fourier transform, FT)如下:

$$x(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)),$$
(14)

式中: x(t)表示信号;  $a_n$ ,  $b_n$ 表示傅里叶系数; n = 0, ±1,±2;  $\omega_0 = \frac{2\pi}{T_c}$ .

完成对于数据的傅里叶变换后,可以得到数据的 低频特征,然后将数据输入合理的数字滤波器,就可 以在保证有效信息不被破坏的情况下,将高频环境噪 声除去.数字滤波器又可以分为FIR数字滤波器和 IIR数字滤波器,本文采用的是IIR数字滤波器,也称为 递归型的数字滤波器,即滤波器当前的输出不仅与当 前和以前的输入值有关,而且还与以前的各个输出值 有关<sup>[11]</sup>.

#### 2.2.3 数据预处理效果

根据第2.2.1节中FF-ARMAX模型的建立过程, 建立修正模型用于消除T25温度数据中存在的测量误 差. T25变量的原始数据如图5所示, 从图5中可以看 出在T25温度数据中不同位置分布着零下温度数据, 这印证了表1中T25的数据信息中最小值为-12.5℃ 这一现象.T25变量经过FF-ARMAX模型消除测量 误差,修正结果如图6所示,从图6中可以看出T25原 始数据中大量存在的-12.5℃数据被消除, T25变 量数据恢复到正常的高炉温度范畴,证明了FF-ARMAX 模型可以有效消除数据中存在的测量误差. 但是,同样可以发现,消除测量误差的高炉数据中仍 然存在着大量由环境因素产生的高频噪声,如图6 中黑色虚线圈处所示,因此需要去噪处理.根据第 2.2.2节,采用傅里叶变换对经过修正的T25温度数据 进行数据分析,然后采用IIR数字滤波器进行去噪处 理,处理结果如图7所示.从图7中可以看出,经过去 噪处理,图6中T25温度数据存在的环境噪声被有效 去除.









图8为经过FF-ARMAX模型修正和去噪处理的 T25温度数据与T25原始温度数据的对比图.由图8可 知,经过数据预处理后的T25温度数据,在保留数据 变化趋势不变的情况下,完全消除了叠加在原始数据 中的测量误差,并且最大限度的消除了数据中的环境 噪声,使数据能够更加真实的反映高炉的生产过程, 为接下来的煤气流预测工作做好了准备.



# 3 基于 RFMLS-RELM 的高炉煤气流多步 预测模型

由于影响高炉煤气流分布的因素众多,并且和其他机器学习算法相比,RFMLS-RELM算法更适用于处理大规模复杂实时环境中的大量数据,对之前的历史数据依赖性比较小,具有比较好的鲁棒性<sup>[12]</sup>.本章首先详细介绍RFMLS-RELM算法,然后介绍采用RFMLS-RELM算法所建立的高炉煤气流模型.

#### 3.1 基于限定记忆的正则化极限学习机

由于高炉内部炉况复杂多变,内部情况无法精确 判断,为了应对十字测温数据在不同时刻的变化,煤 气流预测模型需要实时读入新数据进行模型训练.但 是传统的极限学习机会随着训练样本的增加,历史数 据逐渐占据训练数据的大部分,新数据所占比重越来 越低,而逐渐失去学习能力,即"数据饱和"问题.即 使是权值更新的在线贯序极限学习机(an online sequential extreme learning machine algorithm based on incremental weighted average, WOS-ELM)<sup>[13-14]</sup>也只 能给靠给历史数据添加遗忘因子的办法降低"历史数 据"在模型训练过程中的影响,然而此举并不能完全 排除历史数据的影响,也就是说,随着时间的推移模 型最终还是会出现"饱和问题".为此本文提出了一 种RFMLS-RELM算法. 该算法在模型训练过程中每 读入一个新样本就会去除一个历史样本,隐含层到输 出层的权值矩阵岛的学习仅依赖于限定个数的最新样 本,从而避免"数据饱和"问题的出现.

#### RFMLS-RELM算法推导过程如下:

给定包含N个训练样本的数据集 $\{(x_j, y_j)\}_j^N$ ,假 设模型完美地描述了 $x_j$ 和 $y_j$ 之间的关系,那么隐含层 中L个隐藏节点和激活函数g(x)可以建立数学模型为

$$\sum_{i=1}^{L} \beta_i g(w_i \cdot x_j + b_i) = y_j, \ j = 1, 2, \cdots, N,$$
(15)

其中:  $w_i = [w_{i1} \ w_{i2} \cdots \ w_{iN}]^T$ 是随机选择的输入 权重, 连接第*i*个隐藏节点和输入节点的权值矢量;  $b_i$ 是第*i*个隐藏节点随机选择的偏置向量;  $\beta_i$ 是连接第*i* 个隐藏节点和输出节点的权值矢量;  $w_i \cdot x_j$ 表示 $w_i$ 和  $x_i$ 的内积;  $g(w_i \cdot x_j + b_i)$ 是第*i*个隐藏节点的输出.

式(15)可以转化成矩阵形式:

$$H\beta = y, \tag{16}$$

H称为隐藏层的输出矩阵, β为隐含层输出权重.为更 好的平衡训练误差和输出权重,防止算法陷入"过拟 合", RFMLS-RELM算法引入正则化系数C<sup>[15-16]</sup>, 用于平衡训练误差和输出权重β.因此在求解式(16) 中输出权重β时,采用如下公式:

$$\hat{\beta} = (H^{\mathrm{T}}H + \frac{I}{C})^{-1}H^{\mathrm{T}}y.$$
(17)

设模型训练数据长度为Q,则RFMLS-RELM模型的 隐含层输出矩阵H数学表达式为

$$H_{(k,k+Q-1)} = \begin{bmatrix} g(w_1x_k + b_1) & \cdots & g(w_Lx_k + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1x_{k+Q-1} + b_1) & \cdots & g(w_Lx_{k+Q-1} + b_L) \end{bmatrix}_{Q \times L},$$
(18)

$$\begin{cases} \beta = [\beta_1 \ \beta_2 \ \cdots \ \beta_L]^{\mathrm{T}}, \\ y = [y_k \ y_{k+1} \ \cdots \ y_{k+Q-1}]^{\mathrm{T}}, \end{cases}$$
(19)

其中k表示采样时刻,

 $h(k) = [g(w_1x_k + b_1) \cdots g(w_Lx_k + b_L)].$  (20) 首先向模型中添加一个新数据 $y_{k+Q}$ ,则当前模型中包 含 $k \le k + Q$ 时刻共计(Q+1)训练样本,根据公式(17) 可推导出基于限定记忆的正则化极限学习机的输出 层权值矩阵为,

$$\beta(k, k + \mathbf{Q}) = [H_{(k,k+\mathbf{Q})}^{\mathrm{T}} H_{(k,k+\mathbf{Q})} + \frac{I}{C}]^{-1} H_{(k,k+\mathbf{Q})}^{\mathrm{T}} Y(k, k + Q).$$
(21)

设

$$P_{(k,k+Q)} = H_{(k,k+Q)}^{\mathrm{T}} H_{(k,k+Q)} + \frac{I}{C} = \\ [H_{(k,k+Q-1)}^{\mathrm{T}} h^{\mathrm{T}}(k+Q)] \begin{bmatrix} H_{(k,k+Q-1)} \\ h(k+Q) \end{bmatrix} + \frac{I}{C} = \\ H_{(k,k+Q-1)}^{\mathrm{T}} H_{(k,k+Q-1)} + h^{\mathrm{T}}(k+Q)h(k+Q) + \frac{I}{C} = \\ P_{(k,k+Q-1)} + h^{\mathrm{T}}(k+Q)h(k+Q).$$
(22)

由递推最小二乘法的原理可推导出

$$\beta(k, k+Q) = \beta(k, k+Q-1) + P_{(k,k+Q-1)}^{-1} h^{\mathrm{T}}(k+Q) \cdot [y(k+Q) - h(k+Q)\beta(k, k+Q-1)], \quad (23)$$

$$P_{(k,k+Q)}^{-1} = [I - \frac{P_{(k,k+Q)}^{-1} h^{\mathrm{T}}(k+Q)h(k+Q)}{1 + h(k+Q)P_{(k,k+Q)}^{-1} h^{\mathrm{T}}(k+Q)}]P_{(k,k+Q-1)}^{-1}. \quad (24)$$

以上,式(23)-(24)就是加入新训练样本时的递推公式.接下来推导剔除旧训练样本时的递推公式.

当矩阵 $\beta$ 剔除旧训练样本 $y_k$ 时,模型中包含k+1至k + Q时刻共计Q个训练样本,则有

$$\beta(k+1, k+Q) = [H_{(k+1,k+Q)}^{\mathrm{T}}H_{(k+1,k+Q)} + \frac{I}{C}]^{-1} \cdot H_{(k+1,k+Q)}^{\mathrm{T}}Y(k+1, k+Q).$$
(25)

则式(22)改写为

$$P_{(k,k+Q)} = H_{(k,k+Q)}^{\mathrm{T}} H_{(k,k+Q)} + \frac{I}{C} = [h^{\mathrm{T}}(k) H_{(k+1,k+Q)}^{\mathrm{T}}] \begin{bmatrix} h(k) \\ H_{(k+1,k+Q)} \end{bmatrix} + \frac{I}{C} = H_{(k+1,k+Q)}^{\mathrm{T}} H_{(k+1,k+Q)} + h^{\mathrm{T}}(k)h(k) + \frac{I}{C} = [h^{\mathrm{T}}(k) H_{(k+1,k+Q)}] + h^{\mathrm{T}}(k)h(k) + h^{\mathrm{T}(k)}(k) + h^{\mathrm{T}}(k)$$

$$P_{(k+1,k+Q)} + h^{\mathrm{T}}(k)h(k).$$
(26)

同样由递推最小二乘法的原理可推导出

$$\beta(k+1, k+Q) = \beta(k, k+Q) - P_{(k+1,k+Q)}^{-1} h^{\mathrm{T}}(k) \cdot [y(k) - h(k)\beta(k, k+Q)], \qquad (27)$$
$$P_{(k+1,k+Q)}^{-1} =$$

$$[I + \frac{P_{(k,k+Q)}^{-1}h^{\mathrm{T}}(k)h(k)}{1 - h(k)P_{(k,k+Q)}^{-1}h^{\mathrm{T}}(k)}]P_{(k,k+Q)}^{-1}.$$
 (28)

以上,式(27)-(28)就是剔除旧训练样本时的递推公式.

# **3.2** 基于RFMLS-RELM的高炉煤气流多步预 测模型

基于RFMLS-RELM高炉煤气流多步预测模型的 建模过程如下所示:

步骤1 确定RFMLS-RELM的网络结构.

由第2.1节的分析可知,以十字测温装置T23, T22,T21,T24,T32,T37,T33和T36作为RFMLS -RELM高炉煤气流预测模型输入层的输入变量,即 输入层有8个节点.以T25作为该预测模型的输出变 量,即输出层有1个节点.

步骤2 建立RFMLS-RELM预测模型.

1) 隐含层节点的激活函数为Sigmoid函数;

 2)隐含层节点个数和正则化因子通过试凑法,经 过反复试验确定,并且对输入层到隐含层的权值矩阵 和隐含层的偏置向量进行随机赋值;

3) 选取训练样本长度Q,不宜太大;

4) 每当有新样本进入到模型当中时,先利用式 (23)-(24)计算 $\beta(k, k+Q)$ 和 $P_{(k,k+Q)}^{-1}$ ,增加新训练样 本的信息; 再利用式(27)-(28)计算 $\beta(k+1, k+Q)$ 和  $P_{(k+1,k+Q)}^{-1}$ ,剔除旧训练样本的信息;

5) 不断迭代直至获得理想的权值矩阵β.

步骤3 多步预测.

将经过数据预处理的十字测温数据作为测试样本 导入已经训练好的RFMLS-RELM预测模型当中,进 行多步预测.

**步骤4** 验证基于FF-ARMAX和RFMLS-RELM的高炉煤气流多步预测模型.

如果运用该模型所得到的预测值与实际测量值的 相对误差(relative errors, RE)达到要求,则模型建立完 毕.如果不符合要求,则返回步骤2重新建立模型.

# 4 实例验证和结果分析

引言中提到传统高炉煤气流温度预测模型未能解 决的两个问题:

 由十字测温装置本身产生的测量误差叠加在 温度数据中,这些异常值严重影响了高炉煤气流预测 模型的性能,而传统的去噪手段无法有效排除数据中 这类由测量误差产生的异常值;

2) 传统高炉煤气流温度预测模型的重点主要集中在优化模型,提高模型预测精度上,而忽视了在工程中的实用性.它们大多属于单步预测,虽然具有很高的精度,但是只能为高炉操作人员提供一分钟甚至几十秒的预测结果,而这对实际操作的帮助十分有限.

针对上述问题,本文基于真实的高炉十字测温数 据,将FF-ARMAX和RFMLS-RELM两种算法融合 建立高炉煤气流在线多步预测模型.为解决问题1), 采用FF-ARMAX模型消除数据中由测量误差产生的 异常值.为检验模型消除异常值的效果,设置检测实 验. 实验内容为: 将经过处理的数据与未经过处理的 原始数据分别导入一个基于RFMLS-RELM算法的 高炉煤气流温度预测模型,根据其预测结果判断 FF-ARMAX 模型对异常值消除效果. 为解决问题2), 本文采用RFMLS方法对RELM算法进行优化,提出了 一种全新的算法—RFMLS-RELM. 接下来首先建立 基于RFMLS-RELM算法的高炉煤气流温度预测模 型(单步预测模型),检验该算法在预测煤气流温度 上的性能.然后与传统预测模型进行比较,证明 RFMLS-RELM预测模型在预测煤气流温度上的优 势. 最后建立基于RFMLS-RELM算法的高炉煤气流 温度多步预测模型,借此解决问题2).

以上所有实验均在具有AMD Ryzen5 1600X 3.6 GHz和8 GB内存的计算机上的MATLAB-2017b 中实现.为衡量实验结果的优劣,采用均方根误差 RMSE、标准差SD及训练时间和测试时间作为各预测 模型的评价标准.以RMSE最小的为最优结果,若 RMSE相同,则比较SD,最后比较时间.

## 4.1 FF-ARMAX模型对异常值的处理效果实验

为了验证所提方法,论证FF-ARMAX模型对异常值处理的有效性,首先进行FF-ARMAX模型性能检验实验.根据表1可知,T25和T29的原始数据中存在异常值.因此以下检验实验中,模型的输入变量中应当至少包含T25和T29其中一个.

根据上述分析,此次试验选取武钢某高炉实际生产数据.该数据是采样频率为60s的离散时间序列数据.实验模型的输入输出变量,根据第2.1节的方法,选取T36为输出变量,T27,T26,T24,T25,T37,

T23, T33 和T29 为 输 入 变 量. 实 验 模 型 采 用 RFMLS-RELM 算法建立, 参数配置为限定长度 Q = 50, 隐含层节点L = 10, 正则化系数 $C = 2^{10}$ , 激 励函数为sigmoid函数. 为验证FF-AMRAX模型对异 常值的处理效果, 实验数据分为两类: 经过处理的数 据和未经处理的原始数据,各选取1000组数据.其中 900组数据作为训练数据,100组数据作为测试数据. 为此需要建立两个模型,分别命名为实验模型(处理数 据)和实验模型(原始数据).表2为实验模型(处理数 据)和实验模型(原始数据)的预测结果.

表 2 实验模型(处理数据)和实验模型(原始数据)的实验结果 Table 2 Experimental results of experimental models (processing data) and experimental models (raw data)

Left Trai	训练时间/s		测试时	寸间/s	ᆒ	练	测	试
<b></b>	MEAN	SD	MEAN	SD	RMSE	SD	RMSE	SD
实验模型(处理数据)	8.4391	2.3079	0.1373	0.0295	0.0136	0.0020	0.0553	0.0022
实验模型(原始数据)	8.6729	2.2168	0.1395	0.0301	0.0763	0.0024	0.1082	0.0087

由表2可知, 在训练和测试时间上二者相差不多, 但是训练和测试RMSE却有很大差距. 实验模型(处理 数据)无论是在训练阶段还是测试阶段, 它的性能表现 都强于实验模型(原始数据), 说明原始数据在经过 FF-ARMAX模型去除异常值后, 对提高预测模型预 测精度有很大帮助. 证明了FF-ARMAX模型对异常 值处理的有效性, 解决了上文提到的问题1).

# 4.2 基于RFMLS-RELM的高炉煤气流温度预 测模型实验

# 4.2.1 RFMLS-RELM性能检测

为测试RFMLS-RELM算法在预测高炉煤气流温 度时的性能,以及检验RFMLS方法对传统ELM算法 和RELM算法的提升作用,本文在建立RFMLS-RELM预测模型的同时,采用ELM算法、RELM算法、 WOS-ELM算法和M-ARMAX算法作为对比算法 分别建立预测模型.但是RFMLS-RELM算法在实际 应用中的性能表现由参数配置决定,因此为了确保预 测模型的性能达到最佳,本文先进行调参实验,确定 预测模型最佳参数配比.

建立预测模型需要确定RFMLS-RELM的4个参数,即正则化系数C、限定长度Q、隐含层节点个数L 和激励函数.正则化系数C分别选取2<sup>10</sup>,2<sup>15</sup>,2<sup>20</sup>, 2<sup>25</sup>,2<sup>30</sup>,限定长度Q分别选取50,100,隐含层节点L 分别选取10,20,50,100,激励函数选用sigmoid函数.

根据上述分析笔者对RFMLS-RELM预测模型进行调参实验,预测模型经过30次重复试验,取实验结果的RMSE和方差作为预测模型的实验结果,针对不同参数配置的RFMLS-RELM预测模型的实验结果如表3所示.由表3可知,当 $Q = 50, L = 50, C = 2^{15}$ 时,预测结果的RMSE均值最小.故本文选取限定长度Q = 50,隐含层节点L = 50,正则化系数 $C = 2^{15}$ ,激励函数选用sigmoid函数的参数配置作为RFMLS-RELM预测模型的最终配置.

			$2^{1}$	0	$2^{1}$	$2^{15}$		$2^{20}$		$2^{25}$		$2^{30}$	
Q L	L	训练时间/s	RMSE	SD	RMSE	SD	RMSE	SD	RMSE	SD	RMSE	SD	
	10	8.4391	0.0553	0.0022	0.0545	0.0025	0.0552	0.0024	0.0576	0.0027	0.0581	0.0031	
50	20	16.4755	0.0466	0.0047	0.0473	0.0039	0.0484	0.0045	0.0492	0.0049	0.0501	0.0051	
	50	16.0714	0.0456	0.0049	0.0444	0.0055	0.0454	0.0051	0.0467	0.0049	0.0473	0.0046	
	100	21.9958	0.0676	0.0110	0.0669	0.0093	0.0671	0.0119	0.0677	0.0117	0.0682	0.0115	
	10	9.5427	0.0783	0.0039	0.0797	0.0048	0.0789	0.0058	0.0801	0.0039	0.0799	0.0042	
100	20	15.1510	0.0794	0.0077	0.0793	0.0075	0.0790	0.0057	0.0830	0.0066	0.0880	0.0065	
100	50	15.1766	0.1071	0.0090	0.1068	0.0101	0.1027	0.0134	0.1059	0.0147	0.1076	0.0128	
	100	20.2057	0.1904	0.0106	0.1912	0.0118	0.1920	0.0107	0.1920	0.0109	0.1923	0.0105	

表 3 RFMLS-RELM模型的调参实验结果 Table 3 Results of the adjustment experiments of RFMLS-RELM

在确定RFMLS-RELM预测模型的参数配置后, 进行不同预测模型的性能对比. 各预测模型的设置如 下:

1) ELM模型: 建立预测模型需要确定隐含层节点 个数*L*和激励函数. 经过试验, 最终确定*L* = 10, 激励 函数为sigmoid函数.

2) RELM模型: 建立预测模型需要确定正则化系数C、隐含层节点个数L和激励函数. 经过试验, 最终确定 $L = 10, C = 2^{30},$  激励函数为sigmoid函数建立 RELM预测模型.

3) WOS-ELM模型: 建立预测模型确定遗忘因 子 $\lambda$ 、隐含层节点个数L和激励函数. 经过试验, 最终 确定 $L = 10, \lambda = 0.9, 激励函数为sigmoid函数.$ 

4) M-ARMAX模型: 该模型的建立过程与第2.2 节中FF-ARMAX模型的建立过程相似,因此M-ARMAX模型无需参与调参实验,M-ARMAX模型 的阶次选择过程和系统辨识过程使用第2.2节中的结 果,即 $n_p = 3$ ,  $n_q = 3$ ,参数向量 $\theta$ 一致.

不同预测模型实验对比结果如表4所示.由表4可 知,ELM与RELM的训练和测试RMSE大致相同,说 明在预测煤气流温度的问题上两者性能接近.然而, 在训练时间和测试时间上,RELM均值和方差更小, 这说明正则化系数虽然没能提升ELM的性能,但是成 功解决了ELM因过度追求训练误差而导致训练时间 过长的问题.WOS-ELM的训练和测试RMSE均低于 ELM和RELM,说明遗忘因子λ的加入对ELM算法有 显著提升.M-ARMAX作为一种常见的系统建模方 法,在性能上与ELM和RELM相似,但是不太稳定(测 试RMSE的SD高于ELM,说明测试结果变动较大),但 是训练时间远大于两者,因此在实际应用中不具备竞 争力.

最后, RFMLS-RELM的训练和测试RMSE是所 有预测模型中最好的,证明RFMLS方法对RELM算法 的性能有很大提升,而且更加稳定(SD较RELM更小). 虽然训练时间和测试时间更长,但是从性能提升的角 度上来说是可以接受的,而且高炉本身就是大滞后系 统,所以在实际应用中这样的时间成本也是可以接受 的. 图9为运用RFMLS-RELM预测模型得到的煤气 流温度预测结果与实际数据的对比图.图10给出 RFMLS-RELM 预测煤气流温度的相对误差. 从图9 和图10可知RFMLS-RELM能够很好的预测煤气流 温度,为了进一步比较各预测模型的性能,图11给出 了RFMLS-RELM和4种对比模型的煤气流温度预测 值与真实值的对比结果,图12给出预测误差的对比结 果. 由图 11 可以看出, 相比其他预测模型 RFMLS-RELM具有更好的预测效果,并且能够更好的预测煤 气流温度的变化趋势.从图12可以看出, RFMLS-RELM预测模型的误差明显小于其他4种预测模型, 由此可知RFMLS-RELM具有较高的预测精度.

Table 4         Comparison of experimental results of different prediction models										
算法	训练时	间/s	测试时	†间/s	ᆒ	练	测	试		
	MEAN	SD	MEAN	SD	RMSE	SD	RMSE	SD		
ELM	0.0489	0.0156	0.0298	0.0078	0.1364	0.0021	0.1389	0.0054		
RELM	0.0258	0.0068	0.0200	0.0036	0.1365	0.0025	0.1376	0.0085		
WOS-ELM	0.0541	0.0385	0.0217	0.0146	0.0511	0.0118	0.0648	0.0090		
M-ARMAX	4.4016	0.1640	0.0657	0.0093	0.1287	0.0036	0.1389	0.0859		
RFMLS-RELM	15.9870	0.5120	0.1094	0.0334	0.0329	0.0040	0.0444	0.0055		

表 4 不同预测模型实验对比结果



















# 4.2.2 基于RFMLS-RELM算法的多步预测模型 性能检测

根据第4.2.1节的实验可知,采用RFMLS-RELM 方法建立的高炉煤气流预测模型具有较好的建模精 度,对高炉煤气流温度数据具有很高的拟合精度. 但是该模型的多步预测精度仍然有待考量,为此本 文将建立RFMLS-RELM多步预测模型与由ELM, RELM, WOS-ELM和M-ARMAX算法建立的多步 预测模型进行对比,验证RFMLS-RELM模型的多步 预测效果. 以上多步预测模型使用的数据是采样周期 为60s的离散时间序列数据,因此各多步预测模型的 预测步长也为60s与高炉实际生产数据采集周期相同. 为衡量多步预测模型对高炉煤气流温度的预测效果, 本文采用相对误差(relative errors, RE)比较各多步预 测模型的预测效果,计算每一步的预测值与真实值的 RE, 当RE < 10%时, 认为该步预测结果为有效预测 结果. 图13为RFMLS-RELM多步预测模型的预测结 果,图14为不同多步模型的预测结果对比图.

由图13可知, RFMLS-RELM单步预测模型的预测结果与真实值非常接近, 而多步预测模型会因为预测值替代真实值的缘故, 导致每一步预测结果都受到之前预测误差的影响, 预测误差不断叠加, 故随着预测步数的增加而预测误差越来越大. 由图14可知, 各多步预测模型在第1步至第3步的预测结果与真实值还比较接近, 除了M-ARMAX模型. 但是从第4步开始出现较大误差, 且随着预测步数的增加误差越来越

大,只是不同模型误差增大的速度不同,RFMLS-RELM与WOS-ELM增速较ELM和RELM而言相对 较慢,而M-ARMAX多步预测模型的预测误差增速 远大于其他模型.单纯从图像上,只能看出各多步预 测模型的大致表现,为衡量各模型性能需要详细信息. 因此表5展示了各多步预测模型的详细预测结果.



图 13 RFMLS-RELM多步预测模型的预测结果 Fig. 13 Prediction results of RFMLS-RELM multi-step

prediction model



图 14 不同多步预测模型的预测结果



由表5可知, RFMLS-RELM模型的多步预测结果 在第1步至第3步时还与单步预测结果接近,但是从 第4步开始误差开始不断加大,直至第8步相对误差超 过10%. 按上文定义, RFMLS-RELM多步预测模型 的预测结果中有效预测结果为7步.而ELM多步预 测模型与RELM多步预测模型性能表现相近,虽然变 化趋势不同但是有效预测结果均为5步,而且随着步 数增长,预测结果越偏离真实值,到第10步时甚至有 40%以上的误差,这对于实际应用而言是不可接受. WOS-ELM多步预测模型的预测结果相对较好,有效 预测结果有6步之多,比较接近RFMLS-RELM模型 的性能,但是它的误差增长速度同样很高,第10步时 同样也有35%以上的误差. 至于M-ARMAX多步预 测模型效果较差,有效预测结果只有3步,而且误差增 长速度很高,第10步时相对误差达到86%,这证明了 该模型不适用于高炉煤气流温度多步预测.

	Table 5 Prediction results of different multi-step prediction models												
上彩 古白仕	志 合 生 %		RFMLS-RELM		EI	ELM		RELM		WOS-ELM		M-ARMAX	
步致	一致 具实值/C	单步	RE/%	多步	RE/%	多步	RE/%	多步	RE/%	多步	RE/%	多步	RE/%
1	541.6	536.9	0.9	538.9	0.5	536.6	0.9	536.6	0.9	539.6	0.4	529.3	2.3
2	528.1	524.7	0.7	532.4	0.8	533.6	1.0	522.6	1.1	533.6	1.0	501.3	5.1
3	511.7	516.2	0.9	519.0	1.4	521.6	1.9	519.6	1.5	520.6	1.7	462.3	9.6
4	492.6	497.1	0.9	507.0	2.9	509.6	3.5	476.6	3.2	507.6	3.1	437.4	11.2
5	471.1	475.7	1.0	496.6	5.4	503.6	6.9	442.5	6.1	441.6	6.3	391.4	16.9
6	448.0	444.3	0.8	483.0	7.8	496.6	10.9	401.5	10.4	407.6	9.0	353.4	21.1
7	423.7	427.3	0.8	461.9	9.0	482.6	13.9	367.5	13.3	376.6	11.1	276.5	34.8
8	399.1	402.7	0.9	446.7	11.9	502.6	25.9	304.6	23.7	321.6	19.4	174.5	56.3
9	374.8	377.4	0.7	449.4	19.9	519.6	38.6	236.6	36.9	289.6	22.7	91.5	75.6
10	351.6	354.9	0.9	445.3	26.7	492.6	40.1	190.6	45.8	226.6	35.6	46.6	86.8

由上述实验可知,在多步预测模型中,由于每一步 预测结果都会参与到接下来的预测工作中,导致每一 步预测产生的预测误差也会叠加到之后的预测结果 中,所以多步预测结果的精度必然会随着预测步数的 增加而不断增加.本文所提出的RFMLS-RELM算 法,在高炉煤气流温度多步预测上,相对其他传统多 步预测模型实现了更好的预测效果.虽然无法改变预 测误差不断叠加的问题,但是依然取得较好的结果, 因此RFMLS-RELM多步预测模型更适用于高炉煤 气流温度的多步预测.

## 5 结论

针对高炉煤气流传统预测模型的缺陷,本文提出 了一种基于FF-ARMAX和RFMLS-RELM的高炉 煤气流多步预测模型. 在数据修正方面, 建立FF-ARMAX模型消除原始数据中因十字测温装置失效或 损坏导致的测量误差,同时采用傅里叶变换法消除数 据中叠加的环境噪声,最后采用RELM算法进行多步 预测,并且运用RFMLS优化RELM隐含层到输出层的 权值矩阵, 消除了历史数据对模型的影响, 避免了"数 据饱和"问题的出现.对比试验表明,该算法在应用 于高炉煤气流预测时,预测精度更高,因此该算法适 用于对高炉煤气流分布状况的多步预测.同时多步预 测实验结果表明,该模型虽然仍旧无法完全解决预测 误差随预测步数的增加而不断叠加的问题,但相较于 其他传统预测模型能够获得更多的有效预测步数,实 现更高的预测精度,为建立更加精确的高炉料面温度 场提供了可能,为高炉操作人员分析炉况提供了有效 的帮助和支持.

# 参考文献:

[1] ZHANG Haigang, ZHANG Sen, YIN Yixin, et al. Feature extraction and clustering of blast furnace burden surface. *Control Theory & Ap*- plications, 2017, 34(7): 938 – 946. (张海刚,张森, 尹怡欣,等.高炉料面特征提取与聚类分析. 控制理 论与应用, 2017, 34(7): 938 – 946.)

- [2] YANG Y L, YIN Y X, WUNSCH D, et al. Development of blast furnace burden distribution process modeling and control. *The Iron and Steel Institute of Japan International*, 2017, 57(8): 1350 – 1363.
- [3] LI Yanjiao, ZHANG Sen, YIN Yixin, et al. Research on optimizaion model of blast furnace burden surface based on data driven. *Control Theory & Applications*, 2018, 35(3): 324 – 334.
  (李艳姣, 张森, 尹怡欣, 等. 基于数据驱动的高炉料面优化决策模型 研究. 控制理论与应用, 2018, 35(3): 324 – 334.)
- [4] LI M Y, PENG X D, WANG W. Research on identification of gas flow distribution with wavelet feature extraction. *The 2011 Fourth International Symposium on Knowledge Acquisition and Modeling*. Piscataway, NJ: IEEE, 2011: 129 – 132.
- [5] ZHANG K X, WU M, AN J Q, et al. Relation model of burden operation and state variables of blast furnace based on low frequency feature extraction. *The 20th IFAC World Congress*. Amsterdam: Netherlands, Elsevier, 2017, 50(1): 13796 – 13801.
- [6] SHI L, WEN Y B, ZHAO G S, et al. Recognition of blast furnace gas flow center distribution based on infrared image processing. *Journal* of Iron and Steel Research International, 2016, 23(3): 203 – 209.
- [7] AN Jianqi, LIU Yang, WU Min, et al. Reconstruction of 3D model of gas flow distribution in blast furnace throat. *Control Theory & Applications*, 2014, 31(5): 624 631.
  (安剑奇, 刘洋, 吴敏, 等. 高炉炉喉煤气流形态三维模型重建. 控制 理论与应用, 2014, 31(5): 624 631.)
- [8] ZHOU Ping, LIU Jiping. Data-driven multi-output ARMAX modeling for online estimation of central temperatures for cross temperature measuring in blast furnace ironmaking. *Acta Automatica Sinica*, 2018, 44(3): 552 – 561.
  (周平,刘记平. 基于数据驱动多输出ARMAX建模的高炉十字测温 中心温度在线估计. 自动化学报, 2018, 44(3): 552 – 561.)
- [9] XU Meiling, HAN Min. Factor echo state network for multivariate chaotic time series prediction. *Acta Automatica Sinica*, 2015, 41(5): 1042 – 1046.

(许美玲,韩敏.多元混沌时间序列的因子回声状态网络预测模型. 自动化学报,2015,41(5):1042-1046.)

- [10] CHAURASIA A, HAREL O. Using AIC in multiple linear regression framework with multiply imputed data. *Health Services and Outcomes Research Methodology*, 2012, 12(2/3): 219 233.
- [11] LU Xiaobo, LING Xiaojing, LIU Bin. License plate character recognition based on the combined features. *Chinese Journal of Scientific Instrument*, 2006, 27(7): 698 701.
  (路小波, 凌小静, 刘斌. 基于组合特征的车牌字符识别. 仪器仪表学报, 2006, 27(7): 698 701.)
- [12] SU Xiaoli, YIN Yixin, ZHANG Sen. Prediction model of improved multi-layer extreme learning machine for permeability index of blast furnace. *Control Theory & Applications*, 2016, 33(12): 1674 – 1684.
  (苏晓莉, 尹怡欣, 张森. 高炉透气性指数的改进多层超限学习机预 测模型. 控制理论与应用, 2016, 33(12): 1674 – 1684.)
- [13] ZHANG Mingyang, WEN Yingyou, YANG Xiaotao, et al. An incremental weighted average based online sequential extreme learning machine algorithm. *Control & Decision*, 2017, 32(10): 1887 1893.
  (张明洋,闻英友,杨晓陶,等. 一种基于增量加权平均的在线序贯极 限学习机算法. 控制与决策, 2017, 32(10): 1887 1893.)

- [14] MIRZA B, LIN Z, TOH K A. Weighted online sequential extreme learning machine for class imbalance learning. *Neural Processing Letters*, 2013, 38(3): 465 – 486.
- [15] ZHANG K, LUO M X. Outlier-robust extreme learning machine for regression problems. *Neurocomputing*, 2015, 151(3): 1519 – 1527.
- [16] RAKHA M A. On the Moore-Penrose generalized inverse matrix. Applied Mathematics and Computation, 2004, 158(1): 185 200.

#### 作者简介:

**吴晓阳** 硕士研究生,目前研究方向为复杂系统建模与控制、机器学习和流程工业自动化, E-mail: s20170614@xs.ustb.edu.cn;

**张 森** 教授,目前研究方向为复杂系统建模与控制和现代大型 高炉布料系统, E-mail: zhangsen@ustb.edu.cn;

**陈先中** 教授,目前研究方向为电磁场与微波技术、工业雷达探测与成像、工业物联网与软件开发,E-mail: cxz@ustb.edu.cn;

**尹怡欣** 教授,目前研究方向为复杂系统建模与控制、人工智能和流程工业自动化, E-mail: yyx@ies.ustb.edu.cn.