

分段线性回归和动态加权神经网络融合的高炉料位预测

蒋朝辉, 李晔月[†], 桂卫华, 谢永芳, 阳春华

(中南大学 信息科学与工程学院, 湖南 长沙 410083)

摘要: 针对高炉料位难以连续高精度测量的问题, 提出了一种基于分段线性回归和动态加权神经网络的高炉料位信息预测方法. 首先, 通过分析高炉布料机制和料位检测数据特点, 提出了一种面向雷达和机械探尺检测数据时间序列的联合划分方法, 用于提取高炉料位的周期性变化特征; 然后, 利用该变化特征构建分段线性回归模型, 获得能准确描述料位变化的回归曲线; 最后, 以回归统计指标为权重调节系数, 利用动态加权径向基神经网络对料位信息进行预测. 实例验证表明, 该方法融合了机械探尺检测数据精度高以及雷达检测数据连续性好的特点, 实现了高炉料位信息的实时有效预测.

关键词: 高炉; 料位; 预测; 分段线性回归; 动态加权; 神经网络

中图分类号: TP273

文献标识码: A

Blast furnace stockline prediction by segmented linear-regression and dynamic weighting neural network

JIANG Zhao-hui, LI Xi-yue[†], GUI Wei-hua, XIE Yong-fang, YANG Chun-hua

(School of Information Science and Engineering, Central South University, Changsha Hunan 410083, China)

Abstract: Because of difficulties in continuously measuring the blast furnace stockline precisely, we propose a new prediction method based on segmented linear-regression (SLR) and dynamic weighting neural network (DWNN). According to blast furnace burden distribution schemes and data characteristics of blast furnace stockline, we design a combined division method based on the time series of radar and the mechanical stock rod data to extract periodical variation features of the blast furnace stockline. Then, a segmented linear-regression model is built based on the periodical variation features which help to obtain regression curves that precisely reflect the change of stockline. Finally, we take regression statistical indexes as coefficients of the regulation weights, and construct a dynamic weighting radial basis function (RBF) neural network model to predict the information of the blast furnace stockline. Case study indicates that the proposed method combines the high-precision of mechanical stock rod data with the continuity of radar data, and provides with real-time prediction for blast furnace stockline information effectively.

Key words: blast furnace; stockline; prediction; segmented linear-regression; dynamic weighting; neural network

1 引言(Introduction)

高炉料位是高炉布料操作中的一个重要调控参数, 其测量数据的准确性和连续性不仅是提高生产效率和保证高炉炉况顺行必要条件, 而且是改善料面分布形状和布料精度的重要依据. 因此, 实时准确地获取高炉料位信息对改善高炉布料、优化炉内煤气流分布、提高高炉产量、降低能耗、保证高炉顺行具有重要意义.

目前广泛使用的高炉料位检测手段主要有机械探尺检测和雷达检测. 机械探尺检测是利用机械探尺在高炉内对料面进行直接接触式测量, 精度和稳定性高;

但存在测量周期长、不能对料位进行连续测量、且加料时无法工作等缺陷^[1]. 雷达检测是利用电磁波测距原理对料面进行非接触式测量, 不受加料过程限制, 可实现料位的连续测量, 因此现有技术大多采用多雷达数据融合的方法对料位进行分析^[2-3], 但雷达波易受炉内粉尘、气流等因素的干扰, 且检测数据是微波覆盖面积的平均值, 存在检测精度不高, 稳定性差的缺点. 此外, 高炉料位检测手段还有激光检测、红外成像、视频监控等, 这些方法或直接测量, 或间接推算, 但由于高炉炉内环境恶劣, 其检测结果也极易受炉顶高温、高粉尘、煤气流等因素影响, 导致这些方法均难

收稿日期: 2015-01-12; 录用日期: 2015-03-19.

[†]通信作者. E-mail: lixychn@163.com; Tel.: +86 18774940442.

国家自然科学基金重大项目(61290325), 国家自然科学基金创新研究群体科学基金项目(61321003)资助.

Supported by Major Program of the National Natural Science Foundation of China (61290325) and Foundation for Innovative Research Groups of National Natural Science Foundation of China (61321003).

以连续获得高精度的料位检测信息,只能作为开炉试验的辅助手段或用于料面分布的定性分析^[4-6].建模方面,由于高炉冶炼过程复杂,系统各参数间存在不完全确定的复杂耦合关系,特别是炉料消耗下降过程,存在较多的未知影响因素,使得料位的变化具有一定的随机性和不确定性,因此目前通过分析炉料运动规律建立的布料模型难以实现料位的精确计算^[7-8].高炉料位检测和建模精度不高,增加了高炉料位预测的难度.

综上所述,现有的检测方法和建模方法难以同时实现对高炉料位连续且高精度的检测,使得布料操作具有一定的盲目性,影响了布料操作的合理性和安全性.为此本文结合机械探尺检测精度高和雷达连续测量的特点,提出了一种基于机械探尺和雷达检测数据融合的高炉料位信息实时预测方法.其功能框图如图1所示.

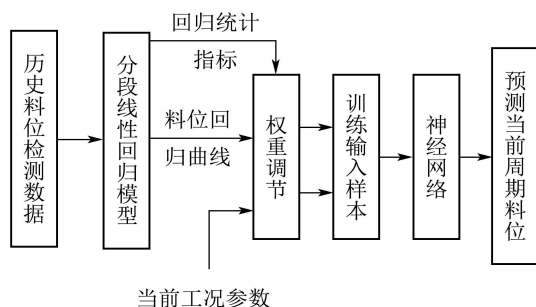


图1 料位信息预测功能框图

Fig. 1 Flowchart of stockline information prediction

首先,通过对机械探尺和雷达的检测数据进行分析,结合高炉布料制度提出了一种针对雷达检测数据非线性时间序列和机械探尺检测数据离散时间序列的联合划分方法,提取出高炉料位的周期性变化特征;其次,将线性回归理论与料位的周期性变化特征相结合,构建分段线性回归模型,获得料位回归曲线;然后,以料位回归曲线和当前的工况参数为输入,以回归统计指标为权重调节参数,构建训练输入样本,利用动态加权RBFN预测方法对料位信息进行实时预测.

2 分段线性回归模型 (Segmented linear-regression model)

高炉按批布料,每批料分为加料和等待两个阶段.每批料的加料时间通常为1.5 min左右,加完料后进入数分钟的等待时间,等待结束后又进入加料阶段,依次循环.因此,高炉料位呈周期性变化.以某钢铁企业高炉为例,通过分析其料位检测数据可知,料位可高度近似为具有周期性的分段线性特征,如图2所示,其纵坐标是料面到炉顶的距离,所以料位数值下降代表料面上升即加料阶段,反之为实际料面下降即等待阶段.图2所示的这一规律性布料制度为本文的研究提供了重要依据.

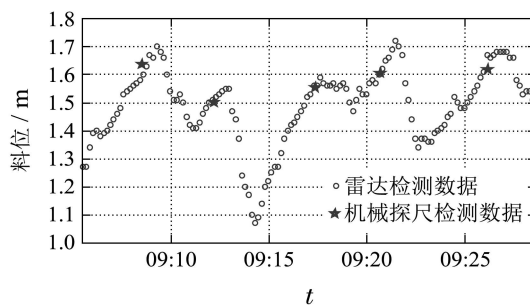


图2 料位检测数据图

Fig. 2 Industrial data of stockline

加料阶段和等待阶段的料位变化主要受加料速度和料耗速度的影响.在理想工况下,加料速度和料耗速度近似匀速^[9],这与料位检测数据中呈现出的近似分段线性特征相吻合.因此,本文通过提取料位检测数据中的周期性变化特征来建立高炉料位变化模型.

2.1 料位周期性变化特征提取 (Extraction of periodical variation features of stockline)

在高炉实际生产过程中,雷达可实现料位的近似连续测量(采样周期为10 s),机械探尺只在每批布料的等待时间测量一次.因此,雷达检测数据为一个周期性变化的时间序列,而机械探尺检测数据则是一个间隔较大的离散时间序列,雷达和机械探尺检测数据时间序列的每个数据单元可以抽象为一个二元组 (t, l) ,其中: t 为时间变量; l 为料位检测值.给出如下定义:

定义1 雷达检测数据时间序列 Q 为有限集 $\{(t_1, l_1), (t_2, l_2), \dots, (t_p, l_p)\}$,且满足 $t_j < t_{j+1} (j = 1, 2, \dots, p-1)$;机械探尺检测数据时间序列 R 为有限集 $\{(t_{R_0}, l_{R_0}), (t_{R_1}, l_{R_1}), \dots, (t_{R_n}, l_{R_n})\}$,满足 $t_{R_i} < t_{R_{i+1}} (i = 0, 1, \dots, n-1)$.其中: p 为雷达数据采样个数; n 为机械探尺数据采样个数.

雷达和机械探尺检测数据形成的两个时间序列在时间上具有一定的关联度,因此通过联合划分,可以获得料位的周期性分段线性特征.联合划分的具体步骤如下:

步骤1 以机械探尺检测数据时间点 $t_{R_i} (i = 0, 1, \dots, n)$ 为基准,在雷达检测数据时间序列中划分出 n 个探尺检测周期;

步骤2 在每个探尺检测周期中,确定雷达检测数据时间序列的极值;

步骤3 根据当前布料操作制度和机械探尺测量时刻确定料位达到最大值和最小值的时间范围,并对步骤2所得的雷达检测数据极值进行筛选,获得每个探尺检测周期中雷达检测数据的最大值点和最小值点,记为 $(t_{\max_i}, l_{\max_i})$ 和 $(t_{\min_i}, l_{\min_i})$, $i = 1, 2, \dots, n$;

步骤4 对于一个周期中出现多个相邻且等值的极值点的情况, 只保留最前和最后两个时刻的极值点; 经步骤3和步骤4筛选出的最大值点和最小值点即为每个周期加料的起始点和终止点;

步骤5 以加料起始点划分出 n 个布料周期, 得到料位的周期性分段线性特征.

料位分段线性特征表示如下:

集合 $U_i = \{(t_j, l_j) \in Q, (t_{R_i}, l_{R_i}) \in R \mid t_{\min_i} \leq t_j \leq t_{\max_i}\}$, 记为第 i 个周期的上升段料位序列;

集合 $D_i = \{(t_j, l_j) \in Q \mid t_{\max_i} \leq t_j \leq t_{\min_{i+1}}\}$, 记为第 i 个周期的下降段料位序列;

集合 $L = \{L_i \mid L_i = \{U_i, D_i\}, i = 1, 2, \dots, n\}$, 记为雷达和机械探尺检测数据时间序列的联合划分集. 联合划分如图3所示.

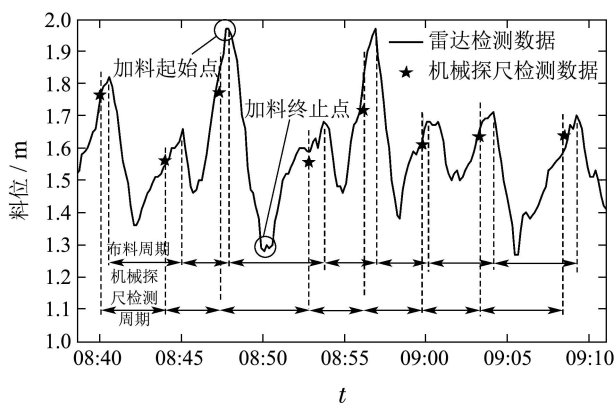


图3 雷达和机械探尺检测数据联合划分示意图

Fig. 3 Combined division for the data of radar and mechanical stock rod

如图3所示, 每个探尺检测周期中有一个加料起始点和一个加料终止点; 以加料起始点划分布料周期, 每周期内下降段为加料阶段, 上升段为等待阶段, 探尺检测时刻在等待阶段中.

2.2 基于布料制度的分段线性回归建模(Segmented linear-regression modeling based on burden distribution schemes)

机械探尺为直接接触式测量设备, 精度和稳定性高, 其检测到的料位信息是高炉实际生产中用于定量调控布料的主要参考依据. 因此, 本文将机械探尺的检测数据视为无误差的料位真值, 将其作为基准点进行建模. 但一个布料周期内机械探尺检测数据只有1个, 无法反映布料周期内料位的变化趋势, 因此, 需要结合雷达检测数据的周期性分段线性特征, 提取雷达检测数据时间序列中的周期模式, 建立料位变化模型.

分析一个周期内的雷达检测数据时间序列可知, 雷达的料位检测值与时间变量呈近似线性关系, 但由

于雷达检测数据精确度不高, 且易引入噪声干扰等随机误差, 这种关系难以用完全确定的函数来表达. 因此, 本文采用回归分析^[10]对单个周期的雷达检测数据时间序列进行分析, 以消除雷达数据中随机误差带来的影响, 并将线性回归理论与料位的周期性变化特点相结合, 构建分段线性回归模型, 其过程如下:

1) 建立回归观测方程.

设单个周期内的雷达检测数据时间序列为 $Q_i = \{(t_1, l_1), (t_2, l_2), \dots, (t_m, l_m)\}$, 则单个周期内分段线性回归方程为

$$l = \begin{cases} a_d + b_d t, & t_1 \leq t \leq t_{\min}, \\ a_u + b_u t, & t_{\min} \leq t \leq t_m, \end{cases} \quad (1)$$

其中: a_d, b_d 为加料段回归系数, a_u, b_u 为等待段回归系数, $t_{\min}$ 为一个周期内料位最小值时刻点, 即加料终止时刻.

设雷达料位检测值 l 的测量随机误差为 ε , 雷达检测时间 t 的测量误差不计, 则单个周期内雷达检测数据样本的分段线性回归的观测方程为

$$l_j = \begin{cases} a_d + b_d t_j + \varepsilon_j, & t_1 \leq t_j \leq t_{\min}, \\ a_u + b_u t_j + \varepsilon_j, & t_{\min} \leq t_j \leq t_m, \end{cases} \quad (2)$$

其中 $j = 1, 2, \dots, m$. 根据料位时间序列联合划分集, 将单个周期内回归方程(1)扩展至多周期, 即对雷达检测数据时间序列 Q , 构建 n 个周期的分段线性回归方程:

$$l = \sum_{i=1}^n l_i = \sum_{i=1}^n \begin{cases} a_{d_i} + b_{d_i} t, & t_{\max_i} \leq t \leq t_{\min_i}, \\ a_{u_i} + b_{u_i} t, & t_{\min_i} \leq t \leq t_{\max_{i+1}}, \end{cases} \quad (3)$$

其中: $t_{\max_i}, t_{\min_i}$ 分别为第 i 个周期内料位最大值和最小值时刻点, 即加料的起始时刻和终止时刻; a_{d_i}, b_{d_i} 分别为第 i 个周期内加料段回归系数, a_{u_i}, b_{u_i} 分别为第 i 个周期内等待段回归系数.

由回归方程(3), 可得多周期雷达检测数据时间序列的分段线性回归观测方程为

$$l_{ij} = \begin{cases} a_{d_i} + b_{d_i} t_{ij} + \varepsilon_{ij}, & t_{\max_i} \leq t_{ij} \leq t_{\min_i}, \\ a_{u_i} + b_{u_i} t_{ij} + \varepsilon_{ij}, & t_{\min_i} \leq t_{ij} \leq t_{\max_{i+1}}, \end{cases} \quad (4)$$

其中: $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m_i$; m_i 为第 i 个周期的雷达检测数据序列点数; (t_{ij}, l_{ij}) 则为第 i 个周期的第 j 个雷达检测数据序列点, ε_{ij} 为该点的测量误差.

为准确描述料位的实际变化, 需综合分析雷达和机械探尺检测数据. 将机械探尺检测数据序列 (t_{R_i}, l_{R_i}) , $i = 1, 2, \dots, n$, 作为基准参考点加入到式(4)中, 构建如式(5)所示的基于雷达和机械探尺检测数据的时间分段线性回归模型:

$$l_{ij} = \begin{cases} b_{u_i} t_{ij} - b_{u_i} t_{R_i} + l_{R_i} + \varepsilon_{ij}, \\ t'_{\min_i} \leq t_{ij} \leq t_{\max_{i+1}}; \\ a_{d_i} - a_{u_i} + b_{d_i} t_{ij} - b_{u_i} t_{R_i} + l_{R_i} + \varepsilon_{ij}, \\ t_{\max_{i+1}} \leq t_{ij} \leq t'_{\min_{i+1}}, \end{cases} \quad (5)$$

其中: $i = 1, 2, \dots, n$; $j = 1, 2, \dots, m_i$; $t'_{\min_i}$ 为周期划分点, 定义第 i 个周期的上升段函数为 $l = l_{\text{up}_i}(t)$, 下降段函数为 $l = l_{\text{dn}_i}(t)$, 则 $t'_{\min_i}$ 满足 $l_{\text{dn}_{i-1}}(t'_{\min_i}) = l_{\text{up}_i}(t'_{\min_i})$.

式(5)是经过融合修正后的料位分段线性回归模型, 为有效融合机械探尺和雷达检测数据, 在不改变分段线性特征的前提下, 将等待阶段定义为周期的开始, 并且在对加料终止时刻 $t_{\min_i}$ 微调的基础上确定周期划分点 $t'_{\min_i}$, 使各周期保持连续.

分段线性回归模型示意图如图4所示.

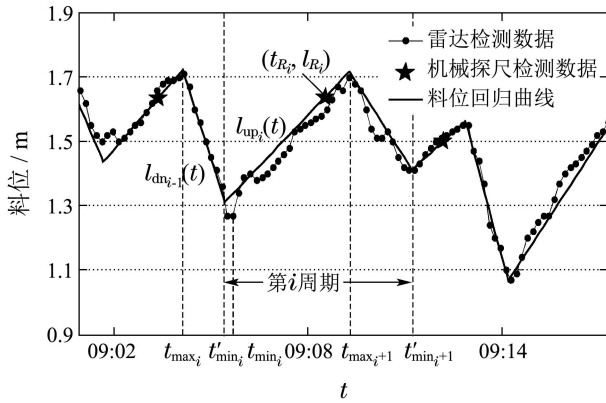


图4 分段线性回归模型示意图

Fig. 4 Schematic diagram of the segmented linear regression model

2) 估计模型参数.

采用最小二乘法估计式(5)中的参数 a_{u_i} , b_{u_i} , a_{d_i} , b_{d_i} , $i = 1, 2, \dots, n$. 以上升段为例, 令

$$\begin{cases} Q(a_{u_i}, b_{u_i}) = \sum_{j=1}^{m_{\text{up}_i}} (l_{ij} - a_{u_i} - b_{u_i} t_{ij})^2, \\ Q(\hat{a}_{u_i}, \hat{b}_{u_i}) = \min_{a_{u_i}, b_{u_i}} Q(a_{u_i}, b_{u_i}), \end{cases} \quad (6)$$

其中 m_{up_i} 为第 i 个周期的上升段雷达检测数据序列点数.

对式(6)进行正规方程组求解, 可求得 a_{u_i} , b_{u_i} 的最小二乘无偏估计值为

$$\begin{cases} \hat{b}_{u_i} = \frac{\sum_{j=1}^{m_{\text{up}_i}} [(t_{ij} - \bar{t}_i)(l_{ij} - \bar{l}_i)]}{\sum_{j=1}^{m_{\text{up}_i}} (t_{ij} - \bar{t}_i)^2}, \\ \hat{a}_{u_i} = \bar{l}_i - \hat{b}_{u_i} \bar{t}_i, \end{cases} \quad (7)$$

其中: $\bar{t}_i = \frac{1}{m_{\text{up}_i}} \sum_{j=1}^{m_{\text{up}_i}} t_{ij}$, $\bar{l}_i = \frac{1}{m_{\text{up}_i}} \sum_{j=1}^{m_{\text{up}_i}} l_{ij}$.

同理可求得 a_{d_i} , b_{d_i} 的最小二乘无偏估计值 $\hat{a}_{d_i}$, $\hat{b}_{d_i}$.

3) 分段线性回归模型的显著性检验.

以回归模型的第 i 周期为例, 在式(5)中, t_{ij} 处的回归值记为 $\hat{l}_{ij}$, 则 t_{ij} 处的残差为 $l_{ij} - \hat{l}_{ij}$, 其回归平方和 S_{R_i} 及残差平方和 S_{e_i} 分别如下:

$$S_{R_i} = \sum_{j=1}^{m_i} (\hat{l}_{ij} - \bar{l}_i)^2, \quad S_{e_i} = \sum_{j=1}^{m_i} (l_{ij} - \hat{l}_{ij})^2,$$

则线性回归分析的 F 检验统计量、 R^2 检验统计量 (R 是相关系数) 和标准差的无偏估计值为

$$\begin{cases} F_i = \frac{S_{R_i}}{S_{e_i}/(m_i - 2)}, \\ R_i^2 = \frac{F_i}{F_i + (m_i - 2)}, \\ \hat{\sigma}_i = \sqrt{\frac{S_{e_i}}{(m_i - 2)}}. \end{cases} \quad (8)$$

对于给定的显著性水平 α , 回归模型具有显著性的判断条件为 $F_i \geq F_{1-\alpha}(1, m_i - 2)$. R_i^2 反映了雷达检测数据落入回归模型置信区间的命中率. 回归模型在估计值 $\hat{l}_{i0}$ 处的 $1 - \alpha$ 的置信区间为 $[\hat{l}_{i0} - \delta_o, \hat{l}_{i0} + \delta_o]$, 其中:

$$\delta_o = t_{1-\alpha/2}(m_i - 2) \hat{\sigma}_i \sqrt{\frac{1}{n} + \frac{(t_{i0} - \bar{t}_i)^2}{\sum_{j=1}^{m_i} (t_{ij} - \bar{t}_i)^2}}, \quad (9)$$

F 和 R^2 是分段线性回归模型的两个主要的检验统计量, 其值越高, 则料位检测数据的分段线性特征越显著, 即所建模型与料位检测数据时间序列的吻合度越高.

2.3 分段线性回归建模结果分析(Analysis of segmented linear-regression modeling result)

式(5)描述的是融合了雷达和机械探尺检测数据的分段线性回归模型, 既保证了机械探尺的料位精度, 又有效融合了雷达检测数据中连续的料位信息.

对建模结果进一步分析, 式(5)实质上是一个回归分析的观测方程, 描述了每个检测时刻与其对应料位值的关系, 其中包含了每个时刻点的回归均值和随机误差两部分. 当只考虑回归均值而不考虑随机误差时, 可得对应的回归方程如下:

$$l_i = \begin{cases} b_{u_i} t_i - b_{u_i} t_{R_i} + l_{R_i}, \\ t'_{\min_i} \leq t_i \leq t_{\max_{i+1}}; \\ a_{d_i} - a_{u_i} + b_{d_i} t_i - b_{u_i} t_{R_i} + l_{R_i}, \\ t_{\max_{i+1}} \leq t_i \leq t'_{\min_{i+1}}. \end{cases} \quad (10)$$

再令 $a'_{u_i} = l_{R_i} - b_{u_i} t_{R_i}$, $a'_{d_i} = a_{d_i} - a_{u_i} + l_{R_i} - b_{u_i} t_{R_i}$, 则式(10)可以改写为

$$l_i = \begin{cases} a'_{u_i} + b_{u_i} t_i, & t'_{\min_i} \leq t_i \leq t_{\max_{i+1}}, \\ a'_{d_i} + b_{d_i} t_i, & t_{\max_{i+1}} \leq t_i \leq t'_{\min_{i+1}}. \end{cases} \quad (11)$$

式(11)中, a'_{u_i} 与 a'_{d_i} 分别表征等待阶段和加料阶段料位的高低水平, b_{u_i} 与 b_{d_i} 分别表征等待阶段和加料阶段料位的变化速度, $t'_{\min_i}$ 与 $t_{\max_{i+1}}$ 分别表征加料的终止时刻和起始时刻。

由此分析可知, 式(5)所示的分段线性回归模型对应的回归方程体现了料位变化的各个关键信息, 同时有效地融合了雷达和机械探尺的检测数据。该分段线性回归模型给出了每个周期内等待阶段和加料阶段的料位变化描述, 并以表征料位高低水平、料位变化速度和加料起止时刻的6个特征参数(a'_{u_i} , a'_{d_i} , b_{u_i} , b_{d_i} , $t'_{\min_i}$, $t_{\max_{i+1}}$)作为模型的最终输出结果。此外, 根据式(9)可得出料位波动的置信区间。

3 基于动态加权神经网络的高炉料位信息预测(Blast furnace stockline information prediction based on DWNN)

由于实际工况中影响料位变化的因素十分复杂, 且存在一定的随机性和不确定性, 因此本文在建立高炉料位分段线性回归模型的基础上, 通过权重调节构建输入样本, 并结合神经网络对料位进行预测。

神经网络, 特别是径向基神经网络(radial basis function networks, RBFN)被广泛用于时间序列预测^[11-13]。近年来, 有学者把神经网络和一些传统模型相融合, 提出了变权重的神经网络, 如文[14]中提出了一种带回归权重的径向基神经网络模型, 该模型中RBFN的输出权重不是单一的常量, 而是输入变量的线性回归函数。针对高炉料位的变化特点, 本文通过对料位分段线性回归模型的参数进行时滞性分析, 结合高炉生产的相关工况参数, 利用回归分析的显著性检验统计量作为样本权重调节系数, 以RBFN为基础, 提出动态回归加权RBFN料位预测方法, 对高炉料位信息进行实时预测。

3.1 训练样本结构(Structure of training sample)

3.1.1 料位模型参数的时滞性分析(Time-lag correlation analysis of the stockline model parameters)

对于式(11)中的6个特征参数, 定义矩阵 $P_i = (a'_{u_i}, a'_{d_i}, b_{u_i}, b_{d_i}, t'_{\min_i}, t_{\max_{i+1}})^T_{6 \times 1}$ 为第*i*批次的料位模型特征参数。 P_i 中表征料位高低和料位变化速度的前4个特征参数存在一定的时滞性, 原因在于: 1) 在高炉生产过程中, 布料按批次进行, 历史批次的料位检测信息直接决定了当前批次的加料量和加料速度; 2) 高炉冶炼是个惯性过程, 当前时期的工况受前一时期工况的影响, 因此当前批次等待段的料位高低和料位变化速度(即料耗速度)与其历史值有很大关联。因此, 本文对所建料位模型的 a'_{u_i} , b_{u_i} , a'_{d_i} , b_{d_i} 4个参数进行时滞相关性分析, 分别计算它们当前批次数据与

其历史批次数据的Pearson相关性, 利用历史料位特征参数对当前料位进行预测。时滞性分析结果如表1所示。

表1 料位模型特征参数时滞性分析结果

Table 1 Time-lag analysis of the stockline model parameters

Pearson 相关性	$i-1$	$i-2$	$i-3$	$i-4$	$i-5$
a'_{u_i}	0.701	0.712	-0.632	-0.554	0.439
a'_{d_i}	-0.718	0.721	-0.673	0.583	-0.564
b_{u_i}	0.789	0.737	-0.643	-0.579	0.516
b_{d_i}	-0.910	0.839	-0.806	0.786	-0.676

相关性分析结果表明, 前3列均在0.05水平(双侧)上显著相关, 且相关系数绝对值大于0.6。因此, 选取前3个批次特征参数作为预测输入变量, 记为 $\{P_{i-3}, P_{i-2}, P_{i-1}\}$ 。此外, 机械探尺检测数据 $J_i = (t_{R_i}, l_{R_i})^T_{2 \times 1}$ 作为基准点, 对料位的预测具有决定性作用, 因此也选取前3个批次的机械探尺数据作为预测输入变量, 记为 $\{J_{i-3}, J_{i-2}, J_{i-1}\}$ 。输出样本由当前批次能够表征料位变化速度和高低水平的4个料位模型特征参数构成, 定义为矩阵 $P_{out_i} = (a'_{u_i}, a'_{d_i}, b_{u_i}, b_{d_i})^T_{4 \times 1}$ 。

3.1.2 工况参数的相关性分析(Correlation analysis of working condition parameters)

加料速度与布料操作有直接关系, 而高炉的生产工况也影响着料耗速度。因此, 进行料位预测时分析高炉实时工况参数对料位的影响有助于提高预测精度。与布料时间相关的参数主要有: 布料圈数和布料起始时刻; 与加料速度相关的参数主要有: 料流阀开度和溜槽转速; 与料耗速度相关的参数主要有: 炉顶温度、炉顶压力、风量、风温等。将这些工况参数分别与同批次输出样本中的特征参数 b_u , b_d 进行相关性分析, 结果如表2所示。

表2 工况参数相关性分析结果

Table 2 Correlation analysis of working condition parameters

Pearson 相关性	布料 圈数	布料 时刻	料流阀 开度	溜槽 转速
b_u	0.405	0.128	0.627	0.431
b_d	0.234	0.376	0.818	0.647
Pearson 相关性	顶温	顶压	风量	风温
b_u	-0.741	-0.640	0.717	0.602
b_d	-0.615	-0.574	0.589	0.392

根据分析结果, 选取相关性较大的料流阀开度

v_i 、炉顶温度 t_i 、风量 f_i 3个当前工况参数为预测输入变量,并定义矩阵 $O_i = [v_i \ t_i \ f_i]_{3 \times 1}^T$ 为第 i 周期的工况参数。

综上可知,料位预测的训练输入样本可由3部分组成: 1) 历史料位检测数据的特征参数 $\{P_{i-3}, P_{i-2}, P_{i-1}\}$; 2) 历史的机械探尺基准点 $\{J_{i-3}, J_{i-2}, J_{i-1}\}$; 3) 高炉当前工况参数 O_i 。输出样本为包含当前批次4个料位模型特征参数的矩阵 P_{out_i} 。训练样本的结构如图5所示。

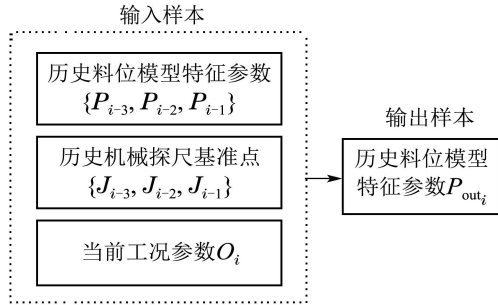


图5 训练样本结构示意图

Fig. 5 The structure of training sample

3.2 训练样本的权重调节 (Weights regulation of training sample)

当出现工况波动或存在突发性干扰时,雷达检测数据引入的随机误差增大,表现为数据的线性回归方程显著性降低,此时基于雷达检测数据序列的回归模型特征参数可信度较低,料位的预测更依赖工况参数;当工况较稳定时,雷达检测数据的随机误差较小,表现为数据的线性特征显著,此时回归模型特征参数可信度较高,工况参数的作用可适当降低。由于表征回归方程显著性的检验统计量 F 和 R^2 反映了雷达检测数据引入的随机误差大小,因此可将其作为衡量回归模型特征参数可信度的指标。由式(8)可知,相关系数 R 越接近1,则线性回归方程显著性越高,反之 R 越趋近0,则线性回归方程显著性越低,因此本文以相关系数 R 的线性函数作为权重调节系数,对训练输入样本的各组成部分进行权重调节。

对于式(5)所示的回归模型,第 i 周期的特征参数为 P_i ,记对应的相关系数为 r_i ,并作如下定义:特征参数矩阵 $P = (P_{i-3}, P_{i-2}, P_{i-1})_{18 \times 1}^T$,基准点矩阵 $J = (J_{i-3}, J_{i-2}, J_{i-1})_{6 \times 1}^T$,调整后的特征参数矩阵为 $P' = (P'_{i-3}, P'_{i-2}, P'_{i-1})_{18 \times 1}^T$,调整后的工况参数为 O'_i 。有权重调节公式如下:

$$\begin{cases} P'_{i-j} = r_{i-j} P_{i-j}, j = 1, 2, 3, \\ O'_i = w_o O_i, \\ w_o = 1 - \frac{1}{n} \sum_{j=1}^n r_{i-j}, n = 3, \end{cases} \quad (12)$$

其中: r_{i-j} 为特征参数 P_{i-j} 的权重调节系数, w_o 为工

况参数 O_i 的权重调节系数。调整后的料位模型特征参数 P' 、工况参数 O'_i 和基准点 J 构成料位预测的训练输入样本。训练样本权重调节如图6所示。

利用实时求解的相关系数 R 对输入变量进行动态加权,可降低因干扰和工况波动对料位预测带来的影响。

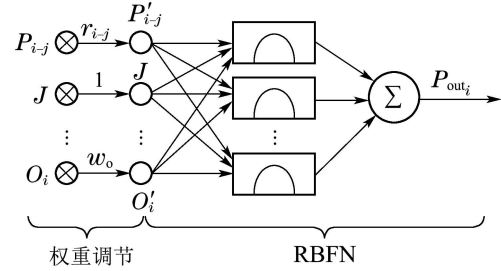


图6 训练样本权重调节示意图

Fig. 6 Schematic diagram of regulating training sample weight

3.3 动态回归加权RBFN料位预测 (Stockline prediction by dynamic regression weighting RBFN)

在复杂工况中,由于存在很多未知的影响因素,料位的变化存在随机性和不确定性;且由于高炉按批布料的工艺特点,料位的检测信息存在滞后。因此,利用历史料位检测数据和当前工况参数,采用动态加权料位预测方法对当前料位信息进行预测,能为高炉生产布料操作提供指导作用,减少布料盲目性。动态回归加权RBFN料位预测方法示意图如图7所示。

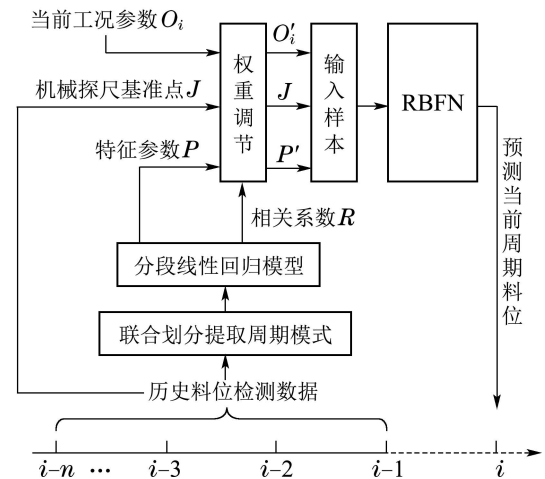


图7 动态回归加权RBFN料位预测方法示意图

Fig. 7 The stockline prediction method based on dynamic regression weighting RBFN

该方法首先对历史料位检测数据进行联合划分提取料位的周期模式,再通过分段线性回归建立料位变化模型,实现对雷达和机械探尺检测数据的融合,获得历史批次的料位模型特征参数 P 以及相应的显著性检验统计量(F 和 R^2),进而得到统计意义上精确且连

续的历史料位回归曲线和置信区间. 在得到历史批次的回归模型特征参数 P 后, 以历史机械探尺检测数据 J 为基准点, 加上当前的工况操作参数 Q_i , 并以回归建模得到的相关系数 R 构建权重调节函数, 构建训练输入样本, 以RBFN为基础, 对当前料位信息进行实时预测.

该预测方法的输出为当前周期的料位模型特征参数, 以此特征参数可以得到当前周期的料位回归曲线, 该回归曲线即为料位的预测回归曲线. 而通过对雷达和机械探尺检测数据分段线性回归建模获得的料位回归曲线则是料位的期望回归曲线. 将预测回归曲线与期望回归曲线进行对比, 计算每个检测时刻点的绝对误差和相对误差, 并统计相对误差的概率分布, 即可得出预测方法的精度评价指标, 精度越高则预测效果越好.

4 实例验证(Example verification)

以某炼铁厂2650 m³高炉为例来验证所提方法的合理性和有效性.

4.1 料位分段线性回归建模(Segmented linear-regression modeling)

选取2013年12月1日至2013年12月31号的数据进行分析. 首先, 基于本文提出的联合划分方法对机械探尺和雷达检测数据进行划分, 部分料位时间序列联合划分结果如图8所示, 可以看出料位具有周期性分段线性特征.

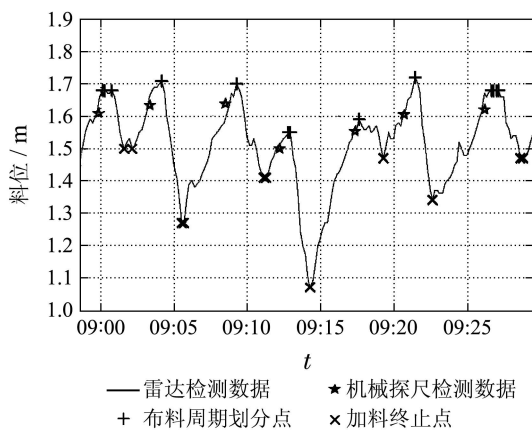


图8 料位时间序列联合划分结果

Fig. 8 Combined division result of stockline time-series

其次, 对雷达检测数据时间序列进行分段线性回归分析, 得到雷达检测数据时间序列的回归方程和置信区间(显著性水平 $\alpha = 0.05$). 通过残差分析找出异常点并排除, 以保证回归结果的精确度. 以一个等待阶段为例, 如图9所示的残差分析图, 残差的置信区间在零点以外的为异常点, 用三角形符号标记; 回归分析结果如图10所示, 置信区间涵盖了95%的雷达检测数据.

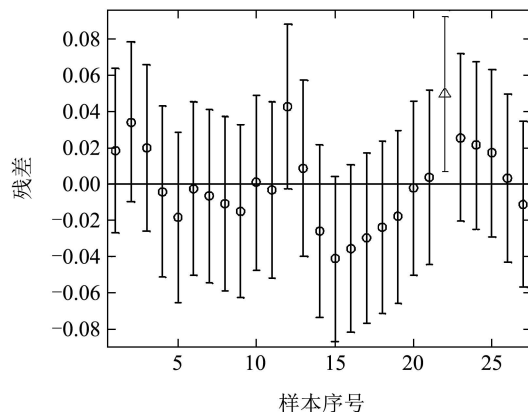


图9 雷达检测数据序列残差分析图

Fig. 9 Residual analysis of radar data

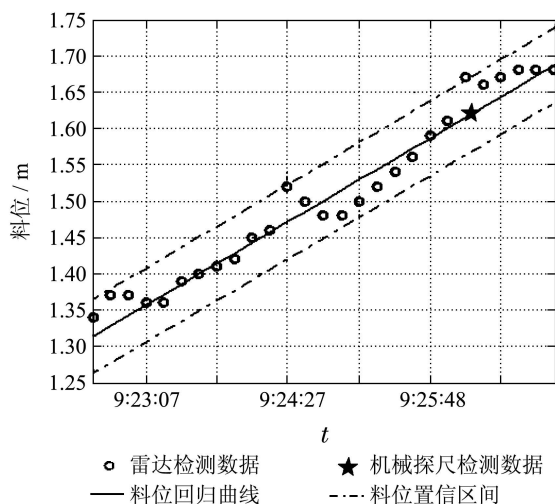


图10 雷达检测数据序列回归分析结果

Fig. 10 Regression analysis of radar data

然后, 结合机械探尺检测数据序列, 对雷达检测数据序列的回归模型进行修正, 并调整周期划分点使各周期保持连续. 同样地, 料位的置信区间也做合理的调整. 调整后的料位分段线性回归结果如图11-13所示.

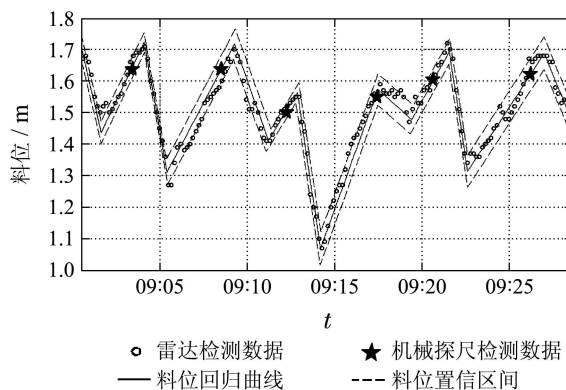


图11 工况稳定的料位分段线性回归建模结果

Fig. 11 Segmented linear-regression modeling result under stable working condition

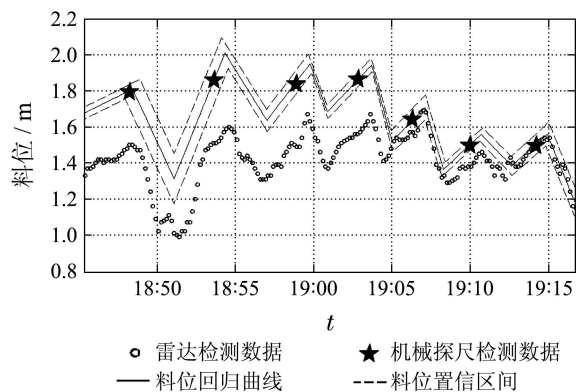


图12 工况波动的料位分段线性回归建模结果

Fig. 12 Segmented linear-regression modeling result under fluctuant working condition

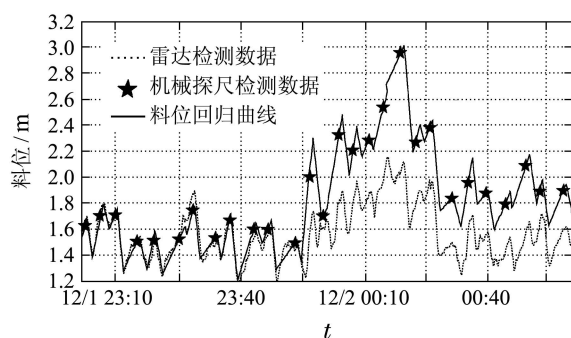


图13 工况异常的料位分段线性回归建模结果

Fig. 13 Segmented linear-regression modeling result under abnormal working condition

由图11可知, 工况稳定时(机械探尺在1.5~1.6 m附近), 雷达检测数据较为准确, 97%的数据都落入了模型给出的置信区间内, 可见所建的分段线性回归模型能够准确地描述料位的变化; 由图12和图13可知, 工况波动或者异常时(机械探尺检测数据大于1.8 m), 雷达检测数据偏差较大, 此时所建模型能以机械探尺检测数据为基准点, 对雷达检测数据进行修正, 实现了对料位信息的准确描述。

4.2 料位信息预测(Prediction of stockline information)

选取历史特征参数、机械探尺检测数据和当前工况参数构建训练样本1500组, 测试样本500组, 基于所提出的动态加权神经网络预测方法对料位信息进行预测, 预测结果如图14所示。

由图14所示的预测结果可得, 预测回归曲线与期望回归曲线变化趋势保持一致, 加料阶段(下降段)两条曲线基本吻合, 等待阶段(上升段)预测曲线稍有偏差。其原因是由于加料阶段的料位变化主要受加料速度影响, 而训练样本中选取的料流阀开度这一工况参数直接反映了加料速度, 因此保证了加料阶段的预测精度; 等待阶段的料位变化主要受料耗速度影响, 而影响料耗速度的因素除了训练样本中包含的顶温、

风量等工况参数外, 还存在其他复杂的耦合因素, 故等待阶段的预测精度比加料阶段略低。

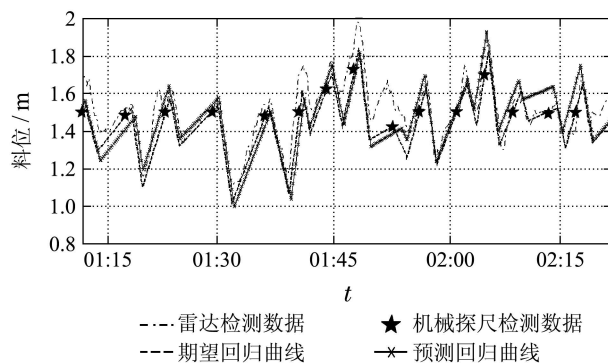


图14 料位信息预测结果

Fig. 14 Prediction results of stockline information

预测的绝对误差、相对误差和误差概率分布图分别如图15—17所示。

由图15至图17可得, 绝对误差的绝对值均小于0.16 m, 相对误差小于11%, 概率密度分布图中, 95%的预测值的相对误差低于9%, 表明本文的动态回归加权RBFN料位预测方法在工况复杂的情况下能够实现对高炉料位信息的有效预测, 减少了高炉布料操作的盲目性, 为高炉生产加料操作提供了实际指导作用。

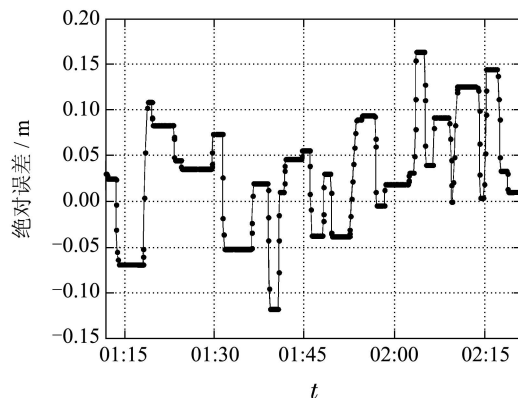


图15 预测的绝对误差

Fig. 15 The absolute errors of prediction results

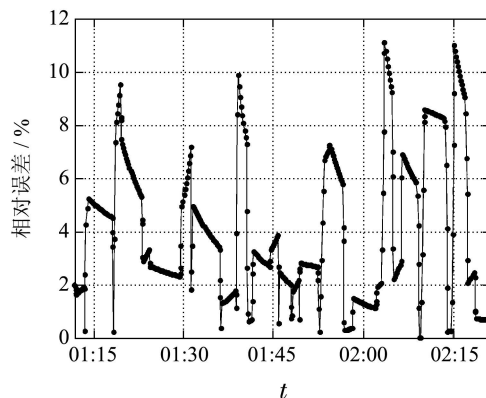


图16 预测的相对误差

Fig. 16 The relative errors of prediction results

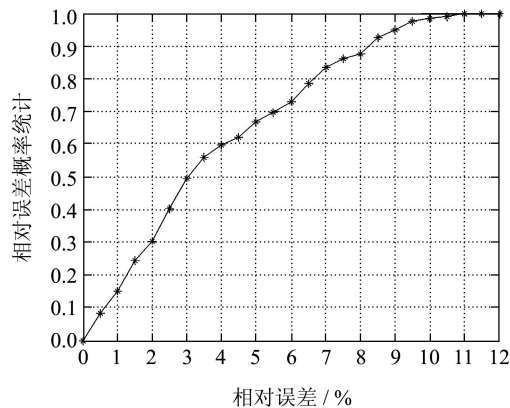


图 17 相对误差概率分布

Fig. 17 Probability distribution of relative error

经实例验证, 在高炉料位信息预测中, 基于数据的神经网络预测方法是可行的. 由于目前针对高炉料位信息预测的研究较少, 本文提出了分段线性回归和动态加权神经网络融合的高炉料位预测方法, 其他基于数据的预测方法还有支持向量机、极限学习机等, 其是否也可用于高炉料位信息预测, 还有待进一步研究.

5 结论(Conclusions)

本文针对高炉料位难以连续且高精度检测的问题, 提出了一种基于分段线性回归和动态加权神经网络的高炉料位信息实时预测方法. 首先, 通过分析高炉布料制度和料位检测数据特点, 在对雷达和机械探尺检测数据时间序列的联合划分基础上, 构建分段线性回归模型, 获得能准确描述料位变化的回归曲线; 然后, 基于分段线性回归模型, 以回归统计指标为权重调节系数, 加入工况参数, 利用动态回归加权RBF神经网络, 实现了料位信息的实时预测. 理论分析和实例验证均表明, 所提方法充分融合了机械探尺检测数据精度高和雷达检测数据连续性好的优点, 可以得到准确反映料位变化的回归曲线, 并实现了料位信息的实时预测.

参考文献(References):

- [1] 何良华. PLC在高炉探尺控制上的优化应用 [J]. 自动化与仪器仪表, 2009, 29(6): 74 – 75.
(HE Lianghua. PLC optimized application on blast furnace stock rod control [J]. *Automation and Instrumentation*, 2009, 29(6): 74 – 75.)
- [2] 张贺顺, 马洪斌, 欧阳普照. 雷达料位仪在首钢2号高炉的应用 [J]. 炼铁, 2009, 28(1): 47 – 48.
(ZHANG Heshun, MA Hongbin, OUYANG Puzhao. The application of radar in blast furnace #2 of Shougang Group [J]. *Ironmaking*, 2009, 28(1): 47 – 48.)
- [3] 刘德馨, 李晓理, 丁大伟, 等. 基于雷达观测数据的高炉料面多模型控制 [J]. 控制理论与应用, 2012, 29(10): 1277 – 1283.
(LIU Dexin, LI Xiaoli, DING Dawei, et al. Multi-model control of blast furnace burden surface based on observed data of radars [J]. *Control Theory & Applications*, 2012, 29(10): 1277 – 1283.)
- [4] ZHU Q, LÜ C, YIN Y, et al. Burden distribution calculation of bell-less top of blast furnace based on multi-radar data [J]. *Journal of Iron and Steel Research(International)*, 2013, 20(6): 33 – 37.
- [5] 杜鹏宇, 程树森, 滕召杰. 激光技术在无钟炉顶布料和料面监测中的应用 [J]. 中国冶金, 2011, 21(5): 1 – 6, 12.
(DU Pengyu, CHENG Shusen, TENG Zhaojie. Application of laser technology in the burden charging and burden profile monitoring for bell-less top [J]. *China Metallurgy*, 2011, 21(5): 1 – 6, 12.)
- [6] TENG Z, CHENG S, DENG P, et al. Mathematical model of burden distribution for the bell-less top of a blast furnace [J]. *International Journal of Minerals Metallurgy and Materials*, 2013, 20(7): 33 – 37.
- [7] 刘云彩. 高炉布料规律 [M]. 北京: 冶金工业出版社, 2005.
(LIU Yuncui. *The Regulation on Burden Distribution* [M]. Beijing: Metallurgical Industry Press, 2005.)
- [8] HELLE H, HELLE M, SAXÉN H. Nonlinear optimization of steel production using traditional and novel blast furnace operation strategies [J]. *Chemical Engineering Science*, 2011, 66(24): 6470 – 6481.
- [9] XU J, WU S, KOU M, et al. Circumferential burden distribution behaviors at bell-less top blast furnace with parallel type hoppers [J]. *Applied Mathematical Modelling*, 2011, 35(3): 1439 – 1455.
- [10] 黄伟平, 徐毓, 王杰. 综合运用模式识别与回归分析的数据融合算法 [J]. 控制与决策, 2010, 25(1): 110 – 114, 120.
(HUANG Weiping, XU Yu, WANG Jie. Algorithm of data fusion based on mode recognition and regression analysis [J]. *Control and Decision*, 2010, 25(1): 110 – 114, 120.)
- [11] PONTES F, PAIVA A, BALESTRASSI P, et al. Optimization of radial basis function neural network employed for prediction of surface roughness in hard turning process using Taguchi's orthogonal arrays [J]. *Expert Systems with Applications*, 2012, 39(9): 7776 – 7787.
- [12] GAN M, PENG H, DONG X. A hybrid algorithm to optimize RBF network architecture and parameters for nonlinear time series prediction [J]. *Applied Mathematical Modelling*, 2012, 36(7): 2911 – 2919.
- [13] DU H, ZHANG N. Time series prediction using evolving radial basis function networks with encoding scheme [J]. *Neurocomputing*, 2008, 71(7/8/9): 1388 – 1400.
- [14] INOUSA G. *Nonlinear time series modeling and prediction using modified neural networks-based autoregressive models* [D]. Changsha: Central South University, 2012.

作者简介:

蒋朝辉 (1978–), 男, 博士, 副教授, 主要研究方向为复杂工业过程建模与优化控制、广义大系统控制理论与应用, E-mail: jzh0903@csu.edu.cn;

李晞月 (1989–), 女, 硕士研究生, 主要研究方向为工业过程建模与控制、信息检测系统研究, E-mail: lixychn@163.com;

桂卫华 (1950–), 男, 中国工程院院士, 教授, 主要研究方向为复杂工业过程建模与优化控制、工业大系统控制理论与应用, E-mail: gwh@csu.edu.cn;

谢永芳 (1972–), 男, 博士, 教授, 研究方向为复杂工业过程建模与控制、分散鲁棒控制, E-mail: yfxie@csu.edu.cn;

阳春华 (1965–), 女, 博士, 教授, 主要研究方向为复杂工业过程建模与优化控制、智能自动化控制系统, E-mail: ychh@csu.edu.cn.