

控制与决策

Control and Decision

基于Stacking的高炉铁水质量区间预测

周平, 刘进进

引用本文:

周平, 刘进进. 基于Stacking的高炉铁水质量区间预测[J]. *控制与决策*, 2021, 36(2): 335–344.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0786>

您可能感兴趣的其他文章

Articles you may be interested in

基于偏态深度分类的高炉硅含量及波动预测

Prediction of blast furnace silicon content and fluctuation based on skewness depth classification

控制与决策. 2021, 36(2): 491–497 <https://doi.org/10.13195/j.kzyjc.2019.1116>

基于多维泰勒网的超前d步预测模型

d-step-ahead predictive model based on multi-dimensional Taylor network

控制与决策. 2021, 36(2): 345–354 <https://doi.org/10.13195/j.kzyjc.2019.0722>

基于高炉炉况评价和反馈补偿的喷煤量决策优化

Decision-making optimization of coal injection volume based on evaluation of blast furnace condition

控制与决策. 2020, 35(11): 2803–2809 <https://doi.org/10.13195/j.kzyjc.2018.1553>

考虑时间序列的动态大群体应急决策方法

Dynamic large group emergency decision-making method considering time series

控制与决策. 2020, 35(11): 2609–2618 <https://doi.org/10.13195/j.kzyjc.2019.0088>

基于不变网络模型和故障注入的分布式信息系统故障溯源方法

Fault source location algorithm for distributed information system based on invariant network and fault injection

控制与决策. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

基于 Stacking 的高炉铁水质量区间预测

周平, 刘进进[†]

(东北大学 流程工业综合自动化国家重点实验室, 沈阳 110004)

摘要: 高炉炼铁多元铁水质量的实时准确预报是高炉内部状态进行实时监测和有效控制的重要手段,但预报结果存在准确度不高和缺乏可信度表征的问题,特别是在炉况不稳定和高炉数据波动大的情况下,多元铁水质量的预报结果存在较大偏差和较低可信度,不能为高炉日常操作和调节提供指导. 针对这一工程难题,提出一种基于 Stacking 的高炉铁水质量区间预测建模算法,用于建立铁水质量区间预报模型. 首先,为了提高建模的准确度,建立基于 Stacking 的铁水质量模型;然后,为了表征预测结果的可信度,引入区间预测的方法,用于多元铁水质量值和预测区间的同时预报;最后,通过工业实验表明,所提出算法能够实现对高炉多元铁水质量的准确预报,并且预测区间拥有较高的可信度,对实际生产操作具有较好的指导意义.

关键词: 高炉炼铁; Stacking 模型; 区间预测; 置信度

中图分类号: TP18

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0786

开放科学(资源服务)标识码(OSID):



引用格式: 周平,刘进进. 基于 Stacking 的高炉铁水质量区间预测[J]. 控制与决策, 2021, 36(2): 335-344.

Prediction of blast furnace hot metal quality interval based on stacking

ZHOU Ping, LIU Jin-jin[†]

(State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110004, China)

Abstract: Real-time accurate prediction of multi-ferrite quality in blast furnace ironmaking is an important means for real-time monitoring and effective control of blast furnace internal state, but the prediction results are not accurate and lack credibility, especially in unstable furnace conditions and in the case of large fluctuations in blast furnace data, the prediction results of multi-component molten iron quality have large deviations and low credibility, and cannot provide guidance for daily operation and adjustment of blast furnace. Aiming at this engineering problem, this paper proposes a Stacking-based blast furnace hot metal quality interval prediction modeling algorithm for establishing the molten iron quality interval prediction model. Firstly, in order to improve the accuracy of modeling, a Stacking-based hot metal quality model is established. Then, in order to characterize the credibility of the prediction results, an interval prediction method is introduced to simultaneously predict the multi-ferrite water quality value and the prediction interval. Finally, industrial experiments show that the proposed algorithm can accurately predict the quality of blast furnace multi-metal hot metal, and the prediction interval has high credibility, which has a good guiding significance for actual production operations.

Keywords: blast furnace ironmaking; Stacking model; interval prediction; confidence

0 引言

钢铁是国民经济的重要支柱,高炉炼铁则为钢铁生产的重要方式. 典型的高炉炼铁系统如图 1 所示,分为高炉本体、送料系统、煤粉喷吹系统、热风系统、出铁系统以及高炉煤气处理系统等子系统,其中高炉本体从上至下又由炉喉、炉身、炉腰、炉腹、炉缸等部分组成^[1-3]. 在进行高炉炼铁时,上料系统的焦

炭、铁矿石和熔剂从炉顶装入,受重力的作用向下运动;同时下部的热风系统和煤粉喷吹系统组合作业,由风口鼓入热风 and 燃料,生成大量还原性气体向上运动,最终生成高炉煤气回收再利用;炉料则经过加热、还原、融化、造渣、渗碳、脱硫等一系列物理化学过程,从上至下经历块状带、软熔带、滴落带、燃烧带、渣铁盛聚带,最后生成液态炉渣和生铁由铁口排出^[4-5].

收稿日期: 2019-06-04; 修回日期: 2019-09-24.

基金项目: 国家自然科学基金项目 (61890934, 61473064, 61790572); 中央高校基本科研业务费专项基金项目 (N180802003).

责任编辑: 柴利.

[†]通讯作者. E-mail: 1780608555@qq.com.

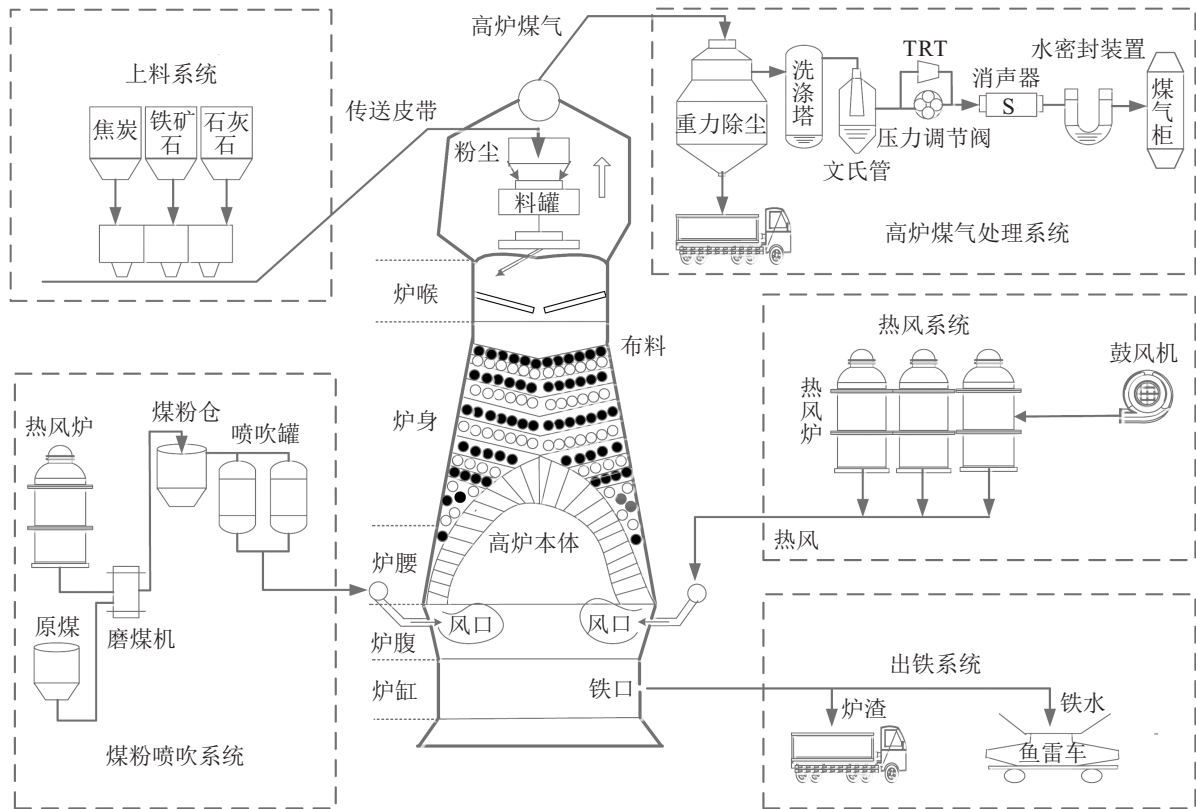


图1 高炉炼铁过程

为了实现高炉生产的高效益、低能耗和长寿稳定运行的目标,需要对高炉内部状态进行实时监测和有效控制^[1-2]。目前,综合性的铁水质量指标能够间接反映高炉内部状态,因而被广泛应用于高炉内部状态的监测和控制中。铁水质量指标通常采用铁水温度 (molten iron temperature, MIT)、硅含量 ([Si])、磷含量 ([P]) 和硫含量 ([S]) 衡量。MIT 代表高炉炼铁过程的物理热,能够反映炉内能量消耗和热状态;铁水 [Si] 反映铁水的化学热,是铁水质量参数中最为重要的指标;铁水 [P] 和铁水 [S] 是危害铁水质量的主要因素,过高会影响生铁品质。但是,高炉内部环境恶劣、外部环境存在较大干扰,且受到现有检测技术的限制,直接对高炉铁水质量进行准确的在线检测存在困难,且耗时耗力,通常的离线化验需要 1~2 h,存在较长的滞后。因此,建立准确可靠的铁水质量模型反映当前及预测的内部温度和指标参数变化显得尤为重要。

目前,常见的铁水质量模型有机理模型^[6-7]、推理模型^[8-9]以及数据驱动的智能模型^[10-11]。近些年来,随着计算机的快速发展,数据驱动的建模方法已被广泛应用于解决铁水质量的建模问题。基于数据驱动的铁水质量建模方法并不需要深入了解高炉内部复杂的机理变化,仅通过数学工具和智能算法对过程数据进行处理,便可建立所需的铁水质量模型。现有数据驱动建模方法主要有多元统计分析方法、神

经网络 (NN) 建模方法和支持向量回归 (SVR) 建模方法。文献 [12-13] 采用多元统计分析方法对铁水质量进行建模,文献 [14-15] 运用 NN 方法对铁水质量进行预测,文献 [16-17] 使用 SVR 方法对铁水质量加以分析,取得了较好的建模效果。但是,这些模型只是单一地对铁水质量进行预测^[18-19],而在高炉炼铁过程中,由于工况经常波动,致使这些高炉铁水质量预测模型结果的不确定特征明显,难以表征真实的高炉内部状态,带来实时监测和控制方向性的错误。由此可见,只是单一地针对铁水质量的值进行预测,在面对炉况不稳定的情况存在较大不确定性,难以解决高炉炉温精确调控的难题^[20]。

针对上述问题,本文从高炉炼铁实际应用的角度出发,为了提高多元铁水质量建模的准确性和进一步表征预测结果的可信度,结合 Stacking 模型和区间预测方法,提出一种基于 Stacking 的高炉铁水质量区间预测建模算法。所提出算法的基本思想是: 1) 为了提高建模的准确度,采用 Stacking 模型策略进行多元铁水质量的建模;同时为了提高 Stacking 模型的建模速度,统一采用具有极快建模速度的随机神经网络 (random vector functional-link networks, RVFLNs) 为子模型。2) 为了解决单一多元铁水质量预测模型不能表征预测结果的可信度问题,在 Stacking 算法建模的基础上引入区间预测方法,同时预报高炉多元铁水质

量值及一定置信度的预测区间. 随后, 进行的工业实验表明, 所提出基于Stacking的高炉铁水质量区间预测建模算法, 不仅提高了多元铁水质量预测模型的准确性, 而且能够克服单一多元铁水质量模型存在预测结果可靠性低的问题, 对可信度低的预测值给出警告, 为高炉内部状态进行实时监测和有效控制提供更有效的信息, 使得现场操作人员可以有选择地参考预测结果, 对提高现场操作人员的操作准确率具有重要意义.

1 建模策略

高炉炼铁系统控制和优化的主要对象是铁水质量参数, 为此, 建立服务于高炉炼铁系统控制与优化的铁水质量模型是不可或缺的重要环节^[21-22]. 针对前述铁水质量建模存在的诸多问题, 本文从高炉炼铁工程应用的角度出发, 提出如图2所示的基于Stacking的高炉铁水质量区间预测建模策略:

1) 针对高炉炼铁系统, 为了提高建模数据质量, 对高炉原始数据进行数据预处理操作. 首先, 由于各个过程变量是由不同采样频率的传感器测量, 在建模前必须进行时间粒度的统一, 将高炉铁水的出铁时间与本体参数的时间点进行对应; 然后, 进行正常数据的筛选, 包括休风数据的剔除和采集异常值的剔除; 最后, 考虑到不同建模数据的量纲不一致问题, 对高炉建模数据进行归一化处理.

2) 在高炉炼铁系统的数据建模中, 模型输入选择大多根据经验主观确定, 容易造成特征冗余或关键特征的丢失, 影响模型效率和建模精度. 在现有数据建模研究中, 辅助变量的选取方法主要为主成分分析(PCA)方法. 但是PCA技术只针对输入空间进行降维, 未考虑输入输出变量之间的关联性. 虽然典型相关性分析(CCA)方法可以计算输入变量和输出变量的相关性^[23], 但CCA分析的是输入输出变量间的线性相关性, 这对于具有强非线性的高炉数据而言, 并不是一个好的处理方法. 为此, 引入灰色关联分析法, 提取与多元铁水质量指标关联度最强的几个关键变量作为建模的输入变量.

3) 我国高炉炼铁的入炉矿源及其品位波动频繁, 导致其冶炼过程中工况经常波动, 致使现有的基于数据的高炉铁水硅含量预测模型的结果不确定性特征明显, 难以准确地表征真实炉况, 带来炉温调控方向错误. 针对单一高炉多元铁水质量预测不能表征预测结果的可信度的问题, 结合Stacking模型和区间预测方法, 提出一种基于Stacking的高炉铁水质量区间预测建模算法, 实现多元铁水质量预测值和预测区间

的同时预报.

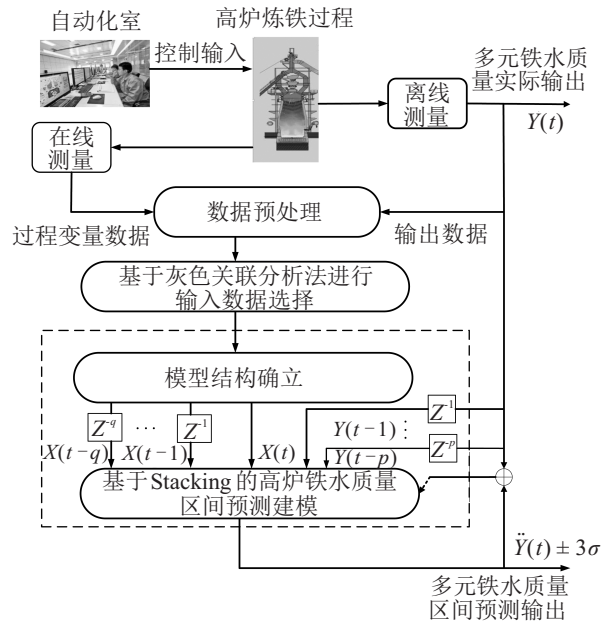


图2 基于Stacking的高炉铁水质量区间预测建模策略

2 数据的预处理及输入变量的选取

为了提高建模的数据质量, 对高炉原始数据进行数据的预处理操作, 具体过程如下.

step 1: 时间粒度的统一.

由于采集各个过程变量的传感器对应不同的采样频率, 采集的高炉数据必然存在时间粒度不统一的问题. 为此, 基于数据时间标注及最近邻时间原则, 采用人工匹配操作, 使得高炉本体参数的时间点与高炉铁水的出铁时间一一对应.

step 2: 正常数据的筛选.

正常数据的筛选包括休风数据的剔除和采集异常值的剔除. 由于每个月对于炼铁设备进行检修而进行休风操作, 此段时间没有出铁信息, 属于可预见性的异常工况, 该段时间的数据无法作为建模数据, 应该筛选剔除. 具体作法为: 依据交班记录确定高炉休风时间段并剔除此段异常工况数据.

由于高炉炼铁过程高炉内部环境恶劣, 外部环境存在干扰, 采集的数据存在异常值. 为了保证建模数据的合理性和有效性, 本文选用计算简单、可靠性高的 3σ 准则(拉依达准则)方法剔除异常值, 即如果数据偏差大于 3σ , 则对该数据剔除. 其中 σ 为所述高炉本体数据的标准差, 求取如下:

$$\sigma = \sqrt{\sum_{i=1}^n e_i^2 / (n - 1)} = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1)}. \tag{1}$$

step 3: 数据归一化.

由于高炉本体数据存在量纲的差异, 若不消除,

则会减缓算法的收敛速度并在一定程度上增大模型误差. 为此, 从高炉实际应用角度出发, 选用最小、最大归一化进行高炉数据的归一化操作^[1], 公式转化如下:

$$\tilde{x}_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} - 1. \quad (2)$$

其中: x_i 、 \tilde{x}_i 分别为第 i 个变量归一化前、后的取值, $\max(x_i)$ 、 $\min(x_i)$ 分别为第 i 个变量的最大值、最小值. 归一化处理得到的数据 $x_i \in (0, 1)$.

为了提高高炉多元铁水质量的建模数据, 本文引入灰色关联分析法, 从众多过程变量中选择最有效的建模输入数据. 在灰色关联分析法中, 反映系统行为特征的数据序列称为参考数列, 影响系统行为因素组成的数据序列称为比较数列. 该方法通过求取参考数列和比较数列之间的关联度, 通过量化分析得出参考数列与比较数列的几何关系^[24-25]. 关联度大, 表明比较数列与参考数列的变化趋势相似程度高, 反之则低. 具体实现步骤如下.

step 1: 确定分析数列.

设参考数列为

$$Y = \{Y(k) | k = 1, 2, \dots, n\}. \quad (3)$$

比较数列为

$$X_i(k) = \{X_i(k) | k = 1, 2, \dots, n, i = 1, 2, \dots, m\}. \quad (4)$$

对于高炉炼铁多元铁水质量建模系统, Si 含量、P 含量、S 含量、铁水温度分别作为系统的参考数列, 比较数列为 m 个影响多元铁水质量指标的关键过程变量.

step 2: 变量的无量纲化.

由于高炉数据之间存在量纲上的差异, 若不统一量纲则会对灰色关联度分析结果造成影响, 为此, 选用归一化处理进行数据的无量纲化操作.

step 3: 计算关联系数.

关联系数是比较数列和参考数列在各个时刻的关联程度, 即高炉多元铁水质量与 m 个过程变量的各对应点的关联程度, 其数学表达式为

$$\xi_i(k) = \frac{\min_i \min_k |y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho \max_i \max_k |y(k) - x_i(k)|} + \frac{\rho \max_i \max_k |y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho \max_i \max_k |y(k) - x_i(k)|}. \quad (5)$$

其中: $|y(k) - x_i(k)|$ 为绝对值之差, $\min_i \min_k |y(k) - x_i(k)|$ 、 $\max_i \max_k |y(k) - x_i(k)|$ 分别为两级最小差和两级最大差; $\rho \in (0, \infty)$ 为分辨系数. 当 $\rho \leq 0.5463$

时, 分辨力最好, 本文取 $\rho = 0.5$.

step 4: 计算关联度.

将各时刻的关联系数求平均, 用于衡量比较数列与参考数列的相似程度, 即关联度, 有

$$r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k). \quad (6)$$

step 5: 关联度排序.

求取关联度后按大小排序, 若 $r_1 > r_2$, 则表明参考数列 Y 与比较数列 X_1 几何关系更相似.

3 基于 Stacking 的区间预测建模算法

3.1 基于 Stacking 的方法

Stacking 是 Wolpert^[26] 于 1992 年提出的一种通过训练将基本学习器组合起来集成学习算法. Stacking 建模的基本思想是通过将基学习器进行相关集成处理, 从而提高算法建模的准确度. 其中基本学习器称为第 1 层学习器, 组合学习机称为第 2 层学习器 (second-level learner) 或者元学习器 (meta-learner). Stacking 先用原始数据训练一阶学习器, 用第 1 层学习器的输出作为第 2 层学习器的输入特征, 原始数据的标签作为新特征的标签, 然后用新的数据集训练第 2 层学习器. 第 1 层学习器中可以包括不同的学习算法, Stacking 集成方法 (第 2 层学习器) 通常也包括不同的学习器, 当然其基本学习器也可以是同类型的. 表 1 为一般化的 Stacking 算法伪代码.

可以从两方面看待 Stacking 方法: 1) 它是许多集成方法的推广; 2) 它是通过学习得到的集成方法. 在 Stacking 的训练阶段, 从第 1 层学习器中得到新的数据集. 如果用同一份完全相同的数据训练第 1 层学习器, 并用该份数据在第 1 层学习器上的输出作为第 2 层学习器的训练数据, 则有过拟合风险. 因此, 训练第 1 层学习器的数据不能作为构造第 2 层学习器的数据, 所以在构造第 2 层学习器的训练数据时需要用交叉验证或留一法选取第 2 层学习器的训练数据.

Stacking 算法的伪代码如下.

输入: 数据集 $D = (x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$.

一阶学习算法: $\text{model}_1, \text{model}_2, \dots, \text{model}_T$;

二阶学习算法: model .

训练过程:

1 for $t = 1, 2, \dots, T$

2 $h_t = \text{model}_t(D)$

3 end

4 $D' = \Phi$

5 for $i = 1, 2, \dots, m$

6 for $t = 1, 2, \dots, T$

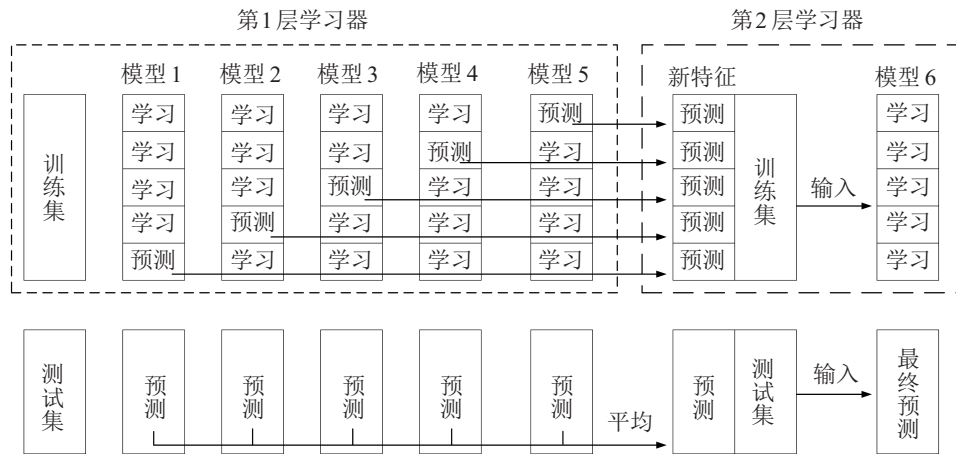


图3 基于五折模型的Stacking算法流程图

7 $z_{it} = h_t(x_i)$

8 end

9 $D' = D' \cup ((z_{i1}, z_{i2}, \dots, z_{iT}), y_i)$

10 end

11 $h' = \text{model}(D')$

输出: $H(x) = h'(h_1(x), h_2(x), \dots, h_T(x))$.

本文选取基于五折模型的Stacking算法建立多元铁水质量模型,其流程如图3所示. 具体步骤如下.

step 1: 子模型的选择.

基于Stacking的集成模型算法能够提高建模精度,但是集成模型具有融合多个模型进行建模的特性,势必在整体建模上会牺牲一定的建模速度,因此选择具有快速建模速度的随机神经网络为子模型进行建模^[27-28].

step 2: 第1层学习器训练数据的划分.

为了使训练数据存在差异性,防止过拟合现象,对训练集进行k-fold切分,分为k份,每一份包括一个验证集和测试集,每次取其中k-1份训练,另外1份用来验证,其中k选为5,即五折模型.

step 3: 第1层学习器的训练.

如图3所示,对训练集进行五折模型的划分. 其中:模型1学习时,使用前4份数据用于模型的学习,第5份作为模型的测试数据;与此类似,模型2、模型3、模型4、模型5分别使用第4份、第3份、第2份、第1份数据做预测数据,其他4份作为学习数据. 这样,既保证了第1层学习器学习数据的差异性,又使得这5个模型的预测值组合正好覆盖训练集所有要预测的值.

step 4: 第2层学习器的训练.

将第1层学习器得到的预测值作为第2层学习器新的输入数据,将原始训练集的输出数据作为第2层学习器的标签数据,从而进行第2层学习器的学习

过程.

step 5: 第1层学习器的预测.

将测试数据输入学习好的模型1~模型5中得到5份预测值,并将这5份预测值进行平均化处理.

step 6: 第2层学习器的预测.

将step 5得到的平均后的预测数据作为第2层学习器的输入数据,对学习好的模型6进行预测得到最终的预测值.

3.2 区间预测方法

传统铁水质量建模方法大多只是单一地针对铁水质量值进行预测,并没有表征预测结果的可信度问题,为此本文引入区间预测方法^[29-30]. 所提出方法的主要思想如图4所示.

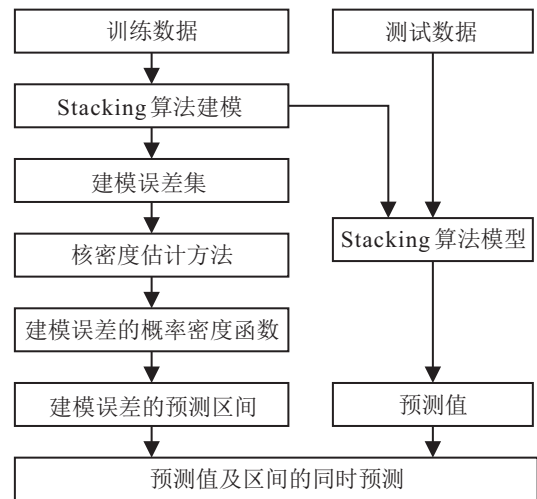


图4 基于Stacking的区间预测建模算法

首先,对多元铁水质量进行建模,使用核密度估计方法估计出建模误差的概率密度函数(probability distribution function, PDF),将具有不确定性的建模误差用确定性的概率密度函数表征;然后,由估计出的概率密度函数曲线求出符合设定概率的建模误差区间,该区间与预测值一起构造随后的预测区间;最

后,同时对多元铁水质量值和建模误差区间进行预测.这里重点讲解核密度函数方法及建模误差预测区间的求解过程.

1) 核密度估计方法.

核密度估计(kernel density estimation, KDE)是求解给定随机变量集合分布密度问题的非参数估计方法,由Parzen^[31]首次提出.假设 $x_i \in R, i = 1, 2, \dots, n$ 为独立同分布随机变量,服从分布密度函数 $f(x), x \in R$,则 $f(x)$ 的核密度估计 $\hat{f}_h(x)$ 定义为

$$\hat{f}_h(x) = \frac{1}{nh_p} \sum_{i=1}^n \phi\left(\frac{x_i - x}{h_p}\right), x \in R. \quad (7)$$

其中: $\phi\left(\frac{x_i - x}{h_p}\right)$ 为核函数; h_p 通常称为窗口宽度或光滑参数,是预先给定的正数.

由上述定义可知,分布密度函数 $f(x)$ 的核密度估计 $\hat{f}_h(x)$ 不仅与给定的数据样本集有关,还与核函数和窗口宽度参数 h_p 的选择有关.利用KDE法对给定建模误差样本集 $\{e_k | k = 1, 2, \dots, K\}$ 进行PDF估计,得到估计误差概率密度函数 Γ_e 为

$$\Gamma_e = \frac{1}{K} \sum_{k=1}^K \frac{1}{h_p} \phi\left(-\frac{e - e_k}{h_p}\right), \quad (8)$$

其中 e 为误差概率密度函数的自变量.

所述 $\phi\left(-\frac{e - e_k}{h_p}\right)$ 核函数的选择:从实用角度出发,因为待估函数为概率密度函数,要求核函数符合概率密度函数性质,即满足以下条件:

$$\begin{aligned} \phi(e) &\geq 0; \\ \int \phi(e) du &= 1. \end{aligned} \quad (9)$$

本文选取高斯核函数,其表达式为

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\|x\|^2}{2}}. \quad (10)$$

所述窗宽的选择:窗宽 h_p 的选择对核函数的密度估计起着局部光滑的作用,若 h_p 过大,则会使模型误差概率密度函数形状很光滑,使其主要部分的某些特征(如多峰性)被掩盖起来,从而增加估计量的偏差;若 h_p 过小,则整个密度函数表现粗糙,尤其是密度估计的尾部会出现较大干扰.于是,窗宽 h_p 设置为 $h_p = 1.06\theta K^{-1/5}$.其中: θ 由 $\min\{S, 0.746Q\}$ 估计, S 为样本标准差, Q 为四分位数间距, K 为建模误差样本个数.

2) 建模误差的预测区间的求解.

建模误差是具有随机特性的,如何由不确定性的建模误差建立表征预测值可信度的预测区间是本节讨论的重点.对于该问题,从建模误差概率密度函数入手,使用核密度估计方法估计出建模误差的概率密

度函数曲线,在理想条件下,建模误差的概率函数分布曲线符合高斯分布,之后根据高斯分布特性求出建模误差的预测区间.建模误差的预测区间具体求取步骤如下.

step 1: 估计建模误差的概率密度曲线.

使用核密度估计方法估计出建模误差集的概率密度曲线.

step 2: 求取建模误差的标准差.

通过统计学的方法求取建模误差集的标准差 σ_1 ,有

$$\sigma_1 = \sqrt{\sum_{j=1}^n e_j^2 / (n - 1)}. \quad (11)$$

其中: e_j 为第 j 个建模误差, n 为建模误差的个数.

step 3: 定义建模误差的预测区间.

由于建模误差的概率密度函数近似符合高斯分布,当建模误差的预测区间选为 $\pm 3\sigma_1$ 时,建模误差以0.9974的概率落在该区间.而0.9974的概率几乎接近于1,有很高的置信度,所以本文定义建模误差正负3倍的标准差作为模型建模误差的预测区间.

4 工业实验

4.1 数据的预处理及输入变量的选取

采用我国柳钢2号高炉的自体数据与铁水质量数据(采样间隔为1h)对所提出方法进行数据测试.首先,根据高炉炼铁工艺及相关仪器仪表设置,确定影响Si含量、P含量、S含量、铁水温度等铁水质量指标的19个关键过程变量,分别为:冷风流量、送风比、热风压力、顶压、压差、顶压风量比、透气性、阻力系数、热风温度、富氧流量、富氧率、设定喷煤量、鼓风湿度、理论燃烧温度、标准风速、实际风速、鼓风动能、炉腹煤气量、炉腹煤气指数^[1,4];然后,调用自动化室数据库中当前及历史时刻的高炉生产实际数据;最后,为了提高建模数据的质量,对高炉数据进行数据的预处理.具体作法包括对高炉自体原始数据进行时间粒度的统一、正常数据的筛选和数据的归一化等预处理操作.

考虑到影响柳钢2号高炉多元铁水质量的过程变量有19个之多,若将它们全部作为输入变量引入模型,则会存在变量维数过高和可能引入较多干扰的问题,相应会导致训练时间长和影响模型的精度.为此,本文采用灰色关联分析法进行输入变量的选取.首先,对19个关键过程变量分别与作为输出变量的[Si] y_1 (%)、[P] y_2 (%)、[S] y_3 (%)、铁水温度MIT y_4 (°C)进行灰色关联度分析,并计算出各自的关联度;然后,将这些关联度进行求和得到如表1所示

的结果;最后,确定与铁水质量关联度最大的关键过程变量阻力系数 x_1 、顶压风量比 x_2 、透气性 x_3 、压差 x_4 、鼓风动能 x_5 、热风压力 x_6 、实际风速 x_7 为影响多元铁水质量指标的主要因素,并将这7个过程变量作为铁水质量模型的输入变量。

表1 灰色关联度分析结果

	[Si]/%	[S]/%	[P]/%	[MIT]/°C	总关联度
冷风流量	0.6299	0.5988	0.6093	0.7956	2.6336
送风比	0.6252	0.5947	0.6057	0.7931	2.6187
热风压力	0.6944	0.6562	0.6583	0.8251	2.8340
顶压	0.4400	0.4209	0.4304	0.5731	1.8644
压差	0.7100	0.6734	0.6711	0.8129	2.8674
顶压风量比	0.7352	0.7396	0.7310	0.7186	2.9244
透气性	0.7299	0.7272	0.7400	0.7270	2.9241
阻力系数	0.7815	0.7600	0.7295	0.7291	3.0001
热风温度	0.5305	0.5008	0.5121	0.7212	2.2646
富氧流量	0.4516	0.4323	0.4423	0.5910	1.9172
富氧率	0.5033	0.4788	0.4905	0.6839	2.1565
设定喷煤量	0.5461	0.5150	0.5244	0.7374	2.3229
鼓风湿度	0.6261	0.6554	0.6530	0.4635	2.3980
燃烧温度	0.4898	0.4681	0.4799	0.6577	2.0955
标准风速	0.6312	0.5999	0.6105	0.7970	2.6386
实际风速	0.6865	0.6613	0.6685	0.8078	2.8241
鼓风动能	0.7022	0.6742	0.6767	0.8116	2.8647
炉腹煤气量	0.5916	0.5645	0.5789	0.7833	2.5183
煤气指数	0.5991	0.5714	0.5855	0.7981	2.5541

为了更好地反映高炉非线性动态特性和输入输出变量的时序和时滞关系,引入非线性自回归模型(NARX),取时序系数 $p = 1, q = 1$.将前一采样时刻的输入测量值 $X(t - 1)$ 以及前一采样时刻铁水质量指标值 $Y(t - 1)$ 连同当前采样时刻的输入测量值 $X(t)$ 作为动态模型的综合输入,即所建立的动态软测量模型用于实现的非线性动态映射关系为 $Y(t) = f_{\text{NARX}}(Y(t - 1); X(t), X(t - 1))$.

4.2 建模及区间预测效果

将预处理得到的520组数据分为二个样本集 D_1 和 D_2 .其中: D_1 为训练样本集,用来训练并建立铁水质量模型,取前400组数据; D_2 为测试样本集,用来测试铁水质量模型,取后120组数据.对于所提出基于Stacking的区间预测多元铁水质量建模算法,将从以下几点进行说明:1)基于Stacking算法多元铁水质量模型的建立;2)预测区间的求取;3)区间预测效果.

1) 基于Stacking算法多元铁水质量模型的建立.

按照第3.1节中基于五折模型的Stacking算法的step 1~step 6建立多元铁水质量模型,其中训练集数据为预处理的前400组数据.为了提高整体集成模型的建模速度,模型1~模型6统一选用具有快速建模能力的RVFLNs作为Stacking算法的子模型.图5为所提出算法的多元铁水质量训练效果.可以看出,基于Stacking训练模型的多元铁水质量输出值能够很好地拟合多元铁水质量的实际值.

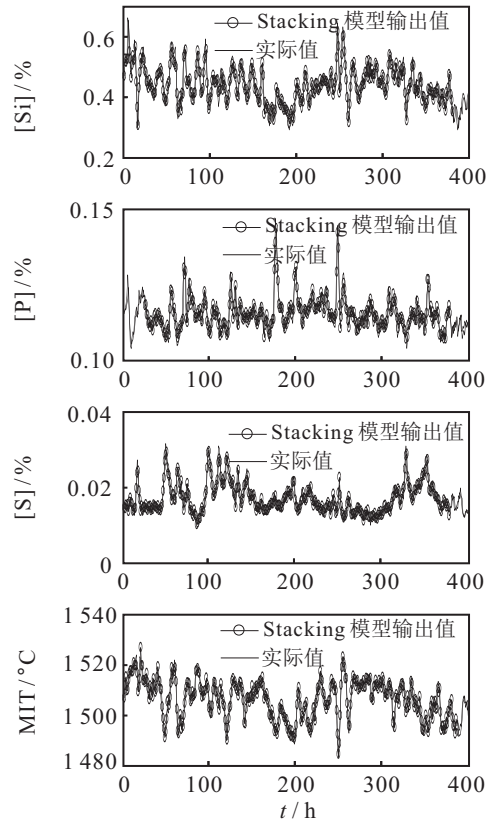


图5 所提出算法下的多元铁水质量效果

为了进一步表明模型的精度问题,构造建模误差曲线如图6所示.由图6可见,多元铁水质量指标[S_i]、[P]、[S]、MIT的建模误差分别集中在±0.1、±0.015、±0.005、±10区间内,进一步计算,求得铁水质量指标[S_i]、[P]、[S]、MIT的建模平均相对误差分别为5.6%、2.2%、8.0%、1.7%,表明所提出算法建立的多元铁水质量模型拥有较高的精度.

2) 预测区间的求取.

为了求取多元铁水质量预测区间,引入核密度估计方法,将图6所示的具有随机特性的建模误差表示成具有确定特性的概率密度函数曲线的形式.图7为多元铁水质量的建模误差PDF曲线.由图7可见,多元铁水质量指标[S_i]、[P]、[S]、MIT的建模误差PDF曲线符合高斯分布.因此,结合统计学概率密度函数置信区间的求取原理,定义建模误差正负3倍的标准

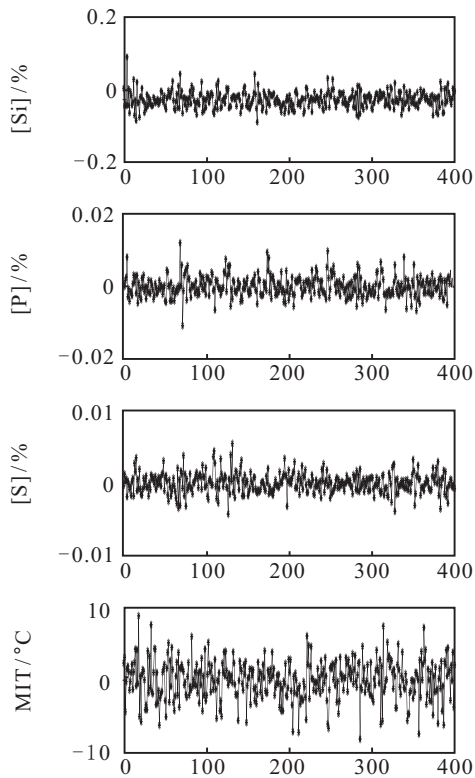


图6 建模误差曲线

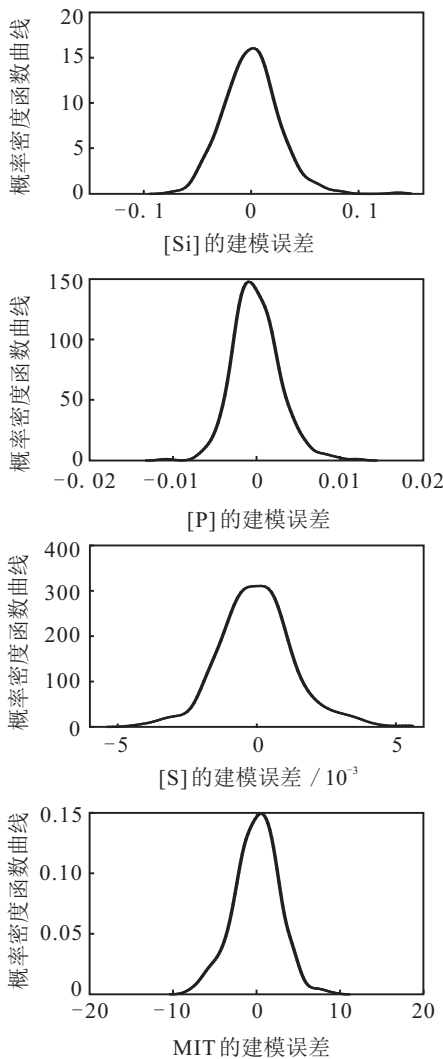


图7 建模误差PDF曲线

差为模型建模误差的预测区间. 预测值的预测区间为预测值加减3倍的建模误差的标准差, 其中, 模型预测值加3倍建模误差的标准差为预测上界, 模型预测值减3倍建模误差的标准差为预测下界. 所求得的多元铁水质量指标 [Si]、[P]、[S]、MIT 建模误差的标准差(即 σ_1) 分别为 0.025 2、0.002 8、0.001 4、2.685 5. 最终, 求得 [Si]、[P]、[S]、MIT 建模误差的预测区间(即 $\pm 3\sigma_1$) 分别为 $\pm 0.075 6$ 、 $\pm 0.008 4$ 、 $\pm 0.004 2$ 、 $\pm 8.056 5$.

3) 区间预测效果.

采用新的后 120 组工业数据对模型进行测试, 除了对多元铁水质量值进行预测外, 还会对 [Si]、[P]、[S]、MIT 的预测区间进行预报, 即实现多元铁水质量值和预测区间的同时预报. 最终得到的多元铁水质量区间预测效果如图 8 所示. 由图 8 可见, 所提出算法不仅使得预测值很好地拟合了多元铁水质量的实际值, 而且多元铁水质量的实际值均包含于预测上界和预测下界之间.

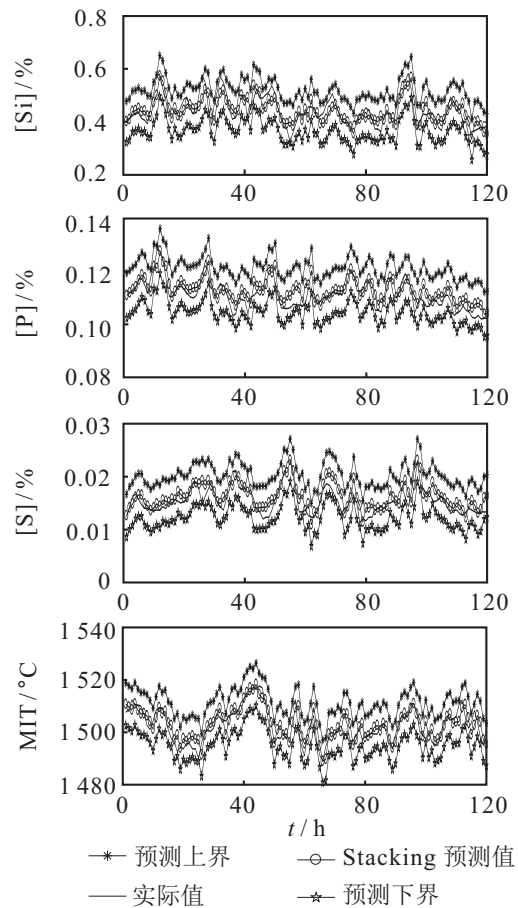


图8 区间预测散点图

为了进一步表明该特性, 将多元铁水质量的实际值和预测值分别作为横、纵坐标, 画出在所提出算法下各个多元铁水质量指标的预测值与实际值的散点图, 如图 9 所示. 由图 9 可见, 点基本分布在过原点且斜率为 1 的曲线上及附近, 这很好地表明了预测值能

够跟踪实际值的变化,模型精度较高,而所有散点都几乎落在预测区间内,也表明了所给预测期间具有较高的可信度。

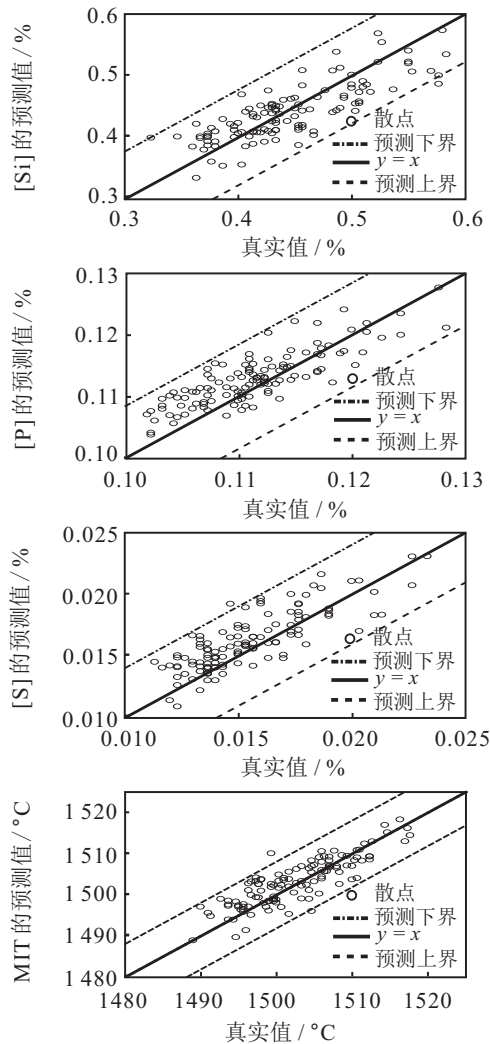


图9 区间预测效果

5 结论

本文结合高炉炼铁的实际过程,旨在建立一种建模快速、精度较高且易于工程实现的高炉多元铁水质量区间预测模型.主要工作包括:1)为了提高建模数据的质量,对高炉实际数据进行预处理,从而得到高质量的高炉建模数据;2)为了提高建模的效率,采用灰色关联分析法筛选出与多元铁水质量关联度最强的7个过程变量作为输入变量;3)为了提高多元铁水质量建模的准确性和进一步表征预测结果的可信度,结合Stacking模型和区间预测方法,提出一种基于Stacking的高炉铁水质量区间预测建模算法,用于建立多元铁水质量区间预测模型,实现多元铁水质量值和预测区间的同时预报.工业实验分析表明:所提出算法具有较高的估计精度,且能够根据高炉炼铁过程实时输入数据的变化,实现对多元铁水质量预测值和预测区间进行同时预报。

参考文献(References)

- [1] Zhou P, Song H D, Wang H, et al. Data-driven nonlinear subspace modeling for prediction and control of molten iron quality indices in blast furnace ironmaking[J]. IEEE Transactions on Control Systems Technology, 2017, 25(5): 1761-1774.
- [2] Kuang S B, Li Z Y, Yan D L, et al. Numerical study of hot charge operation in ironmaking blast furnace[J]. Minerals Engineering, 2014, 63: 45-56.
- [3] Chen Z P, Jiang Z H, Gui W H, et al. A novel device for optical imaging of blast furnace burden surface: Parallel low-light-loss backlight high-temperature industrial endoscope[J]. IEEE Sensors Journal, 2016, 16(17): 6703-6717.
- [4] 姜慧研, 许桂清, 周建常. 高炉煤气流分布模式识别与布料操作指导专家系统[J]. 控制与决策, 2001, 16(6): 930-933.
(Jiang H Y, Xu G Q, Zhou J C. Pattern recognition of gas fluid distribution and expert system for charging operation in blast furnace[J]. Control and Decision, 2001, 16(6): 930-933.)
- [5] 邵富群, 颜华, 王师. 高炉计算机断层成像技术成像算法的研究[J]. 控制与决策, 2000, 15(2): 201-204.
(Shao F Q, Yan H, Wang S. Image reconstruction algorithm for blast furnace imaging[J]. Control and Decision, 2000, 15(2): 201-204.)
- [6] Das S K, Kumari A, Bandopadhyay D, et al. A mathematical model to characterize effects of liquid hold-up on bosh silicon transport in the dripping zone of a blast furnace[J]. Applied Mathematical Modeling, 2011, 35(9): 4208-4221.
- [7] Nogami H, Chu M S, Yagi J I. Multi-dimensional transient mathematical simulator of blast furnace process based on multi-fluid and kinetic theories[J]. Computers and Chemical Engineering, 2005, 29(11/12): 2438-2448.
- [8] Zarandi M H F, Ahmadpour P. Fuzzy agent-based expert system for steel making process[J]. Expert Systems with Applications, 2009, 36(5): 9539-9547.
- [9] Zhou P, Lu S W, Chai T Y. Data-driven soft-sensor modeling for product quality estimation using case-based reasoning and fuzzy-similarity rough sets[J]. IEEE Transactions on Automation Science and Engineering, 2014, 11(4): 992-1003.
- [10] Chao Y C, Su C H, Huang H P. The adaptive autoregressive models for the system dynamics and prediction of blast furnace[J]. Chemical Engineering Communications, 1986, 44(1/2/3/4/5/6): 309-330.
- [11] 邵传厚, 渐令, 陈积明, 等. 复杂高炉炼铁过程的数据驱动建模及预测算法[J]. 自动化学报, 2009, 35(6): 725-730.
(Gao C H, Jian L, Chen J M, et al. Data-driven modeling and predictive algorithm for complex blast furnace ironmaking process[J]. Acta Automatica Sinica, 2009, 35(6): 725-730.)
- [12] Shi L, Li Z L, Li J P. Model of hot metal silicon content

- in blast furnace based on principal component analysis application and partial least square[J]. *Journal Iron and Steel Research, International*, 2011, 18(10): 13-16.
- [13] Saxen H, Gao C H, Gao Z. Data-driven time discrete models for dynamic prediction of the hot metal silicon in the blast furnace — A review[J]. *IEEE Transactions on Industrial Informatics*, 2013, 9(4): 2213-2225.
- [14] Wang W, Chen W L, Ye Y, et al. Application of neural network to predict sulphur content in hot metal[J]. *Iron & Steel*, 2006, 41(10): 19-22.
- [15] Chen W, Wang B X, Han H L. Prediction and control for silicon content in pig iron of blast furnace by integrating artificial neural network with genetic algorithm[J]. *Ironmaking & Steelmaking*, 2010, 37(6): 458-463.
- [16] Gao C H, Jian L, Luo S H. Modeling of the thermal state change of blast furnace hearth with support vector machines[J]. *IEEE Transactions on Industrial Electronics*, 2012, 59(2): 1134-1145.
- [17] Zhou P, Guo D, Wang H, et al. Data-driven robust M-LS-SVR-based NARX modeling for estimation and control of molten iron quality indices in blast furnace ironmaking[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2018, 29(9): 4007-4021.
- [18] Tang X L, Zhuang L, Jiang C J. Prediction of silicon content in hot metal using support vector regression based on chaos particle swarm optimization[J]. *Expert Systems with Applications*, 2009, 36(9): 11853-11857.
- [19] Liu Y, Gao Z L. Enhanced just-in-time modelling for online quality prediction in BF ironmaking[J]. *Ironmaking and Steelmaking*, 2015, 42(5): 321-330.
- [20] 刘学艺, 刘祥官, 王文慧. 贝叶斯网络在高炉铁水硅含量预测中的应用[J]. *钢铁*, 2005, 40(3): 17-20.
(Liu X Y, Liu X G, Wang W H. Application Of bayesian network to predicting silicon content in hot metal[J]. *Iron and Steel*, 2005, 40(3): 17-20.)
- [21] Hou Z S, Wang Z. From model-based control to data-driven control: Survey, classification and perspective[J]. *Information Sciences*, 2013, 235: 3-35.
- [22] Zeng J S, Gao C H, Su H Y. Data-driven predictive control for blast furnace ironmaking process[J]. *Computers and Chemical Engineering*, 2010, 34(11): 1854-1862.
- [23] Yuan M, Zhou P, Li M L, et al. 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.
- [24] Gidde R R, Shinde A B, Pawar P M, et al. Design optimization of a rectangular wave micromixer (RWM) using Taguchi based grey relational analysis (GRA)[J]. *Microsystem Technologies*, 2018, 24(9): 3651-3666.
- [25] Kuo Y, Yang T, Huang G W. The use of grey relational analysis in solving multiple attribute decision-making problems[J]. *Computers & Industrial Engineering*, 2008, 55(1): 80-93.
- [26] Wolpert D H. Stacked generalization[J]. *Neural Networks*, 1992, 5(2): 241-259.
- [27] Pao Y H, Takefuji Y. Functional-link net computing[J]. *IEEE Computer*, 1992, 25(5): 76-79.
- [28] 乔俊飞, 李凡军, 杨翠丽. 随机神经网络研究现状与展望[J]. *智能系统学报*, 2016, 11(6): 758-767.
(Qiao J F, Li F J, Yang C L. Review and prospect on neural networks with random weights[J]. *CAAI Transactions on Intelligent Systems*, 2016, 11(6): 758-767.)
- [29] Khosravi A, Nahavandi S, Creighton D, et al. Comprehensive review of neural network-based prediction intervals and new advances[J]. *IEEE Transactions on Neural Networks*, 2011, 22(9): 1341-1356.
- [30] Quan H, Srinivasan D, Khosravi A. Particle swarm optimization for construction of neural network-based prediction intervals[J]. *Neurocomputing*, 2014, 127: 172-180.
- [31] Parzen E. On estimation of a probability density function and mode[J]. *Annals of Mathematical Statistics*, 1962, 33(3): 1065-1076.

作者简介

周平(1980—), 男, 教授, 博士生导师, 从事工业过程运行反馈控制、数据驱动智能建模等研究, E-mail: zhouping@mail.neu.edu;

刘进进(1990—), 男, 硕士生, 从事数据驱动智能建模的研究, E-mail: 1780608555@qq.com.

(责任编辑: 郑晓蕾)