基于最优工况迁移的高炉铁水硅含量预测方法

蒋朝辉[1,2] 许 川[1] 桂卫华[1,2] 蒋 珂[1]

摘 要 高炉铁水硅含量是铁水品质与炉况的重要表征, 冶炼过程关键参数频繁波动及大时滞特性给高炉铁水硅含量预测带来了巨大挑战. 提出一种基于最优工况迁移的高炉铁水硅含量预测方法. 首先, 针对过程变量频繁波动问题, 提出基于邦费罗尼指数的自适应密度峰值聚类算法, 实现对高炉冶炼过程变量的工况划分, 并建立不同工况硅含量预测子模型. 其次, 针对冶炼过程的大时滞特性, 定义相邻时间节点间的硅含量工况迁移代价函数, 并提出多源路径寻优算法, 实现冶炼过程中硅含量最优工况迁移路径及当前时刻硅含量最优预测值的求解. 最后, 基于工业现场数据验证了所提方法的有效性与准确性.

关键词 高炉炼铁, 铁水硅含量, 预测, 工况迁移, 密度峰值聚类

引用格式 蒋朝辉, 许川, 桂卫华, 蒋珂. 基于最优工况迁移的高炉铁水硅含量预测方法. 自动化学报, 2022, **48**(1): 194−206

DOI 10.16383/j.aas.c200980

Prediction Method of Hot Metal Silicon Content in Blast Furnace Based on Optimal Smelting Condition Migration

JIANG Zhao-Hui[1,2] XU Chuan[1] GUI Wei-Hua[1,2] JIANG Ke[1]

Abstract The hot metal silicon content in blast furnace can characterize the hot metal quality and the condition of blast furnace. It poses a great challenge to the online prediction of silicon content because of the frequent fluctuation of smelting parameters and the existence of large time delay during the ironmaking process. This paper proposes an algorithm for predicting the hot metal silicon content in blast furnace based on optimal smelting condition migration. Firstly, arming at the frequent fluctuation of smelting process variables, an adaptive density peak clustering algorithm based on the Bonferroni index to dynamically cluster the process variables of blast furnace ironmaking process is proposed, which can obtain clusters of different smelting conditions, and establish sub-models for different smelting conditions. Secondly, to mitigate the large time delay of blast furnace ironmaking process, this paper defines the silicon content migration cost function between adjacent time nodes, and proposes a multi-source path optimization algorithm to solve the optimal migration path of silicon content during the smelting process and the optimal prediction value of silicon content at the current time. Finally, the effectiveness and accuracy of the proposed method are verified based on the industrial field data.

Key words Blast furnace ironmaking, hot metal silicon content, prediction, smelting condition migration, density peak clustering

Citation Jiang Zhao-Hui, Xu Chuan, Gui Wei-Hua, Jiang Ke. Prediction method of hot metal silicon content in blast furnace based on optimal smelting condition migration. *Acta Automatica Sinica*, 2022, **48**(1): 194−206

高炉炼铁是钢铁生产的核心流程[1], 其冶炼过程如图 1 所示[2]. 布料系统从高炉炉顶投入烧结矿、焦炭、熔剂等炼铁物料, 热风系统与煤粉喷吹系统从高炉下部风口鼓入高温热风及煤粉, 在炉内高温条件下生成氢气、一氧化碳等还原性气体. 伴随冶炼过程进行, 下降的炉内物料与上升的高温煤气流接触, 经还原、软化、熔融、脱碳, 形成铁水, 最终与炉渣从出铁口一起排出.

高炉铁水硅含量是冶炼过程的关键性能指标, 表征了高炉炉缸热状态及其变化趋势, 同时与高炉炉况和铁水质量密切相关[3]. 目前, 工业现场铁水硅含量主要依靠人工采样后离线化验的方式获取. 采样过程具有一定危险性, 化验过程耗时耗力, 数据不具有时效性, 不能为高炉冶炼提供及时的反馈信

收稿日期 2020-11-25 录用日期 2021-03-19
Manuscript received November 25, 2020; accepted March 19, 2021
国家自然科学基金 (61773406, 61988101), 中南大学中央高校基本科研任务业务费专项资金 (2020zzts572) 资助
Supported by National Natural Science Foundation of China (61773406, 61988101), and Central South University Central University Basic Scientific Research Task Business Expenses Special Funds (2020zzts572)
本文责任编委 杨涛
Recommended by Associate Editor YANG Tao
1. 中南大学自动化学院 长沙 410000 2. 鹏城实验室 深圳 518000
1. School of Automation, Central South University, Changsha 410000 2. Peng Cheng Laboratory, Shenzhen 518000

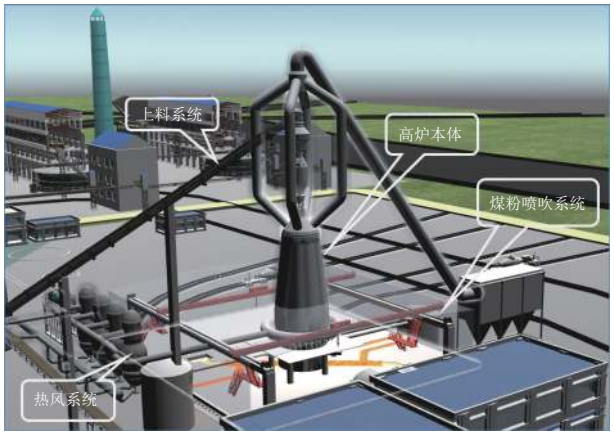

图 1　高炉炼铁工艺

Fig. 1　Blast furnace ironmaking process

息. 因此, 实现铁水硅含量实时高精度预测, 能够帮助现场及时评估高炉冶炼状态以及调整布料、送风等操作制度, 进而为优化冶炼过程操作、实现精细化调控提供数据支撑[4], 对保证炉况稳定、提升产品质量、降低冶炼能耗具有重大实际意义.

当前硅含量预测模型主要分为机理模型与数据驱动模型. 机理模型根据反应过程中的热状态、动力学特性、物料守恒特性等进行分析建模[5−6], 为认识冶炼过程及炉内现象起到了积极作用. 但机理模型建模复杂, 模型性能受高炉自身参数、运行状态、传感器分布及精度的影响[7], 不具有良好的普适性, 难以应用于工业现场硅含量在线预测. 随着集散控制系统和工业互联网的发展, 反映高炉冶炼过程的海量数据存储在历史数据库中, 为基于数据驱动的铁水硅含量预测提供了数据支持. 数据驱动模型成为当前的研究热点, 线性自回归模型[8]、状态空间模型[9]、偏最小二乘模型[10]、T-S 模糊模型[11]、支持向量机[12−13]、Wiener 模型[14]、神经网络模型[15−19]、随机权神经网络[20] 等广泛应用于硅含量预测, 但均存在一定局限性. 例如, 偏最小二乘模型、最小二乘支持向量机[21] 主要基于统计学指标进行优化, 预测结果依赖于指标的选取, 模型难以适应复杂的高炉炉况, 预测精度较差; Wiener 模型具有短期记忆特性, 但在一定时间后, 模型的预测精度会显著下降, 模型需重新训练[22], 模型的稳定性较差; 神经网络模型没有充分利用冶炼过程历史数据, 忽视了硅含量的渐变性, 不能克服冶炼过程工况复杂多变及大时滞特性, 在变量相关性不稳定及过程参数频繁波动情况下稳定性较差. 文献 [23] 融合神经网络与 Bootstrap 预报方法, 实现了硅含量数值预报并给出可信度; 文献 [24−25] 提出集成建模方法, 对高炉铁水硅含量、铁水温度等多元铁水指标进行了综合预报. 现有数据驱动模型在硅含量预测上取得了一定的积极效果. 但是, 当冶炼过程中的过程变量频繁波动时, 硅元素物理、化学反应环境剧烈变化, 导致硅含量数据波动大以及过程变量相关性不稳定, 过程变量分布呈现显著非平衡性, 硅含量预测结果存在较大不确定性, 在过程参数频繁波动情况下模型的稳定性不足; 并且, 冶炼过程存在大时滞特性, 高炉出铁周期为 2 h ~ 3 h, 单一时刻的过程变量难以全面反映整个冶炼过程的信息, 炉况波动时, 预测精度显著下降.

针对上述问题, 本文提出了一种基于最优工况迁移的铁水硅含量预测方法. 首先, 为提高高炉炉况波动时模型的稳定性, 本文提出基于邦费罗尼指数的自适应密度峰值聚类算法, 对过程变量进行动态聚类, 划分为不同工况下的过程变量数据集, 并对不同工况变量进行单独建模. 划分工况能够避免数据的非平衡特性给模型带来的影响, 进而显著提高工况波动时的硅含量预测精度. 其次, 为充分利用历史硅含量信息以克服冶炼过程的大时滞特性, 提出相邻时间节点间硅含量工况迁移代价函数与多源路径寻优算法. 通过历史硅含量预测值与化验值对当前时刻硅含量预测值进行寻优, 在复杂工况条件下取得了良好的预测效果. 最后, 通过工业实验表明, 所提基于最优工况迁移的硅含量预测方法能够降低高炉运行过程参数波动给模型预测精度带来的影响, 在提升模型预测精度的同时显著增加了模型预测性能的稳定性.

1　建模策略

本文从实际冶炼过程出发, 为提高模型对过程变量波动的适应性与稳定性, 提出如图 2 所示的基于最优工况迁移的硅含量预测动态建模策略.

1) 数据预处理是建立数据驱动模型的重要环节, 输入参数选择不当会造成特征缺失或信息冗余, 不利于后续建模. 本文首先引入最大信息系数对过程变量与硅含量进行相关性分析, 根据相关性筛选输入变量; 随后对筛选的过程变量数据进行异常值剔除; 最后针对各参数量纲不一致问题, 进行归一化处理.

2) 针对过程变量波动特性, 提出基于邦费罗尼指数的自适应密度峰值聚类算法, 对筛选的过程变量进行动态聚类, 自动求解最优截断距离与聚类中心, 从而将数据划分为不同工况的聚类簇. 选择具有时变适应特性的 Elman 网络对不同工况聚类簇的过程变量进行建模, 采用基于模拟退火的 Levenberg-Marquardt 算法更新模型参数, 求解得到不同工况下的硅含量预测子模型.

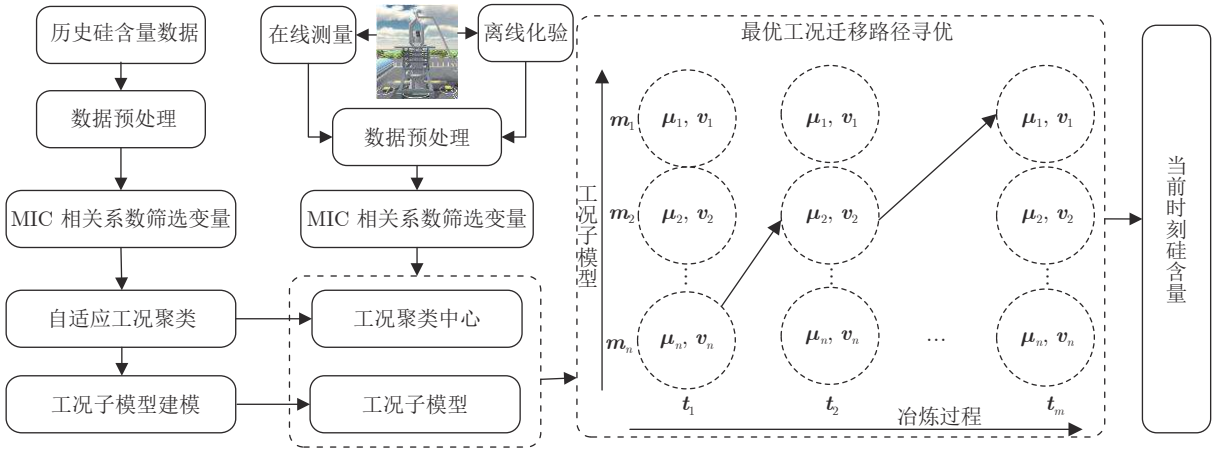

图 2 基于最优工况迁移的建模策略

Fig. 2 Modeling strategy based on optimal smelting condition migration

3) 冶炼过程具有大时滞特性, 在硅元素反应的时间段内, 过程变量参数频繁波动, 单一时刻过程变量难以表征整个反应时间段内整体冶炼情况, 导致现有模型硅含量预测结果具有不稳定性. 由于工业现场过程变量通常每 10 s 更新一次, 在此时间间隔内, 硅含量不会发生突变. 因此, 本文建立包含当前时刻及一系列历史时间节点过程变量的滑动窗口, 根据滑动窗口中各时间节点工况隶属度以及各子模型预测值, 定义相邻时间节点间硅含量工况迁移代价函数, 并提出多源路径寻优算法, 从而根据历史硅含量预测值与化验值求解当前时刻最优工况子模型硅含量预测值.

2 自适应密度峰值聚类的工况子模型建模

冶炼过程中关键参数频繁波动, 造成硅含量数据波动大、相关性不稳定且呈现显著非平衡特性, 对模型稳定性提出极高的要求, 单个数据驱动模型难以适应复杂多变的工况. 本节提出基于邦费罗尼指数的自适应密度峰值聚类算法, 将过程变量划分为不同工况类型的聚类簇, 并分别建立不同工况硅含量预测子模型, 作为后续硅含量预测寻优的基础.

2.1 基于邦费罗尼指数的自适应密度峰值工况聚类

2.1.1 过程变量数据特点分析

高炉冶炼过程变量主要存在以下特点:

1) 数据维度较高. 从机理和数据角度分析可知, 透气性指数、冷风流量、鼓风动能、炉腹煤气指数等数十维过程变量与冶炼过程中硅元素物理、化学反应相关.

2) 数据存在波动性. 在炉况稳定情况下, 各过程变量在某一水平附近保持相对稳定, 但随着布料周期性的进行、反应物料特性的改变、人为操作及炉况不稳情况出现, 各过程变量会发生剧烈波动. 过程变量一定程度上能够反映炉况优劣并影响硅元素物理、化学反应环境, 造成铁水硅含量的频繁波动.

3) 呈现聚堆现象. 在不同的高炉冶炼状态下, 高炉的操作制度存在差异, 硅元素物理、化学反应环境存在着剧烈变化, 导致硅含量水平存在巨大的差异. 通过大量过程变量分析发现, 伴随冶炼进程及操作的进行, 过程变量与硅含量的分布存在着显著的聚堆现象.

由于以上特性的存在, 过程变量与硅含量的相关性呈现出不稳定状态, 单一模型建模精度较差. 本文通过聚类, 将过程变量划分为不同工况类型, 并对不同工况条件下过程变量数据集分别建立硅含量预测子模型. 同一工况下的过程变量数据集具有较强的稳定性, 使模型能专注于不同工况条件下的硅含量预测, 由此减少过程变量波动给模型带来的扰动, 进而提高模型的预测精度和在过程变量频繁波动情况下模型的稳定性.

2.1.2 基于邦费罗尼指数的自适应密度峰值聚类

密度峰值聚类算法[26] (Density peak clustering algorithm, DPC) 继承了基于密度聚类算法的优点, 且具有思想简洁、参数少、适用于高维数据等优势. 但该算法的截断距离 d_c 与聚类中心需要人为给定, 难以保证最优聚类效果. 因此, 本文引入邦费罗尼指数[27] (Bonferroni index) 实现截断距离自适应, 并提出聚类中心自动选取策略.

1) 最优截断距离

对于标准的 DPC 算法, 截断距离的选择依赖

人工经验, 对于不同特点、不同规模的数据集缺乏普适性, 影响聚类中心的选取、聚类簇的数量以及最终的聚类效果. 标准密度峰值聚类算法中, ρ_i 为各数据点的局部密度, δ_i 为数据点至更高密度点的最小距离, 两参数乘积 $\rho_i\delta_i$ 是评判数据点能否成为聚类中心的关键参数, $\rho_i\delta_i$ 越大则越有可能成为聚类中心. 本文将邦费罗尼指数引入密度峰值聚类算法, 用于衡量参数 $\rho_i\delta_i$ 的有序程度, 邦费罗尼指数越大, 数据不确定度越小, 越有利于聚类. 对所有数据点参数 $\rho_i\delta_i$ 进行归一化处理, 再进行升序排列, 各数据点参数 $\rho_i\delta_i$ 的下标 $(i=1,\ 2,\ 3,\cdots,\ N)$ 依次递增, 则离散邦费罗尼指数为:

$$B_N=\frac{1}{N-1}\sum_{i}^{N-1}\left(\frac{P_i-Q_i}{P_i}\right) \tag{1}$$

式中, B_N 为离散邦费罗尼指数, N 为数据划分单元总数. P_i 与 Q_i 定义为:

$$\begin{cases} P_i=\dfrac{i}{N} \\ Q_i=\dfrac{\sum\limits_{j=1}^{i}\rho_j\delta_j}{\sum\limits_{j=1}^{N}\rho_j\delta_j} \end{cases} \tag{2}$$

通过改变截断距离, 计算不同截断距离下系统邦费罗尼指数, 绘制邦费罗尼指数曲线. 当邦费罗尼指数取得全局极大值时, 系统参数 $\rho_i\delta_i$ 有序程度最高, 最有利于聚类, 此时 d_c 即为最优截断距离.

2) 自动选取聚类中心

在获得最优截断距离 d_c 条件下, 需要进一步确定聚类中心, 以完成不同工况聚类簇的划分. 首先, 求解聚类参数的均值 $\overline{\rho\delta}$, 将大于 $\overline{\rho\delta}$ 的聚类参数 $\rho_i\delta_i$ 进行降序排列, 绘制聚类决策图. 定义聚类中心截断系数 ω_m, 以确定聚类中心的最终选取. 聚类中心截断系数定义如下:

$$\begin{cases} \omega_m=\dfrac{\left(\frac{1}{m}\sum\limits_{i=1}^{m}\rho_i\delta_i\right)-\rho_{m+1}\delta_{m+1}}{\left(\frac{1}{m}\sum\limits_{j=1}^{m}d_{j,\,m+1}\right)},\ 1<m\le N-1 \\ \omega_{\max}=\max\limits_{1<m<N-1}(\omega_m) \end{cases} \tag{3}$$

式中, ω_m 为降序排列的第 m 个数据点的截断系数, 当 ω_m 取得全局最大值 $\omega_{\max}$ 时, 此时的数据点为聚类中心的截止点, $\omega_{\max}$ 及之前的数据点确定为聚类中心. 将样本中非聚类中心的数据点划分到距离最近的聚类中心所属聚类簇中, 每种聚类簇数据代表一种工况, 获得聚类中心与聚类簇如下:

$$\begin{cases} \boldsymbol{c}=[c_1,\ c_2,\ c_3,\cdots,\ c_n]^{\mathrm{T}} \\ \boldsymbol{g}=[g_1,\ g_2,\ g_3,\cdots,\ g_n]^{\mathrm{T}} \end{cases} \tag{4}$$

式中, c_i 为第 i 种工况聚类簇 g_i 的聚类中心, n 为工况聚类中心总数. 对于非聚类中心数据点, 定义该数据点与各聚类中心距离的倒数作为与各聚类中心的相关系数. 相关系数归一化后为该组数据对各工况聚类簇隶属度, 如式 (5):

$$\mu_{ij}=d_{ij}^{-1}\left(\sum_{k=1}^{n}d_{ik}^{-1}\right)^{-1} \tag{5}$$

式中, μ_{ij} 为非聚类中心的第 i 个数据点对第 j 种工况聚类簇的隶属度, d_{ij}^{-1} 为非聚类中心的第 i 个数据点与第 j 种工况聚类中心 c_j 的距离.

2.2　基于 Elman 神经网络的工况子模型

通过上述自适应聚类过程, 将硅含量历史数据集划分为 n 种工况聚类簇, 并采用 Elman 网络对不同工况数据进行建模. Elman 网络是一种具有局部记忆单元与局部反馈连接的递归神经网络, 其网络结构如图 3 所示, 承接层的出现使得网络具备了时变适应特性, 能够较好适应冶炼过程中硅含量变化特点. Elman 网络的数学模型为:

$$\begin{cases} y_t=g(\boldsymbol{W}^3x(t)) \\ x(t)=f(\boldsymbol{W}^2x_c(t)+\boldsymbol{W}^1u(t-1)) \\ x_c(t)=x(t-1) \end{cases} \tag{6}$$

式中, y_t 表示 t 时刻网络输出, $\boldsymbol{W}^1$、$\boldsymbol{W}^2$、$\boldsymbol{W}^3$ 分别为承接层、输入层、输出层的权值矩阵, $g(\bullet)$ 为输出层传递函数, $f(\bullet)$ 为隐含层神经元激励函数. 本文采用基于模拟退火的 Levenberg-Marquardt 算法[28] 对网络进行权值更新, 能够加快网络训练速度并避免陷入局部最优解, 从而收敛到更高精度. 将聚类簇 $\boldsymbol{g}=[g_1,\ g_2,\ g_3,\cdots,\ g_n]^{\mathrm{T}}$ 中数据分别输入 Elman 神经

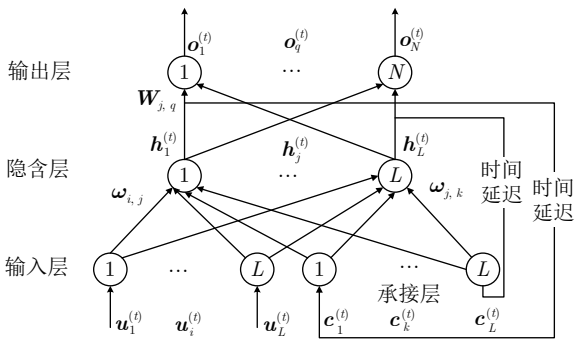

图 3　Elman 神经网络结构

Fig. 3　Structure of Elman neural network

网络, 训练后获得不同工况的子模型 $\boldsymbol{m}=[m_1,\ m_2,\ m_3,\cdots,\ m_n]^{\mathrm{T}}$.

3 基于工况子模型的最优工况迁移硅含量预测

对于同一时刻的过程变量, 不同工况子模型会同时求解出多个硅含量预测值. 仅通过当前时刻数据难以判断各子模型的预测效果, 而相邻时间节点间隔为 10 s 左右, 此时间间隔内硅含量不会发生突变. 因此, 通过一系列连续时间节点的硅含量工况迁移代价函数寻优, 能够确定当前时刻硅含量最优工况子模型预测值. 本节首先确定一个包含了一系列连续时间节点高炉运行过程变量的滑动窗口, 对各单一时间节点求解其不同工况子模型的硅含量预测值与工况隶属度; 然后定义相邻时间节点间的硅含量迁移代价函数与滑动窗口中硅含量最优工况迁移路径目标函数; 最后通过多源路径寻优算法求解当前时刻硅含量最优工况子模型预测值.

3.1 冶炼硅含量工况迁移矩阵

3.1.1 定义滑动窗口

为建立当前时刻硅含量预测值与历史值的联系, 首先建立一个包含一系列连续时间节点过程变量的动态滑动窗口, 如图 4 所示. 滑动窗口中各时间节点表征该时刻的过程变量检测值. 滑动窗口长度取决于最新两个硅含量化验时刻之间过程变量检测样本数, 滑动窗口长度随化验值的更新而动态变化, 以确保滑动窗口中至少包含一个硅含量化验值, 作为后续预测寻优的校验值. 滑动窗口确定了数据样本的采样范围, 是后续硅含量最优工况迁移矩阵建立与硅含量最优工况子模型预测值求解的基础.

3.1.2 建立硅含量工况迁移矩阵

工况迁移矩阵一共 m 列, 与滑动窗口中包含的时间节点数相同, 每一列与滑动窗口中一个时间节点相对应, 代表某时刻的过程变量对应的所有工况隶属度以及所有工况子模型的硅含量预测值. 对单一时间节点的过程变量, 分别与各聚类中心 $\boldsymbol{c}=[c_1,\ c_2,\ c_3,\cdots,\ c_n]^{\mathrm{T}}$ 匹配, 计算该节点对应不同工况子模型的隶属度 $\boldsymbol{\mu}=[\mu_1,\ \mu_2,\ \mu_3,\cdots,\ \mu_n]^{\mathrm{T}}$, 同时根据工况子模型 $\boldsymbol{m}=[m_1,\ m_2,\ m_3,\cdots,\ m_n]^{\mathrm{T}}$ 求解各子模型的硅含量预测值 $\boldsymbol{v}=[v_1,\ v_2,\ v_3,\cdots,v_n]^{\mathrm{T}}$, 得到单一时间节点的数据 $[(\mu_1,\ v_1)_t,\cdots,\ (\mu_n,\ v_n)_t]^{\mathrm{T}}$, 其运算过程如图 5 所示. 对滑动窗口中所有时间节点过程变量进行工况隶属度计算及工况子模型预测, 获得工况迁移矩阵如式 (7):

$$\begin{bmatrix} (\mu_1,v_1)_1,\ (\mu_1,v_1)_2,\cdots,\ (\mu_1,v_1)_t,\cdots,\ (\mu_1,v_1)_m \\ (\mu_2,v_2)_1,\ (\mu_2,v_2)_2,\cdots,\ (\mu_2,v_2)_t,\cdots,\ (\mu_2,v_2)_m \\ (\mu_3,v_3)_1,\ (\mu_3,v_3)_2,\cdots,\ (\mu_3,v_3)_t,\cdots,\ (\mu_3,v_3)_m \\ \cdots \\ (\mu_n,v_n)_1,\ (\mu_n,v_n)_2,\cdots,\ (\mu_n,v_n)_t,\cdots,\ (\mu_n,v_n)_m \end{bmatrix} \tag{7}$$

式中, $(\mu_n,v_n)_t$ 表示时间节点 t 对应的过程变量对第 n 种工况聚类簇的隶属度与子模型预测值. 工况迁移矩阵为相邻时间节点间工况迁移代价函数定义及当前时刻最优工况子模型硅含量预测值求解奠定了基础.

3.2 硅含量预测值最优迁移路径求解

3.2.1 硅含量工况迁移代价函数与目标函数

考虑到高炉冶炼过程的时序性与硅含量的渐变性, 硅含量预测值最优迁移路径为工况迁移矩阵中最平稳的一条. 为详细说明最优路径的求解过程, 将上述工况迁移矩阵扩展为工况迁移图, 如图 6 所示. 其中工况迁移图中每列与工况迁移矩阵中每列一一对应, 硅含量可以从一个时间节点的任意工况模型向下一时间节点的任意工况模型迁移, 从工况迁移图第一列源节点到工况迁移图最后一列终节点为一条完整的工况迁移路径, 图中虚线表示一条完整的工况迁移路径. t_a 时刻为硅含量化验时间节点, 设定该时刻的硅含量为化验值 v_a.

工况迁移图中两相邻时间节点间, 不同工况子模型硅含量预测值迁移连接如图 7 所示. 定义两相邻时间节点间的硅含量工况迁移代价函数为:

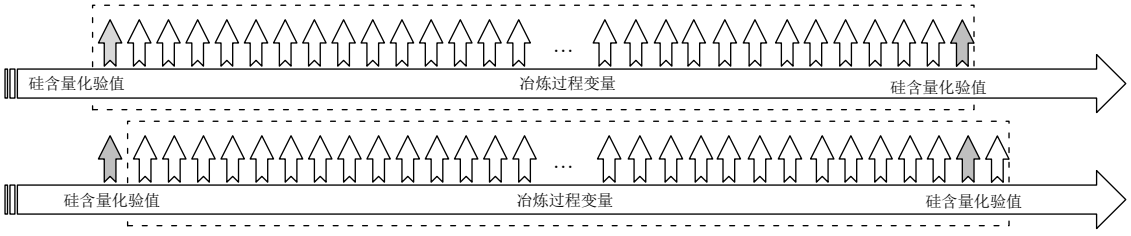

图 4 滑动窗口采样

Fig. 4 Sliding window sampling

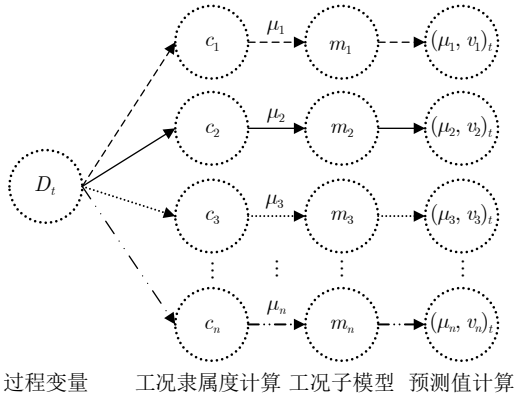

图 5　过程变量工况隶属度与模型预测

Fig. 5　Process variable membership degree matching and model prediction

$$f(x_{p(i-1),\,qi}) = \frac{1}{2}\left(\frac{1}{\mu_{p(i-1)}} + \frac{1}{\mu_{qi}}\right)(|x_{p(i-1)}| - |x_{qi}|) \tag{8}$$

式中, $f(x_{p(i-1),\,qi})$ 定义为从节点 $i-1$ 的第 p 种工况子模型预测值迁移到节点 i 的第 q 种工况子模型预测值的代价函数, $\mu_{p(i-1)}$ 为节点 $i-1$ 对应的第 p 种工况子模型预测值 $x_{p(i-1)}$ 对第 p 种工况聚类簇的隶属度. $f(x_{p(i-1),\,qi})$ 表明硅含量在两相邻时间节点间从节点 $i-1$ 工况子模型硅含量预测值 $x_{p(i-1)}$ 迁移到节点 i 工况子模型硅含量预测值 x_{qi} 的代价, 代价越低, 则硅含量预测值越容易从 $x_{p(i-1)}$ 迁移到 x_{qi}. 由此, 可根据硅含量历史预测值与迁移特性求解当前时刻最优工况子模型硅含量预测值.

为通过工况迁移矩阵求解当前时刻最优工况子模型硅含量预测值, 应求解最小代价迁移路径. 故定义硅含量最优迁移路径目标函数为:

$$cost = \min\left(\sum_{i=2}^{m}(f(x_{p(i-1),\,qi}))\right)$$

$$\text{s. t.}\begin{cases} f(x_{p(i-1),\,qi}) = \dfrac{1}{2}\left(\dfrac{1}{\mu_{p(i-1)}} + \dfrac{1}{\mu_{qi}}\right)(|x_{p(i-1)}| - |x_{qi}|) \\ \mu_1 + \mu_2 + \cdots + \mu_n = 1 \\ 0 \le \mu_i \le 1;\ 1 \le p \le n;\ 1 \le q \le n \\ \text{if}\ (i == t_a) \Rightarrow\ v_i = v_a \end{cases} \tag{9}$$

式中, m 代表工况迁移矩阵中所包含的时间节点数; n 代表工况子模型数目; t_a 表示化验时间节点. 设定此时刻的硅含量为对应的硅含量化验值, 将其作为寻优的基准, 确保最优路径求解的可靠性.

3.2.2　求解算法

为求解硅含量最优工况子模型预测值, 本文提出一种多源路径寻优算法. 首先构建邻接矩阵, 同时更新多条最优路径, 并建立记忆矩阵, 保存最优节点及最小代价. 为求解最优迁移路径, 需要求解相邻时间节点的最优迁移方式, 相邻时间节点间的所有硅含量预测迁移方式如图 8 所示. 两相邻时间节点的硅含量预测值的邻接矩阵 $\boldsymbol{A}_{ij}$ 的构造为:

$$\boldsymbol{A}_{ij} = \begin{bmatrix} f(x_{i1,\,j1}),\ f(x_{i1,\,j2}),\cdots,\ f(x_{i1,\,jn}) \\ f(x_{i2,\,j1}),\ f(x_{i2,\,j2}),\cdots,\ f(x_{i2,\,jn}) \\ \cdots \\ f(x_{in,\,j1}),\ f(x_{in,\,j2}),\cdots,\ f(x_{in,\,jn}) \end{bmatrix} \tag{10}$$

式中, 时间节点 i 与节点 j 相邻, $f(x_{in,\,jn})$ 为节点 i

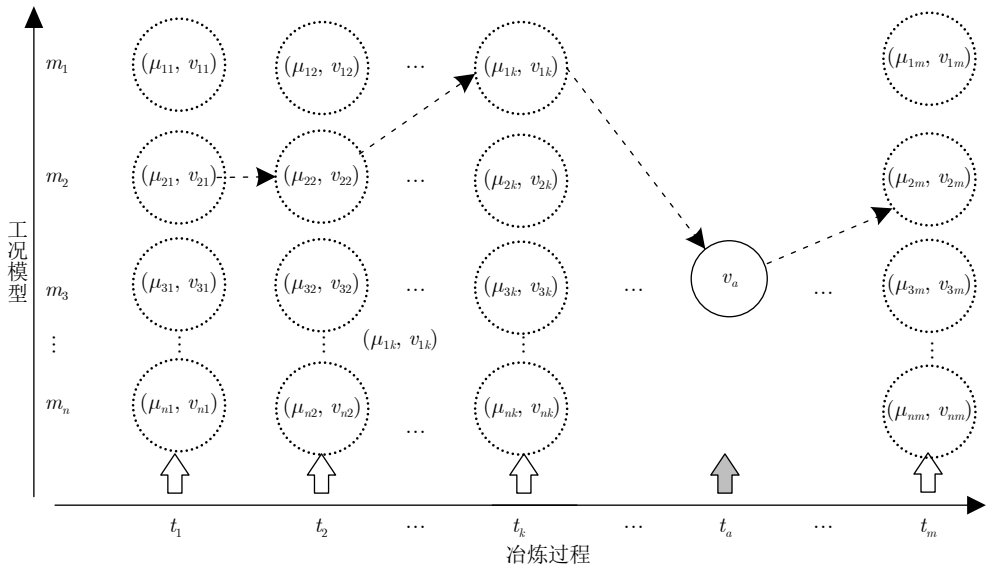

图 6　铁水硅含量工况迁移图

Fig. 6　Smelting condition migration diagram of hot metal silicon content

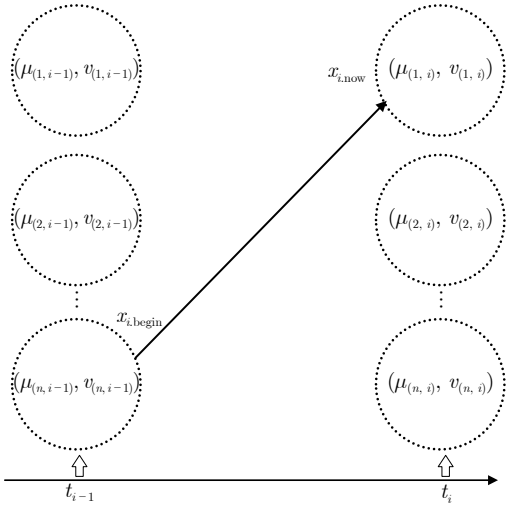

图 7 相邻时间节点工况迁移代价函数

Fig. 7 Smelting condition migration cost function of adjacent node

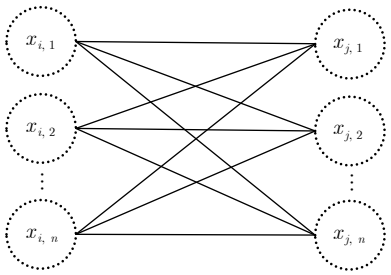

图 8 相邻时间节点连接图

Fig. 8 Connection graph of adjacent node

的第 n 种子模型预测值与节点 j 的第 n 种子模型预测值间的迁移代价函数. 节点 i 到节点 j 中第一种工况子模型预测值的最小迁移代价 min^i_{j1} 为:

$$min^i_{j1} = \min[f(x_{i1,\,j1}),\ f(x_{i2,\,j1}), \cdots,\ f(x_{in,\,j1})] \tag{11}$$

为记录完整迁移路径, 需要记录各节点对应的硅含量预测值. 定义 $x^{\min}_{jn}$ 为两相邻时间节点最小迁移代价 min^i_{jn} 在节点 i 中所对应硅含量预测数据. 为实现最优迁移曲线求解, 定义记忆矩阵 $\boldsymbol{R}_j$ 为:

$$\boldsymbol{R}_j = \begin{bmatrix} min^i_{j1},\ x^{\min}_{j1},\ cost_{j1} \\ min^i_{j2},\ x^{\min}_{j2},\ cost_{j2} \\ \cdots \\ min^i_{jn},\ x^{\min}_{jn},\ cost_{jn} \end{bmatrix} \tag{12}$$

式中, 第一列为节点 i 到节点 j 中各子模型预测值的最小迁移代价, 第二列为节点 i 到节点 j 中各子模型预测值最短路径中所对应的节点 i 中的预测值, 第三列为从源节点到节点 j 中各子模型预测值的最小迁移代价数值. 多源路径寻优算法的具体步骤如下:

步骤 1. 确定源节点 i, 并设置为当前节点, 初始化最小代价矩阵 $[cost_{i1}, cost_{i2}, \cdots, cost_{in}]^{\mathrm{T}}$, 初值设为 0, 表示源节点到其本身的最短路径为 0;

步骤 2. 构造节点 i 与下一相邻时间节点 j 的邻接矩阵, 计算节点 i 到节点 j 各子模型预测值的最小迁移代价 $[min^i_{j1},\ min^i_{j2}, \cdots,\ min^i_{jn}]^{\mathrm{T}}$, 计算节点 i 到节点 j 各子模型预测值的最小迁移代价所对应的节点 i 中子模型预测值 $[x^{\min}_{j1},\ x^{\min}_{j2}, \cdots, x^{\min}_{jn}]^{\mathrm{T}}$, 源节点到节点 j 各子模型预测值的最小迁移代价 $[cost_{j1}, cost_{j2}, \cdots,\ cost_{jn}]^{\mathrm{T}}$, 将上述三向量融合为记忆矩阵 $\boldsymbol{R}_j$;

步骤 3. 若节点 j 为工况迁移矩阵的终节点, 转到步骤 4, 否则对节点 i, j 进行赋值, $i = i + 1$, $j = j + 1$, 再转到步骤 2 继续执行;

步骤 4. 搜索已经到达终节点 j, 选取全局最短路径 $cost_{\min} = \min[cost_{j1}, cost_{j2}, \cdots, cost_{jn}]^{\mathrm{T}}$, $cost_{\min}$ 即为全局最小迁移代价, 并将终节点指向的硅含量预测数值设置为当前数据点 $x_{\text{si.now}}$, 即为当前时刻最优工况子模型硅含量预测值 x_{si};

步骤 5. 通过记忆矩阵 $\boldsymbol{R}_j$ 搜索 $x_{\text{si.now}}$ 的前驱节点 $x_{\text{si.begin}}$, 并记录已搜索的节点, 压入栈 P;

步骤 6. 判断 $x_{\text{si.begin}}$ 是否到达工况迁移矩阵源节点, 若未到达源节点, 则将 $x_{\text{si.begin}}$ 设置为新的当前数据点 $x_{\text{si.now}}$, 转到步骤 5 继续搜索前驱节点, 否则, 结束搜索. 栈 P 中保存的即为最短路径数据点, 最短距离数值为 $cost_{\min}$.

通过上述步骤, 在工况迁移矩阵中求解获得最优工况子模型硅含量预测路径, 以及当前时刻硅含量最优预测值 x_{si}. 通过寻优, 求解出各时间节点对应的最优工况子模型, 保证两相邻时间节点间硅含量预测值的平稳性, 符合冶炼过程硅含量变化特性, 减少过程变量波动给模型精度带来的影响.

4 工业实验验证与结果分析

为验证本文所提最优工况迁移硅含量预测方法有效性, 采用某钢铁集团 2 号高炉 3 000 组数据进行实验仿真, 实验结果证明本文方法预测精度较高, 且在过程参数频繁波动情况下模型有较强的稳定性.

4.1 数据分析与建模

高炉历史数据库中保存了大量重要数据, 但由于某些特定的原因 (如高炉休风或减压) 会使数据出现异常, 不能反映真实的炉况. 因此, 需要对数据进行预处理来提高数据质量, 进而提高硅含量预测模型的精度和性能. 本文采用最大信息系数[20] (Maximal information coefficient, MIC) 求解过程变量与硅

含量间的相关性, MIC 相关性系数定义如下:

$$\mathrm{MIC}=\max_{|x||y|<B}\frac{\sum\limits_{x,\,y}p(x,\,y)\log_2\frac{p(x,\,y)}{p(x)p(y)}}{\log_2(\min(|x|,\,|y|))}\tag{13}$$

式中, $p(x)$ 为数据点落在 x 列的概率; $p(y)$ 为数据点落在 y 行的概率; $p(x,\,y)$ 为变量 x 与变量 y 联合概率; B 为变量最优推荐值, 为数据量的 0.6 次方. 铁水红外温度与硅含量之间有较强的相关性[30], 是反映铁水硅含量的重要过程变量, 因此选定铁水红外温度作为模型候选输入变量. 各过程变量与硅含量的 MIC 相关性系数, 如表 1 所示.

表 1　过程变量 MIC 相关性系数

Table 1　MIC correlation coefficient of process variables

过程变量	MIC 系数	过程变量	MIC 系数
富氧率	**0.291**	总压差	0.204
透气性指数	**0.270**	**炉腹煤气指数**	**0.278**
标准风速	**0.275**	**热风压力**	**0.268**
富氧流量	0.218	实际风速	0.173
冷风流量	**0.264**	冷风温度	0.209
鼓风动能	0.204	热风温度	0.213
设定喷煤量	**0.241**	顶温下降管	0.209
理论燃烧温度	**0.248**	**铁水红外温度**	**0.291**
顶压	0.195	**顶温**	**0.292**
富氧压力	0.229	鼓风湿度	0.179
冷风压力	0.197	阻力系数	0.204

根据各变量相关性系数分析, 选定富氧率、透气性指数、标准风速等 10 个过程变量作为模型输入. 其中, 铁水红外温度数据依据文献 [30] 方法测得, 其余为现场检测数据. 将样本按照时序分为 D_1、D_2 两个集合, 其中 D_1 包含 2 800 组数据, 用于过程变量的聚类与建模; D_2 包含 200 组数据, 为测试样本. 对于密度峰值聚类算法, 需定义数据样本间的距离, 本文定义过程变量的加权欧氏距离为:

$$d_{ij}=\sqrt{\sum_{k=1}^{10}r_k(x_{ik}-x_{jk})^2}\tag{14}$$

式中, x_{ik} 为第 i 组数据的第 k 项过程变量, x_{jk} 为第 j 组数据的第 k 项过程变量, r_k 为第 k 项过程变量与硅含量的相关性系数.

为求解聚类过程的最优截断距离, 绘制不同截断距离下的邦费罗尼指数曲线, 如图 9 所示. 当邦费罗尼指数取得最大值 0.99931 时, 系统有序程度最高, 最有利于聚类, 此时刻的截断距离为最优截断距离, 绘制此截断距离下的 $\rho_i\delta_i$ 聚类中心决策图, 如图 10 所示, $\rho_i\delta_i$ 越大则越可能成为聚类中心. 为

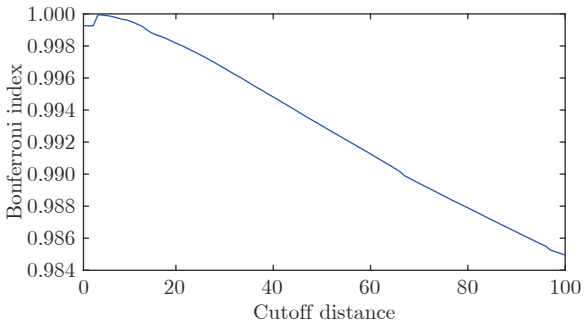

图 9　邦费罗尼指数曲线

Fig. 9　Bonferroni index curve

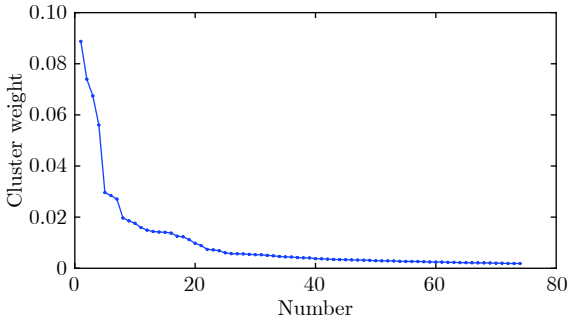

图 10　聚类中心决策图

Fig. 10　Decision diagram of cluster center

进一步选定聚类中心, 再根据聚类中心决策图求解各数据点截断系数, 截断系数数值如表 2 与图 11 所示. 数据点在第 4 点取得最大值 52.50, 选定前 4 个数据点为聚类中心, 即可将数据集 D_1 中数据按照与各聚类中心距离划分为 4 种不同工况聚类簇, 降维后如图 12 所示 (多维过程变量映射在三维空间中位置坐标). 采用 Elman 网络对 4 种不同工况聚类簇数据进行训练, 获得 4 种工况子模型. 对于测试数据集 D_2, 求解样本对不同工况聚类中心隶属度与各子模型预测值, 进而构造工况迁移矩阵, 最后根据最优路径求解算法求解当前时刻最优工况子模型硅含量预测值.

表 2　聚类中心截断标志

Table 2　Cluster center truncation flag

序号	1	2	3	4	5	6
截断系数	3.00	4.02	42.30	52.50	28.02	24.34

4.2　结果分析

为验证本文最优路径算法有效性, 采用 Floyd 路径寻优算法与本文寻优算法进行对比. 在寻优节点数相同的情况下, 本文寻优算法耗时显著低于

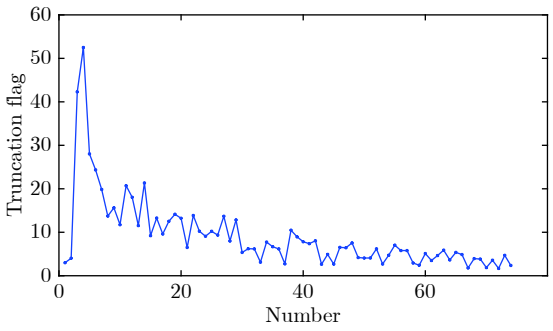

图 11 聚类中心截断系数

Fig. 11 Truncation coefficient of cluster center

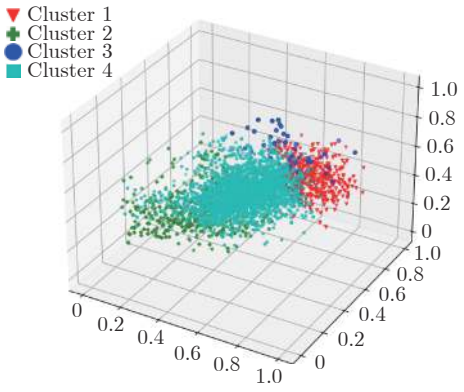

图 12 4 种工况聚类簇

Fig. 12 Clusters of 4 smelting conditions

Floyd 算法的寻优耗时. 根据节点数目不同, Floyd 算法与本文算法运行耗时对比如表 3 所示. 若寻优数据的节点数为 m , 各节点包含 n 个硅含量工况子模型的预测数据. 本文节点寻优情况下, Floyd 算法时间复杂度为 $\mathrm{O}(m^3n^5)$, 本文寻优算法时间复杂度为 $\mathrm{O}(mn^2)$. 在实际预测过程中, m 通常较大, 在 200 左右; n 为工况子模型数, 通常较小. 本文算法在保证寻优结果的同时显著提升了计算效率, 能够满足现场实时性要求.

表 3 寻优算法耗时对比

Table 3 Comparison of the time consumption of optimization algorithms

寻优算法	节点数				
	40	80	120	160	200
Floyd 算法耗时 (ms)	3.20×10^4	2.72×10^5	8.96×10^5	2.09×10^6	4.05×10^6
本文算法耗时(ms)	3	8	11	13	18

为验证本文所提基于最优工况迁移的硅含量预测方法有效性, 选择单一 Elman 网络进行对比分析. 此外, 为进一步验证本文工况分类与硅含量工况寻优模型效果, 与同样采用 Elman 网络作为子模型的集成学习模型 Elman-Adaboost[31]、FEEMD-Adaboost-Elman[32] 进行对比实验. 硅含量数值预测命中率是衡量模型精度的重要指标, 此外, 硅含量变化趋势能够为现场实际操作提供指导, 在实际冶金过程中, 趋势命中率更能反映模型的应用价值[33]. 为全面分析最优工况迁移硅含量预测模型, 通过硅含量数值预测命中率 (预测值与实际值误差绝对值小于等于 0.1 的样本占总测试样本的比例)、硅含量趋势预测准确率 (预测趋势与实际化验趋势相同的样本点占总数的比例)、硅含量预测均方误差三项指标来评价硅含量预测模型的性能. 硅含量数值预测命中率 H 定义如下:

$$H=\frac{1}{N}\left(\sum_{i=1}^{N}h_i\right)\times 100\ \%$$
$$h_i=\begin{cases}1, & e_i\le 0.1\\ 0, & \text{otherwise}\end{cases} \tag{15}$$

式中, N 为测试样本总数; e_i 为预测误差绝对值, 即硅含量化验值 $\hat{y}_i$ 与预测值 y_i 的误差绝对值 $|\hat{y}_i-y_i|$. 当 e_i 的绝对值小于或等于 0.1 时, 认为模型的硅含量预测值是准确可靠的, 此时 h_i 记为 1, 否则为 0. 硅含量趋势预测准确率 T 定义如下:

$$T=\frac{1}{N-1}\left(\sum_{i=1}^{N-1}t_i\right)\times 100\ \%$$
$$t_i=\begin{cases}1, & (\hat{y}_{i+1}-\hat{y}_i)\cdot(y_{i+1}-y_i)>0\\ 0, & \text{otherwise}\end{cases} \tag{16}$$

当两相邻时刻硅含量化验值之差 $(\hat{y}_{i+1}-\hat{y}_i)$ 与相邻时刻硅含量预测值之差 $(y_{i+1}-y_i)$ 的乘积大于 0 时, t_i 等于 1, 否则等于 0. 硅含量预测均方误差 MSE 定义如下:

$$MSE=\frac{1}{N}\sum_{i=1}^{N}e_i^2 \tag{17}$$

目前, 许多学者致力于集成学习建模研究, 以进一步改善单一模型在硅含量预测上难以克服数据波动及冶炼过程的大时滞特性的问题. Elman-Adaboost 模型、FEEMD-Adaboost-Elman 模型选取 Elman 网络作为弱预测器, 结合数据集经验模态分解、Adaboost 算法, 相较于单一 Elman 网络在硅含量预测上有较好的表现. 选取本文模型与上述集成预测模型在现场数据集上进行仿真实验, 测试硅含量数值预测命中率、硅含量趋势预测准确率及硅含量预测均方误差. 其中, Elman 网络的参数设置依

据经验公式推荐值及多次实验结果分析, 设置 Elman 网络输入层、隐藏层、输出层节点数分别为 10、10、1, 最大迭代次数为 300 次, 学习率为 0.01, 其结果如表 4 所示. 本文模型硅含量数值预测命中率能够达到 88 %, 硅含量预测均方误差为 0.0043, 相较于单一 Elman 网络及 Elman 网络集成学习模型有一定的精度优势并且预测结果具有较高的稳定性. 与此同时, 本文模型硅含量趋势预测准确率能够达到 82 %, 相较于对比网络的结果, 本文模型在硅含量趋势预测准确率、趋势的预测精度上有极大的提升. 表明本文模型具有较高的预测精度和较强的泛化能力, 并且具备了在工业现场进行应用的基础.

表 4　模型性能对比

Table 4　Model performance comparison

模型类别	性能指标		
	数值预测命中率 (%)	趋势预测准确率 (%)	预测均方误差
工况迁移预测模型	88	82	0.0043
Elman 网络	79	69	0.0069
Elman-Adaboost	85	71	0.0054
FEEMD-Adaboost-Elman	86	74	0.0049

为进一步对比模型精度以及细节预测效果, 分别绘制本文模型、Elman 网络、Elman-Adaboost 模型、FEEMD-Adaboost-Elman 模型预测曲线与硅含量化验曲线进行对比分析. 本文模型硅含量预测曲线与实际化验曲线对比如图 13 所示. 通过对比发现, 本文所提方法硅含量预测曲线与硅含量实际化验曲线在数值与趋势上吻合较好, 尽管在第40 ~ 80 组与第 160 ~ 200 组数据区间, 硅含量剧烈波动, 本文方法仍能较好地预测硅含量数值与趋势. 本文方法的预测值受到数据波动影响较小, 预测性能较为稳定, 不会跟随过程变量及硅含量数据的波动而剧烈变化. 单一 Elman 网络的硅含量预测值与硅含量化验值的对比曲线如图 14 所示. Elman 网络具有较好的时变适应特性, 但是在硅含量数值波动较大时, 模型的预测精度会显著低于本文方法. 这是由于当硅含量数值剧烈波动时, 炉内工况发生了剧烈变化, 预测模型的结构参数不能良好地适应波动工况, 并且当前时刻的硅含量与冶炼过程的历史数据有紧密联系, 单一模型的输入信息有限, 忽略了大量历史信息, 直接影响了模型的预测精度. 而本文模型对过程参数进行了聚类并划分不同工况类型, 针对不同工况类型的过程变量进行单独建模, 在工况剧烈波动下, 根据历史时刻的硅含量预测值及化验值求解最优的硅含量工况迁移路径, 确保了当前时刻预测值的可靠性, 使得模型仍然具有较高

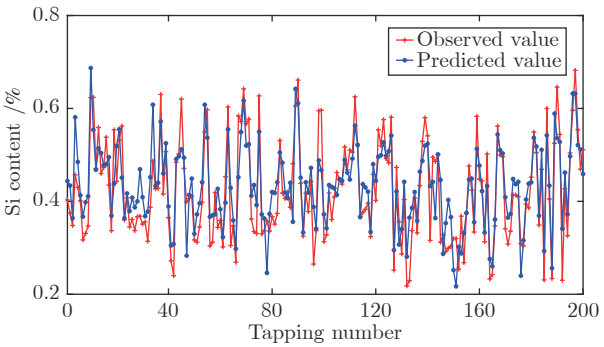

图 13　最优工况迁移模型硅含量预测结果

Fig. 13　Prediction of silicon content in hot metal based on optimal smelting condition migration model

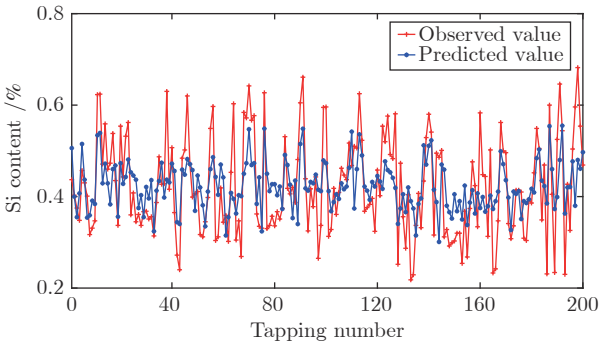

图 14　Elman 网络预测结果

Fig. 14　Prediction of Elman network

预测精度. Elman-Adaboost 建模方法集成了多个 Elman 弱预测器, 以提升模型整体预测精度, 本文通过多次实验设定弱预测器数目为 5 个, 预测曲线与硅含量化验曲线如图 15 所示. 通过曲线分析, Elman-Adaboost 集成模型预测精度相较于单一 Elman 网络有一定提升, 但是在数据波动区间其预测性能提升较小, 预测值显著偏离了化验值. FEEMD-Adaboost-Elman 模型对硅含量化验值时间序列进行 FEEMD 模态分解为 24 维数据, 设定弱预测器数目为 10 个, Elman 网络输入层、隐藏层、输出层节点数分别为 24、16、24, 最大迭代次数为 300 次, 学习率为 0.01, 采用 Elman-Adaboost 算法进行集成建模, 其预测曲线与硅含量化验曲线如图 16 所示. 模型的预测精度相较于单一 Elman 网络有了进一步提升, 但模型的数据处理过程较为复杂, 不利于模型在线更新. 并且由于该算法的输入数据仅包含硅含量化验值时间序列, 没有考虑高炉冶炼过程变量, 使得该模型的预测结果对历史化验值有一定的依赖, 而无法将较短时间内过程变量的波动情况及时反映在硅含量的预测结果上, 使得该模型的趋势预测准确率较低, 在数据频繁波动情况下模型稳定

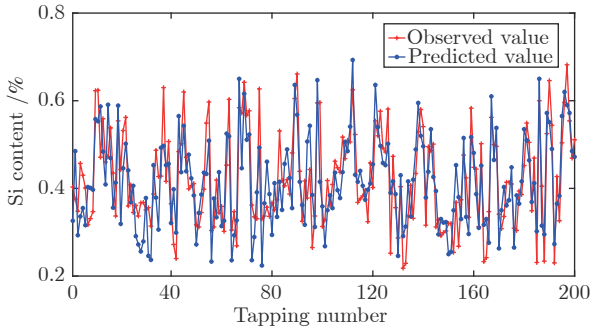

图 15 Elman-Adaboost 网络预测结果

Fig. 15 Prediction of Elman-Adaboost network

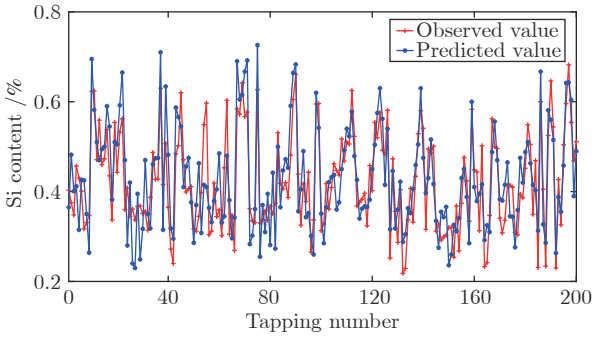

图 16 FEEMD-Adaboost-Elman 网络预测结果

Fig. 16 Prediction of FEEMD-Adaboost-Elman network

性较差. 而本文模型预测值是通过工况迁移代价函数寻优求解得到, 相邻时间节点的间隔仅为 10 s, 此时间间隔内硅含量数值不会发生突变, 从而保证了滑动窗口中硅含量趋势预测的可靠性, 提升了本文模型的数值预测与趋势预测精度. 通过以上对比, 验证了最优工况迁移模型针对工况波动频繁、趋势预测困难等问题的有效性.

为更直观地了解本文所提方法的预测误差分布情况, 根据实验结果绘制基于本文模型的预测误差分布曲线, 如图 17 所示. 本文模型预测误差普遍分布在 [−0.1 ~ 0.1] 范围内, 受过程变量波动影响较小, 整体分布较为平稳, 在过程参数频繁波动情况下有较强的稳定性. 为进一步说明此特性, 分别以本文模型硅含量预测值与实际化验值作为横纵坐标, 绘制铁水硅含量预测值与化验值的散点分布图, 结果如图 18 所示. 从图 18 中可以发现, 散点普遍分布在 $y = x$ 附近, 当实测硅含量数值偏小时, 预测硅含量数值普遍偏大; 当实测硅含量数值偏大时, 预测硅含量数值普遍偏小. 这是由于训练集中做了数据预处理, 剔除了极端异常的数据样本, 预测模型不会输出极端的预测值. 硅含量数据在不同数值区间预测结果偏向不一致, 是因为工况状态发生了改变, 各工况子模型的结构参数不一致, 求解的预

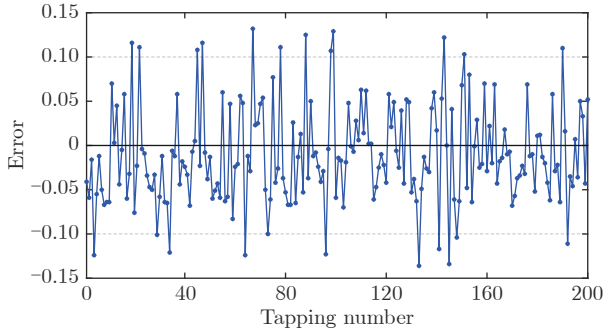

图 17 模型预测误差

Fig. 17 Model prediction error curve

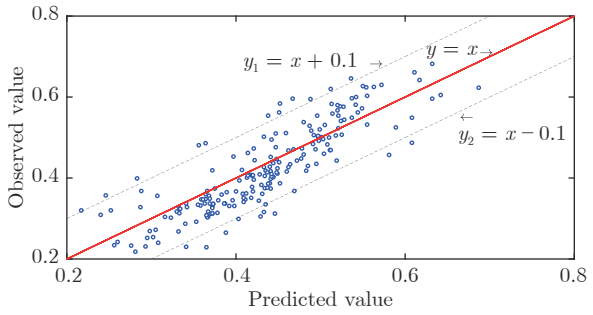

图 18 硅含量预测值与实际值散点图

Fig. 18 The scatter plot of observed and predicted

测值具有不同的偏向. 经过最优路径算法寻优后, 求解出当前时刻最优工况子模型硅含量预测值.

5 结论

高炉铁水硅含量是高炉炼铁过程中的关键技术指标. 针对高炉过程变量频繁波动以及硅含量的大时滞特性, 本文提出一种基于最优工况迁移的高炉铁水硅含量预测方法, 给出过程变量自适应工况聚类、建模方法, 并对最优工况迁移路径进行求解. 基于最优工况迁移的硅含量预测命中率可达 88 %, 硅含量趋势预测准确率可达 82 %. 模型预测精度受过程变量波动影响较小, 趋势预测准确率基本满足现场需求. 基于最优工况迁移矩阵的最优曲线求解使得硅含量预测曲线更为平稳, 符合实际冶炼过程的物理、化学反应特性, 提升了模型的可靠性, 弥补了单一模型难以适应工况频繁变化的缺点. 但是本文的工况聚类及建模求解过程时间复杂度较高, 模型的训练需要较高的时间成本. 有望通过对单个模型的选取、优化以及对工况迁移代价函数进行优化等手段以进一步提升模型的精度, 使得基于最优工况迁移的硅含量预测方法性能进一步提高.

References

1 Biswas A K. *Principles of Blast Furnace Ironmaking —— Theory*

and Practice. Brisbane: Cootha Publishing House, 1981. 1−12

2 Zhou P, Zhang S, Dai P. Recursive learning based bilinear subspace identification for online modeling and predictive control of a complicated industrial process. *IEEE Access*, 2020, **8**: 62531−62541

3 Saxén H, Gao C H, Gao Z W. Data-driven time discrete models for dynamic prediction of the hot metal silicon content in the blast furnace —— A revieww. *IEEE Transactions on Industrial Informatics*, 2012, **9**(4): 2213−2225

4 Onorin O P, Spirin N A, Istomin A S, Lavrov V V, Pavlov A V. Features of blast furnace transient processes. *Metallurgist*, 2017, **61**(1): 121−126

5 Spirin N A, Onorin O P, Istomin A S, Lavrov V V, Gurin I A. Study of transition processes of blast-furnace smelting by the mathematical model method. In: Proceedings of the 2018 IOP Conference Series, Materials Science and Engineering. Novokuznetsk, Russia: Institute of Physics Publishing, 2018. 012−073

6 Spirin N, Onorin O, Alexander I. Prediction of blast furnace thermal state in real-time operation. *Solid State Phenomena*, 2020, **299**: 518−523

7 Spirin N A, Polinov A A, Gurin I A, Beginyuk V A, Pishnograev S N, Istomin A S. Information system for real-time prediction of the silicon content of iron in a blast furnace. *Metallurgist*, 2020, **63**(9): 898−905

8 Östermark R, Saxen H. VARMAX-modelling of blast furnace process variables. *European Journal of Operational Research*, 1996, **90**(1): 85−101

9 Saxen H, Östermark R. State realization with exogenous variables —— A test on blast furnace data. *European Journal of Operational Research*, 1996, **89**(1): 34−52

10 Bhattacharya T. Prediction of silicon content in blast furnace hot metal using partial least squares. *ISIJ International*, 2005, **45**(12): 1943−1945

11 Li J P, Hua C C, Yang Y N, Guan X P. Bayesian block structure sparse based T-S fuzzy modeling for dynamic prediction of hot metal silicon content in the blast furnace. *IEEE Transactions on Industrial Electronics*, 2017, **65**(6): 4933−4942

12 Xu X, Hua C C, Tang Y G, Guan X P. Modeling of the hot metal silicon content in blast furnace using support vector machine optimized by an improved particle swarm optimizer. *Neural Computing and Applications*, 2016, **27**(6): 1451−1461

13 Han Y, Li J, Yang X L, Liu W X, Zhang Y Z. Dynamic prediction research of silicon content in hot metal driven by big data in blast furnace smelting process under hadoop cloud platform. *Complexity*, DOI: 10.1155/2018/8079697

14 Xu X, Hua C C, Tang Y G, Guan X P. Wiener model identification of blast furnace ironmaking process based on laguerre filter and linear programming support vector regression. In: Proceedings of the 2014 International Joint Conference on Neural Networks. Beijing, China: IEEE Press, 2014. 2198−2204

15 Zhou P, Li W P, Wang H, Li M J, Chai T Y. Robust online sequential rvflns for data modeling of dynamic time-varying systems with application of an ironmaking blast furnace. *IEEE Transactions on Cybernetics*, 2019, **50**(11): 4783−4795

16 Gao Chuan-Hou, Jian Ling, Chen Ji-Ming, Sun You-Xian. Data-driven modeling and prediction algorithm for complex blast furnace ironmaking process. *Acta Automatica Sinica*, 2009, **35**(6): 725−730
(郜传厚, 渐令, 陈积明, 孙优贤. 复杂高炉炼铁过程的数据驱动建模及预测算法. 自动化学报, 2009, **35**(6): 725−730)

17 David S F, David F F, Machado M L P. Artificial neural network model for predict of silicon content in hot metal blast furnace. *Materials Science Forum*, 2016, **869**: 572−577

18 Song Jing-Hua, Yang Chun-Jie, Zhou Zhe, Liu Wen-Hui, Ma Shu-Yan. Application of improved EMD-Elman neural network in prediction of silicon content in molten iron. *CIESC Journal*, 2016, **67**(3): 729−735
(宋菁华, 杨春节, 周哲, 刘文辉, 马淑艳. 改进型 EMD-Elman 神经网络在铁水硅含量预测中的应用. 化工学报, 2016, **67**(3): 729−735)

19 Jiang K, Jiang Z H, Xie Y F, Chen Z P, Pan D, Gui W H. Classification of silicon content variation trend based on fusion of multilevel features in blast furnace ironmaking. *Information Sciences*, 2020, **521**: 32−45

20 Zhou Ping, Zhang Li, Li Wen-Peng, Dai Peng, Chai Tian-You. Modeling of blast furnace multi-element molten iron quality with random weight neural network based on self-encoding and PCA. *Acta Automatica Sinica*, 2018, **44**(10): 1799−1811
(周平, 张丽, 李温鹏, 戴鹏, 柴天佑. 集成自编码与 PCA 的高炉多元铁水质量随机权神经网络建模. 自动化学报, 2018, **44**(10): 1799−1811)

21 Jian L, Song Y Q, Shen S Q, Wang Y, Yin H Q. Adaptive least squares support vector machine predictor for blast furnace ironmaking process. *ISIJ International*, 2015, **55**(4): 845−850

22 Zeng J S, Liu X G, Gao C H, Luo S H, Jian L. Wiener model identification of blast furnace ironmaking process. *ISIJ International*, 2008, **48**(12): 1734−1738

23 Jiang Zhao-Hui, Dong Meng-Lin, Gui Wei-Hua, Yang Chun-Hua, Xie Yong-Fang. Two-dimensional prediction for silicon content of hot metal of blast furnace based on Bootstrap. *Acta Automatica Sinica*, 2016, **42**(5): 715−723
(蒋朝辉, 董梦林, 桂卫华, 阳春华, 谢永芳. 基于 Bootstrap 的高炉铁水硅含量二维预报. 自动化学报, 2016, **42**(5): 715−723)

24 Li Wen-Peng, Zhou Ping. Blast furnace hot metal quality robust regularization random weight neural network modeling. *Acta Automatica Sinica*, 2020, **46**(4): 721−733
(李温鹏, 周平. 高炉铁水质量鲁棒正则化随机权神经网络建模. 自动化学报, 2020, **46**(4): 721−733)

25 Wen Liang, Zhou Ping. Model free adaptive control of molten iron quality based on multi-parameter sensitivity analysis and GA optimization. *Acta Automatica Sinica*, 2021, **47**(11): 2600−2613
(温亮, 周平. 基于多参数灵敏度分析与遗传优化的铁水质量无模型自适应控制. 自动化学报, 2021, **47**(11): 2600−2613)

26 Rodriguez A, Laio A. Clustering by fast search and find of density peaks. *Science*, 2014, **344**(6191): 1492−1496

27 Bárcena-Martin E, Silber J. The Bonferroni index and the measurement of distributional change. *Metron*, 2017, **75**(1): 1−16

28 Sun Tian, Ling Wei-Xin. Application of Levenberg-Marquardt algorithm based on simulated annealing in neural network. *Science Technology and Engineering*, 2008(18): 5189−5192
(孙甜, 凌卫新. 基于模拟退火的 Levenberg-Marquardt 算法在神经网络中的应用. 科学技术与工程, 2008(18): 5189−5192)

29 Reshef D N, Reshef Y A, Finucane H K, Grossman S R, Mcvean

G, Turnbaugh P J, et al. Detecting novel associations in large data sets. *Science*, 2011, **334**(6062): 1518−1524

30 Pan D, Jiang Z H, Chen Z P, Gui W H, Xie Y F, Yang C H. Temperature measurement and compensation method of blast furnace molten iron based on infrared computer vision. *IEEE Transactions on Instrumentation and Measurement*, 2018, **68**(10): 3576−3588

31 Zhuang Tian, Yang Chun-Jie. Prediction method of silicon content in molten iron based on Elman-Adaboost strong predictor. *Metallurgical Automation*, 2017, **41**(04): 1−6, 17
(庄田, 杨春节. 基于 Elman-Adaboost 强预测器的铁水硅含量预测方法. 冶金自动化, 2017, **41**(04): 1−6, 17)

32 Wang Kai, Bi Gui-Hong, Gao Han, Pu Xian-Yi, Chen Shi-Long. Short-term wind speed prediction method based on improved fast ensemble empirical mode decomposition and Elman-Adaboost. *Electric Power Science and Engineering*, 2020, **36**(05): 32−39
(王凯, 毕贵红, 高晗, 蒲娴怡, 陈仕龙. 基于改进快速集合经验模态分解和 Elman-Adaboost 的短期风速预测方法. 电力科学与工程, 2020, **36**(05): 32−39)

33 Jiang Ke, Jiang Zhao-Hui, Xie Yong-Fang, Pan Dong, Gui Wei-Hua. Intelligent prediction of silicon content change trend in molten iron of large blast furnace. *Control Engineering*, 2020, **27**(03): 540−546
(蒋珂, 蒋朝辉, 谢永芳, 潘冬, 桂卫华. 大型高炉铁水硅含量变化趋势的智能预报. 控制工程, 2020, **27**(03): 540−546)

蒋朝辉 中南大学自动化学院教授, 鹏城实验室研究员. 2011 年获得中南大学博士学位. 主要研究方向为光电信息感知, 图像处理, 人工智能, 工业 VR 和智能优化控制.
E-mail: jzh0903@csu.edu.cn
(**JIANG Zhao-Hui** Professor at the School of Automation, Central South University. Professor at the Peng Cheng Laboratory. He received his Ph. D. degree from Central South University in 2011. His research interest covers photoelectric information perception, image processing, artificial intelligence, industrial VR, and intelligent optimization control.)

许 川 中南大学自动化学院博士研究生. 主要研究方向为复杂工业过程建模, 数据分析和机器学习. 本文通信作者.
E-mail: csuxuchuan@csu.edu.cn
(**XU Chuan** Ph. D. candidate at the School of Automation, Central South University. His research interest covers complex industrial process modeling, data analysis, and machine learning. Corresponding author of this paper.)

桂卫华 中国工程院院士, 中南大学自动化学院教授, 鹏城实验室研究员. 1981 年获得中南矿冶学院硕士学位. 主要研究方向为复杂工业过程建模与最优控制, 分布式鲁棒控制和故障诊断. E-mail: gwh@csu.edu.cn
(**GUI Wei-Hua** Academician of Chinese Academy of Engineering, and professor at the School of Automation, Central South University. Professor at the Peng Cheng Laboratory. He received his master degree from Central South Institute of Mining and Metallurgy in 1981. His research interest covers modeling and optimal control of complex industrial process, distributed robust control, and fault diagnoses.)

蒋 珂 中南大学自动化学院博士研究生. 2019 年获得中南大学硕士学位. 主要研究方向为数据驱动的工业过程建模与控制, 过程数据分析和机器学习.
E-mail: jiangke@csu.edu.cn
(**JINAG Ke** Ph. D. candidate at the School of Automation, Central South University. She received her master degree from Central South University in 2019. Her research interest covers data-based modeling and control of industrial process, process data analysis, and machine learning.)