

控制与决策

Control and Decision

基于WKLS-C-LWKL相似性度量策略的转炉炼钢终点碳温软测量方法

杨路, 刘辉, 熊倩

引用本文:

杨路, 刘辉, 熊倩. 基于WKLS-C-LWKL相似性度量策略的转炉炼钢终点碳温软测量方法[J]. *控制与决策*, 2022, 37(11): 2869–2879.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.0744>

您可能感兴趣的其他文章

Articles you may be interested in

[考虑边界样本邻域归属信息的粗糙K-means增量聚类算法](#)

Rough K-means incremental clustering algorithm considering neighborhood belonging information of boundary samples
控制与决策. 2022, 37(11): 2968–2976 <https://doi.org/10.13195/j.kzyjc.2021.0624>

[基于WGRA-FCM样本相似性度量的转炉炼钢终点碳温软测量方法](#)

End point carbon temperature measurement method based on WGRA-FCM for sample similarity measurement
控制与决策. 2021, 36(9): 2170–2178 <https://doi.org/10.13195/j.kzyjc.2020.0128>

[基于相互邻近度的密度峰值聚类算法](#)

Density peaks clustering based on mutual neighbor degree
控制与决策. 2021, 36(3): 543–552 <https://doi.org/10.13195/j.kzyjc.2019.0795>

[基于相异性度量选取初始聚类中心改进的K-means聚类算法](#)

Improved K-means clustering algorithm for selecting initial clustering centers based on dissimilarity measure
控制与决策. 2021, 36(12): 3083–3090 <https://doi.org/10.13195/j.kzyjc.2020.0554>

[基于边缘峰度度量的特征缩减模糊聚类算法](#)

Feature-reduction fuzzy clustering algorithm based on marginal kurtosis measure
控制与决策. 2021, 36(11): 2665–2673 <https://doi.org/10.13195/j.kzyjc.2020.0220>

基于 WKLSC-LWKL 相似性度量策略的 转炉炼钢终点碳温软测量方法

杨路, 刘辉[†], 熊倩

(1. 昆明理工大学信息工程与自动化学院, 昆明 650500;
2. 昆明理工大学云南省人工智能重点实验室, 昆明 650500)

摘要: 转炉炼钢终点控制的关键是实现碳温准确预测. 针对炉次样本间波动性较大, 导致即时学习的样本相似性度量困难而造成预测精度不高的问题, 提出一种基于改进谱聚类算法构建的相似性度量策略. 首先, 根据过程变量和关键变量间的耦合关系构造全局加权 KL 度量准则的谱聚类算法, 获得类间方差较大、类内方差较小的聚类子集以消除炉次样本间的波动; 其次, 根据类簇间差异信息, 融入局部加权 KL 度量准则计算待测样本隶属于各类的后验概率, 从而构造出一种适合描述转炉炼钢过程复杂特性的相似性度量策略; 最后, 采用该度量策略度量出与待测炉次工况特性更加相似的样本子集, 建立相关向量机回归模型进行终点碳温预测. 实际转炉炼钢生产过程数据仿真结果表明, 碳含量在 $\pm 0.02\%$ 的预测误差范围内精度达到 89%, 温度在 $\pm 10^\circ\text{C}$ 的误差范围内精度达到 92%.

关键词: 转炉炼钢; 即时学习; 加权 KL 度量准则; 谱聚类; 后验概率

中图分类号: TP273 文献标志码: A

DOI: 10.13195/j.kzyjc.2021.0744

引用格式: 杨路, 刘辉, 熊倩. 基于 WKLSC-LWKL 相似性度量策略的转炉炼钢终点碳温软测量方法[J]. 控制与决策, 2022, 37(11): 2869-2879.

Soft sensor method of endpoint carbon content and temperature based on WKLSC-LWKL similarity measurement strategy

YANG Lu, LIU Hui[†], XIONG Qian

(1. Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China; 2. Yunnan Key Laboratory of Artificial Intelligence, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: Accurate prediction of carbon content and temperature is the crucial to the endpoint control of converter steelmaking. For the large sample fluctuation, it is difficult to measure the similarity of samples in just-in time learning (JITL), which leads to the problem of low prediction accuracy, therefore, this paper proposes a similarity measurement strategy based on an improved spectral clustering algorithm. Firstly, according to the coupling relationship between process variables and dominant variables, a spectral clustering algorithm with global weighted KL measurement standards is constructed, thus, the clustering subsets with large between clusters variance and small intra-cluster variance are obtained to eliminate the fluctuation among the furnace samples. Secondly, according to the difference information between class clusters, the local weighted KL metric criterion is integrated to calculate the posterior probability of the predicted samples belonging to various clusters, then, a similarity measurement strategy suitable for describing the complex characteristics of converter steelmaking process is constructed. Finally, this measurement strategy is used to calculate a subset of samples that are more similar to the properties of the new furnace, and the RVM model is established to predict the end point carbon content and temperature. The simulation results of actual converter steelmaking process show that the prediction accuracy of carbon content within $\pm 0.02\%$ error range reaches 89%, temperature within $\pm 10^\circ\text{C}$ error range reaches 92%.

Keywords: converter steelmaking; JITL; weighted KL measurement standards; spectral cluster; posteriori probability

收稿日期: 2021-04-28; 录用日期: 2021-08-09.

基金项目: 国家自然科学基金项目(61863018); 云南省科技厅面上项目(202001AT070038).

责任编辑: 毛志忠.

[†]通讯作者. E-mail: liuhui621@126.com.

0 引言

钢铁产业对国家工业化进程的提高发挥着不可替代的作用,转炉炼钢作为炼钢的主要技术手段,通过钢铁废料、铁合金等金属材料和其他非金属材料,在熔池中进行大量复杂的物化反应产生热量以完成炼钢过程,其中对熔池内钢液碳含量和温度的准确控制是转炉炼钢生产终点控制的关键步骤。因此,实现终点碳温的准确预测,对提高转炉炼钢生产的钢铁质量和生产效率、最大限度降低生产成本、提升经济效益及节能减排具有重要的意义^[1]。

目前,转炉终点碳温测量方法主要有接触式测量和非接触式测量。接触式测量主要有副枪检测法和炉气分析法。副枪检测法是采用副枪设备直接检测熔池内的碳温^[2],该方法由于设备价格昂贵且长期工作在高温腐蚀环境下,使用成本较高,此外,由于副枪在测量过程中只是间断式测量,不能连续测量碳温;炉气分析法使用专用的测量仪器对管道中的炉气进行测量分析^[3],而炉气测量仪器的安装位置使得测量存在一定的时间滞后。随着信息技术的进步和对炼钢质量要求的提高,基于光谱分析和火焰图像处理这类非接触式测量得以发展^[4]。文献[5]采用模糊支持向量机识别炉口光辐射状态进行炼钢终点判断,但实际炼钢环境恶劣,干扰因素较多,使得采集的光谱图片有一定干扰,使用该方法常会产生较大误差。文献[6]通过提取炉口火焰图像多特征建立预测模型取得较好效果,但炼钢过程的复杂性使得炉口火焰在不同时刻的形态和特征都有所不同,导致该方法预测精度不理想。

近年来,工业过程数据测量较困难,常使用软测量方法进行实时预报。文献[7-8]通过将深度学习引入软测量方法,提高了工业加氢裂化过程中航空煤油沸点的预测性能。在转炉炼钢过程中,原料的加入量、吹氧量及对应的吹氧时间等生产过程数据与炼钢终点的碳含量和温度有密切关系^[9],这使得基于转炉炼钢生产过程数据的软测量方法得以发展。首先,研究者基于转炉运行的相关原理和软测量相关技术建立了转炉终点预测模型,文献[10]建立径向基神经网络模型实现了对转炉炼钢终点碳温的预测,克服了传统检测方法只能根据边界条件来确定相关参数的不足;其次,考虑工艺复杂性,过程变量与转炉终点碳温之间存在大量的冗余信息导致模型性能下降,文献[11]提出了改进的遗传算法对转炉炼钢过程数据进行特征选择,去除生产过程中冗余变量,得到与终点碳温最相关的特征。但在实际生产过程中,受检测传感器

老化、人工操作失误以及原材料品质差异等影响,采集到的过程数据存在严重的非线性及数据波动性大的特点^[12],这些因素使得之前构建的软测量模型可能不匹配当前实际工况,导致预测值与实际终点碳温产生较大偏差。

考虑到采用全局样本建模无法保证模型能够自适应复杂工况变化,研究者们常采用即时学习(just-in-time learning, JITL)的方法^[13]对复杂工况数据建模以解决上述问题。JITL是在历史数据中通过选择与待预测样本相似的样本来构建模型,而JITL建模方法的核心是相似性度量策略的使用^[14-15]。例如,文献[16]将历史样本经过高斯混合模型训练后加权至马氏距离进行相似性度量,用于解决非高斯工业过程的软测量;文献[17]为了获得更合理的子样本训练集,充分考虑输入输出间的相关性,采用互信息加权相似性度量的方法;文献[15]在一阶相似度基础上构建的二阶相似度策略,为待测样本选择了更近邻的相似样本,改进了传统一次度量的距离缺陷。由于实际转炉炼钢是个复杂工业生产过程,过程数据分布差异较大,多数相似性度量只能计算两个点的变量值之间的相似度,若不考虑样本当前工况特性可能会选择不合适的样本进行局部建模。

基于上述分析,将JITL方法引入终点碳温的预测,通过相似性度量策略从历史样本中找出与待测样本最相似的样本子集进行建模预测。然而,为了在JITL方法中获得相似性更高、类内方差更小的相似样本集,本文从考虑待预测样本的当前工况特性出发,提出一种基于全局加权KL度量准则的谱聚类(weighted KL spectral clustering, WKLSC)算法,将历史训练样本聚类后采用局部加权KL度量准则(Local Weighted KL, LWKL)计算后验概率,以构建新的局部相似性度量策略,用于面向转炉炼钢过程数据即时学习的相似样本选择。通过终点碳温预测的仿真实验及对比分析,结果表明采用本文提出的度量策略度量的最佳训练样本集更符合当前待测炉次样本特性,从而有效解决了炼钢复杂过程数据波动性大导致碳温预测精度不高的问题。

本文创新点体现在以下几个方面:1)提出适合描述炼钢过程数据复杂特性的WKLSC-LWKL度量策略;2)所提出的加权KL度量准则既实现了历史样本的聚类,又实现了待预测样本后验概率的计算;3)为了使待测样本隶属类别计算更加合理,先通过加权了样本总体特性的聚类算法将历史样本聚类获得合理质心后,通过加权类簇间的差异信息计算待测样本

的后验概率。

1 WKLSC-LWKL相似性度量

在实际转炉炼钢生产中,由于过程数据测量目标不同,对原始数据集进行标准化处理从而消除数据间的量纲影响,预处理后的数据集为 $D' = \{(x_i, y_i) \in \mathbf{R}^{m+1}; i = 1, 2, \dots, n\}$,其中 $x_i \in \mathbf{R}^m$ 是炼钢过程数据的输入特征, $y_i \in \mathbf{R}$ 是终点碳含量或者温度标签, n 是样本总数, m 是输入特征维数。

本文提出的相似性度量策略如下:首先,为了将波动性较大的历史炉次样本聚类成类间方差较大、类内方差较小的样本子集,采用基于全局加权KL度量策略的谱聚类算法实现;其次,为了计算新来测试样本隶属于每个类的概率,使得后续度量出的相似样本集类内方差更小,采用基于局部加权KL度量准则的后验概率算法实现;最后,建立适合描述转炉炼钢过程复杂特性的WKLSC-LWKL度量策略,度量出与当前炉次工况特性更加相似的样本子集用于即时学习建模。其度量策略思路流程如图1所示。

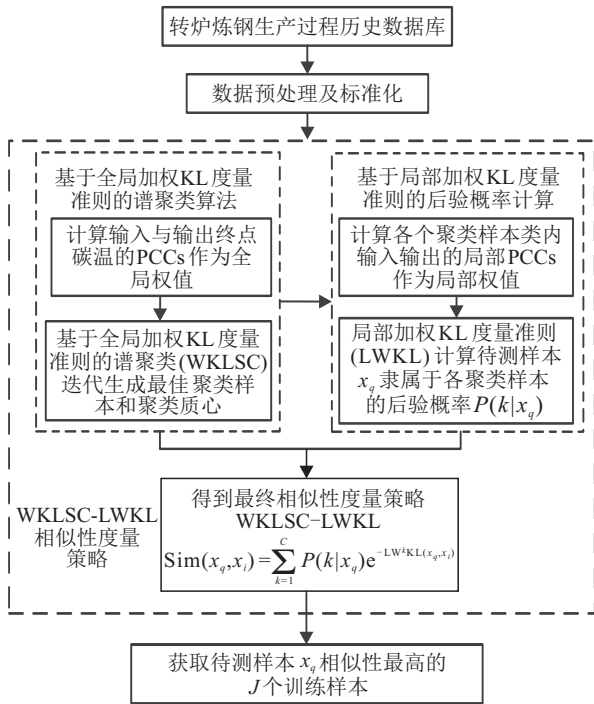


图1 本文相似性度量策略思路流程

1.1 基于全局加权KL度量准则的谱聚类算法

谱聚类(spectral clustering, SC)算法^[18]是一种基于图论的聚类算法,将数据作为空间内的点集 \mathbf{V} ,相似性矩阵 \mathbf{S} 作为点与点之间连接边,对所有数据点组成的图模型 $G = (\mathbf{V}, \mathbf{S})$ 进行切分达到聚类的目的。由于SC的结果直接依赖于相似性度量,在已有的谱聚类算法中,常采用欧氏距离核函数作为相似性度量函数,但实验结果表明该度量方法在表达转炉炼

钢过程数据的空间特性时存在缺陷,使得聚类效果不佳,进而影响后续炉次待预测样本的度量。

已有文献证明,将非线性相似性度量代替传统的欧氏距离获得了类内样本更相似的度量结果,非线性度量中的KL距离(Kullback-Leibler divergence)常用于度量两个分布之间的相似度,也被研究者们用于度量样本间的相似度^[19-20],其中两个分布 $p_1(x)$ 和 $p_2(x)$ 之间的KL距离定义为

$$D(p_1(x)||p_2(x)) = \int p_1(x) \log \frac{p_1(x)}{p_2(x)} dx. \quad (1)$$

从KL距离出发,考虑到炼钢过程数据非线性变化的特性,在其基础上提出加权KL度量准则的相似性度量概念,既保留了KL距离用于度量的稳定性,又利用加权方式实现了对输入特征和输出碳温的综合评价,将该炉次样本度量准则与谱聚类算法进行结合,利用非线性度量代替传统线性度量,解决了传统度量方法只对数据在特征空间呈球形或椭球形分布度量有效、对复杂分布数据度量失效的情况。设两个炉次样本 $x_i, x_j \in D'$ 的全局加权KL度量准则(WKL)的矢量定义如下:

$$\begin{cases} D(x_i, x_j) = e^{-WKL(x_i, x_j)}; \\ WKL(x_i, x_j) = \sum_{h=1}^m w_{gh} \cdot \left| x_{ih} \ln \frac{x_{ih}}{x_{jh}} + x_{jh} \ln \frac{x_{jh}}{x_{ih}} \right|. \end{cases} \quad (2)$$

其中: x_{ih} 为样本 x_i 的第 h 维特征, w_{gh} 为全局加权系数 w_g 的第 h 维权值。

首先,确定KL度量准则的 w_g ,若使得与输出碳温较为相关的特征在计算相似度时获得权重较大,则计算所得相似样本在输出碳温上的相关性也较高;其次,通过式(2)计算 x_i, x_j 两样本的相似度, $D(x_i, x_j)$ 为两样本的全局加权KL度量准则。

定义1 基于皮尔逊相关系数(Pearson correlation coefficient, PCCs)确定的全局权值定义如下。

1) 计算历史数据库的输入特征和输出碳温间的PCCs,得到炉次样本第 h 维特征 x_h 与输出 y 之间的相关系数

$$\rho(x_h, y) = \frac{\text{cov}(x_h, y)}{\sigma_{x_h} \sigma_y}, h \in (1, m). \quad (3)$$

其中: $\text{cov}(x_h, y)$ 为 x_h 与 y 之间的协方差, σ_* 为计算该变量的标准差。

2) 由下式计算相关比例系数:

$$w_{gh} = \left| \frac{\rho(x_h, y)}{\sum_{h=1}^m \rho(x_h, y)} \right|, h \in (1, m), \quad (4)$$

则 $w_g = (w_{g1}, w_{g2}, \dots, w_{gm})$ 为输入特征的全局权重。

基于上述度量准则的谱聚类算法能够较好地反映转炉炼钢过程特性,从而能合理地将炉次划分为不同阶段类别. WKLSC算法步骤如下.

step 1: 输入数据集 D' , 依据式(2)得出两两样本间的WKL值, 依据下式计算其相似度矩阵 S , 构造图模型 $G = (D', S)$:

$$S_{ij} = \begin{cases} e^{-\frac{\text{WKL}(x_i, x_j)}{\sigma_s^2}}, & i \neq j, i, j \in (1, n); \\ 0, & i = j. \end{cases} \quad (5)$$

其中 σ_s^2 表示两两样本间WKL值的方差.

step 2: 根据炉次样本相似度矩阵 S 利用 Ncut^[18] 算法对图 G 进行切图.

step 3: 得到聚类结果 $X = \{X_k\}_{k=1}^C$ 和聚类质心 $\{x_k^i\}_{i=1}^C$, 其中 C 表示聚类数目, X_k 表示第 k 个类的样本集, 聚类质心的计算如下所示, $N(X_k)$ 为第 k 个类下的样本数目:

$$x_k^i = \frac{\sum_{i=1}^{N(X_k)} x_i}{N(X_k)}, \quad x_i \in X_k. \quad (6)$$

一般谱聚类算法的聚类数目需要人为确定, 而不同聚类数目的聚类结果也不相同. 对炼钢过程数据来说, 其聚类效果越好, 最终度量出的相似样本集越合理. 通过聚类有效性指标中的SD有效性指数^[21], 本文基于WKL准则定义了衡量类内同一性和类间分离性的评价函数, 以此确定最佳聚类数目.

定义2 确定最佳聚类数目的评价函数定义为:

1) 以计算类内样本与该类质心的WKL值的方差作为衡量类内方差的评价指标, 方差越小说明类簇内相似度越高, 此为类内同一性评价, 计算如下:

$$\text{IC} = \frac{1}{C} \sum_{k=1}^C \sqrt{\frac{1}{N(X_k)} \sum_{x_i \in X_k} \text{WKL}^2(x_i, x_k^i)}. \quad (7)$$

2) 以计算各类质心间WKL相似度的均值作为衡量类间方差的评价指标, 此为类间分离性的评价, 类间分离性与WKL成反比, 计算方式如下:

$$\text{BC} = \frac{1}{C(C-1)} \sum_{i=1}^C \sum_{j=1, j \neq i}^C e^{-\text{WKL}^2(x_k^i, x_k^j)}. \quad (8)$$

根据式(8)计算, BC值越小, 说明类间方差越大.

3) 为综合评价聚类效果, 结合式(7)和(8)定义一种确定最佳聚类数目的评价函数 $E(C)$, 目的是最大化类内相似性和最小化类间相似性, 定义如下:

$$E(C) = \xi \text{IC} + (1 - \xi) \text{BC}. \quad (9)$$

其中: $\xi = 0.5$ 时表示两种评价占比相等, 不断更新聚类数 C , 使得 $E(C)$ 达到最小值, 说明类间方差相对最小、类内方差相对最大, 此时的 C 为最佳聚类数.

利用本节提出的WKLSC聚类算法能够更好地对转炉炼钢复杂过程数据进行聚类, 由于全局权值的引入提高了类内相似样本在输出碳温上的相关性, 克服了传统谱聚类应用于转炉炼钢数据时只能获得在空间距离上相似的聚类样本的局限性.

1.2 基于局部加权KL度量准则的后验概率计算

由于历史样本经过聚类后, 不同类簇间与终点碳温相关性高的特征变量不同, 在计算待预测样本与各个聚类样本的相似性以确定隶属度时, 需要考虑类内样本的局部信息. 因此, 本节将每个类下的输入与输出的PCCs作为局部权值. 针对待预测样本分类的问题, 考虑获得的质心样本在一定程度上能够代表该类别的样本, 以式(2)定义的加权KL度量准则为基础计算待预测样本和聚类质心的局部加权KL度量准则(LWKL), 以确定待预测样本隶属于各类簇的后验概率. 其中, LWKL算法步骤如下.

step 1: 考虑到类簇间样本的不同输入特征对输出贡献度可能不一样, 首先, 确定KL度量准则的局部加权系数 $\{w_l^k\}_{k=1}^C$, 见如下定义.

定义3 基于PCCs确定的KL度量准则的类内局部权值定义如下:

1) 由式(3)计算各个类内样本的输入特征和输出碳温间的相关系数 ρ , 得到每个类下的样本第 h 维特征与输出之间的相关系数 $\{\rho_k(x_h, y)\}_{k=1}^C$;

2) 将 $\{\rho_k\}_{k=1}^C$ 经式(4)计算得到类内局部加权系数 $w_l^k = (w_{l1}^k, w_{l2}^k, \dots, w_{lm}^k), k \in (1, C)$.

step 2: 在KL度量准则的加权方式上采取局部信息加权, 得到测试样本 x_q 在第 k 个聚类样本下的LWKL局部相似度

$$\text{LW}^k \text{KL}(x_q, x_i) = \sum_{h=1}^m w_{lh}^k \cdot \left| x_{qh} \ln \frac{x_{qh}}{x_{ih}} + x_{ih} \ln \frac{x_{ih}}{x_{qh}} \right|. \quad (10)$$

其中: $x_i \in X_k, w_{lh}^k$ 为第 k 个类内局部加权系数 w_l^k 的第 h 维权值.

按此度量准则计算 x_q 和聚类质心 x_k 的LWKL, 按下式得到似然函数:

$$P(x_q|k) = \frac{\exp[-\text{LW}^k \text{KL}(x_q, x_k)]}{\sum_{k=1}^C \exp[-\text{LW}^k \text{KL}(x_q, x_k)]}. \quad (11)$$

按照如下公式计算 x_q 隶属于各类的后验概率:

$$P(k|x_q) = \frac{P(x_q|k)P(k)}{\sum_k P(x_q|k)P(k)}, \quad (12)$$

$$P(k) = \frac{n(X_k)}{n(X)}. \quad (13)$$

其中: $P(k)$ 为第 k 类的先验概率, $n(x_k)$ 和 $n(\mathbf{X})$ 分别表示 k 类和整个训练集中的样本个数.

1.3 基于WKLSC-LWKL相似性度量策略

JITL的关键是相似性度量策略的选取, 常常采用欧氏距离和余弦相似度结合作为其度量策略^[13], 但该度量策略只考虑样本间的距离信息, 难以用于复杂过程在线建模^[16]. 也就是说, 在转炉炼钢复杂工况下该方法度量出的局部历史样本与当前炉次特性差别较大, 从而导致终点碳温预测精度不高.

为了考虑新来样本当前工况特性, 度量出与其类内方差更小的历史样本集, 根据 1.1 和 1.2 节, 经过 WKLSC 对历史炼钢过程数据进行有效划分, 依据聚类结果计算新来样本隶属各类簇的后验概率, 将其加权在 LWKL 局部相似度上, 新构建的相似度准则可以表示为

$$\text{Sim}(x_q, x_i) = \sum_{k=1}^C P(k|x_q) e^{-LW^k \text{KL}(x_q, x_i)}. \quad (14)$$

其中: $P(k|x_q)$ 表示待预测样本 x_q 属于第 k 个聚类样本的后验概率, 依据式 (14), $P(k|x_q)$ 越大, 说明 x_q 属于第 k 个类簇的权重越大; Sim 越大, 两样本越相似.

算法1 WKLSC-LWKL度量策略.

输入: 历史数据库 D' , 新来待预测样本 x_q ;

输出: 与测试样本最相似的训练子集.

- 1) 初始化聚类数目 C ;
- 2) for $C = 2$ to 8 do
- 3) for $i = 1$ to n do
- 4) 利用 WKLSC 将历史样本进行聚类;
- 5) 得到聚类结果 $\{\mathbf{X}_k\}_{k=1}^C$ and $\{x_k^i\}_{i=1}^C$;
- 6) 计算评价函数 $E(C)$;
- 7) end for
- 8) end for
- 9) $C \leftarrow \text{find min } E(C)$
- 10) for $k = 1$ to C do
- 11) 计算类内局部加权系数 $\{w_l^k\}_{l=1}^C$
- 12) end for
- 13) for $k = 1$ to C do
- 14) 计算 x_q 与各聚类样本隶属度 $P(k|x_q)$;
- 15) for $i = 1$ to n do
- 16) 计算 x_q 与历史聚类样本的 LWKL 相似度;
- 17) end for
- 18) $\text{Sim} \leftarrow \sum P(k|x_q) \exp[-LW^k \text{KL}(x_q, x_i)]$
- 19) end for
- 20) 将 Sim 值从大到小排列, 取前 J 个样本作为

最佳训练样本子集.

2 面向转炉炼钢生产过程数据的WKLSC-LWKL即时学习终点碳温软测量模型

针对全局样本建模无法自适应当前炉次工况进行有效预测, 且在转炉炼钢复杂工况下数据波动性大导致传统 JITL 度量失效的问题, 构建面向转炉炼钢生产过程数据的即时学习软测量模型. 在离线建模部分, 利用 WKLSC 算法对历史数据库进行聚类, 将数据波动性大的样本库聚类成为类间方差更大、类内方差更小的样本子集, 基于该聚类结果得到聚类样本类内的局部权重; 在线预测部分, 通过计算待测样本隶属于各类簇的后验概率及 LWKL 局部相似度, 根据式 (14) 度量出最相似的训练样本集, 建立相关向量机 (relevance vector machine, RVM) 终点碳温回归模型, 输出碳温预测结果; 当下一个样本到来后, 为其重新获取训练样本集进行实时更新. 建模流程如图 2 所示.

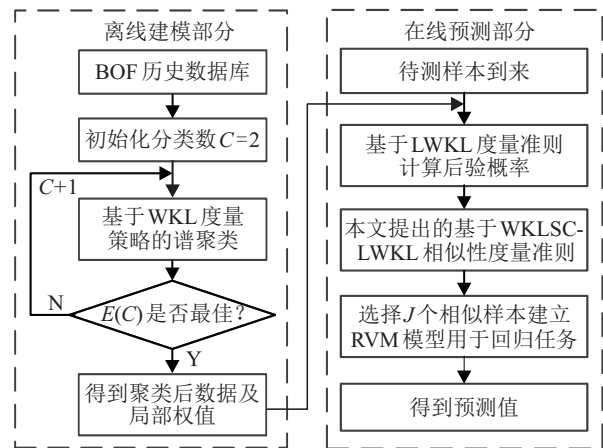


图2 面向转炉炼钢过程数据即时学习软测量建模流程

预测模型选用 RVM^[22], RVM 是一种基于概率贝叶斯框架的稀疏回归模型, 有着与支持向量机 (support vector machines, SVM) 许多相同的特点, 具有较快训练速度, 适用于在线建模和小样本训练, 且泛化能力好于 SVM^[22]. 对于 N 个训练样本 $(\mathbf{X}, \mathbf{y}) = \{(x_i, y_i)\}_{i=1}^N$, RVM 的数学模型可表示为

$$\mathbf{y} = f(\mathbf{X}; \mathbf{w}) + \varepsilon = \sum_{i=1}^N \omega_i K(x', x_i) + \omega_0 + \varepsilon. \quad (15)$$

其中: 权重 $\mathbf{w} = [w_0, w_1, \dots, w_i, \dots, w_N]^T$, ε 为 $\mathcal{N}(0, \sigma^2)$ 分布的样本误差, x' 为相关向量. 训练过程中, 对模型的超参数 α 和方差 σ^2 设置初值, 而后通过最大似然估计迭代求解. 由于高斯核函数具有局部学习能力强的特点, 适用于小样本建模的情况, 本文选择高斯核函数作为此回归模型的核函数.

3 实验及结果分析

3.1 实验数据介绍

对熔池内碳含量和温度的准确预测是转炉炼钢终点控制的关键. 本文实验数据来源于某钢厂实际转炉炼钢生产过程数据, 终点碳含量和温度为主导变量, 即输出变量, 辅助变量包括铁水温度、矿石量、硅锰量、氧枪位置等126维数据. 为了构建软测量模型, 通过特征选择, 选取与输出碳温密切相关的6个辅助变量作为特征^[11], 如表1所示. 其中, 特征名称中含有的数字表示传感器对该变量第几次的测量, 采样方式为点式测量, 传感器每15 s进行一次采样.

表1 转炉炼钢生产过程数据变量表

输出	辅助变量	输出	辅助变量
终点钢水 碳含量/%	氧压31	终点钢水 温度/°C	装入生铁量
	兑铁时长		铁水P
	兑铁结束到开氧时间		枪位22
	铁水温度		氧压29
	兑铁开始与上炉出钢时长		氧压11
	装入废钢量		枪位16

实验过程中一共采集2 100炉次正常工况数据, 为了预测图更加清晰地表达预测值与真实值的拟合效果, 选取其中2 000炉次作为历史库样本, 100炉次作为测试样本, 实验参数设置如表2所示.

表2 实验参数设置

参数名称	参数值
JITL相似样本个数 J	min = 2, max = 100
WKLS C 聚类数目 C	min = 2, max = 8
RVM 预测模型的核函数	'Gaussian'
RVM 训练初值设置	$(\alpha, \sigma^2) = (0.001, 0.001)$
碳含量预测误差范围(Th)/%	{±0.01, ±0.02, ±0.03}
温度预测误差范围(Th)/°C	{±5, ±10, ±15}

为了评价不同软测量模型的预测性能, 本文采用预测精度(prediction accuracy, PA)、均方根误差(RMSE)和平均绝对百分比误差(MAPE)作为指标, PA值越大预测精度越高, RMSE和MAPE值越小预测性能越好, 其计算方式如下:

$$RMSE = \sqrt{\frac{1}{N_{test}} \sum_{i=1}^{N_{test}} (y_{test,i} - y_{pre,i})^2}; \quad (16)$$

$$MAPE = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left| \frac{y_{test,i} - y_{pre,i}}{y_{test,i}} \right|; \quad (17)$$

$$PE_i = \begin{cases} 1, & |y_{test,i} - y_{pre,i}| \leq Th; \\ 0, & |y_{test,i} - y_{pre,i}| > Th; \end{cases} \quad (18)$$

$$PA = \frac{\sum_{i=1}^{N_{test}} PE_i}{N_{test}} \cdot 100\%. \quad (19)$$

其中: N_{test} 表示测试样本的数量; $i \in (1, N_{test})$; y_{pre}

和 y_{test} 表示预测值和真实值; PA表示终点碳温在误差范围内的预测精度, 计算方式见式(18)和(19); PE_i 表示误差在误差范围(Th)内时即为预测命中.

3.2 本文方法的性能比较

根据第2节的建模流程, 进行转炉炼钢终点碳温软测量模型仿真实验, 共进行5组独立的实验: 通过 $E(C)$ 确定训练集的最佳聚类数目, 并改变 C 进行建模预测以验证定义2的有效性; 改变建模相似样本数 J , 比较实验结果; 计算待测样本隶属度时不同计算方式对实验结果的影响; 比较不同回归模型下的预测效果; 通过10折交叉验证本文相似性度量算法的无偏性.

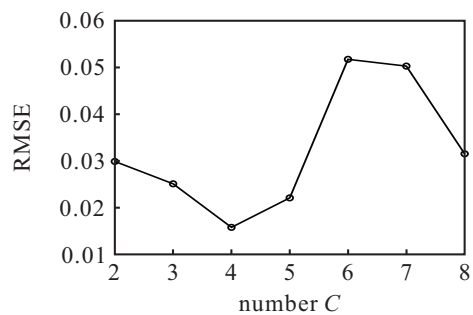
1) 最佳聚类数目 C 值的确定及验证.

$E(C)$ 定义了炼钢过程数据聚类效果的评价函数, 在离线阶段设置聚类数目 C 从2到8进行计算, 表3展示了碳含量和温度两个数据集不同 C 下的 $E(C)$ 值, 通过结果比较得到在预测终点碳含量和温度时历史炉次样本合理的聚类数目 C 分别为4和6.

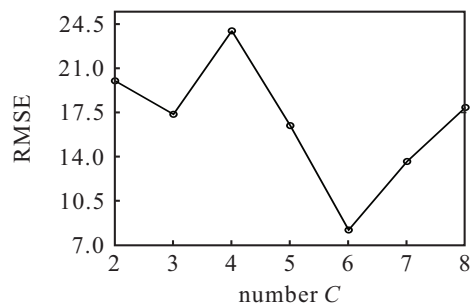
表3 计算不同 C 下的 $E(C)$ 以确定最佳聚类数

C	2	3	4	5
碳含量 $E(C)$	4.957 768	2.497 71	1.848 923	2.472 308
温度 $E(C)$	1.237 964	1.173 9	0.989 24	0.830 86

C	6	7	8
碳含量 $E(C)$	2.696 908	2.651 39	2.555 704
温度 $E(C)$	0.528 567	0.944 719	1.035 697



(a) $J = 40$ 时, 碳含量预测在不同聚类数目 C 对应的均方根误差



(b) $J = 48$ 时, 温度预测在不同聚类数目 C 对应的均方根误差

图3 不同聚类数目 C 对应的均方根误差值

通过第2节的软测量建模方法改变不同的聚类数目进行终点碳温预测,验证 $E(C)$ 可以针对炼钢过程数据聚类的有效性进行合理评估. 分别以最终建模样本子集个数 $J = 40$ 和 $J = 48$ 为例,改变 C 分析其对实验结果的影响,不同聚类数目对应预测结果的均方根误差如图3所示.

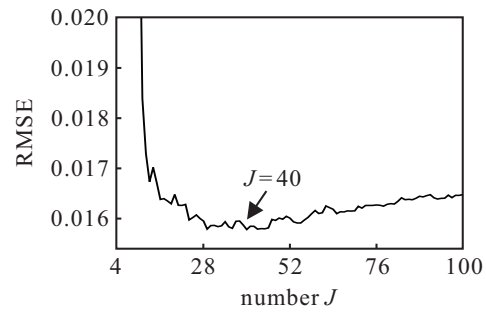
对比图3可知,在同样的最终建模样本子集数 J 情况下,碳含量和温度的 C 值分别在取4和6时的RSME最小,实验表明预测效果最佳的聚类数目与通过 $E(C)$ 计算得到的最佳聚类数吻合.

2) 局部建模相似样本数 J 对实验结果的影响.

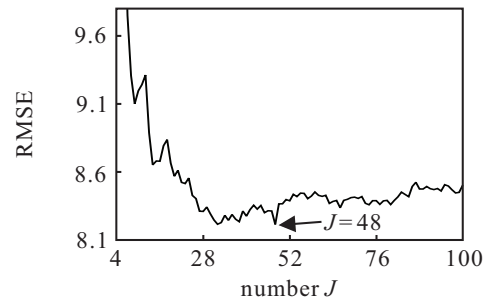
JITL方法中,待预测样本与历史样本经度量后,选取最相似的 J 个样本建立RVM局部模型进行终点碳温的预测,而参数 J 对预测精度影响较大,取 J 为4到100,分析在不同局部建模样本数 J 下预测结果,如图4所示,终点碳温的预测指标分别在 J 取40和48时达到最佳.

3) 计算待预测样本隶属度时度量准则加权参数对实验结果的影响.

由1.2节可知,待预测样本后验概率的算法主要是利用类簇内输入输出的相关系数加权给KL度量准则,由于引入局部信息,使得在计算时扩大了类间方差,缩小了类内方差. 通过改变计算后验概率和样本度量的加权参数部分进行仿真实验,进而分析加权参数对实验结果的影响. 度量策略分别是:①不进行加权的WKLSC-KL度量策略;②利用定义1得到全



(a) 碳含量预测在不同局部建模样本数 J 下对应的均方根误差



(b) 温度预测在不同局部建模样本数 J 下对应的均方根误差

图4 不同局部建模样本数 J 下对应的均方根误差值

局权值加权给KL度量准则计算的WKLSC-WKL (weighted KL)度量策略;③利用定义3计算局部权值加权的WKLSC-LWKL (local weighted KL)度量策略. 实验结果取参数表现最佳的情况,碳含量和温度预测结果曲线见图5,各自的评价指标如表4所示,其中表内参数一行中 J 表示待预测样本的最佳训练样本个数, C 代表最佳聚类数目.

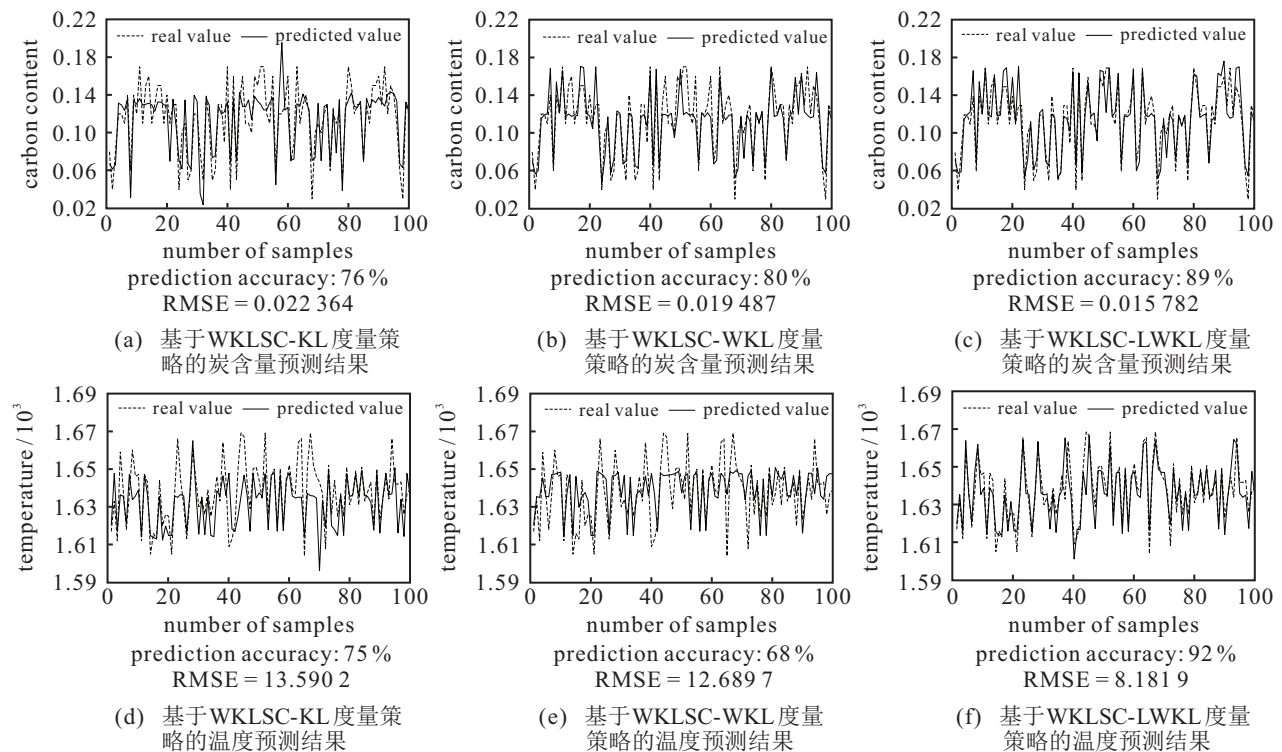


图5 不同加权参数下对应的终点碳含量和温度预测结果曲线

表4 不同加权参数下终点碳含量和温度预测评价指标

datasets	Th	WKLSC-KL	WKLSC-WKL	WKLSC-LWKL
参数		$J, C = (23,4)$	$J, C = (35,4)$	$J, C = (40,4)$
碳含量	0.01 %	43 %	64 %	72 %
	0.02 %	76 %	80 %	89 %
	0.03 %	89 %	90 %	96 %
RMSE		0.022 364	0.019 487	0.015 782
MAPE		0.196 465	0.145 71	0.128 425
参数		$J, C = (28,6)$	$J, C = (28,6)$	$J, C = (48,6)$
温度 PA	5 °C	46 %	44 %	60 %
	10 °C	75 %	68 %	92 %
	15 °C	81 %	80 %	96 %
RMSE		13.590 2	12.689 7	8.181 9
MAPE		5.87e-03	5.86e-03	3.56e-03

由图5和表4可以看出,虽然离线部分已将历史训练样本聚类为类内方差较小、类间方差较大的类簇,但在线预测时待测样本的隶属度计算对最终预测造成很大影响,若计算后验概率时只利用KL度量准则,则碳温预测精度分别在误差±0.02%和±10°C时只有76%和75%;若利用与聚类算法中同样的相似性度量方式的WKL计算后验概率,则温度的预测精度虽然降低为68%,但是RMSE和MAPE值减小,碳温预测的性能都有提高;而考虑了类间差异的LWKL扩大了类间方差,对待预测样本进行有效分类,实验结果表明其能够合理计算待测样本的隶属度,提高了模型预测精度。

4) 本文相似性度量算法下不同回归模型对实验结果的影响。

由于回归模型的好坏影响着预测结果,在这一独立实验中将基于本文相似性度量算法,比较BP神经网络

网络(BP)、极限学习机(ELM)、SVM和RVM的预测效果,所对比方法的模型参数都调至最优。

本文相似性度量算法下回归模型为BP、ELM、SVM和RVM的碳温预测评价指标,见表5。由RMSE指标可知,在小样本建模情况下BP和ELM都不同程度地出现了模型拟合程度不够导致预测值与真实值出现较大偏差的现象,且RMSE相较于SVM和RVM模型都较高;虽然SVM和RVM在各Th下的预测精度相差不大,但RMSE和MAPE的指标均为RVM较好,说明其泛化性能与SVM相比较好。综上,RVM作为终点碳温的预测模型更合适。

表5 本文相似性度量算法下不同回归模型预测评价指标

datasets	Th	BP	ELM	SVM	RVM
碳含量	0.01 %	52 %	59 %	65 %	72 %
	0.02 %	78 %	83 %	86 %	89 %
	0.03 %	91 %	90 %	96 %	96 %
RMSE		0.023 388	0.021 141	0.016 747	0.015 782
MAPE		0.179 8	0.171 005	0.137 622	0.128 425
温度 PA	5 °C	53 %	50 %	54 %	60 %
	10 °C	81 %	83 %	87 %	92 %
	15 °C	88 %	92 %	93 %	96 %
RMSE		10.169 6	9.999 0	8.797 7	8.181 9
MAPE		4.59e-03	4.52e-03	3.89e-03	3.56e-03

5) 10折交叉验证实验。

将实验过程中一共采集的2100炉次进行10折划分,每次用不同的折作为测试集,其余9折作为训练集,按第2节建模流程作10次仿真实验,共获得10折各自作为测试集的预测结果及指标,最后取均值。表6展示了每折作为测试集时碳温的预测指标。

表6 10折交叉验证下碳含量和温度每折预测指标及均值

datasets	Th	第1折	第2折	第3折	第4折	第5折	第6折	第7折	第8折	第9折	第10折	均值
碳含量	0.01 %	60.48 %	68.10 %	68.10 %	66.67 %	65.24 %	64.76 %	68.57 %	68.10 %	69.05 %	68.57 %	66.76 %
	0.01 %	88.10 %	90.95 %	88.57 %	86.19 %	90.48 %	87.62 %	85.71 %	89.05 %	88.57 %	87.14 %	88.24 %
	0.01 %	95.71 %	96.67 %	93.81 %	94.76 %	95.71 %	93.33 %	95.71 %	93.81 %	94.76 %	93.33 %	94.76 %
RMSE(E-02)		1.724	1.739	1.658	1.839	1.764	1.800	1.783	1.814	1.815	1.825	1.776
MAPE		0.129 1	0.122 0	0.112 8	0.131 9	0.131 8	0.150 9	0.130 8	0.117 8	0.133 8	0.120 3	0.1280
温度 PA	5 °C	56.19 %	66.67 %	55.71 %	51.43 %	56.19 %	57.14 %	51.90 %	54.29 %	54.76 %	56.67 %	56.10 %
	10 °C	89.52 %	89.05 %	90.00 %	90.95 %	88.10 %	86.67 %	90.00 %	87.62 %	89.52 %	88.57 %	89.00 %
	15 °C	96.67 %	94.29 %	95.24 %	96.19 %	94.76 %	94.76 %	94.76 %	96.19 %	93.81 %	94.29 %	95.10 %
RMSE(E-02)		7.877 9	8.597 8	8.038 8	8.135 1	8.379 3	8.018 9	8.705 4	8.422 5	9.113 8	9.142 7	8.443 2
MAPE		3.81	3.62	3.67	3.81	3.75	3.77	4.03	3.98	3.98	4.06	3.85

由表6的预测指标可知:碳含量预测在误差±0.02%内精度最高为第2折的90.95%,最低为第7折的85.71%;温度预测在误差±10°C内精度最高为第4折的90.95%,最低为第6折的86.67%,终点碳温的预测精度保持稳定,且相应的RMSE和MAPE值波动较为稳定。观察交叉验证获得的最终平均

预测精度,碳含量预测在误差±0.02%内的精度为88.24%,±0.03%内的精度为94.76%;温度预测在误差±10°C内精度为89%,±15°C内精度为95.10%。

3.3 与其他即时学习算法实验对比及分析

在本节中通过纵向对比文献[13]的传统JITL度量策略,基于传统谱聚类将历史样本聚类后基于欧

氏距离计算后验概率及样本局部相似性的度量策略(SC-ED)和本文算法的预测结果, 见图6和图7中的(a)、(b)和(f), 验证本文所提出方法在转炉炼钢过程数据中炉次样本相似性度量的有效性; 横向对比文献[15]提出的二次相似性度量策略, 文献[16]中高斯混合模型加权马氏距离(GMM-MD)的度量策略和文献[17]提出的互信息加权欧氏距离度量策略, 实验预

测结果见图6和图7中的(c)、(d)和(e). 所对比的实验中预测模型均为RVM, 以验证本文相似性度量策略在转炉炼钢终点碳温预测的优势.

所对比方法均已通过选取最优参数在转炉炼钢数据下的最佳表现效果, 图6和图7分别是终点碳含量和温度的预测结果, 表7呈现了不同即时学习算法下碳含量和温度的预测性能指标.

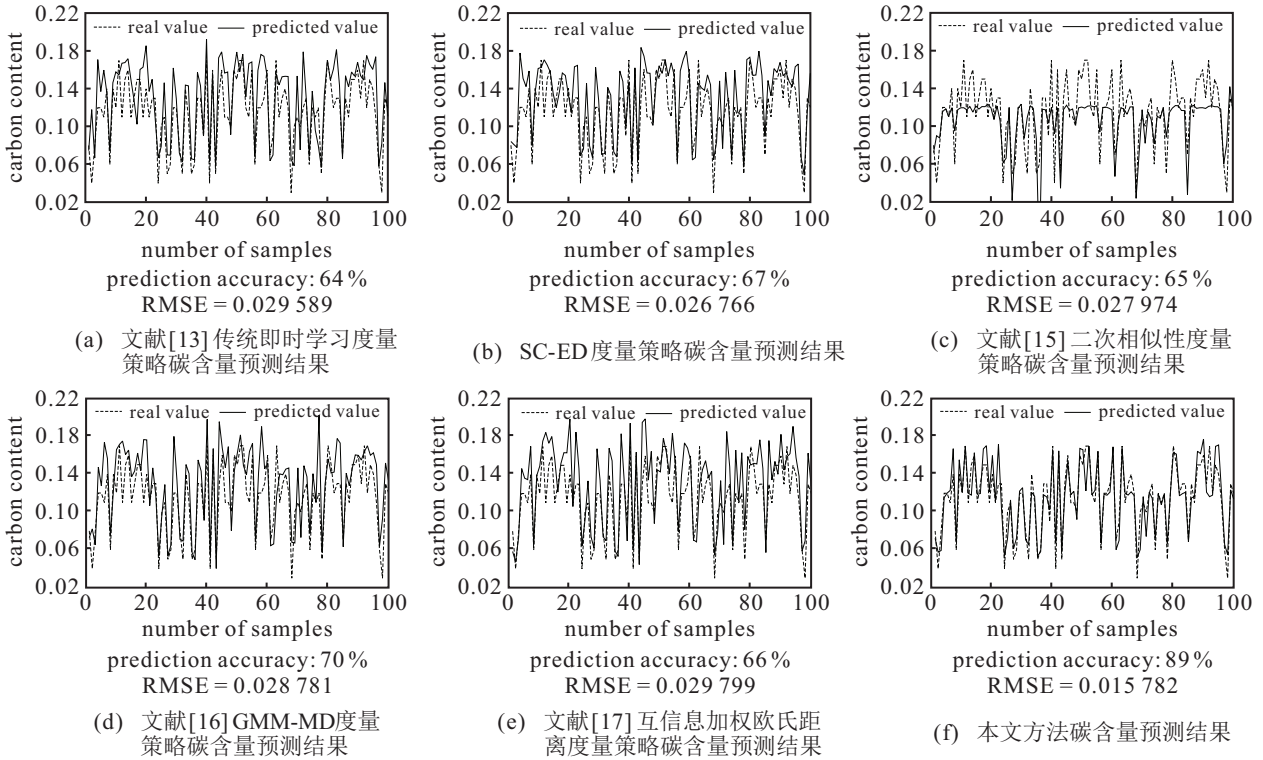


图6 不同即时学习算法的碳含量预测结果

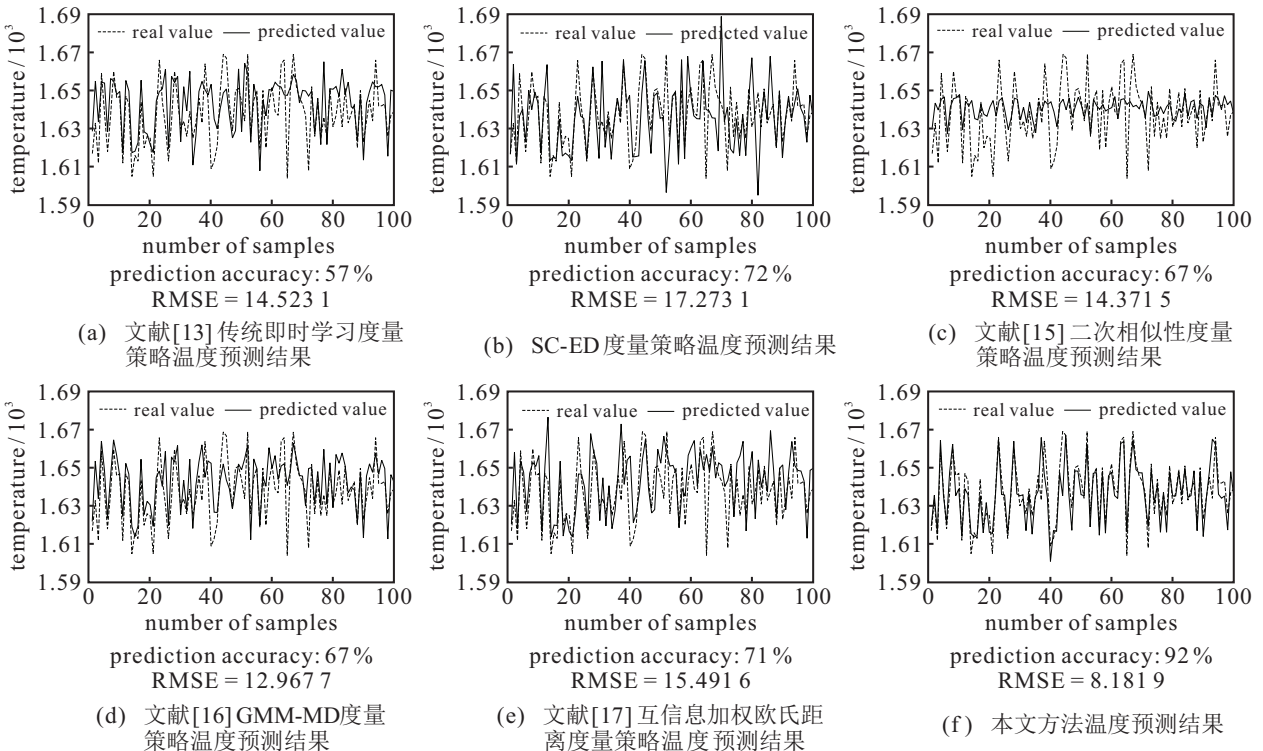


图7 不同即时学习算法的温度预测结果

表7 不同即时学习算法下碳含量和温度的预测性能指标

Datasets	Th	传统JITL 度量策略 ^[13]	SC-ED度量 策略	二次相似性 度量策略 ^[15]	GMM-MD 度量策略 ^[16]	互信息加权欧氏 距离度量策略 ^[17]	本文方法
碳含量	0.01 %	43 %	44 %	50 %	49 %	37 %	72 %
	0.02 %	64 %	67 %	65 %	70 %	66 %	89%
	0.03 %	72 %	79 %	78 %	79 %	75 %	96 %
RMSE		0.029 6	0.026 8	0.027 9	0.028 8	0.029 8	0.015 8
MAPE		0.250 6	0.226 4	0.237 1	0.230 5	0.229 5	0.128 4
温度PA	5 °C	33 %	53 %	39 %	42 %	35 %	60 %
	10 °C	57 %	72 %	67 %	67 %	71 %	92%
	15 °C	76 %	76 %	72 %	80 %	80 %	96 %
RMSE		14.523 1	17.273 1	14.371 5	12.967 7	15.491 6	8.181 9
MAPE		6.96e-03	6.78e-03	6.47e-03	5.95e-03	6.89e-03	3.56e-03

1) 使用文献[13]中基于传统JITL度量策略的碳温即时学习模型预测精度分别在误差 $\pm 0.02\%$ 和 $\pm 10^\circ\text{C}$ 时为64%和57%,而SC-ED度量策略由于先将历史样本进行聚类后进行度量,与传统JITL度量策略相比,预测精度分别提高到了67%和72%,但RMSE和MAPE指标都在增大,观察预测曲线可知预测偏差过大的样本变多,说明SC-ED度量策略度量出的相似样本类内方差过大,并未解决炉次样本波动性大的问题,导致预测结果偏差也过大;而基于本文度量策略的即时学习碳温预测模型精度在误差 $\pm 0.02\%$ 和 $\pm 10^\circ\text{C}$ 时分别达到89%和92%,原因是WKLSC-LWKL既有效平衡了历史训练数据波动性大的问题,又考虑了类簇间差异性的问题,对待预测样本有效分类进而合理度量炉次样本的相似性,得到了不错的预测结果。

2) 由表7的结果可知:本文所提出方法相比于文献[15]方法预测碳含量在误差 $\pm 0.02\%$ 下从65%提高到89%,预测温度在误差 $\pm 10^\circ\text{C}$ 下从67%提高到92%;同样的,文献[16]方法虽然在预测碳温误差 $\pm 0.02\%$ 和 $\pm 10^\circ\text{C}$ 下达到70%和67%,较传统即时学习有一定提高,但仍不能有效度量炉次样本进行预测;文献[17]方法利用输入输出的互信息加权在度量策略上,在温度预测方面有明显提高,在误差 $\pm 10^\circ\text{C}$ 下预测精度达到了71%,但未解决炉次样本波动性大的问题,RMSE和MAPE值并未有所改善。

3) 由图6(f)和图7(f)预测结果曲线可知,本文方法能够有效预测终点,比其他即时学习算法效果更好,在转炉炼钢终点碳温预测方面具有一定优势,原因是本文方法充分考虑了待预测样本当前工况的特性,通过将历史波动性大的训练样本先进行聚类,借助聚类结果,引入类簇间差异计算待测样本的隶属度及其局部KL度量准则,建立合理的相似性度量策略,有效解决了样本波动性大的问题,且引入输入输出相关性权值,使得最终获得的相似样本集类内方差更

小,在RVM建模预测时,训练子集越相似,终点碳温预测精度越高。

4 结论

本文针对转炉炼钢复杂生产过程由于材料品质差异及人工失误造成的样本间波动性大,全局模型无法自适应样本波动进行准确预测,且传统即时学习方法中的度量策略无法对炉次样本有效度量造成预测精度不高的问题,提出一种基于WKLSC-LWKL度量策略的即时学习软测量方法用于转炉炼钢的终点碳温预测,内容如下:

1) 度量策略充分考虑了当前炉次工况特性,通过将历史样本聚类成类内方差较小、类间方差较大的类簇后,通过类簇间差异信息计算待预测样本属于各个类的后验概率和局部相似度,建立了合理的相似度策略。

2) 以类内同一性和类间分离性定义的聚类评价函数,能够对转炉炼钢过程数据进行有效聚类,使得质心样本点更加合理地代表其所在类簇,有利于后续计算待预测样本的后验概率。

3) 对历史样本聚类时考虑的是全局信息,将全局权值加权给KL度量准则从而构建图进行谱聚类,聚类后考虑类簇间差异信息,将局部权值加权给KL度量准则以计算待测样本的后验概率,这种度量准则既实现了历史样本的聚类,又实现了新来样本类别隶属度的计算。

通过对转炉炼钢过程数据的建模仿真,验证了本文方法对终点碳温预测的有效性,较于传统JITL方法,能够选择更为合理的相似样本用于即时学习建模,也为转炉炼钢生产过程数据建模提供了更好的研究方向,在实际应用中具有一定参考价值。

参考文献(References)

- [1] 刘超. 中国转炉炼钢技术的发展、创新与展望[J]. 特钢技术, 2013, 19(4): 6-9.

- (Liu C. Development, innovation and prospect of converter steelmaking technologies in China[J]. Special Steel Technology, 2013, 19(4): 6-9.)
- [2] 王忠刚, 任科社, 刘建. 副枪技术在莱钢120t转炉上的应用[J]. 中国冶金, 2009, 19(11): 30-32.
(Wang Z G, Ren K S, Liu Z J. Sublance technology application of 120t converter at Laiwu steelmaking plant[J]. China Metallurgy, 2009, 19(11): 30-32.)
- [3] 吴令, 姜周华, 龚伟, 等. 通过炉气分析实现转炉连续控制[J]. 炼钢, 2008, 24(1): 12-15.
(Wu L, Jiang Z H, Gong W, et al. Continuous control of BOF by gas analysis[J]. Steelmaking, 2008, 24(1): 12-15.)
- [4] 温宏愿, 周末春. 转炉终点控制技术方法现状及应用[J]. 现代工业经济和信息化, 2019, 9(2): 79-81.
(Wen H Y, Zhou M C. Current status and application of converter end-point control technology[J]. Modern Industrial Economy and Informationization, 2019, 9(2): 79-81.)
- [5] 周末春, 赵琦, 陈延如, 等. 基于FSVM的转炉终点光辐射状态识别[J]. 红外与激光工程, 2018, 47(7): 312-317.
(Zhou M C, Zhao Q, Chen Y R, et al. State recognition of light radiation of BOF end-point based on fuzzy support vector machine[J]. Infrared and Laser Engineering, 2018, 47(7): 312-317.)
- [6] Liu H, Wang B, Xiong X. Basic oxygen furnace steelmaking end-point prediction based on computer vision and general regression neural network[J]. Optik, 2014, 125(18): 5241-5248.
- [7] Yuan X F, Li L, Shardt Y A W, et al. Deep learning with spatiotemporal attention-based LSTM for industrial soft sensor model development[J]. IEEE Transactions on Industrial Electronics, 2021, 68(5): 4404-4414.
- [8] Yuan X F, Ou C, Wang Y L, et al. A layer-wise data augmentation strategy for deep learning networks and its soft sensor application in an industrial hydrocracking process[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 32(8): 3296-3305.
- [9] 王心哲, 韩敏. 基于变量选择的转炉炼钢终点预报模型[J]. 控制与决策, 2010, 25(10): 1589-1592.
(Wang X Z, Han M. Variable selection based BOF endpoint prediction model[J]. Control and Decision, 2010, 25(10): 1589-1592.)
- [10] 谢书明, 陶钧, 柴天佑. 基于神经网络的转炉炼钢终点控制[J]. 控制理论与应用, 2003, 20(6): 903-907.
(Xie S M, Tao J, Chai T Y. BOF steelmaking endpoint control based on neural network[J]. Control Theory & Applications, 2003, 20(6): 903-907.)
- [11] 刘辉, 曾鹏飞, 巫丞顺, 等. 基于改进遗传算法的转炉炼钢过程数据特征选择[J]. 仪器仪表学报, 2019, 40(12): 185-195.
(Liu H, Zeng P F, Wu Q S, et al. Feature selection of converter steelmaking process based on the improved genetic algorithm[J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 185-195.)
- [12] 曾鹏飞, 刘辉. 基于二次相似性度量的即时学习转炉炼钢终点碳温软测量方法[J]. 计算机集成制造系统, 2021, 27(5): 1429-1439.
(Zeng P F, Liu H. Soft-sensing method for end-point carbon temperature of converter steelmaking based on quadratic similarity measurement[J]. Computer Integrated Manufacturing Systems, 2021, 27(5): 1429-1439.)
- [13] Cheng C, Chiu M S. A new data-based methodology for nonlinear process modeling[J]. Chemical Engineering Science, 2004, 59(13): 2801-2810.
- [14] Fujiwara K, Kano M, Hasebe S, et al. Soft-sensor development using correlation-based just-in-time modeling[J]. AIChE Journal, 2009, 55(7): 1754-1765.
- [15] 祁成, 史旭东, 熊伟丽. 基于二阶相似度的即时学习软测量建模方法[J]. 智能系统学报, 2020, 15(5): 910-918.
(Qi C, Shi X D, Xiong W L. A just-in-time learning soft sensor modeling method based on the second-order similarity[J]. CAAI Transactions on Intelligent Systems, 2020, 15(5): 910-918.)
- [16] Fan M, Ge Z Q, Song Z H. Adaptive Gaussian mixture model-based relevant sample selection for JITL soft sensor development[J]. Industrial & Engineering Chemistry Research, 2014, 53(51): 19979-19986.
- [17] 陆荣秀, 饶运春, 杨辉, 等. 基于改进即时学习算法的镨/铈元素组分含量预测[J]. 控制理论与应用, 2020, 37(8): 1846-1854.
(Lu R X, Rao Y C, Yang H, et al. Prediction of Pr/Nd component content based on improved just-in-time learning algorithm[J]. Control Theory & Applications, 2020, 37(8): 1846-1854.)
- [18] Luxburg U. A tutorial on spectral clustering[J]. Statistics and Computing, 2007, 17(4): 395-416.
- [19] Yuan X F, Ge Z Q, Huang B, et al. A probabilistic just-in-time learning framework for soft sensor development with missing data[J]. IEEE Transactions on Control Systems Technology, 2017, 25(3): 1124-1132.
- [20] Guo F, Xie R M, Huang B. A deep learning just-in-time modeling approach for soft sensor based on variational autoencoder[J]. Chemometrics and Intelligent Laboratory Systems, 2020, 197: 103922.
- [21] Halkidi M, Vazirgiannis M, Batistakis Y. Quality scheme assessment in the clustering process[C]. Principles of Data Mining and Knowledge Discovery. London, 2000: 265-276.
- [22] Tipping M E. Sparse Bayesian learning and the relevance vector machine[J]. Journal of Machine Learning Research, 2001, 1: 211-244.

作者简介

杨路(1997—), 男, 硕士生, 从事数据分析、软测量的研究, E-mail: yanglumj@163.com;

刘辉(1984—), 男, 副教授, 博士, 从事图像处理、软测量等研究, E-mail: liuhui621@126.com;

熊倩(1996—), 女, 硕士生, 从事数据分析、软测量的研究, E-mail: 1716186533@qq.com.

(责任编辑: 孙艺红)