

基于JITL-XGBoost的烧结终点预测模型

王金杨, 吴朝霞, 李中正, 康增鑫

(东北大学秦皇岛分校 控制工程学院, 河北 秦皇岛 066004)

摘 要: 烧结终点(burning through point, BTP)位置是烧结过程中重要的参数, 直接影响烧结机效率。由于烧结生产过程具有多工况、时变等特性, 使得全局模型预测性能不足, 为此提出了一种在即时学习框架中使用极端梯度提升(extreme gradient boosting, XGBoost)作为局部模型的烧结终点预测模型, 即JITL(just-in-time learning)-XGBoost。首先采用KL散度(Kullback-Leibler divergence)相似性度量方法提取待测样本的特性, 选出与待测样本最相关的数据集。然后将该数据集作为XGBoost模型的输入来预测烧结终点的位置。此外, 考虑了相关数据集数量对模型预测精度和计算时间的影响。最后与其他模型对比, 结果表明, 所建模型在合理的时间内具有最佳预测精度, 为提高烧结机效率提供新的指导方向。

关键词: 烧结矿; 烧结终点; 预测模型; 即时学习; 极端梯度提升

中图分类号: TF 046.4 文献标志码: A 文章编号: 1005-3026(2025)02-0028-08

Prediction Model of Burning Through Point Based on JITL-XGBoost

WANG Jin-yang, WU Zhao-xia, LI Zhong-zheng, KANG Zeng-xin

(School of Control Engineering, Northeastern University at Qinhuangdao, Qinhuangdao 066004, China. Corresponding author: WU Zhao-xia, E-mail: ysuwzx@126.com)

Abstract: The burning through point (BTP) is an important parameter in the sintering process, which directly affects the efficiency of the sintering machine. Due to the multi-working conditions and time-varying characteristics of the sintering production process, the prediction performance of the global model is insufficient. Therefore, a burning through point prediction model using XGBoost as a local model in the just-in-time learning framework is proposed, namely JITL-XGBoost. Firstly, the KL divergence similarity measurement method is used to extract the characteristics of the sample to be tested, and the most relevant data set of the sample to be tested is selected. Secondly, this dataset is used as input to the XGBoost model to predict the location of the burning through point. In addition, the impact of related dataset numbers on model prediction accuracy and model computation time is considered. Finally, by comparing with other models, the results show that the model built has the best prediction accuracy within a reasonable time, providing new guidance for improving the efficiency of sintering machines.

Key words: sintering ore; burning through point; prediction model; just-in-time learning; extreme gradient boosting

理想的BTP是指原材料混合物在烧结机上进行完全燃烧的位置。如果BTP位于目标位置的前面, 烧结机的有效面积不会得到充分利用, 如果BTP位于目标位置的后面, 烧结卸料时料层无

法完全燃烧, 导致烧结矿返矿率增加、产量减少^[1]。因此, 精准预测BTP位置可以使烧结机机速得到及时、准确的调整, 进而提高烧结机效率和减少经济成本。

收稿日期: 2023-08-31

基金项目: 河北省教育厅科学技术研究项目(BJ2021099)。

作者简介: 王金杨(1997—), 男, 内蒙古通辽人, 东北大学硕士研究生; 吴朝霞(1969—), 女, 浙江嘉兴人, 东北大学教授, 硕士生导师。

由于涉及多种物理和化学现象,烧结过程表现出高度的非线性、不确定性、强耦合和多工况。因此预测BTP的位置非常具有挑战性,并且一直是研究的热点。目前,对BTP进行预测的方法主要有两种,一种是机理模型,另一种是数据驱动模型。机理模型通常是根据第一性原理的静态参数和物理化学反应。Patisson等^[2]以烧结理论为基础,开发了烧结料层温度分布的数学模型,并在实验装置中对模型结果进行了验证。Cumming等^[3]研究了烧结过程中的物理化学反应,建立了烧结过程的数学模型,并在计算机上进行了计算验证。但上述模型都是建立在假设基础上的简化模型,限制了它们在实际烧结过程中的适用性。

近年来,随着烧结领域大量历史数据的积累和大数据技术的发展,数据驱动模型受到广泛关注。与机理模型不同,数据驱动模型基于实际烧结过程数据,不需要考虑难以描述的多个耦合物理化学反应,着重探讨数据之间的相关性^[4]。Du等^[5]提出了基于模糊时间序列的BTP预测模型,为铁矿石烧结过程的稳定控制奠定了基础。Chen等^[6]提出了BTP的Takagi-Sugeno模糊模型,分析了影响时变运行的因素条件。Liu等^[7]提出了一种梯度推进决策树算法来建立BTP和BTP温度的预测模型。Wang等^[8]运用烟花算法和遗传算法优化模型参数,构建了一种基于支持向量机的BTP预测模型。尽管传统的BTP预测模型具有很强的非线性拟合能力,但随着时间的推移,离线训练的模型无法适应工业过程的时变,导致模型性能随时间退化甚至失效。频繁地重新训练模型以保持准确性又将增加时间和经济成本。

针对传统模型存在的局限性,即时学习建模框架采用“边建模、边预测”的方式,通过相似度计算选出与待测样本最相关的样本集构建局部模型,实现短时间内高精度预测^[9]。当前,即时学习算法已在部分领域用于构建预测模型。任明仑等^[10]将即时学习算法引入灰铸铁抗拉强度预测,并与传统全局模型对比验证即时学习算法的有效性。王通等^[11]利用即时学习算法来提高油田液面测量的精度。针对烧结过程中时变导致预测性能不足的问题,本文首先对烧结数据进行预处理,得到可以直接建模的有效数据。其次,通过最大互信息系数(maximal information coefficient, MIC)得到26个与BTP最相关的变量。然后,利用KL散度相似性度量方法计算出查询样本和历史数据库的距离,进而选取相关数据

集。最后将选取的相关数据集作为XGBoost模型的输入,得到查询样本的最终BTP预测值,并将BTP数据库及时更新。烧结生产过程中的实际数据验证了该模型的优越性和可靠性,对BTP的及时、准确预测具有重要意义。

1 数据分析与处理

1.1 数据预处理

实验数据来自某烧结工厂3号烧结机共12 920组样本,可以将全部烧结过程参数分为以下4种:原料参数、混合料参数、操作参数以及状态参数,结合经验和工艺分析,最终在表1列出52种可能和BTP有关的生产参数。铁矿石烧结过程的生产顺序为配料、混合、制粒、加料、点火、烧结、通风、破碎、冷却、筛分^[12]。首先通过混合制粒机将原料搅拌成混合料并送到烧结台车上。当混合料通过点火器时,其表面被点燃,然后在风机的作用下,混合料在台车里自上向下燃烧,废气沿烟道流出。在混合料燃烧过程中,混合料首次完全燃烧的地方被称为BTP。由于没有仪器可以直接测量BTP值,一种基于风箱中废气温度的间接测量方法被提出^[13]。因此,本文采用风箱的数量来衡量BTP,一共22个风箱,BTP的位置一般出现在倒数第2个风箱附近。一般来说,从实际生产现场获得的数据不能直接用于建模,因为原始数据中可能存在传感器检测异常、恶劣环境等条件导致的异常数据。这些数据可能会影响模型的预测精度和可靠性。因此,需要对原始数据进行预处理。本文采用箱形图的异常值检测方法来检测并处理异常值。箱形图的原理是:用3条分位线将图形分为4个区域,每个区域占比1/4。3条分位线分别为下四分位线(Q1)、中位数(median, MD)、上四分位线(Q3),上四分位线与下四分位线之间的距离为四分位距(interquartile range, IQR)。Q1 - 1.5IQR 和 Q3 + 1.5IQR 是正常和异常数据的分界线,如图1所示。边界线内的数据为正常值,边界线外的数据为异常值。图2为箱形图法识别异常数据点的结果。对于缺失值和异常值数据点,本文采用其前后的5个正常数据点的平均值进行替换。由于不同变量之间的量纲不同,如果直接训练模型,会影响模型的结果。因此,要对处理完异常值和缺失值的数据进行标准化处理。

1.2 关键特征选择

烧结过程各变量之间存在高度的非线性,且

有些变量和 BTP 的相关性不大. 因此, 需要对各变量进行特征选择. 常见的特征选择方法有线性相关分析方法和非线性分析方法. 面对高度非线性的烧结数据, 引入可计算非线性相关的最大互信息系数 (C_{MI}), 提取与 BTP 相关性最强的几个关键参数作为输入变量. C_{MI} [14] 的定义为

$$C_{MI} = \max_{|x||y| \leq B} \frac{\sum_{x,y} P(x,y) \ln \frac{P(x,y)}{P(x)P(y)}}{\ln[\min(|x|, |y|)]} \quad (1)$$

式中: $|x|$ 和 $|y|$ 分别为变量 x 和变量 y 划分的区间数量; $P(x)$ 为样本点落在变量 x 上的概率; $P(y)$ 为样本点落在变量 y 上的概率; $P(x,y)$ 为变量 x 与变量 y 联合概率密度; B 为划分的最大网格数量, 是样本量的 0.6 次方. 如果 C_{MI} 值接近 1, 说明该输入变量和输出关系密切; 反之, 当 C_{MI} 值接近 0, 说明它们几乎不相关. 经过计算, 按 C_{MI} 大小顺序选取的 26 个输入变量如表 2 所示.

表 1 烧结过程主要的参数
Table 1 Main parameters of sintering process

种类	No.	参数	No.	参数
原料参数	1	焦粉/(t·h ⁻¹)	2	除尘灰/(t·h ⁻¹)
	3	燃料/(t·h ⁻¹)	4	铁粉/(t·h ⁻¹)
	5	烧结返矿/(t·h ⁻¹)	6	高炉返矿/(t·h ⁻¹)
	7	混合料 TFe 质量分数/%	8	混合料 V ₂ O ₅ 质量分数/%
混合料参数	9	混合料 CaO 质量分数/%	10	混合料 SiO ₂ 质量分数/%
	11	混合料水分质量分数/%		
	12	圆辊转速/(r·h ⁻¹)	13	九辊转速/(r·h ⁻¹)
操作参数	14	烧结机速度/(m·min ⁻¹)	15	点火温度/°C
	16	煤气流量/(m ³ ·h ⁻¹)	17	风机风量/(m ³ ·h ⁻¹)
	18	助燃风流量/(m ³ ·h ⁻¹)	19	助燃风压力/kPa
	20	南烟道温度/°C	21	南烟道负压/kPa
状态参数	22	北烟道温度/°C	23	北烟道负压/kPa
	24~37	14 个风箱废气温度/°C	38~51	14 个风箱负压/kPa
	输出参数	52	BTP	—

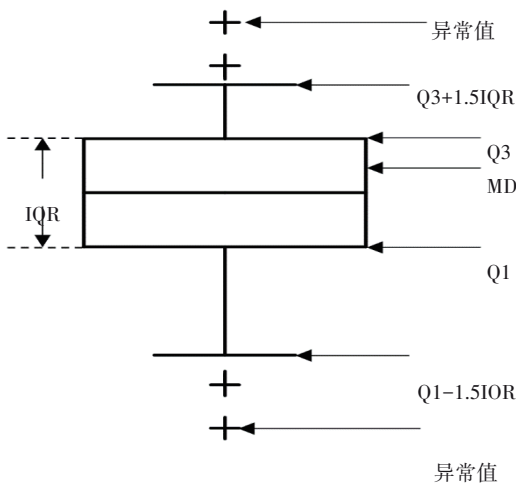


图 1 箱形图
Fig. 1 Box plot

由表 2 可知, 和 BTP 相关的过程参数主要有风箱废气温度、原料下料量和风箱真空度等. 风

箱废气温度可以反映料层表面的温度变化, 是判断 BTP 位置的关键因素. 焦粉等原料在烧结过程中会自上向下燃烧, 下层原料的燃烧需要由上层的燃烧热给予支持, 同时原料下料量的变化会影响到烧结小球的成分和结构, 从而影响 BTP 的位置. 风箱真空度可以影响烧结料层的氧化还原性质和气相组成. 当风箱真空度过大时, 会使料层中的氧气含量减少, 导致燃烧不充分, 从而使 BTP 偏后; 反之则会使 BTP 靠前.

2 BTP 预测模型

2.1 相似性度量方法

即时学习算法之所以可以适应不断变化的过程环境, 原因在于它是从数据库中选择最相似的样本形成相关数据集, 不同于全局模型使用数据库中所有的数据, 这种做法更具有针对性, 对

烧结过程的时变特性具有更强的适应性.在 JITL 方法中,一定数量的输入输出参数存储在数据库中.XGBoost 等模型是通过该数据库开发的.该数据库是由 KL 散度计算出的数据组成的.KL 散度计算公式如下:

$$D_{KL}(P_d|P_q) = \sum_{i=1}^N P_d(x_i) \log \frac{P_d(x_i)}{P_q(x_i)} \quad (2)$$

式中: P_d 和 P_q 为数据库和查询数据的概率分布值; N 为数据集大小.KL 散度获得的值越小,说明样本间相似度越高.

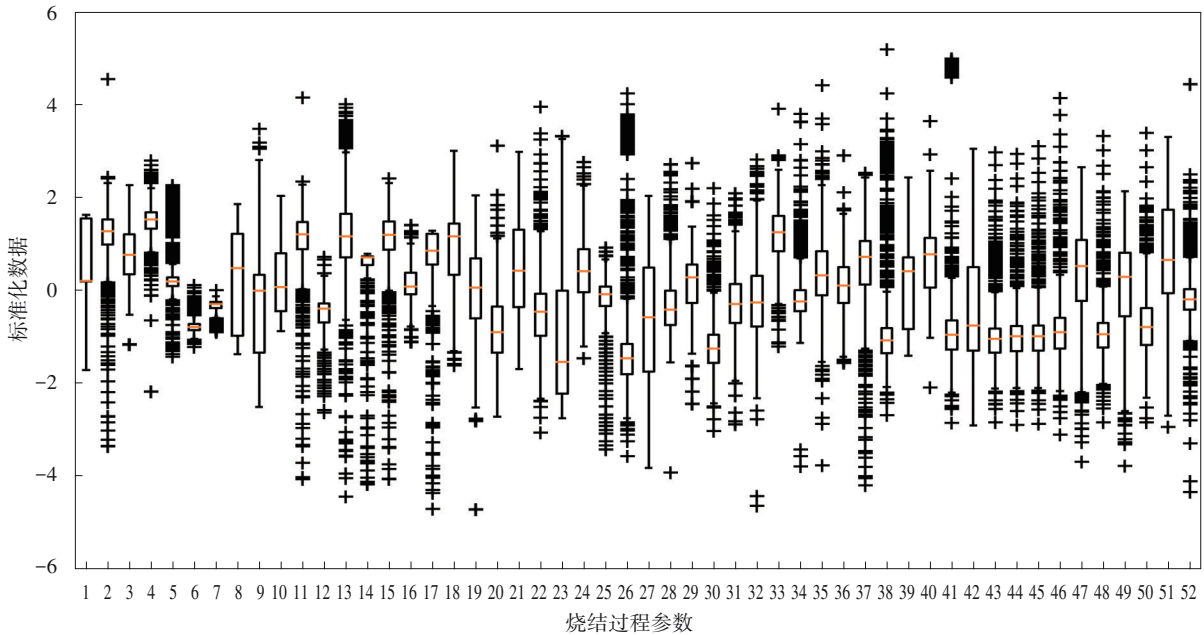


图 2 箱形图法识别异常数据点的结果

Fig. 2 Results of box plot method for identifying abnormal data points

表 2 C_{Mi} 计算结果
Table 2 Calculation results of C_{Mi}

过程参数	C_{Mi} 值	过程参数	C_{Mi} 值
20号风箱废气温度	0.517 500	焦粉	0.419 891
1号风箱废气温度	0.366 053	7号风箱真空度	0.355 897
2号风机风量	0.353 517	11号风箱真空度	0.276 374
7号风箱废气温度	0.273 018	22号风箱废气温度	0.233 880
3号风箱废气温度	0.174 285	混合料 CaO 质量分数	0.172 590
混合料 V_2O_5 质量分数	0.148 207	TFe	0.145 821
9号风箱真空度	0.143 866	2号风箱废气温度	0.141 890
混合料 SiO_2 硅质量分数	0.134 581	烧结机速度	0.133 350
2号风箱真空度	0.129 686	混合料除尘灰质量分数	0.127 677
13号风箱废气温度	0.126 931	混合料水分质量分数	0.125 018
5号风箱真空度	0.121 172	15号风箱废气温度	0.113 292
5号风箱废气温度	0.105 676	11号风箱废气温度	0.105 623
烟道温度	0.103 928	3号风箱真空度	0.093 048

2.2 XGBoost 算法

XGBoost 是一种基于树模型的集成学习算法,通过引入正则化和优化技术来防止过拟合,能够有效地处理烧结过程中复杂的非线性关系,为 BTP 的预测提供有力支持.该算法中对于回归

问题构造平方项损失作为目标函数,用泰勒级数对目标函数进行二项展开以解决目标函数优化困难的问题,并且用梯度下降算法进行优化求解^[15].该算法的目标函数一般形式如下:

$$\hat{y}_i = \sum_{k=1}^T f_k(x_i), \quad (3)$$

$$O_{bj} = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \left(\gamma T + \lambda \frac{1}{2} \sum_{j=1}^T w_j^2 \right). \quad (4)$$

式中： y_i 为第 i 个样本的实际值； \hat{y}_i 为第 i 个样本的预测值； $f_k(x_i)$ 为第 k 个数对第 i 个数据样本的预测结果； $l(y_i, \hat{y}_i)$ 为第 i 个样本的损失函数； γ 和 λ 为模型复杂度变量； T 为叶子节点个数； w_j 为叶子节点的权重； n 为测试样本数。对式(4)目标函数泰勒展开后求偏导并令其等于 0，此时的目标函数为

$$O_{bj} = \frac{1}{2} \sum_{i=1}^T \frac{\sum g_i}{\sum h_i + \lambda} + \gamma T. \quad (5)$$

式中： g_i 和 h_i 分别为目标函数的一阶导数、二阶

导数。

2.3 JITL-XGBoost 模型

本文采用 JITL-XGBoost 模型对 BTP 进行预测，图 3 为预测模型流程图。其基本原理如下：

- 1) 将中国某烧结工厂 3 号烧结机的实际生产数据进行数据预处理和特征向量选择之后，得到 BTP 数据集 (n 个输入、输出)；
- 2) 将从传感器得到的查询数据与 BTP 数据库进行相似度计算，得到相似度值；
- 3) 从 BTP 数据库中挑选 k 个数据集作为最相关数据集 ($k < n$)；
- 4) 采用 k 个数据来训练 XGBoost 模型，得到预测结果；

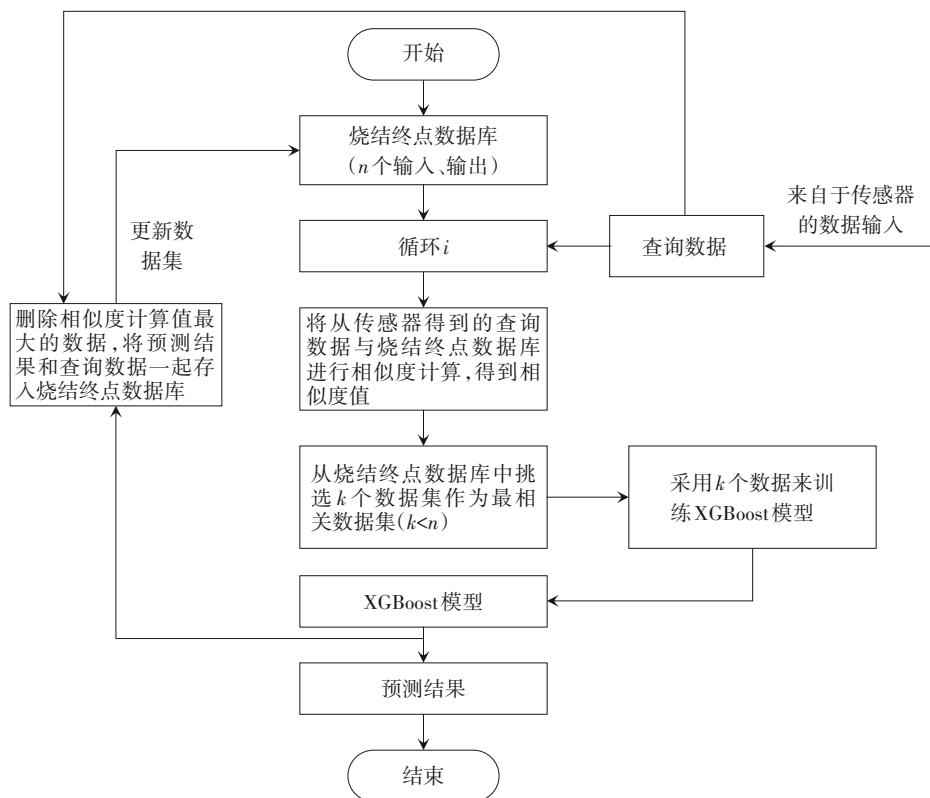


图 3 JITL-XGBoost 模型流程图

Fig. 3 Flowchart of the JITL-XGBoost model

5) 将预测结果和查询数据一起存入烧结终点数据库，并且删除相似度最小的数据。

为了避免数据库数量随着时间增加而扩大，应该摒弃数据库中较坏的数据。因此，本文采用数据库更新技术，通过这种方式使 BTP 数据集大小固定，同时递归地更新数据来解决烧结过程中的时变问题。

2.4 模型评价指标

本文介绍了 4 种不同的评价指标，即相关系

数 (C_c)、平均绝对误差 (e_{MA})、均方误差 (e_{MS}) 和均方根误差 (e_{RMS})，用来评价该方法的准确性。其中，相关系数表示一个变量与另一个依赖关系的量化。相关系数越接近一个稳定数值、误差越小，模型的准确度越高。其计算公式如下：

$$C_c = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}; \quad (6)$$

$$e_{MA} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|; \quad (7)$$

$$e_{MS} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2; \quad (8)$$

$$e_{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (9)$$

式中: \bar{y}_i 表示所有预测值的平均值.

3 模型预测性能对比与分析

本文将数据比例划分为 8:2, 前部分用来训练数据, 后部分用来测试模型的训练能力. 本文使用的软件环境是 python 3.8 和 scikit learn 1.0.2, 硬件环境是 Intel (R) Core (TM) i5-7200U CPU 和 8 GB 内存.

在 JITL 框架中, 确定最佳相关数据集的数量是决定 BTP 预测性能好坏的一个重要因素.

为了确定最佳数据集数量, 采用网格搜索法基于 1 000 到 9 000 的数据集数量开发了局部模型 (XGBoost, SVR, GPR), 得到相关系数 (C_c) 和相关数据集数量之间的依赖关系, 如图 4 所示.

从图 4 可以看出, 最佳的数据集数量在 3 000 附近, 预测精度较高, 在数据集大于 3 000 时虽然预测精度值继续上升, 但是预测精度的提高可以忽略不计. 因此, 将最佳数据集数量定为 3 000. 同样可以看出, JITL-XGboost 的性能均优于其他算法模型的性能.

为了更好地展示所提模型解决烧结过程中时变的能力, 本文通过不同方法在不同时间段的

数据子集进行建模, 来评估所建模型的预测优势. 即将所建模型与 JITL-SVR, JITL-GPR 和 XGBoost 模型进行对比, 图 5 为过程状态发生变化前 BTP 预测结果, 图 6 为过程状态发生变化后 BTP 预测结果. 从图 5 可以看出, 与其他模型相比, JITL-XGBoost 模型预测值与实际值之间误差最小, XGBoost 模型的预测值也与实际值差距不大. 当过程状态发生变化时, 从图 6 可以看出, 与 JITL-XGBoost 模型相比, 全局模型 XGBoost 预测性能降低, 而 JITL-XGBoost 模型的预测结果依然非常准确, 这说明 JITL-XGBoost 模型可以较好地处理烧结过程中时变特性导致预测精度不足的问题.

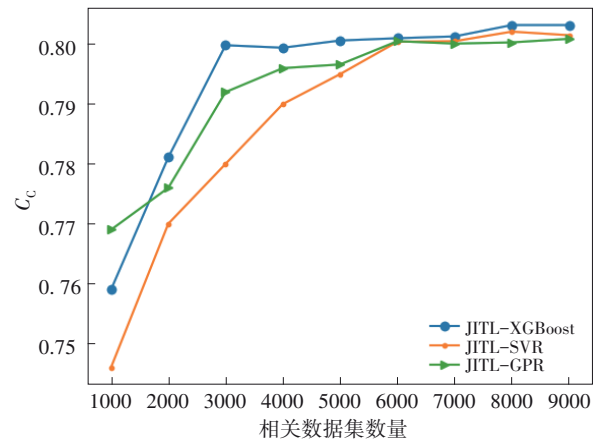


图 4 相关数据集数量对模型预测精度的影响

Fig. 4 Influence of the numbers of relevant data sets on prediction accuracy of the model

另外, 模型计算时间的长短决定模型是否可

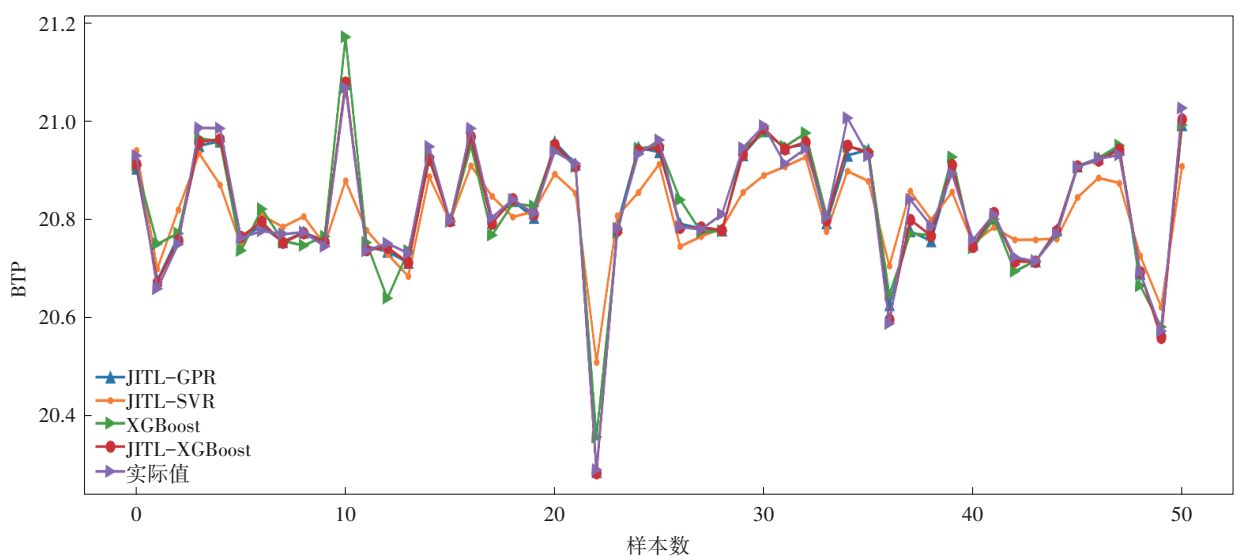


图 5 过程状态发生变化前 BTP 预测结果

Fig. 5 BTP prediction results before the change of process state

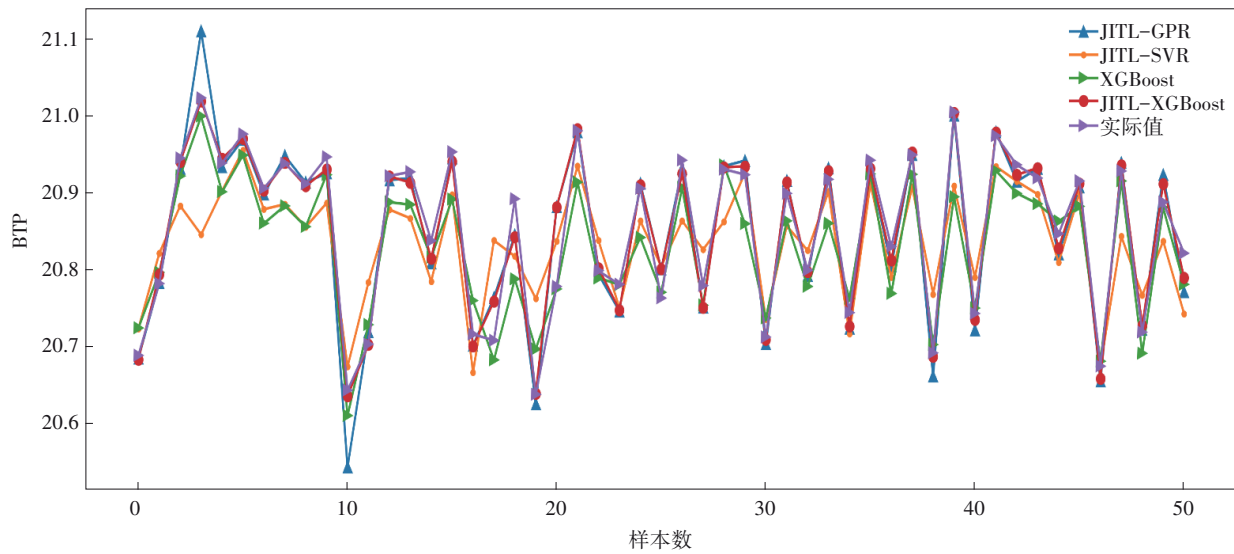


图6 过程状态发生变化后BTP预测结果

Fig. 6 BTP prediction results after the change of process state

以快速地得到BTP预测结果.模型的计算时间涉及相似度计算和局部模型开发,本文采用每个模型运行3次的结果,取这3次的平均值用于比较.经计算,计算时间在1.0 s和1.4 s之间,这种计算时间说明相关数据集数量为3 000用于预测是可以接受的,因为其远小于传感器测量的时间间隔.

不同预测模型的预测性能对比如表3所示.XGBoost模型预测精度较低,预测时间较长,是因为其采用全部数据,易被其他无关样本干扰,导致其对时变的BTP适应性较差;JITL-SVR模型对于高维数据易出现处理困难,导致模型预测性能下降;JITL-GPR模型对超参数选择较为敏感,选择不当也会导致预测精度下滑;而JITL-XGBoost是即时学习与梯度提升决策树相结合的模型,能够很好地处理以上模型存在的问题,因此,其预测性能较高.

表3 不同预测模型的预测性能对比
Table 3 Comparison of prediction performance of different prediction models

模型	e_{MA}	e_{MS}	e_{RMS}
JITL-GPR	0.064 9	0.013 8	0.117 4
JITL-SVR	0.057 5	0.006 9	0.083 3
XGBoost	0.034 2	0.027 3	0.165 2
JITL-XGBoost	0.029 9	0.003 8	0.061 6

注:黑体表示本文模型结果.

4 结 语

针对烧结合实际生产过程中的时变特性所导致预测精度不足的问题,本文采用JITL-XGBoost模型进行BTP预测.将可以适应烧结合实际生产过程的JITL和可以处理复杂非线性的XGBoost算法相结合来解决上述问题,并在此基础上,考虑了相关数据集对预测性能和预测时间的影响.通过与其他模型对比,本文所建模型的 e_{MA} , e_{RMS} 和 e_{MS} 均比其他3种模型低,说明所建模型预测能力较强.综上所述,本文所建模型为烧结合实际生产提供了一种新的评判方法,后续可以根据模型预测结果来调整烧结机参数,从而减少资源浪费.

参考文献:

- [1] Du S, Wu M, Chen L F, et al. A fuzzy control strategy of burn-through point based on the feature extraction of time-series trend for iron ore sintering process[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(4): 2357-2368.
- [2] Patisson F, Bellot J P, Ablitzer D, et al. Mathematical modelling of iron ore sintering process [J]. *Ironmaking & Steelmaking*, 1991, 18(2): 89-95.
- [3] Cumming M J, Thurlby J A. Developments in modelling and simulation of iron ore sintering [J]. *Ironmaking & Steelmaking*, 1990, 17(4): 245-254.
- [4] Chen X X, Shi X H, Tong C D. Multi-time-scale TFe prediction for iron ore sintering process with complex time delay[J]. *Control Engineering Practice*, 2019, 89: 84-93.
- [5] Du S, Wu M, Chen L F, et al. Prediction model of burn-through point with fuzzy time series for iron ore sintering process [J]. *Engineering Applications of Artificial Intelligence*, 2021, 102: 104259.

(下转第41页)