

引文格式:魏凯锋,祁凤琴,侯静茹,等.基于TCN-XGBoost的镍冶炼过程镍钼温度预测方法[J].中国有色冶金,2025,54(6):26-38.

WEI Kaifeng, QI Fengqin, HOU Jingru, et al. A nickel matte temperature prediction method in nickel smelting process based on TCN-XGBoost[J]. China Nonferrous Metallurgy, 2025, 54(6): 26-38.

基于TCN-XGBoost的镍冶炼过程镍钼温度预测方法

魏凯锋¹, 祁凤琴¹, 侯静茹¹, 李荻妍¹, 康建强², 高宇希³

(1. 镍钴共伴生资源开发与综合利用全国重点实验室, 甘肃 金昌 737100;

2. 金川集团镍钴股份有限公司 镍冶炼厂, 甘肃 金昌 737100; 3. 中南大学 自动化学院, 湖南 长沙 410083)

[摘要] 镍冶炼过程中出口镍钼温度的精准预测是工艺优化、设备安全保障与产品质量稳定的关键,但该过程存在多变量非线性耦合、参数动态波动的问题,制约了传统方法的应用效果。针对此,本文提出一种融合时序卷积网络(TCN)与XGBoost动态误差补偿的镍钼温度预测方法。首先,通过皮尔逊相关系数与F-score融合策略,从入炉物料组分、鼓风量、氧浓度等多维度工业参数中筛选关键特征,降低冗余干扰;其次,利用TCN的膨胀因果卷积捕捉特征间长短期时序依赖与非线性关系,生成温度初始预测;再引入XGBoost并构建一阶/二阶时间差分特征,学习预测残差演变规律以实现动态误差补偿。依托XGBoost分位数回归,能够实现镍钼温度的区间预测。实验结果表明,在点预测性能上,TCN-XGBoost模型较基线TCN模型精度显著提升,MAE从7.6277降至7.4941,MAPE从0.5928降至0.5842, RMSE从9.7913优化至9.5732, R^2 提升至0.4516,且优于LSTM、AGCRN等对比方法;在区间预测性能上,TCN-XGBoost的90%预测区间展现出“覆盖可靠、宽度紧凑”的平衡优势,对比TCN-LightGBM与TCN-KDE,其既能在温度平稳区间紧密包裹真实值,又能在剧烈波动区段可靠包容真实值变化,避免区间过度宽泛或覆盖缺失的问题。该方法对镍钼出口温度具有高精度预测能力,并能有效适应工业生产中的动态波动,可为镍冶炼过程的实时监测与工艺调控提供科学支撑。

[关键词] 镍冶炼; 镍钼温度; 精准预测; 时序卷积网络(TCN); XGBoost动态误差补偿; 点预测; 区间预测

[中图分类号] TF815; TP183 [文献标志码] A [文章编号] 1672-6103(2025)06-0026-13
DOI:10.19612/j.cnki.cn11-5066/tf.2025.06.003

0 引言

镍凭借其优异的耐腐蚀性、耐高温性及机械强度,在现代工业体系中扮演着不可替代的角色^[1]。在镍冶炼火法工艺中,镍钼温度是表征炉内反应程度以及工艺稳定性的重要指标,准确监测和预测镍钼温度,对于优化操作、提高产品质量以及降低能耗

都具有显著意义^[2]。

受工艺自身特性与现场环境制约,当前镍钼温度预测存在以下挑战:一是温度随历史状态变化,存在明显时序依赖;二是多源变量耦合复杂,非线性特征突出;三是高温粉尘环境下噪声和偏差显著。在这种复杂工况下实现镍钼温度的精准预测温度面临较大难度。

为实现冶金炉窑的温度预测,国内外学者进行了很多研究。例如徐吉等^[3]针对VD炉重点钢液温度预测任务,提出了结合冶金机理与贝叶斯优化的XGBoost算法(MM-BO-XGBoost)。王子昂等^[4]针对热镀锌带钢感应加热过程,利用相关性筛选输入特征,通过TPE-贝叶斯优化调参的GBDT模型,实现

[收稿日期] 2025-09-20 [修订日期] 2025-11-04

[录用日期] 2025-11-05 [出版日期] 2025-12-01

[第一作者] 魏凯锋(1996—),男,硕士,工程师,研究方向为有色冶金工艺建模与优化控制。

[基金项目] 镍钴共伴生资源开发与综合利用全国重点实验室科研项目(KY-YJ-04-2024)。

了满足生产要求的目标温度高精度预测。施树蓉等^[5]针对 RH 炉终点钢液温度预测,提出了一种递归特征消除结合 PSO 优化的 MLP 网络,提高了钢水终点温度预测精度和鲁棒性。Yanhua Yin 等^[6]面向锌冶炼过程,将滑动窗口与学习机制融入 PLS,构建了自适应温度预测模型。Yaozu Wang 等^[7]针对回转窑的温度场预测,利用遗传算法优化随机森林(GA-RF)建立了软测量模型。然而,镍冶炼过程数据特性复杂且易发生突变,现有机器学习方法在捕捉非线性耦合特征方面仍存在不足。

近年来,随着深度学习的迅速发展,研究者进一步将其应用于冶金过程建模与软测量^[8-10]。例如,韩洁等^[11]针对氧化铝蒸发过程的结垢参数预测问题,提出了一种基于时序约束模糊 C 均值聚类的多工况软测量方法。罗世华等^[12]提出基于 VMD-TCN-GRU 的组合模型,实现了对高炉透气性指数的高精度提前预测。尹刚等^[13]提出了一种基于 PID 搜索优化的 CNN-LSTM-Attention 预测方法,用于铝电解槽电解温度的精准预测。Jiayang Dai 等^[14]针对铝冶炼过程,提出了一种结合图网络与 LSTM 的炉温预测模型。在工业过程建模中,误差补偿作为提升模型性能的重要手段,已被广泛应用。例如,杜文莉等^[15]则结合 LS-SVM 与 ARMA 模型,提出了一种具有时序误差补偿能力的软测量建模方法。Kai Guo 等^[16]针对高炉过程建立了灰色系统预测模型并引入了误差补偿机制,有效实现了高炉煤气利用率预测。然而,现有基于误差补偿的预测方法仍存显著局限:该类方法聚焦于点预测,仅能输出单一确定性结果,难以量化原料成分波动、设备工况漂移及传感器噪声等固有不确定性所带来的风险。

冶金工业具有强耦合、高安全、连续生产等特点,对预测结果的不确定性边界量化要求更高。区间预测作为点预测的重要延伸,可为工业操作提供具备工程参考价值的“决策弹性空间”和“风险边界”,从而保障冶炼过程稳定运行。

针对上述挑战与工业需求,为实现镍冶炼过程中出口镍钨温度的高精度、高可靠性预测,本文提出一种融合时序卷积网络(TCN)时序特征提取与 XGBoost 动态误差补偿的镍钨温度预测方法。TCN 凭借卓越的时序特征提取能力,可深度刻画复杂工况下多参数的非线性耦合与动态演化规律;XGBoost

误差补偿模块则针对初步预测残差的变化特性进行实时校正,进一步提升整体预测精度与鲁棒性。此外,依托 XGBoost 的分位数回归能力,所提方法不仅可实现镍钨温度的确定性点预测,还能输出具备统计可靠性的区间预测结果,为实际生产中的工艺参数微调、异常风险预警与优化控制决策提供更全面的依据。

1 方法

1.1 总体框架

本文构建了基于 TCN 与 XGboost 误差补偿的镍冶炼过程镍钨温度预测方法。其总体框架如图 1 所示。首先采用融合皮尔逊相关系数与 F-score 的变量筛选方法,对镍钨冶炼过程中入炉物料组分、风量、氧浓度等多维度过程变量进行重要性评估,构建高区分度的输入变量集合。继而利用 TCN 网络对时序特征序列进行深度特征提取,通过膨胀因果卷积操作捕捉多变量间的非线性耦合关系及动态时序特征,输出初始预测结果。

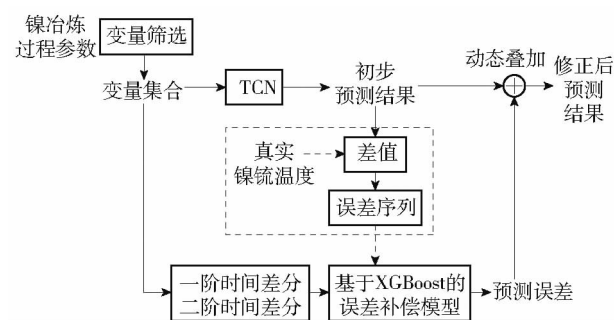


图 1 TCN-XGboost 镍钨温度预测框架

Fig. 1 Framework for nickel matte temperature prediction using TCN-XGboost

针对初始预测值与真实值的偏差,引入基于时间差分特征的 XGBoost 动态误差补偿模型:通过构建包含一阶/二阶时间差分的动态特征空间,挖掘预测误差在时间维度上的演变规律,实现对非线性误差模式的自适应学习。最终将 TCN 网络的初始预测结果与 XGBoost 输出的补偿误差进行动态融合,形成多模型协同优化的预测机制,有效抑制冶炼过程中物料波动、工艺参数漂移等因素对预测精度的干扰。

1.2 皮尔逊-Fscore 融合特征筛选

镍冶炼生产线的 DCS 系统实时监测并记录上

百个过程变量,涵盖加料、供氧、除尘等多个子系统,例如镍精矿量、压缩空气总管流量与温度、氧气总管温度与计量值等。高维、异质的数据虽蕴含镍钼温度变化的丰富信息,但同时带来特征冗余、噪声传播、计算开销和“维度灾难”问题;其中相当一部分参数对温度几乎无显著影响,若直接全部输入模型,势必降低预测精度并增加训练难度。为此,本文采用皮尔逊-Fscore 融合去冗(CF³R)特征筛选策略,在保证关键信息完整性的前提下,系统压缩输入维度,为后续建模提供高质量、低冗余的特征子集。具体流程如下所述。

1) 皮尔逊相关性筛选。对每一候选变量 x_i 计算其与目标变量(镍钼温度) y 的皮尔逊相关系数 r_i 。考虑到正负相关对预测贡献等价,取绝对值 $|r_i|$ 并按降序排列,保留排名前 50% 的特征集 S_{corr} ;经统计,本文数据集中该筛选比例对应的最小相关系数绝对值为 0.2007,即仅保留与镍钼温度关联强度不低于该阈值的特征,初步剔除弱关联变量。

2) F-Score 线性统计筛选。对所有特征计算 F-Score(衡量特征对 y 的线性解释强度),同样选取排名前 50% 的特征集 S_F 。经统计,本文数据集中该筛选比例对应的最小 F-Score 值为 102.203,以此作为线性贡献度的筛选阈值,进一步剔除对温度解释能力较弱的冗余特征,确保筛选后的特征具备较强的线性区分能力。

3) 综合评分。令交集 $S_{cand} = S_{corr} \cap S_F$,对 S_{cand} 中的每个特征,分别记录其在皮尔逊相关系数与 F-Score 两个指标上的综合得分,计算排名,保留前 25% 的集合 S_{core} 。

4) 高相关去冗余。计算 S_{core} 内两两特征的相关系数矩阵 $R = [r_{ij}]$ 。若 $[r_{ij}] \geq 0.95$ 则判定为高度共线。按照个特征与目变量的 r_i 从小到大依次遍历,若当前特征与已保留特征存在高共线则舍弃,否则保留。得到去冗集合 S_{CFR} 。

5) 业务关键特征强制保留。基于冶金机理与生产经验,对于“精矿料量”等关键变量 S_{key} ,无论筛选结果如何均予以保留。最终特征集合定义为 $S_{final} = S_{CFR} \cap S_{key}$ 。

通过上述五步,原始 N 维特征被压缩为 $S_{final} \ll N$,既保留了对镍钼温度最具解释力的信息,又显著降低了冗余度,为模型训练提供了高质量输入。

1.3 TCN

TCN 是一类专门针对时序数据设计的卷积网络,由 Bai 等^[17]在 2018 年提出。与循环网络不同,TCN 完全由卷积层构成,以一维因果卷积(Causal Conv)确保输出时刻 t 仅依赖于 t 及之前的输入,同时借助残差连接(Residual Block)堆叠深层网络,兼顾长时依赖建模与训练稳定性。其核心思想是用空洞卷积快速扩大感受野,使模型在不增加参数量的情况下捕获远距离依赖。TCN 结构如图 2 所示。

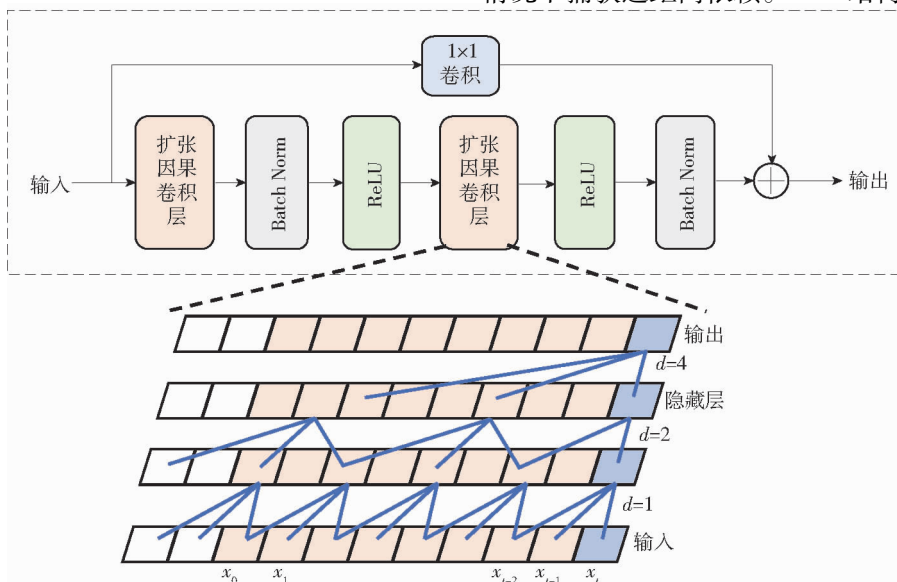


图2 TCN 框架

Fig. 2 TCN framework

1.3.1 因果卷积

在时序建模里,必须保证网络在计算某一时刻 t 的输出时,只能访问到时刻 t 及其之前的输入。因果卷积正是为此设计,见式(1)。

$$y_t = \sum_{i=0}^{k-1} w_i z_{t-i}, t=1, \dots, T \quad (1)$$

式中: y_t 表示时间步 t 的输出; k 为卷积核大小; z_t 表示时间步 t 输入特征; w_i 为卷积核的第 i 个参数。

与循环结构相比,因果卷积仍保持卷积一次性并行计算整段序列的优势;与普通一维卷积相比,其能满足严格的时序因果性。

1.3.2 空洞卷积

为了让网络看到更长的历史,直接增大卷积核或在时域下采样都会增加参数或损失分辨率。空洞卷积的做法是把卷积核内的采样间隔“拉伸”成给定的膨胀率 d ,见式(2)。

$$y_t = \sum_{i=0}^{k-1} w_i z_{t-id}, d \geq 1 \quad (2)$$

当 $d=1$ 时,其会等效为普通因果卷积;当 $d>1$ 时,相邻卷积权重对应的输入距离被放大为 d 个时间步,感受野增长到 $(k-1)d+1$ 。

1.3.3 残差块

深层空洞卷积层层堆叠时,梯度在反向传播过程中可能急剧衰减或爆炸,训练容易失稳。为化解这一难题,TCN 采用残差连接:将原始输入 z 与经过卷积得到的输出 $F(z)$ 相加。若输入和输出通道数不同,再借助一层 1×1 卷积做维度对齐,得到式(3)。

$$y = F(z) + \text{Conv}_{1 \times 1}(z) \quad (3)$$

式中: z 为残差块的输入序列; $F(z)$ 表示经过一系列卷积后的非线性变换; $\text{Conv}_{1 \times 1}(z)$ 为 1×1 卷积; y 为残差块的输出。

残差连接使网络在需要时可以轻松学习恒等映射,使梯度能够毫无阻碍地跨越多层传播,同时保留 $F(z)$ 提供的复杂特征变换,实现深而稳的时序建模。

1.4 XGBoost 误差补偿

在镍冶炼工业生产场景中,侧吹炉的运行是一个复杂且动态的过程。入炉物料成分的多变性、入炉量的不稳定以及风量的实时波动,使得整个镍冶炼过程充满了不确定性。这些因素的波动范围难以精确预测,进而导致镍硫温度的预测结果存在不可

忽视的偏差。为了有效应对这一问题,进一步提升镍硫温度预测的准确性,本文提出了基于 XGBoost 的动态误差补偿模型,旨在对预测结果进行精准修正。

本文基于主干网络对镍冶炼关键指标进行初步预测。通过收集训练集与测试集的预测值和真实值,构建了模型误差的训练样本集。对于训练过程中产生的预测结果 $\hat{y}_i$ 和真实值 y_i ,提取 $\hat{y}_i$ 与 y_i 的差值作为后续 XGBoost 误差模型的目标标签,为后续补偿提供依据。

在误差补偿模型构建过程中,为了深入挖掘镍冶炼过程数据中的动态特征,本文采用了一阶和二阶时间差分的方法进行特征提取,并将提取后的特征用作误差补偿模型的输入。由于镍冶炼过程中的数据不仅包含当前时刻的状态信息,还蕴含相邻时刻以及跨时刻的变化趋势,故通过差分运算可获得更丰富的动态特征。设原始时间序列数据表示为式(4)。

$$X = x_1, x_2, \dots, x_T \quad (4)$$

式中: x_t 表示第 t 时刻的数据值; T 为时间序列的长度。

一阶差分 Δx_t 可表示为式(5),其能够突出数据在相邻时刻的变化情况,反映出数据的即时变化率。

$$\Delta x_t = x_t - x_{t-1}, t=2, \dots, T \quad (5)$$

Δx_t 能够反映相邻时刻间的即时变化率。基于一阶差分的二阶差分则进一步定义为式(6),二阶差分用于捕捉数据变化的加速度信息。

$$\begin{aligned} \Delta^2 x_t &= \Delta x_t - \Delta x_{t-1} = \\ &= (x_t - x_{t-1}) - (x_{t-1} - x_{t-2}) = \\ &= x_t - 2x_{t-1} + x_{t-2}, t=3, \dots, T \end{aligned} \quad (6)$$

将一阶和二阶差分特征相结合,即可构建更全面、细致的特征向量(式(7)),为后续模型训练提供更丰富、有效的信息。

$$F = [\Delta x_1, \Delta^2 x_1, \Delta x_2, \Delta^2 x_2, \dots, \Delta x_T, \Delta^2 x_T] \quad (7)$$

XGBoost,即极端梯度提升树,是一种基于梯度提升框架的先进机器学习模型^[18],隶属于集成学习中的提升树(Boosting Tree)范式。与传统 GBDT 相比,XGBoost 在目标函数中结合二阶导数信息以及 L1、L2 正则化项,大幅提升了梯度估计精度并有效抑制了过拟合。通过对训练集样本的迭代学习,XGBoost 能够捕捉各时间段误差的非线性特征,从而为后续补偿提供精确的误差预测。

模型的优化目标函数可表示为式(8)。

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (8)$$

式中: $(\hat{y}_i, y_i)$ 表示第 i 个样本的损失函数; $\hat{y}_i$ 是预测值; y_i 是真实值; $\Omega(f_k)$ 是第 k 棵树的正则化项。

将损失函数进行二阶泰勒展开, 可获得更准确的梯度信息, 见式(9); 其中 g_i 、 h_i 表达式分别见式(10)(11)。

$$L^{(t)} \approx \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_i) \quad (9)$$

$$g_i = \frac{\partial}{\partial \hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)}) \quad (10)$$

$$h_i = \frac{\partial^2}{\partial (\hat{y}^{(t-1)})^2} l(y_i, \hat{y}^{(t-1)}) \quad (11)$$

式中: g_i 为在第 $t-1$ 轮预测下第 i 个样本的一阶梯度; h_i 为在第 $t-1$ 轮预测下, 第 i 个样本的二阶梯度。

为了适应镍冶炼过程在不同时间段内可能出现的大幅波动或相对平稳的情况, 模型引入了基于局部标准差的动态误差补偿策略。通过在训练阶段统计预测误差在滑动窗口内的标准差序列 $\sigma_{\text{local}}(t)$, 并与全局标准差为 $\sigma_{\text{global}}(t)$ 进行比较, 计算得到全局阈值 $\text{threshold}_{\text{value}}$ 。

若某时间段的局部波动程度 ($\sigma_{\text{local}}(t)$) 明显高于全局阈值, 则表明该时段更为剧烈, 模型将启用更高的补偿权重 w_{high} , 否则维持保守的修正系数 w_{low} 。

将主干模型的预测值 $y_{\text{base}}(t)$ 与通过 XGBoost 预测得到的误差 $e_{\text{xgb}}(t)$ 进行融合, 得到最终的补偿预测结果, 见式(12)。

$$y_{\text{final}}(t) = y_{\text{base}}(t) + w(t) \times e_{\text{xgb}}(t) \quad (12)$$

式中: $w(t)$ 根据局部波动情况在 $\{w_{\text{high}}, w_{\text{low}}\}$ 之间动态切换, 体现动态补偿策略对不同波动去预测的差异化处理。

2 实验设置

2.1 数据处理

本文所用数据来源于国内某冶炼厂镍冶炼生产线, 涵盖 2024 年 7 月至 10 月的实时生产过程。在其冶炼过程中, DCS 系统监测的过程变量多达数百项, 其中车间氮气总管压力、循环水总量等对镍钨温度的影响极为有限。为剔除冗余信息并保留与温度预测最为相关的特征, 本文采用皮尔逊-Fscore 融合去冗策略对所有变量进行筛选, 最终得到 10 个关键

过程变量, 作为镍钨温度预测模型的输入, 具体见表 1。

表 1 输入过程变量

Table 1 Input process variables

序号	过程变量名称	序号	过程变量名称
No. 1	块煤入料量	No. 6	车间天然气总管温度值
No. 2	冷料入料量	No. 7	天然气熔炼喷枪流量 1 [#]
No. 3	溶剂入料量	No. 8	天然气熔炼喷枪流量 2 [#]
No. 4	镍精矿入料量	No. 9	熔炼富氧支路流量
No. 5	返料入料量	No. 10	熔炼富氧支路压力

在完成变量筛选后, 进一步对所选特征进行异常值检测。具体方法采用四分位间距 (IQR) 法, 以第 25 百分位数 (Q1) 和第 75 百分位数 (Q3) 为界限, 计算各特征的异常值比例。检测结果如图 3 所示, 其中部分变量的异常比例超过 5%, 如“天然气熔炼喷枪流量 1[#]” (6.22%) 和“熔炼富氧支路压力” (5.19%), 而“冷料入料量”和“天然气熔炼喷枪流量 1[#]”等变量的异常比例较低, 仅为 0.09%。

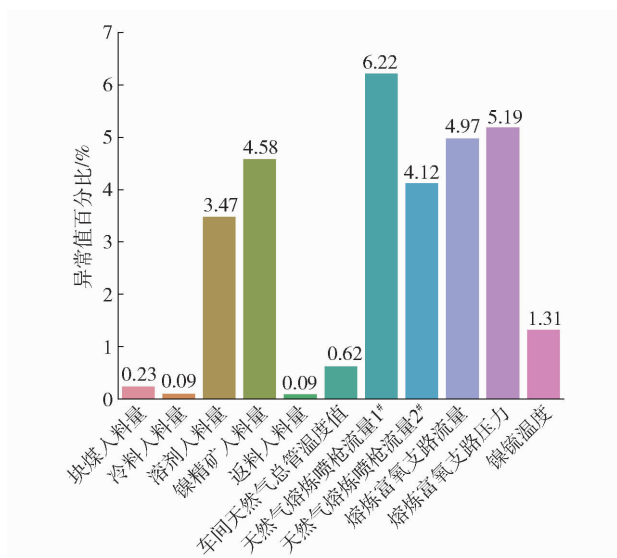


图 3 变量异常值比例

Fig. 3 Proportion of outliers in variables

在后续实验中, 本文并未对检测出的异常值进行剔除或修正。一方面, 异常值本身可能反映实际生产工况下的波动信息; 另一方面, 保留这些异常点有助于验证所提出预测模型在真实复杂环境下的泛化能力和鲁棒性。

镍钨冶炼过程中, 进料量、富氧浓度和鼓风量等

过程变量每 1 min 采集一次,而镍铈温度仅每 1 h 测量一次。考虑到温度演化呈显著的时间自相关性,本文在模型输入中同时引入前 3 h 的温度作为自回归项,并将连续 180 min 内的分钟级过程参数作为时序输入,用于预测当前小时温度 y_t 。设 x_t 为第 t min 的过程变量数据,则输入-输出映射为式(13)。

$$\{x_{t-179}, x_{t-178}, \cdots, x_t, y_{t-1}, y_{t-2}, y_{t-3}\} \rightarrow y_t \quad (13)$$

受限于镍铈温度标签与过程变量存在采样周期不一致的问题,为确保模型输入特征与预测标签在时间维度上精准对齐,本文采用时间窗口划分策略对原始数据进行处理:以 60 min 为一个固定时间窗口,在每个窗口内统一采用该小时内的镍铈温度值作为对应窗口的标签值,从而消除因采样频率差异导致的时间错位问题。同时,为避免不同输出样本之间出现标签重叠、确保每个样本标签的独立性,窗

口滑动步长严格设置为 60 个时间步(即与窗口长度保持一致),每滑动一个步长生成一条独立样本,使得相邻样本的标签分属不同小时段、互不重叠;而在输入特征层面,允许相邻窗口在过程变量数据上存在重叠,以最大程度保留时序数据中的连续依赖关系。

此外,考虑到模型构建过程中需引入 y_{t-1} 、 y_{t-2} 、 y_{t-3} 三个自回归温度,三个历史时刻的镍铈温度作为自回归特征(即利用前 3 h 的温度数据辅助预测当前时刻温度),样本构造的时间起点需从原始数据的第 4 h 开始($t \geq 3$)。具体窗口构建如图 4 所示。本文实验所用数据集包含 89 941 条分钟级原始数据,在完成时间窗口划分前,按 8:2 划分训练集与测试集。经过上述处理最终得到训练集样本数量为 1 197,测试集样本数量为 300。

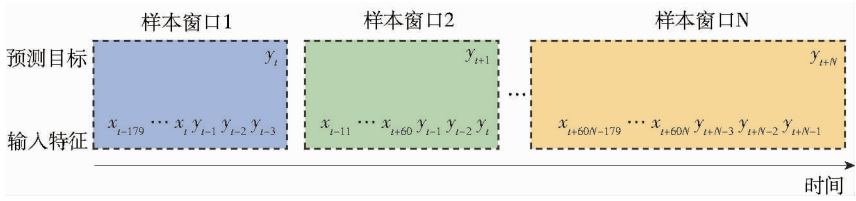


图4 样本窗口划分

Fig.4 Sample window partitioning

实验过程中,本文采用真实数据计算各项性能指标;在预测结果的可视化展示中,由于数据保密性要求,本文对结果进行了数据标准化处理。标准化公式见式(14)。

$$x' = \frac{x - \mu}{\sigma} \quad (14)$$

式中: x 为原始数据; μ 为数据均值; σ 为数据标准差; x' 为标准化后的数据。

2.2 参数设置

本文所有实验在 Windows 平台(Intel Core i9-13900, 32GB RAM)上完成。算法采用 Python 3.9 实现,模型部分基于 PyTorch 1.9.0 构建;TCN 与 XGBoost 的关键超参数见表 2、表 3,模型超参数均通过实验调优得到。为保证实验的科学性与结果的可靠性,模型训练过程中对超参数选择进行了充分考虑;并在训练过程中引入 Dropout 以及早停机制抑制过拟合;此外,Adam 优化器的初始学习率设为 0.001,结合学习率衰减策略,使模型在收敛过程中更加平稳。

表2 TCN 超参数设置

Table 2 TCN hyperparameter settings

参数	数值	参数	数值
输入维度	11	激活函数	ReLU
输出维度	1	Dropout	0.2
残差块数量	4	损失函数	MSE
膨胀率	{1, 2, 4, 8}	优化器	Adam
卷积核大小	3	Bacth_Size	32
通道大小	32	训练次数	100

表3 XGBoost 超参数设置

Table 3 XGBoost hyperparameter settings

参数	数值	参数	数值
最大树深	6	n_estimators	100
学习率	0.05	colsample_bytree	0.6
目标函数	平方误差	subsample	0.6

2.3 评估方法

为全面评估所提 TCN-XGBoost 模型的预测性

能,本文从点预测精度与区间预测可靠性两个维度设计评估指标:点预测聚焦单一预测值与真实值的偏差,区间预测则量化预测不确定性的刻画能力。

点预测采用平均绝对误差(MAE)、平均绝对百分比误差(MAPE)、均方根误差(RMSE)、以及相关系数(R^2)作为模型的评价指标,具体公式如式(15)~(18)所示。

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (15)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (18)$$

式中: y_i 表示真实值; $\hat{y}_i$ 表示预测值; $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$; N 为样本数。

为验证预测区间的可靠性,本文引入平均预测区间覆盖率(PICP)与平均预测区间宽度(PINAW)两个指标,分别衡量区间对真实值的覆盖能力与区间自身的紧凑性。

PICP反映真实值被预测区间覆盖的比例,是衡量区间可靠性的核心指标,定义见式(19)。若PICP越接近设定的名义覆盖概率(本文为90%),则说明预测区间对真实值的覆盖可靠性越高。

$$PICP = \frac{1}{N} \sum_{i=1}^N I(y_i \in [Q_{l,i}, Q_{u,i}]) \quad (19)$$

式中: $Q_{l,i}$ 表示第*i*个样本的预测区间下界; $Q_{u,i}$ 对应预测区间上界; I 为指示函数,当真实值 y_i 落在区间 $[Q_{l,i}, Q_{u,i}]$ 内时,取值为1,否则为0。

PINAW用于衡量预测区间的平均宽度,并通过真实值的取值范围进行标准化,以消除量纲影响,定义见式(20)。PINAW越小,表示预测区间越紧凑,越能准确刻画镍钼温度的波动特征。

$$PINAW = \frac{1}{N \cdot (y_{\max} - y_{\min})} \sum_{i=1}^N (Q_{u,i} - Q_{l,i}) \quad (20)$$

式中: N 表示样本数量; $y_{\max}$ 与 $y_{\min}$ 分别为所有样本真实值的最大值与最小值; $(Q_{u,i} - Q_{l,i})$ 为第*i*个样本的预测区间宽度。

3 结果分析

3.1 异常值处理有效性实验

为提升模型的泛化能力与鲁棒性,本文在数据预处理阶段未对数据中存在的异常值进行处理。为系统验证异常值对模型预测精度的影响,本文设计并开展异常值处理有效性实验。首先,采用四分位间距(IQR)法进行异常值识别,以数据的第25百分位数(Q1)与第75百分位数(Q3)为界定边界,将超出该边界的样本定义为异常值。在模型训练阶段,将训练集中的异常值替换为对应特征的IQR上下限;为模拟真实生产场景,确保评估结果的真实性,测试集未进行异常值处理。实验所得模型性能指标对比结果见表4。由对比结果可知在本研究的数据场景下,训练集保留异常值的处理方案显著优于异常值替换方案。这是由于本实验所采用数据集集中的异常值并非无效噪声,而是数据分布的重要组成部分。保留训练集异常值可使模型充分学习数据完整特征。

表4 异常值处理有效性实验结果

Table 4 Experimental results of outlier handling effectiveness

训练数据处理	MAE	MAPE	RMSE	R^2
去除异常值	8.086 8	0.630 3	10.404 6	0.387 0
未去除异常值	7.494 1	0.584 2	9.573 2	0.451 6

3.2 点预测结果与分析

如图5所示,通过对补偿前后的预测曲线进行细致对比分析,可清晰观察到:在样本80~120、180~220及240~260区段,补偿后的曲线峰谷与真实温度曲线实现了更紧密的贴合。这一现象表明,差分特征在捕捉镍冶炼过程中因突发波动引发的瞬时偏差方面发挥了关键作用,有效提升了模型对动态变化的响应精度。而在样本0~60和140~170等区间,补偿后的曲线几乎与基线完全重合,充分说明该框架在优化动态响应的同时,并未对原有稳态精度造成任何破坏,实现了动态与稳态表现的平衡。

插补前后性能指标见表5,XGBoost残差补偿策略为TCN基线模型带来了稳定且显著的性能增益。MAE从7.627 7降至7.494 1,MAPE从0.592 8降至0.584 2, RMSE从9.791 3降至9.573 2, R^2 从

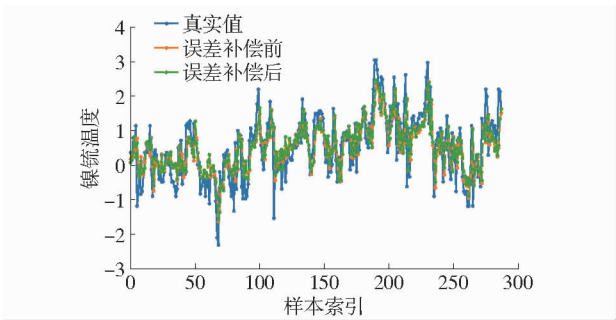


图 5 误差补偿前后预测结果

Fig. 5 Prediction results before and after error compensation

0.441 5 提升至 0.451 6。表明预测误差的显著降低,预测精度得到提升。

表 5 误差补偿前后性能指标

Table 5 Performance metrics before and after error compensation

方法	MAE	MAPE	RMSE	R^2
误差补偿前	7.627 7	0.592 8	9.791 3	0.441 5
误差补偿后	7.494 1	0.584 2	9.573 2	0.451 6

3.3 对比实验

为系统验证 TCN-XGBoost 模型的预测有效性,本文设计多组对比实验,将其与三类典型模型展开性能比较:其一为经典时序预测与传统机器学习模型,包括 LSTM^[19]、XGBoost、及图神经网络模型 AGCRN^[20]进行对比;其二为基于不同基模型构建的集成模型,引入 LSTM-LightGBM、TCN-LightGBM 等作为额外对比对象。各模型的性能指标结果见表 6。从单模型性能对比来看:LSTM 的预测性能表现最差,其 MAE 为 9.836 9、MAPE 为 0.776 3、RMSE 达 12.570 9, R^2 仅 0.105 2,整体呈现显著的预测误差与较弱的数据拟合性能;XGBoost 虽在 LSTM 基础上实现预测性能提升,MAE 降至 8.900 1、MAPE 降至 0.710 1、RMSE 为 11.264 7, R^2 提升至 0.264 9,但预测误差仍处于较高水平;AGCRN 进一步优化预测精度,MAE 降至 7.829 6、RMSE 降至 9.947 7、MAPE 降至 0.614 9, R^2 达到 0.412 2,不过其性能与 TCN-XGBoost 相比仍有差距。

在集成模型对比层面:LSTM-XGBoost 与 LSTM-LightGBM 两类以 LSTM 为基模型的集成模型,其预测误差未得到降低,反而出现明显上升(MAE 为

表 6 对比实验性能指标结果

Table 6 Results of performance metrics from comparative experiments

模型	MAE	MAPE	RMSE	R^2
LSTM	9.836 9	0.776 3	12.570 9	0.105 2
XGBoost	8.900 1	0.710 1	11.264 7	0.264 9
AGCRN	7.895 6	0.614 9	10.073 4	0.412 2
TCN-LightGBM	7.539 1	0.586 2	9.598 0	0.448 7
LSTM-XGboost	10.388 0	0.661 2	12.746 6	0.080 0
LSTM-LightGBM	10.408 3	0.664 1	12.772 6	0.076 3
Ours	7.494 1	0.584 2	9.573 2	0.451 6

10.4, RMSE 为 12.7)。这表明在镍铑温度这类强波动性时序预测任务中,LSTM 在集成框架下的鲁棒性存在不足,难以有效捕捉复杂波动特征,导致模型泛化性能难以保障。相较而言,以 TCN 作为基模型的集成方案可更稳定地提取时序数据中的深层依赖特征,进而使 TCN-XGBoost 在相同串联集成模式下展现出更优的适应性与鲁棒性,充分验证了该模型针对强波动镍铑温度时序预测任务的优异预测性能。

预测结果可视化如图 6 所示。LSTM 模型的绿色预测曲线与蓝色真实值曲线在多段序列区间存在明显偏离,难以精准刻画镍铑温度的动态变化趋势;XGBoost 模型的橙色曲线虽能在部分序列区段捕捉真实值的大致变化走向,但整体偏差仍较为显著,无法精准拟合真实温度的波动特征;AGCRN 模型的红色曲线虽较前两者在拟合精度上有所提升,然而与真实值曲线相比,仍存在一定程度的偏移,未能细致匹配温度的实时动态变化。相比之下,TCN-XGBoost 模型的黄色预测曲线与蓝色真实值曲线贴合程度最高,无论是在温度波动剧烈的区间,还是相对平稳的序列段,均能更精准地响应真实温度的变化趋势,充分展现出对镍铑出口温度动态演化过程更强的特征捕捉能力与场景适应性。

3.4 显著性检验实验

为量化不同模型间预测性能的差异并确保统计结论的可靠性,本文采用配对 Wilcoxon 符号秩检验进行分析:以各模型预测值与真实值的绝对误差作为性能差异的核心量化指标,因所有对比模型均在同一数据集的相同样本上完成预测,各组误差数据具备配对属性;同时,为规避多组模型间多重比较可能导致的假阳性问题,进一步采用 FDR-BH (False

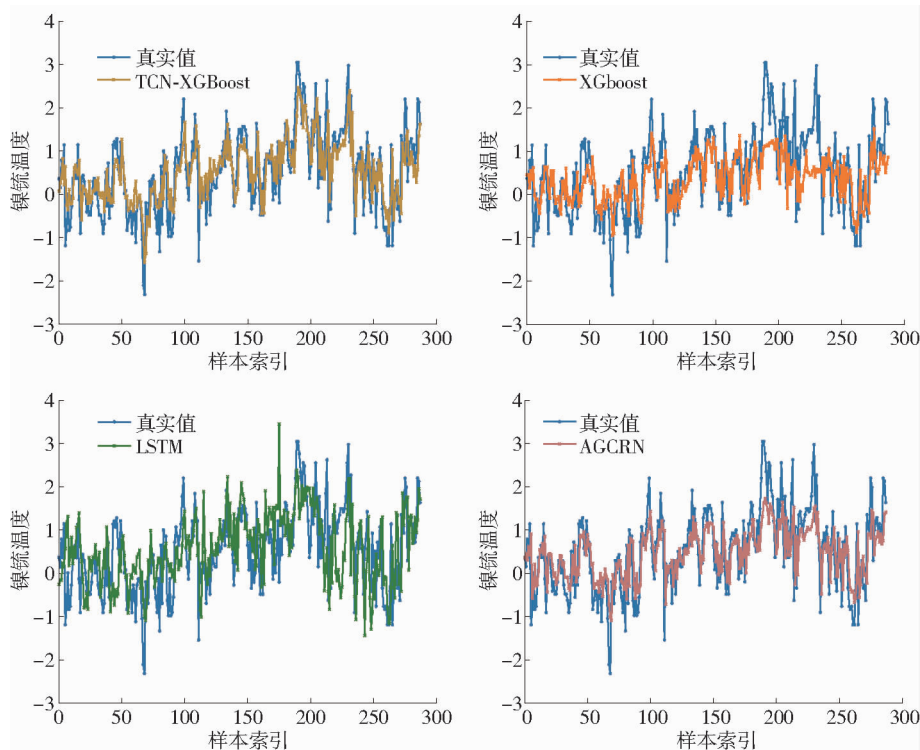


图6 对比实验预测结果

Fig. 6 Prediction results of comparative experiments

Discovery Rate-Benjamini-Hochberg) 方法对原始 P 值进行矫正,最终基于矫正后的 P 值判断模型性能差异的显著性。

Wilcoxon 显著性检验结果如图 7 所示。LSTM 与多数对比模型间的差异具有高度统计显著性 (p 值趋近于 0),直观反映出其预测性能的明显劣势;XGBoost 与基于深度学习的方法 (如 LSTM、AGCRN、TCN) 间的差异同样具有统计显著性,印证了深度学习在时序特征提取环节的优势。AGCRN 与 TCN-XGBoost 模型间的差异无统计显著性 ($p = 0.573$),尽管二者在结构上各有侧重,前者利用图卷积捕获空间依赖,后者通过时序卷积提取关键特征,但都能在本数据集上充分挖掘时空信息,因而误差表现趋于等效,并显著优于 LSTM 与 XGBoost。TCN-LightGBM 与 TCN-XGBoost 模型间的差异接近统计显著性边界 ($p = 0.052$),表明二者性能相近,而 TCN-XGBoost 模型在综合性能指标上略具优势。

3.5 泛化实验

为验证本文所提模型的泛化能力。本文利用该冶炼厂镍冶炼产线于 2024 年 12 月的实时生产数据进行测试。模型复用由 2024 年 7 月至 10 月数据训

练得到的模型。结果见表 7。所提模型在 2024 年 12 月的实时数据上,MAE 为 10.604 6,MAPE 为 0.822 0, RMSE 为 14.276 8, R^2 为 0.342 6,整体呈现出较好的预测精度。模型预测可视化如图 8 所示。其中预测曲线能够紧密跟随真实值曲线的波动趋势。不过在部分样本点处,预测值与真实值仍存在一定偏差,这可能是由于 12 月生产过程中出现了一些训练数据中未充分涵盖的特殊工况或干扰因素,但从整体泛化表现来看,模型仍具备在新数据场景下的有效预测能力,可满足工业生产中镍流温度预测的需求。

3.6 区间预测结果与分析

为量化镍流温度预测的不确定性并实现区间预测,本文针对 TCN-XGBoost 模型,利用 XGBoost 的分位数回归能力,选取分位数 τ 为 0.05, 0.50, 0.95 训练三组误差补偿模型,其中 $\tau = 0.5$ 对应预测分布的中位数 (点预测结果), $\tau = 0.05$ 与 $\tau = 0.95$ 则构成双侧区间的上下界,理论上落在区间 $[Q_{0.05}, Q_{0.95}]$ 内的概率约为 90% (即置信度为 90%);同时,本文还引入 TCN-LightGBM (TCN 提取特征后由 LightGBM 执行分位数回归)、TCN-KDE (基于 TCN 预测残差

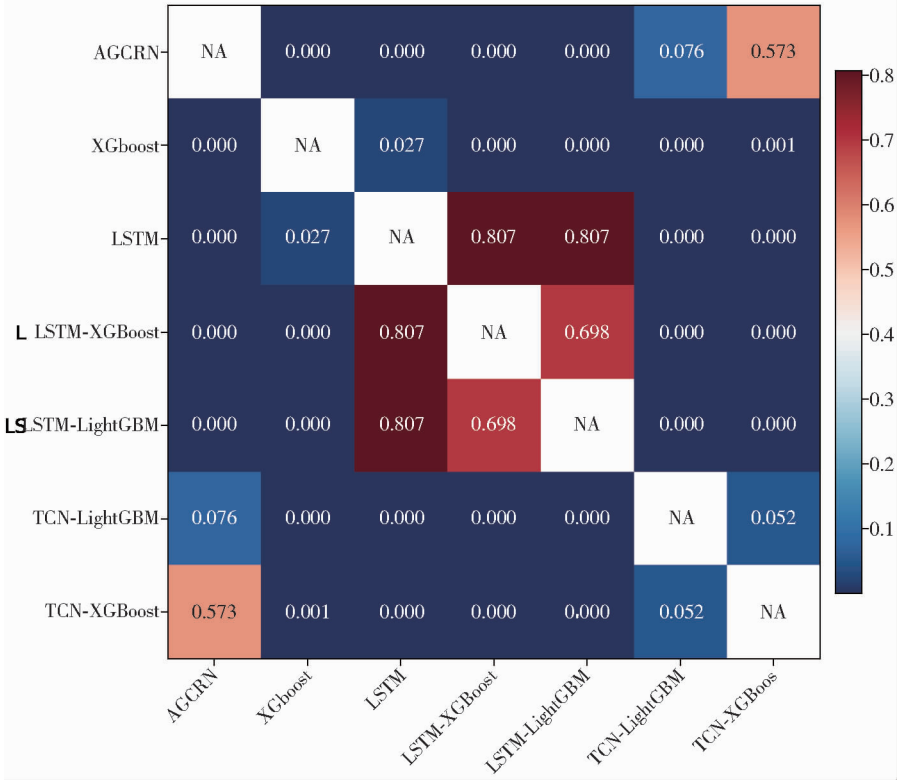


图 7 Wilcoxon 显著性检验矩阵

Fig. 7 Wilcoxon significance test matrix

表 7 泛化实验性能指标结果

Table 7 Performance metrics of generalization experiments				
方法	MAE	MAPE	RMSE	R^2
TCN-XGBoost	10.604 6	0.822 0	14.276 8	0.342 6

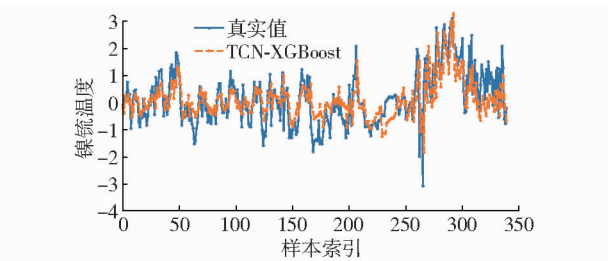


图 8 泛化实验预测结果

Fig. 8 Prediction results of generalization experiment

的核密度估计推导区间)两种方法开展对比,以全面验证所提方法在区间预测上的优越性。

区间预测可视化结果如图 9 所示(图中蓝色曲线为镍钼温度真实值,彩色区域为各模型的 90% 预

测区间)。TCN-LightGBM 的橙色预测区间对真实值的覆盖能力有限,在样本索引 100 附近的温度低谷、200 附近的高峰等波动剧烈区段,真实值多次超出区间范围,显露出对强波动时序的适应缺陷;TCN-KDE 的粉色预测区间虽提升了对真实值的覆盖频次,但区间宽度显著增大,过度宽泛的区间削弱了对温度动态变化的精准刻画能力;相比之下,TCN-XGBoost 的绿色预测区间在多数情况下能紧密包裹真实值,即便在温度剧烈波动的区间,也能较好适配变化趋势,展现出更优的覆盖贴合度。

各方法的定量指标见表 8。由于镍钼呈流体态,现场测温需由工人以热电偶插入流体完成,受测点位置偏差及仪器精度限制,本身存在约 $\pm 10\text{ }^{\circ}\text{C}$ 的测量误差(归一化后为 0.015),该值可视为理想合理的区间宽度参考。然而,镍钼温度同时受入炉原料成分波动、鼓风参数动态调整及设备工况漂移等多重因素耦合影响,数据呈现显著的动态波动性,从而导致三类模型的 PICP 均未能达到 90% 的名义覆盖概率目标,但不同方法在区间预测性能上的差异依然明显。具体而言,TCN-LightGBM 的 PINAW 为

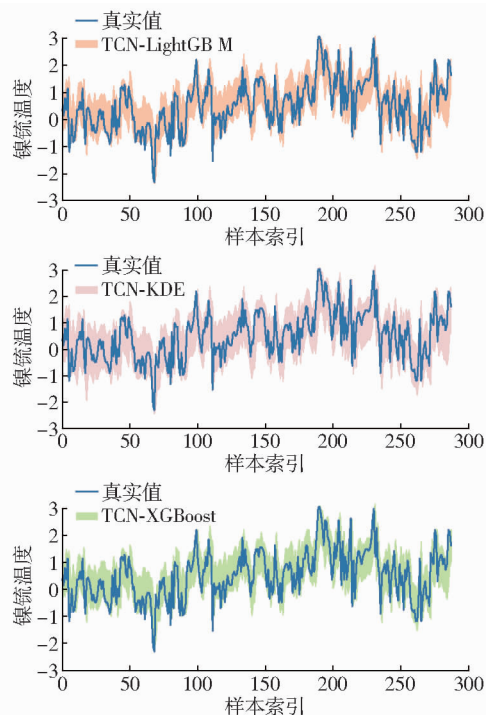


图9 区间预测结果

Fig.9 Prediction interval results

0.233 5,虽实现了最紧凑的区间,但相较于理想区间宽度偏窄,PICP 仅为 0.638 9,覆盖率不足;TCN-KDE 的 PICP 提升至 0.767 4,但其 PINAW 扩展至 0.296 9,明显超过合理区间范围,区间过宽,难以刻画温度的真实波动;相比之下,TCN-XGBoost 的为 0.236 6,与理想值接近,且 PICP 相较 TCN-LightGBM 有所提升,实现了覆盖率与区间宽度之间的更优平衡。由此可见,TCN-XGBoost 在保证 PINAW 合理性的同时,兼顾了区间的可靠性与精准性,整体预测性能更为突出。

表8 区间预测性能指标结果

Table 8 Results of interval prediction performance metrics

模型	PICP	PINAW
TCN-LightGBM	0.638 9	0.233 5
TCN-KDE	0.767 4	0.296 9
TCN-XGboost	0.649 3	0.236 6

4 未来展望与结论

4.1 未来展望

本文所提的 TCN-XGBoost 预测方法,已实现镍

冶炼过程中出口镍硫温度的点预测与区间预测,为镍硫温度预测提供了有效支撑。但结合实际冶炼场景的复杂性与精准化需求,其仍然存在优化与扩展的空间。

首先,在镍冶炼过程中,受物料类型变化与工艺调整的影响,冶炼工况在一定时间内可能会发生较大变化,而当前研究尚未充分考虑工况变迁对镍硫温度预测的影响。在未来,应考虑引入工况识别模块,通过构建工况特征指标体系,结合聚类算法实现对不同冶炼工况的实时划分;在此基础上,设计动态参数调整机制,使模型能够匹配工况变化规律,提升跨工况预测的稳定性。

其次,本文所提方法目前虽已实现镍硫温度的区间预测,但与理想区间结果仍然存在一定差异,尤其在温度剧烈波动区段。未来,需考虑融合多模型区间预测结果,结合 TCN-XGBoost 与其他先进区间预测方法,并通过贝叶斯模型平均(BMA)等方法整合多源区间信息,进一步压缩冗余区间宽度,同时确保覆盖概率稳定达标。

4.2 结论

镍冶炼过程中出口镍硫温度的精准预测是保障冶炼设备安全、优化工艺参数与稳定生产效率的核心环节,然而该过程存在多输入参数非线性耦合强、温度动态波动剧烈(如原料成分波动、设备工况漂移引发的骤升骤降)等建模难题。针对上述挑战,本研究提出一种融合时序卷积网络(TCN)与 XGBoost 动态误差补偿的镍硫温度预测方法,构建“关键特征筛选-时序初始预测-动态误差校正-不确定性量化”的多层次建模框架:首先通过皮尔逊相关系数与 F-score 融合策略,筛选出入炉物料组分、鼓风量等对温度影响显著的关键特征;利用 TCN 的膨胀因果卷积结构,高效捕捉多变量间的长短期时序依赖关系,生成温度初始预测值;引入 XGBoost 构建一阶、二阶时间差分特征,精准学习预测残差的演变规律以实现动态误差补偿;最终依托 XGBoost 的分位数回归能力,构建覆盖概率为 90% 的预测区间,量化温度预测的不确定性。

实验结果从点预测与区间预测两大方面验证了所提模型的优越性。在点预测性能上,TCN-XGBoost 模型较基线 TCN 模型精度显著提升,MAE 从 7.627 7 降至 7.494 1,MAPE 从 0.592 8 降至 0.584 2,RMSE 从 9.791 3 优化至 9.573 2, R^2 提升

至 0.451 6,且优于 LSTM、AGCRN 等对比方法。在区间预测性能上,TCN-XGBoost 的 90% 预测区间展现出“覆盖可靠、宽度紧凑”的平衡优势:对比 TCN-LightGBM 与 TCN-KDE,TCN-XGBoost 既能在温度平稳区间紧密包裹真实值,又能在剧烈波动区段可靠包容真实值变化,避免区间过度宽泛或覆盖缺失的问题。

综上,本研究提出的 TCN-XGBoost 通过多环节协同优化,有效解决了镍钨温度预测中“时序特征捕捉难、动态误差校正弱、不确定性量化缺失”的问题,显著提升了预测精度与鲁棒性,为镍冶炼过程的温度实时监测提供了可行方案。

[参考文献]

- [1] 顾明杰,祁永峰,李智,等. 富氧侧吹熔炼短流程生产金属化镍阳极工艺研究[J]. 绿色矿冶,2023,39(5):9-14.
GU Mingjie, QI Yongfeng, LI Zhi, et al. Research on the process of producing metallized nickel anode plate by oxygen enriched side-blown smelting technology[J]. Green Mining and Metallurgy, 2023, 39(5): 9-14.
- [2] 李擎,杨思琪,陈松路,等. 基于自适应数据增强的转炉炼钢终点碳温预测方法[J]. 冶金自动化,2025,49(2):64-74.
LI Qing, YANG Siqi, CHEN Songlu, et al. Prediction method for endpoint carbon and temperature in converter steelmaking based on adaptive data augmentation[J]. Metallurgical Automation, 2025, 49(2): 64-74.
- [3] 徐吉,信自成,兰模,等. 冶金机理与贝叶斯优化 XGBoost 融合的 VD 炉精炼终点钢液温度预测[J]. 工程科学与技术,2024,56(6):63-72.
XU Ji, XIN Zicheng, LAN Mo, et al. Prediction of molten steel temperature at the VD furnace refining endpoint by fusing metallurgical mechanism with Bayesian optimized XGBoost[J]. Advanced Engineering Sciences, 2024, 56(6): 63-72.
- [4] 王子昂,王伟,马乾伦,等. 面向带钢热镀锌感应加热的改进贝叶斯温度预测算法[J]. 制造业自动化,2024,46(10):157-164.
WANG Ziang, WANG Wei, MA Qianlun, et al. An improved Bayesian temperature prediction algorithm for induction heating in steel strip hot-dip galvanizing[J]. Manufacturing Automation, 2024, 46(10): 157-164.
- [5] 施树蓉,韩啸,胡倩倩,等. 基于 PSO-MLP 的 RH 精炼终点温度预报模型开发与应用[J]. 冶金能源,2025,44(2):60-65.
SHI Surong, HAN Xiao, HU Qianqian, et al. Development and application of an RH refining endpoint temperature prediction model based on PSO-MLP[J]. Metallurgical Energy, 2025, 44(2): 60-65.
- [6] YIN Y, LIU Y, LIANG X, et al. Adaptive data-driven soft-sensor for monitoring and prediction of temperature inside zinc rotary kiln[J]. IEEE Sensors Journal, 2025,25(9):15276-15294.
- [7] WANG Y, XU Y, SONG X, et al. Novel method for temperature prediction in rotary kiln process through machine learning and CFD[J]. Powder Technology, 2024, 439: 119649.
- [8] 尹刚,朱森,颜玥涵,等. DMIFD:一种基于深度学习的多模态工业故障诊断方法[J]. 仪器仪表学报,2025,46(6):215-227.
YIN Gang, ZHU Miao, YAN Yuehan, et al. DMIFD: A deep learning-based method for multimodal industrial fault diagnosis[J]. Chinese Journal of Scientific Instrument, 2025, 46(6): 215-227.
- [9] 尹刚,钱中友,曹文琦,等. 基于 Adaboost-PSO-SVM 的铝电解槽健康状态诊断方法研究[J]. 化工学报,2024,75(1):354-365.
- [10] YIN Gang, QIAN Zhongyou, CAO Wenqi, et al. Health state diagnosis of aluminum electrolytic cells based on Adaboost-PSO-SVM[J]. CIESC Journal, 2024, 75(1):354-365.
- [11] YIN G, LI Y H, YAN F Y, et al. Deep learning for electrolysis process anode effect prediction based on LSTM and SDAE[J]. Rare Metals, 2024, 43(7):1-9.
- [12] 韩洁,赵灼,朱亮,等. 基于改进 FCM 的多工况氧化铝蒸发过程结垢参数软测量方法[J/OL]. 化工进展,1-17[2025-04-23]. <https://doi.org/10.16085/j.issn.1000-6613.2025-0202>.
HAN Jie, ZHAO Zhuo, ZHU Liang, et al. A soft-sensing method for scaling parameters of the alumina evaporation process under multiple operating conditions based on improved FCM[J]. Chemical Industry and Engineering Progress, 1-17[2025-04-23]. <https://doi.org/10.16085/j.issn.1000-6613.2025-0202>.
- [13] 罗世华,董力豪,王栋,等. 基于 VMD-TCN-GRU 组合模型的高炉透气性指数预测[J/OL]. 中国冶金,1-15[2025-04-23]. <https://doi.org/10.13228/j.boyuan.issn1006-9356.20240760>.
LUO Shihua, DONG Lihao, WANG Dong, et al. Blast Furnace Permeability Index Prediction Based on a VMD-TCN-GRU Hybrid Model[J]. China Metallurgy, 1-15[2025-04-23]. <https://doi.org/10.13228/j.boyuan.issn1006-9356.20240760>.
- [14] 尹刚,朱森,全鹏程,等. 基于 PID 搜索优化的 CNN-LSTM-Attention 铝电解槽电解温度预测方法研究[J]. 仪器仪表学报,2025,46(1):324-337.
YIN Gang, ZHU Miao, QUAN Pengcheng, et al. Research on CNN-LSTM-Attention aluminum electrolyzer electrolysis temperature prediction method based on PID search optimization[J]. Chinese Journal of Scientific Instrument, 2025, 46(1): 324-337.
- [15] DAI J, SHI H, CHEN X, et al. Aluminum smelting temperature prediction using GCN with correlation aggregation strategy and LSTM with dynamic attention[J]. International Journal of Machine Learning and Cybernetics, 2025, 16: 6887-6901.

- [16] 杜文莉,官振强,钱锋.一种基于时序误差补偿的动态软测量建模方法[J].化工学报,2010,61(2):439-443.
DU Wenli, GUAN Zhenqiang, QIAN Feng. A dynamic soft-sensing modeling method based on time series error compensation," Journal of Chemical Industry and Engineering [J]. 2010,61(2):439-443.
- [17] GUO K, ZHANG Y, ZHANG S, et al. Grey system prediction model of GUR based on error feedback compensation[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74. DOI: 10.1109/TIM.2025.3556206.
- [18] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. arXiv preprint arXiv:1803.01271, 2018. <https://arxiv.org/pdf/1803.01271>.
- [19] CHEN T, GUESTRIN C. XGBoost: A scalable tree boosting system[C]//Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016: 785-794.
- [20] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: Continual prediction with LSTM [J]. Neural computation, 2000, 12(10): 2451-2471.
- [21] BAI L, YAO L, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[J]. Advances in neural information processing systems, 2020, 33: 17804-17815.

A nickel matte temperature prediction method in nickel smelting process based on TCN-XGBoost

WEI Kaifeng¹, QI Fengqin¹, HOU Jingru¹, LI Diyan¹, KANG Jianqiang², GAO Yuxi³

(1. State Key Laboratory of Ni&Co Associated Minerals Resources Development and Comprehensive Utilization, Jinchang 737100, China;

2. Nickel Smelting Plant, Jinchuan Group Nickel Cobalt Co., Ltd., Jinchang 737100, China;

3. School of Automation, Central South University, Changsha 410083, China)

Abstract: Accurate prediction of the outlet temperature of nickel matte during nickel smelting is crucial for process optimization, equipment safety assurance, and product quality stability. However, the process involves multi-variable nonlinear coupling and dynamic parameter fluctuations, which restrict the application effect of traditional methods. To address this issue, this paper proposes a nickel matte temperature prediction method integrating Temporal Convolutional Network (TCN) with XGBoost dynamic error compensation. First, a fusion strategy of Pearson correlation coefficient and F-score is adopted to select key features from multi-dimensional industrial parameters such as furnace feed composition, blast volume, and oxygen concentration, reducing redundant interference. Second, the dilated causal convolution of TCN is used to capture long- and short-term temporal dependencies and nonlinear relationships among features, generating initial temperature predictions. Furthermore, XGBoost is introduced and first-order/second-order time difference features are constructed to learn the evolution law of prediction residuals for dynamic error compensation. Relying on XGBoost quantile regression, interval prediction of nickel matte temperature can be realized. Experimental results show that in terms of point prediction performance, the TCN-XGBoost model significantly improves accuracy compared with the baseline TCN model: MAE decreases from 7.627 7 to 7.494 1, MAPE reduces from 0.592 8 to 0.584 2, RMSE optimizes from 9.791 3 to 9.573 2, and R^2 increases to 0.451 6. It also outperforms comparative methods such as LSTM and AGCRN. In terms of interval prediction performance, the 90% prediction interval of TCN-XGBoost exhibits a balanced advantage of "reliable coverage and compact width". Compared with TCN-LightGBM and TCN-KDE, it can not only tightly wrap the true values in stable temperature intervals but also reliably accommodate changes in true values in sharply fluctuating sections, avoiding problems of excessively wide intervals or missing coverage. This method possesses high-precision prediction capability for the outlet temperature of nickel matte and can effectively adapt to dynamic fluctuations in industrial production, providing scientific support for real-time monitoring and process regulation of the nickel smelting process.

Key words: nickel smelting; nickel matte temperature; accurate prediction; temporal convolutional network (TCN); XGBoost dynamic error compensation; point prediction; interval prediction