基于特征优选的 TPA-GRU 的烧结矿 FeO 含量预测模型

冯 伟[1,2]，李 斌[2]，李天丁[1]，康学军[2]，陈皓轩[1]，蔺旭辉[1]，邢相栋[1]
（1. 西安建筑科技大学 冶金工程学院，陕西 西安 710055；2. 陕西龙门钢铁有限责任公司，陕西 韩城 715405）

摘要：针对采用化学分析法检测烧结矿中 FeO 含量存在滞后性、难以实时反映烧结过程状态，以及现有预测模型仅关注特定工况下的强相关性参数、难以适应参数相关性的动态变化，导致模型泛化能力与预测精度不足的问题，本文提出一种融合时序注意力机制（TPA）的门控循环单元（GRU）网络预测模型。首先，对工业现场采集的烧结生产数据进行预处理，包括异常值检测、异常值与缺失值处理以及数据标准化等步骤，以保证数据质量与模型训练的稳定性；其次，利用注意力机制对各输入特征参数进行权重分配，实现关键特征的自动筛选；最后，将筛选后的特征值输入 GRU 网络进行训练，构建烧结矿 FeO 含量预测模型。以某烧结厂 2024 年全年实际生产数据为研究对象，通过试验验证该模型的预测效果，并与其他神经网络模型进行对比。结果表明，模型评价指标 R^2 达到 0.923 1，均方根误差（RMSE）为 0.179 1，平均绝对误差（MAE）为 0.203 5，表现出较强的泛化能力与预测精度。该研究可为烧结生产中的配料优化与工艺调控提供实时决策依据，对稳定烧结矿质量具有重要的工程应用价值。

关键词：注意力机制；烧结矿 FeO 含量；门控循环单元；特征选择
中图分类号：TF046.4；TF325.1 **文献标志码**：A **文章编号**：1000－8764（2026）02－0054－09
doi：10.13403/j.sjqt.2026.02.029

FeO content prediction model of sinter based on TPA-GRU with feature optimization

FENG Wei[1,2], LI Bin[2], LI Tianding[1], KANG Xuejun[2], CHEN Haoxuan[1], LIN Xuhui[1], XING Xiangdong[1]
(1. School of Metallurgical Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China; 2. Shaanxi Longmen Iron & Steel Co., Ltd., Hancheng 715405, Shaanxi, China)

Abstract: Aiming at the hysteresis of FeO content in sinter detected by chemical analysis, which is difficult to reflect the sintering process state in real time, and the existing prediction models only focus on strong correlation parameters under specific working conditions, and it is difficult to adapt to the dynamic changes of parameter correlation, resulting in insufficient generalization ability and prediction accuracy of the model, a gated recurrent unit (GRU) network prediction model fused with temporal attention mechanism (TPA) is proposed. Firstly, the sintering production data collected at the industrial site is preprocessed, including outlier detection, outlier and missing value processing, and data standardization to ensure data quality and model training stability. Secondly, the attention mechanism is used to allocate the weight of each input feature parameter to realize the automatic screening of key features. Finally, the screened eigenvalues are input into the GRU network for training, and a prediction model of FeO content in sinter is constructed. Taking the actual production data of a sintering plant in 2024 as the research object, the prediction effect of the model is verified through experiments

收稿日期：2025－08－22；**修回日期**：2025－09－30
基金项目：陕西省创新能力支撑计划资助项目（2023-CX-TD-53）；渭南市秦创原“科学家＋工程师”队伍建设项目（Z20250053）；中央引导地方科技发展资金资助项目（24ZYQB003）
作者简介：冯 伟（1985—），男，高级工程师，博士研究生，从事高炉炼铁、烧结球团生产技术方面的工作。
通信作者：邢相栋（1983—），男，教授，博士，从事铁矿造块及冶金大数据分析等方面的研究。

and compared with other neural network models. The results show that the model evaluation index R^2 reaches 0.923 1, the root mean square error (RMSE) is 0.179 1, and the mean absolute error (MAE) is 0.203 5, showing strong generalization ability and prediction accuracy. This study can provide a real-time decision-making basis for batching optimization and process regulation in sintering production with important engineering application value for stabilizing the quality of sinter.

Key words: attention mechanism; FeO content of sinter; gated recurrent unit (GRU); feature selection

烧结是钢铁工业的关键工序之一[1]，涉及复杂的物理化学反应过程。烧结矿的质量直接影响高炉冶炼的稳定性、能耗水平及铁水品质[2-3]。烧结矿中 FeO 含量作为衡量烧结矿质量的重要指标，与转鼓指数、还原度指数和低温还原粉化指数密切相关，同时也能够反映烧结过程热工制度的合理性，对高炉炼铁工序的间接还原效果及焦比具有一定影响[4-5]。将 FeO 含量控制在适宜水平，不仅有助于提升烧结矿的冶金性能，也有利于高炉的稳定顺行[6-7]。采用化学分析法检测烧结矿 FeO 含量存在滞后性，难以及时指导烧结生产[8]。因此，建立烧结矿 FeO 含量的预测模型，对于实时指导现场调控、稳定烧结矿质量具有重要意义。烧结矿 FeO 含量预测模型中特征选择的合理性对预测性能影响显著，目前围绕特征选择与机器学习算法相结合的研究已形成多种技术路径。李中正等[9]采用最大互信息系数、Kendall 秩相关系数、灰色关联度等三种特征选择方法，结合烧结工艺机理筛选出 FeO 含量预测的最佳特征集，从而提升了 GA-BiLSTM 模型的预测精度与可解释性。李芹芹等[10]对原始数据进行序列化与差分处理后，基于灰色关联度与相关系数分析筛选关键过程特征，增强了 FeO 预测模型的泛化性能与鲁棒性。ZHANG 等[11]利用皮尔逊相关性分析与主成分分析法对原始数据进行降维处理，结合经麻雀搜索算法改进的 BP 神经网络，构建了寻优能力强、收敛速度快的烧结矿 FeO 含量预测模型。

然而，现有研究通常仅选取在正常运行状态下相关性较强的参数作为预测模型的输入特征。在烧结过程中，参数间的相关性会因原料质量波动与操作制度调整而动态变化，若仅关注单一状态下的强相关参数，易导致预测模型遗漏关键信息，从而降低其在长周期运行场景下的预测精度。因此，增强模型在多工况下对特征参数的处理能力，对于提升预测模型的泛化性能与长期稳定性具有重要意义。本文提出一种融合时序注意力机制(TPA)的门控循环单元(GRU)网络模型，用于烧结矿 FeO 含量的预测。收集 2024 年 1 月至 12 月期间的工业现场生产数据，采用局部离群因子(LOF)算法识别极端异常值，并通过线性插值法进行数据填补。完成数据标准化后，利用 TPA 对各输入参数进行权重自动分配，将加权后的特征输入 GRU 网络进行训练，最终构建烧结矿 FeO 含量预测模型。本文将所提模型的预测性能与多种 GRU 建模算法进行对比分析，以验证 TPA-GRU 模型的实际可行性。

1　数据收集与处理

本文数据来源于某烧结厂 2024 年全年的工艺特征日报。由于数据量大、质量参差不齐且存储形式多样，需经处理后方可用于模型训练与测试。结合现场生产经验，共选取 29 个特征参数，如表 1 所示。

1.1　异常值检测

在实际生产过程中，数据异常或缺失往往由传感器故障、设备检修等原因导致[12]。对于缺失比超过 50% 的数据，为避免直接删除破坏时序的连续性，标记为 NaN；当缺失比较低时，在相邻时间段内数据的变化通常不大，实际采用均值填充法进行填补，以保持数据的连续性。

数据异常值箱形图的原理如图 1 所示。图 1 中，箱型图由下四分位数(Q_1)、中位数(Q_2)、上四分位数(Q_3)、最小值、最大值组成。其中，下四分位数、中位数和上四分位数分别对应所选特征参数样本数据中累计占比为 25%、50% 和 75% 的位置特征值。

采用箱形图对 29 个特征参数的异常值进行检测分析，如图 2 所示。图 2 中，上下两侧圆圈标记的数据为极端异常值，这些数据通常由传感器故障或人为记录失误造成。这些异常值不仅无法

表 1 29 个烧结过程特征参数

Table 1 29 characteristic parameters of the sintering process

参数编号	参数名称	参数编号	参数名称	参数编号	参数名称
X_1	w(TFe)/%	X_{11}	烧结台车速度/(m·min^{-1})	X_{21}	w(高硅矿)/%
X_2	w(FeO)/%	X_{12}	环冷机速度/(m·min^{-1})	X_{22}	w(硫酸渣)/%
X_3	w(SiO_2)/%	X_{13}	烧结机料层厚度/mm	X_{23}	w(津巴布韦粉)/%
X_4	w(CaO)/%	X_{14}	环冷机料层厚度/mm	X_{24}	w(高钛矿)/%
X_5	w(MgO)/%	X_{15}	点火温度/℃	X_{25}	w(巴西混合粉)/%
X_6	w(Al_2O_3)/%	X_{16}	煤气压力/Pa	X_{26}	w(巴西粗粉)/%
X_7	w(P)/%	X_{17}	点火煤气流量/(m^3·h^{-1})	X_{27}	w(恩施粉)/%
X_8	w(S)/%	X_{18}	点火炉助燃空气流量/(m^3·h^{-1})	X_{28}	轧钢粉用量/t
X_9	碱度	X_{19}	w(白云石)/%	X_{29}	w(氧化铁皮)/%
X_{10}	w(TiO_2)/%	X_{20}	w(石灰石)/%	—	—

注:X_1 ~ X_8、X_{10}中,w 表示化学成分的质量分数;X_{19} ~ X_{28}、X_{29}中,w 表示物料配比的质量分数。

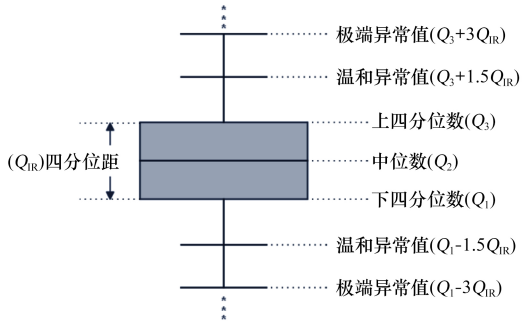


图 1 数据异常值箱形图原理

Fig. 1 The principle of box plot for detecting outliers in data

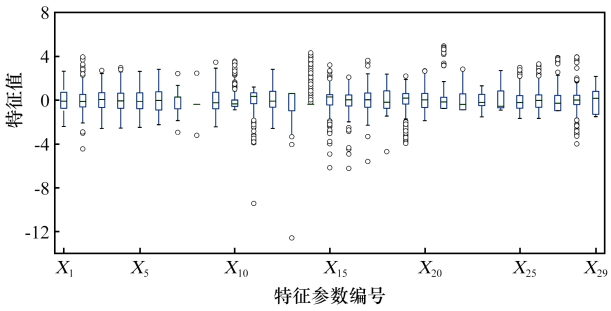


图 2 29 种参数的数据异常值检测结果

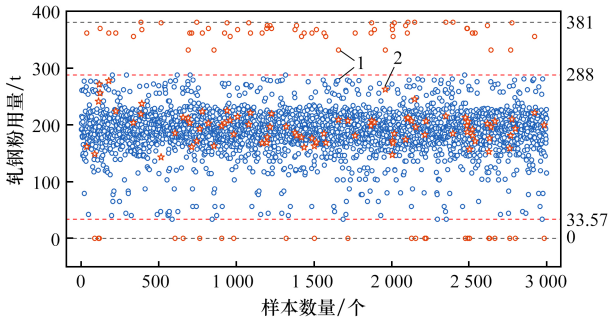
Fig. 2 The detection results of outliers in data for 29 parameters

反映真实工艺状态，还会严重干扰模型训练，需要剔除。而处于 $Q_1 - 1.5Q_{IR}$ 到 $Q_1 - 3Q_{IR}$ 之间或 $Q_3 + 1.5Q_{IR}$ 到 $Q_3 + 3Q_{IR}$ 之间的数据则被认为是温和异常值，这些异常值的出现可能是工艺原因造成的正常波动，其可以保留。

1.2 异常值及缺失值处理

本文采用均方误差(MSE)评估真实数据与重构数据之间的误差，通过计算输入样本在各时间步的重构误差，有效区分正常波动与异常偏离。当真实数据与重构数据误差过大时，将该数据标记为异常。考虑到工业数据具有时变性，静态阈值难以适应整个数据序列，本文采用动态阈值法进行异常判别。动态阈值可根据当前数据窗口内的统计特性自动调整检测标准，增强噪声检测的自适应性，避免固定阈值在数据分布变化时失效。

采用异常检测修复后的轧钢粉用量数据如图 3 所示。由图 3 可以看出，温和异常值得到了修复并重构至数据集中，同时被标记为缺失值的数据也得以补齐。



1—原始数据；2—修复数据。

图 3 轧钢粉用量异常检测与修复效果对比

Fig. 3 Comparison of abnormal detection and repair effects of rolling steel powder consumption

1.3 数据标准化

考虑到烧结数据中各特征单位不一致且数量级差异较大，模型训练前需对数据集进行标准化处理以消除量纲影响[13]。本文采用 Z-score 方法作

为标准化处理规则，计算公式如式(1)所示。

$$y_i = \frac{y_i - \frac{1}{n}\sum_{j=1}^{n} x_{ij}}{\sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(x_{ij} - \bar{x})^2}} \tag{1}$$

式中：y_i为归一化后的输入数据(i 为 1、2、…、n，n 为样本数量)；x_{ij}为输入的原始数据；$\bar{x}$为(x_1，x_2，…，x_n)序列的平均值。对各特征列进行变换，确保新序列被缩放至[0,1]区间。

1.4　特征选取

皮尔逊相关系数(Pearson correlation coefficient，PCC)可反映两个随机变量之间的线性相关程度[14]，其计算如式(2)所示。

$$p(A,B) = \frac{\sum_{j=1}^{n}[(A_j - \bar{A})(B_j - \bar{B})]}{\sqrt{\sum_{j=1}^{n}(A_j - \bar{A})^2}\sqrt{\sum_{j=1}^{n}(B_j - \bar{B})^2}} \tag{2}$$

式中：$p(\,)$为皮尔逊相关系数计算函数；$\bar{A}$和$\bar{B}$表示变量 A 和 B 的平均值。

采用皮尔逊相关系数法对特征参数进行相关性分析，结果如图 4 所示。根据分析结果，最终选取前 14 个特征作为预测模型的输入，分别为 w(MgO)、点火炉助燃空气流量、轧钢粉用量、点火煤气流量、w(氧化铁皮)、w(恩施粉)、w(巴西混合粉)、w(高硅矿)、w(白云石)、w(高钛矿)、煤气压力、w(石灰石)、点火温度、w(硫酸渣)。

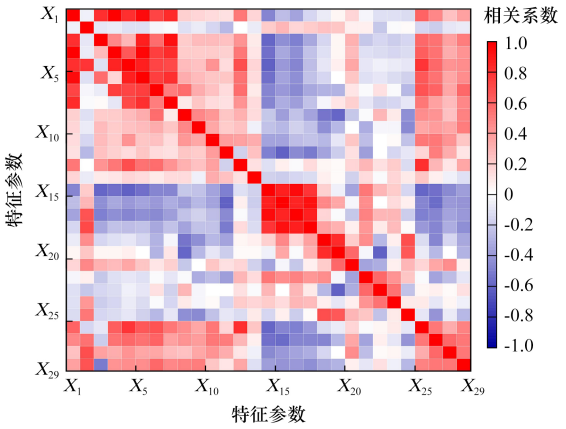


图 4　29 个参数的相关性热力图

Fig. 4　Correlation heatmap of 29 parameters

2　模型原理

2.1　门控循环单元

GRU 是 LSTM 的一种重要变体，两者均通过神经元内的门限控制器判断输入信息的有效性，并控制相关信息的流入[15]。相比 LSTM，GRU 不仅特征数量更少，网络结构也更为简单，如图 5 所示。图 5 中：$\boldsymbol{h}_{t-1}$为上一时刻隐含状态输出；$\boldsymbol{h}_t$为当前时刻隐含状态输出；$\boldsymbol{x}_t$为当前时刻隐含状态输入；$\boldsymbol{r}_t$和$\boldsymbol{z}_t$分别为当前时刻重置门和更新门输出；$\tilde{\boldsymbol{h}}_t$为当前时刻候选向量输出；σ 为 Sigmoid 函数；tanh 为双曲正切函数；$\boldsymbol{W}_r$、$\boldsymbol{W}_z$、$\boldsymbol{W}_h$为网络输入权重向量；$\boldsymbol{U}_r$、$\boldsymbol{U}_z$、$\boldsymbol{U}_h$为隐含权重向量。

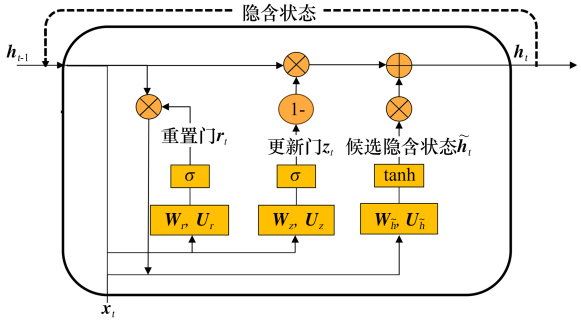


图 5　门控循环单元结构

Fig. 5　Gated recurrent unit network

GRU 网络包含重置门与更新门两个门控制器，以及候选向量，其计算方法分别如式(3)～(5)所示。

$$\boldsymbol{r}_t = \sigma(\boldsymbol{W}_r \boldsymbol{x}_t + \boldsymbol{U}_r \boldsymbol{h}_{t-1} + \boldsymbol{b}_r) \tag{3}$$

$$\boldsymbol{z}_t = \sigma(\boldsymbol{W}_z \boldsymbol{x}_t + \boldsymbol{U}_z \boldsymbol{h}_{t-1} + \boldsymbol{b}_z) \tag{4}$$

$$\tilde{\boldsymbol{h}}_t = \tanh(\boldsymbol{W}_h \boldsymbol{x}_t + \boldsymbol{U}_h(\boldsymbol{r}_t \odot \boldsymbol{h}_{t-1}) + \boldsymbol{b}_h) \tag{5}$$

式中：$\boldsymbol{b}_r$、$\boldsymbol{b}_z$和$\boldsymbol{b}_h$为偏置向量；⊙表示为哈达玛积运算。当前时刻隐含状态 $\boldsymbol{h}_t$如式(6)所示。

$$\boldsymbol{h}_t = (1 - \boldsymbol{z}_t) \odot \boldsymbol{h}_{t-1} + \boldsymbol{z}_t \odot \tilde{\boldsymbol{h}}_t \tag{6}$$

综上所述，GRU 网络通过其门控制器与时序反馈机制，能够实现信息的动态更新、记忆与持久化留存。

2.2　注意力机制

注意力机制是深度学习领域最具影响力的思想之一，其旨在解决 RNN 在输入序列较长时信息损失的问题，而用于存储学习信息的中间向量容量有限[16]。模型结构如图 6 所示。

$\boldsymbol{x}_1$、$\boldsymbol{x}_2$、…、$\boldsymbol{x}_T$表示在连续的 T 个时间步长内模型的输入序列。结构中的网络单元可以是任意的 RNN 网络单元。在注意力机制中，当模型预测步长为 t 时，考虑每个 $\boldsymbol{h}$ 向量对输出的贡献，从而得到相应的编码向量$\boldsymbol{c}_t$。$\boldsymbol{S}_t$为时刻t的隐藏状

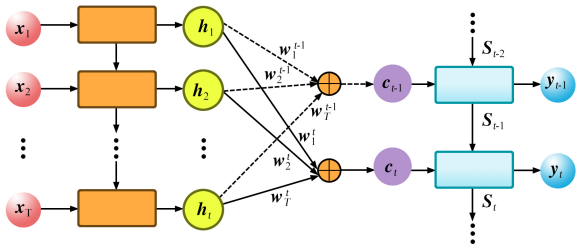


图 6 注意力机制结构

Fig. 6 Attention mechanism structure

态，用于表征当前时刻模型的内部记忆信息；y_t为模型在时刻 t 的输出结果。注意力机制下编码向量 c 的计算如式(7)、(8)所示。

$$h_t = f_{\mathrm{RNN}}(x_t, h_{t-1}) \tag{7}$$

$$c_t = \sum_{i=1}^{T} w_i^t h_i \tag{8}$$

式中：$f_{\mathrm{RNN}}()$为循环神经网络的状态更新函数。

GRU 模型能够有效提取序列特征，但不同时间节点的特征对最终结果的贡献程度存在差异，因此有必要增强模型对关键特征的学习能力。基于此，本文将 GRU 与注意力机制相结合。特征输入至注意力模块后，与对应权重系数相乘并进行累加求和，得到最终的输出状态向量 Y。其计算方法如式(9) ~ (11)所示。

$$s_t = \tanh(w_t h_t + b_t) \tag{9}$$

$$\alpha_t = \frac{\exp(s_t)}{\sum_{i=1}^{t}} s_i \tag{10}$$

$$Y = \sum_{t=1}^{n} \alpha_t s_t \tag{11}$$

式中：w_t和b_t分别表示第 t 个特征向量的权重和偏置。

依据式(9)、(10)和式(11)，将输入的初始状态转换为最终输出的状态向量 Y；随后将全连接层与之相连，并通过 soft max 函数对最终输出进行计算，得到分类结果 $\tilde{y}$。其计算方法如式(12)所示。

$$\tilde{y} = \arg\max[\mathrm{soft\ max}(w_y Y + b_y)] \tag{12}$$

式中：$\arg\max()$为最大值索引的函数；w_y、b_y为输入向量的权重和偏置。

2.3 时序注意力机制(TPA)

本文所提出的模型借助特征关注机制，将不同参数对烧结矿 FeO 含量的影响程度作为权重进行学习。对于一个时间序列输入$[x_1, x_2, \cdots, x_t, \cdots, x_T]$，$x_t = (x_t^1, x_t^2, \cdots, x_t^n) \in \mathbf{R}^n$，其中 t 为时间，x_t表示时间 t 的样本输入。得到特征注意力后，采用式(13)计算权重。

$$\tilde{x} = \alpha_t \odot x_t = [\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \cdots, \alpha_t^n x_t^n]_{n\times 1} \tag{13}$$

式中：α_t为特征注意力；x_t为加权样本。

采用如图 7 所示的时序注意力机制，对 GRU 网络各历史时刻的重要性进行分析，为各历史时刻赋予相应的注意力权重。计算方法如式(14)、(15)所示。

$$\alpha_i = \sigma[f(h_i, h_t)] \tag{14}$$

$$v_t = \sum_{i=1}^{m} \alpha_i h_i \tag{15}$$

式中：$f()$为评分函数；α_i为注意力权重；v_t为中间向量；σ 为激活函数，将任意实数映射到(0,1)区间。

模型的隐含状态h'_t如式(16)所示。

$$h'_t = \tanh(W_h h_t + W_c v_t) \tag{16}$$

式中：W_h、W_c分别为对应的权重矩阵。

TPA 能够提取各变量的时序特征，并根据其重要性分配相应的权重向量。烧结过程属于典型的多变量长时序动态系统，影响 FeO 含量的因素随时间动态变化，不同时段内的特征参数对 FeO 含量的贡献权重存在差异。GRU 仅依靠循环单元的门控机制被动记忆时序信息，为此引入 TPA 注意力机制，以识别对当前状态影响最大的历史信息。将 TPA 输出的加权结果与 GRU 的更新门和重置门联动，通过增强更新门的激活强度，强化注意力权重较高的时序片段信息在隐藏状态中的保留。

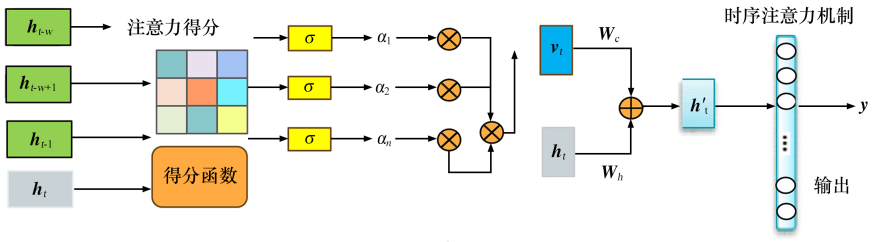


图 7 时序注意力网络结构

Fig. 7 temporal pattern attention mechanism

2.4　TPA-GRU

在 TPA-GRU 算法中，本文采用正则化用于解决多重共线性的问题，将 L1、L2 范数结合，同时加入到 GRU 网络中，实现网络结构的稀疏化，提高算法的泛化能力。L1 范数和 L2 范数结合的特征收缩方法如式(17)所示。

$$\overline{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}}\{\sum_{i=1}^{N}(\boldsymbol{y}_i - \boldsymbol{X}_i\boldsymbol{\beta})^2 + \lambda_1\|\boldsymbol{\beta}\|_2^2 + \lambda_2\|\boldsymbol{\beta}\|_1\},\lambda_1 \geqslant 0,\lambda_2 \geqslant 0 \tag{17}$$

式中：λ_1、λ_2为正则化系数；$\boldsymbol{X}_i = [\boldsymbol{x}_{i1},\boldsymbol{x}_{i2},\cdots,\boldsymbol{x}_{in}]$为第 i 个样本的特征向量；n 为输入变量个数；$\boldsymbol{y}_i$为输出变量；$\boldsymbol{\beta}$ 为回归特征。

TPA-GRU 模型融合了注意力机制与 GRU 模块的优势。以时间序列为基准，采用 Python 语言将所有数据整合为初始样本集，对烧结生产数据进行预处理后，输入至门控循环单元网络。针对小样本问题，引入正则化方法以增强模型的泛化能力。时序注意力机制模块通过计算历史状态对当前时刻的贡献度，实现权重的自适应分配，并采用贝叶斯优化算法自动搜索最优超参数组合，从而在训练过程中动态调整 GRU 网络的层次结构。最后，基于性能评价指标对模型进行测试，得到最终结果。算法流程如图 8 所示。

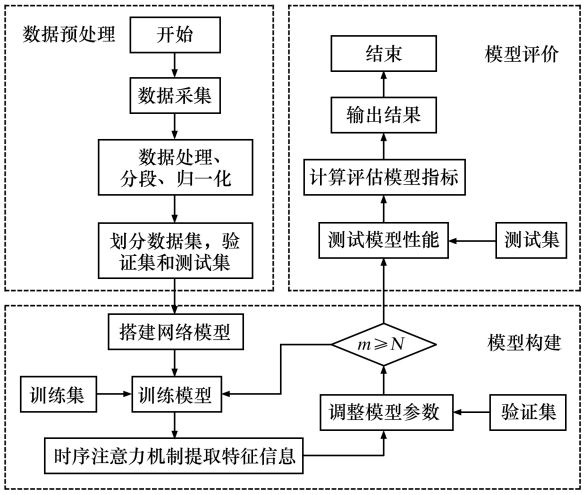


图 8　TPA-GRU 模型预测流程

Fig. 8　TPA-GRU model prediction process

3　工业实例验证与分析

将数据划分为长度为 10 的序列，并按时间顺序依次排列。将划分后的序列随机排序，并按 8∶2 的比例划分为训练集与测试集。该划分方式既可保留数据间的时间关联性，又能提升模型预测的泛化能力。在相同仿真环境下，将 TPA-GRU 模型与 GRU[17]、LASSO-LSTM[18]、TPA-LSTM[19]及 NNG-GRU[20]四种算法进行性能对比，并采用均方根误差(root mean square error，RMSE)、平均绝对误差(mean absolute error，MAE)和确定系数(R-squared，R^2)作为模型性能的评估指标。

3.1　TPA-GRU 模型中各个模块的贡献

TPA-GRU 模型在基础 GRU 上引入了 L1/2 正则化与 TPA 机制。各算法均采用初始学习率 0.005、时间步长 10、最大迭代次数 500 的参数设置。为验证各改进模块对模型性能的影响，设计了消融试验，结果如表 2 所示。由表 2 可以看出，L1/2 正则化与 TPA 机制均能有效提升模型的预测性能。相较于基础 GRU 模型，TPA-GRU 算法的 R^2 由 0.848 7 提升至 0.923 1，提高了 8.77%；均方根误差由 0.234 7 降低至 0.179 1，降低了 23.69%。结果充分验证了各改进模块对模型预测性能提升的有效性。

表 2　不同模型预测误差对比

Table 2　Comparison of prediction errors of different models

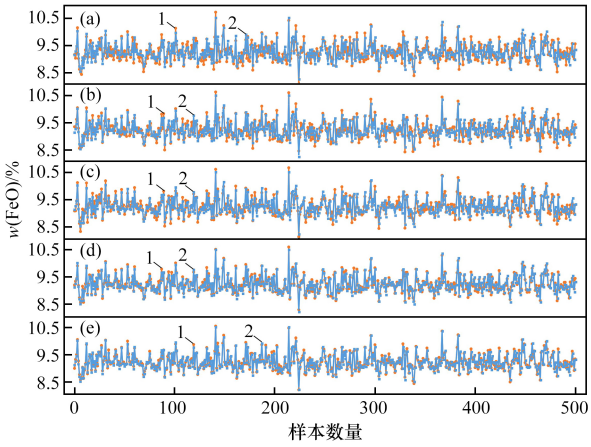
项目	R^2	均方根误差	平均绝对误差
TPA-GRU	0.923 1	0.179 1	0.203 5
GRU	0.848 7	0.234 7	0.427 6
LASSO-LSTM	0.864 9	0.219 1	0.383 4
TPA-LSTM	0.883 7	0.188 4	0.324 5
NNG-GRU	0.905 7	0.181 1	0.283 2
L1/2 正则化 GRU	0.857 2	0.229 5	0.402 3

相较于 GRU、LASSO-LSTM、TPA-LSTM 和 NNG-GRU 模型，TPA-GRU 预测模型的均方根误差为 0.179 1，平均绝对误差为 0.203 5，决定系数 R^2 达到 0.923 1。与上述四种模型相比，TPA-GRU 模型的平均绝对误差分别降低了 0.224 1、0.179 9、0.121 0、0.079 7，均方根误差分别降低了 0.055 6、0.040 0、0.009 3、0.002 0。上述结果表明，该模型不仅能够有效捕捉烧结过程中的非线性动态特征，显著提升预测精度，还能为现场操作人员预判烧结矿 FeO 含量提供可靠依据，从而为后续分析烧结生产状态、优化操作参数奠定坚实基础。

3.2　不同模型性能对比

各模型均采用相同的结构与超参数设置：输

入维度为 30，隐含层数为 1，隐含层节点数为 128，输出维度为 1，学习率为 0.005，最大迭代次数为 200。将 TPA-GRU 模型与 GRU、LASSO-LSTM、TPA-LSTM 及 NNG-GRU 模型进行对比，结果如图 9 所示。由图 9 可知，TPA-GRU 模型在烧结矿 FeO 含量预测中表现出优越的性能，预测精度超过 90%，且误差变异性明显小于其他四种对比模型。烧结过程具有显著的非线性和动态性，引入注意力机制增强的 GRU 模型，其预测效果优于标准 LSTM 与 GRU 模型，充分体现了注意力机制在处理长时序数据中的积极作用，有效缓解了信息在后向传播过程中的丢失问题。



（a）GRU；（b）LASSO-LSTM；（c）TPA-LSTM；（d）NNG-GRU；（e）TPA-GRU

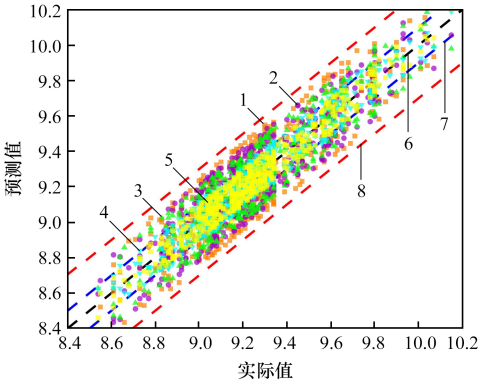
1—预测值；2—实际值。

图 9 不同模型的预测性能对比

Fig. 9 Comparison of prediction performance among different models

3.3 预测精度

各模型预测误差的对比结果如图 10 所示。由图 10 可知，与多种 GRU 建模方法相比，TPA-GRU 网络的预测结果散点更集中于 45°线附近，预测准确率更高。NNG-GRU 作为 GRU 的改进模型，虽增强了特征提取能力，但仍难以有效利用时序特征，导致识别准确率上限偏低，且网络加深后训练难度增加。LSTM 结构预测结果中，较多散点分布于 $y=x\pm0.10$ 与 $y=x\pm0.30$ 区间，其模型结构相对简单，泛化能力不足，收敛速度较慢，因而准确度较低。相比之下，本文提出的 TPA-GRU 模型通过引入注意力机制优化 GRU 模块，增强了当前工况下关键特征的权重，有效降低了训练难度，同时提升了预测精度；其预测值散点普遍集中在 $y=x\pm0.10$ 范围内，分布集中，表明注意力模块对模型具有显著的积极作用。



1—GRU；2—LASSO-LSTM；3—TPA-LSTM；4—NNG-GRU；5—TPA-GRU；6—$y=x$；7—$y=x\pm0.1$；8—$y=x\pm0.3$。

图 10 不同模型预测散点图

Fig. 10 Scatter plot of prediction results of different models

4 结 论

（1）基于 TPA 与 GRU 构建的预测模型能够较好地预测烧结矿中 FeO 含量。模型评价指标均方根误差为 0.179 1，平均绝对误差为 0.203 5，决定系数 R^2 达 0.923 1，表明该模型具有较高的预测精度与良好的泛化能力，能够有效、稳定地预测烧结矿 FeO 含量。

（2）引入注意力机制优化后的 GRU 模型性能显著提升，预测散点主要分布在 $y=x\pm0.10$ 范围内，绝对误差范围与均值均明显减小。相较于优化前，优化后的 GRU 模型在烧结矿 FeO 含量预测中表现出更高的准确度与更优的预测效果。

（3）该预测模型可提前预判烧结矿的化学成分，为烧结配料阶段混合矿成分的调控提供依据，进而稳定烧结矿 FeO 含量，对改善烧结矿质量、保障高炉稳定顺行具有重要支撑作用。

参考文献：

[1] 张建良，刘东辉，王筱留，等. 烧结矿化学成分控制现状及发展方向[J]. 烧结球团，2017，42(4)：1－5；14.
ZHANG Jianliang，LIU Donghui，WANG Xiaoliu，et al. Research status and development trend of chemical composition control of sinter[J]. Sintering and Pelletizing，2017，42(4)：1－5，14.

[2] 康健,宁晓钧,张建良,等. 烧结混匀料液相生成行为对烧结矿质量的影响[J]. 钢铁研究学报,2020,32(10):860 – 867.
KANG Jian, NING Xiaojun, ZHANG Jianliang, et al. Effect of liquid phase generation behavior of sintering mixture on sinter quality[J]. Journal of Iron and Steel Research,2020,32(10):860 – 867.

[3] 刘佳文,陈铁军,周仙霖,等. 烧结料层透气性影响因素及其改善措施解析[J]. 钢铁研究学报,2024,36(7):817 – 826.
LIU Jiawen,CHEN Tiejun,ZHOU Xianlin,et al. Analysis of influencing factors and improving measures of permeability of sinter layer[J]. Journal of Iron and Steel Research,2024,36(7):817 – 826.

[4] 范晓慧,冯婧,陈许玲,等. 烧结矿化学成分预报模型与控制指导专家系统[J]. 矿冶工程,2011,31(4):77 – 80;85.
FAN Xiaohui,FENG Jing,CHEN Xuling,et al. Prediction model and control-guidance expert system of sinter chemical composition[J]. Mining and Metallurgical Engineering,2011,31(4):77 – 80;85.

[5] 吕庆,黄宏虎,曹丽华,等. 原矿中 FeO 含量对烧结矿的影响[J]. 烧结球团,2014,39(5):1 – 4.
LÜ Qing, HUANG Honghu, CAO Lihua, et al. Influence of FeO content of crude ore on sinter[J]. Sintering and Pelletizing,2014,39(5):1 – 4.

[6] 刘昱钤,潘冬,蒋朝辉,等. 钢铁制造烧结过程关键参数检测方法综述[J]. 烧结球团,2023,48(6):10 – 21.
LIU Yuqian,PAN Dong,JIANG Zhaohui,et al. A review of detection methods of key parameters in sintering process of iron and steel manufacturing[J]. Sintering and Pelletizing,2023,48(6):10 – 21.

[7] ZHOU P, GAO B H, ZHAO C H, et al. Heterogeneous data-driven measurement method for FeO content of sinter based on deep learning and tensor decomposition [J]. Control Engineering Practice,2023,134:105479.

[8] 唐佳润,刘月明,张福明,等. 烧结矿 FeO 含量预测技术的探讨[J]. 烧结球团,2025,50(1):21 – 30;37.
TANG Jiarun, LIU Yueming, ZHANG Fuming, et al. Discussion on prediction technology of FeO content in sinter[J]. Sintering and Pelletizing,2025,50(1):21 – 30,37.

[9] 李中正,吴朝霞,王金杨,等. 基于特征优选的 GA-BiLSTM 烧结矿中 FeO 含量预测模型[J]. 东北大学学报(自然科学版),2025,46(6):56 – 65.
LI Zhongzheng,WU Zhaoxia,WANG Jinyang,et al. FeO content prediction model in sinter based on GA-BiLSTM with feature optimization[J]. Journal of Northeastern University (Natural Science),2025,46(6):56 – 65.

[10] 李芹芹,宋宝宇,张兆鑫,等. 基于时间序列模型的烧结矿 FeO 含量预测[J]. 钢铁研究学报,2025,37(5):570 – 578.
LI Qinqin, SONG Baoyu, ZHANG Zhaoxin, et al. Prediction of FeO content in sinter based on time series model[J]. Journal of Iron and Steel Research, 2025, 37(5):570 – 578.

[11] ZHANG N, CHEN X L, HUANG X X, et al. Online measurement method of FeO content in sinter based on infrared machine vision and convolutional neural network [J]. Measurement,2022,202:111849.

[12] WILLIAMS J, HILL R R, PIGNATIELLO J J Jr, et al. Wavelet analysis of variance box plot[J]. Journal of Applied Statistics,2022,49(14):3536 – 3563.

[13] 杨玉好,杨斌,胡军,等. 基于 MHA-BiLSTM 的尾矿坝位移预测[J]. 有色金属工程,2024,14(10):147 – 157.
YANG Yuhao, YANG Bin, HU Jun, et al. Study on prediction of tailings dam deformation based on MHA-BiLSTM[J]. Nonferrous Metals Engineering, 2024, 14(10):147 – 157.

[14] 李洁松,刘韬,伍星. 基于皮尔逊相关系数统计特征滚动轴承 RUL 预测[J]. 航空动力学报,2025,40(7):85 – 94.
LI Jiesong, LIU Tao, WU Xing. Rolling bearing RUL prediction based on Pearson correlation coefficient statistical features[J]. Journal of Aerospace Power, 2025,40(7):85 – 94.

[15] 章诚,申超. 基于 FFT-DC-GRU-NLA 的中长期居民用电量预测模型[J]. 现代电子技术,2025,48(16):88 – 96.
ZHANG Cheng, SHEN Chao. A medium and long-term residential electricity consumption forecasting model based on FFT-DC-GRU-NLA [J]. Modern Electronics Technique,2025,48(16):88 – 96.

[16] 蔡松睿,张仕斌,丁润宇,等. 基于跨模态注意力机制和弱监督式对比学习的虚假新闻检测模型[J]. 信息安全研究,2025,11(8):693 – 701.
CAI Songrui, ZHANG Shibin, DING Runyu, et al. Fake news detection model based on cross-modal attention mechanism and weak-supervised contrastive learning [J]. Journal of Information Security Research, 2025, 11(8):693 – 701.

[17] 杨驰,龚晓峰,雒瑞森. 基于注意力机制的门控密集

卷积网络调制识别算法[J]. 计算机应用与软件,2024,41(10):122 – 127.
YANG Chi, GONG Xiaofeng, LUO Ruisen. A modulation recognition algorithm based on attention mechanism and gated dense convolutional networks[J]. Computer Applications and Software, 2024, 41(10): 122 – 127.

[18] 胡聿文. 基于优化 LSTM 模型的股票预测[J]. 计算机科学,2021,48(增刊1):151 – 157.
HU Yuwen. Stock forecast based on optimized LSTM model[J]. Computer Science, 2021, 48(S1): 151 – 157.

[19] 何锦华,刘洋,朱誉,等. 融合物理信息的 TPA-LSTM 网络应用于台风条件下海上风电场风速多步预测[J]. 电网技术,2023,47(10):4152 – 4160.
HE Jinhua, LIU Yang, ZHU Yu, et al. Physics-informed TPA-LSTM network applied to multi-step wind speed prediction of offshore wind farms under typhoon conditions[J]. Power System Technology, 2023, 47(10): 4152 – 4160.

[20] 李艳,史艳华,戴庆瑜,等. 基于 MIC-NNG-LSTM 的有机废液焚烧 SCR 入口 NO_x 浓度动态预测[J]. 工程科学与技术,2025,57(3):21 – 30.
LI Yan, SHI Yanhua, DAI Qingyu, et al. Dynamic prediction of NO_x concentration at SCR inlet for organic waste liquid incineration based on MIC-NNG-LSTM[J]. Advanced Engineering Sciences, 2025, 57(3): 21 – 30.

(上接第 36 页)

[8] PODLIŃSKI J, NIEWULIS A, MIZERACZYK J. Electrohydrodynamic flow and particle collection efficiency of a spike-plate type electrostatic precipitator [J]. Journal of Electrostatics, 2009, 67(2/3): 99 – 104.

[9] 中国环境保护产业协会电除尘委员会. 电除尘器选型设计指导书[M]. 北京:中国电力出版社,2013.
Electrostatic Precipitator Committee of CAEPI. Guide to type selection and design of electrostatic precipitator [M]. Beijing: China Electric Power Press, 2013.

[10] NOWAK A. Vibration of collecting electrodes in electrostatic precipitators: Modelling, measurements and simulation tests[J]. Journal of Electrostatics, 2012, 70(3): 327 – 332.

[11] CHEN B, LI S Q, GUO Y H, et al. Research on electrostatic shielding characteristics of electrostatic precipitator[J]. Journal of the Air & Waste Management Association, 2022, 72(4): 331 – 345.

[12] 邵毅敏,祝志芳,熊绍武,等. 电除尘器阳极板振打加速度分布规律与仿真设计[J]. 环境工程学报,2010,4(5):1147 – 1152.
SHAO Yimin, ZHU Zhifang, XIONG Shaowu, et al. Simulation of rapping acceleration distribution principle on a positive plate for the design of an electrostatic precipitator[J]. Chinese Journal of Environmental Engineering, 2010, 4(5): 1147 – 1152.

[13] ZHAO H B, YANG G, SHEN J D, et al. Research on the shielding effect of the electrode frames in the electrostatic precipitator[J]. Journal of the Air & Waste Management Association, 2023, 73(6): 462 – 470.

[14] 罗水源,赵海宝,袁伟锋,等. 双 V 型均流式电除尘技术改造及流场优化[J]. 烧结球团,2024,49(5):124 – 129.
LUO Shuiyuan, ZHAO Haibao, YUAN Weifeng, et al. Modification and flow field optimization of double V type current-sharing electrostatic precipitator technology[J]. Sintering and Pelletizing, 2024, 49(5): 124 – 129.