DOI: 10.19666/j.rlfd.202602043

融合物理先验的 PID-CTNet 锅炉状态协同预测

苏国新[1]，巨重阳[2]，杨 原[1]，孙好英[1]，王玉平[1]，
李荣政[1]，高明楷[2]，秦慧平[2]
（1. 国能锦界能源有限责任公司，陕西 榆林 719000；
2. 山西华仁通电力科技股份有限公司，山西 太原 030000）

［摘 要］【目的】针对燃煤机组运行状态强非线性、大迟延及物理耦合显著的特点，传统建模方法难以兼顾动态精度与部署成本，本文提出一种融合物理先验的 PID-CTNet（physics-informed dual-branch convolutional transformer network，PID-CTNet）协同预测模型。【方法】首先，基于燃烧学机理重构多源特征空间，从负荷指令、燃料与风烟系统、工质热力参数及关键指标历史反馈 4 类变量中筛选得到 27 维关键物理量；其次，构建卷积神经网络（convolutional neural network，CNN）-Transformer 混合架构的多任务深度学习框架，前端利用时间卷积神经网络（temporal convolutional networks，TCN）捕捉局部瞬态变化并降维，后端引入 Transformer 自注意力机制以全局感受野处理大迟延长时序依赖；最后，通过参数共享机制实现对 NO_x、SO_2、CO 排放及主/再热蒸汽温度的同步预测。【结果】实验表明，该端到端模型实现了均方根误差为 2.895、平均绝对误差为 1.654、平均绝对百分比误差为 0.565%、决定系数为 0.957 的高精度指标，工业关键参数预测误差低于 3%，显著优于传统时序模型及单任务方法。消融实验进一步验证了物理先验、TCN-Transformer 协同结构及多任务机制对模型精度、稳定性与复杂工况鲁棒性的提升作用。【结论】PID-CTNet 模型融合物理机理与深度学习优势，兼顾预测精度与计算效率，为燃煤机组智能运行提供了轻量化解决方案。

［关 键 词］燃煤机组；CNN-Transformer；物理机理；多任务学习；协同预测

［引用本文格式］苏国新，巨重阳，杨原，等. 融合物理先验的 PID-CTNet 锅炉状态协同预测[J]. 热力发电，2026，55(8)：55.
Su G X, Ju C Y, Yang Y, et al. Physics-informed PID-CTNet for collaborative prediction of boiler operating states[J]. Thermal Power Generation, 2026, 55(8): 55.

Physics-informed PID-CTNet for collaborative prediction of boiler operating states

Su Guoxin[1], Ju Chongyang[2], Yang Yuan[1], Sun Haoying[1], Wang Yuping[1],
Li Rongzheng[1], Gao Mingkai[2], Qin Huiping[2]
(1. CHN Energy Jinjie Energy Co., Ltd., Yulin 719000, China;
2. Shanxi Huarentong Electric Power Technology Co., Ltd., Taiyuan 030000, China)

Abstract: [Objective] Coal-fired power units typically exhibit strong nonlinearity, long time delays, and pronounced physical couplings during operation. These characteristics make it difficult for traditional mechanism-based or purely data-driven modeling approaches to simultaneously achieve high dynamic accuracy, strong generalization capability, and low deployment cost in real industrial environments. To address these challenges, this study proposes a physics-informed dual-branch convolutional transformer network (PID-CTNet) that integrates physical priors with deep learning to achieve collaborative prediction of multiple combustion-related variables.

收 稿 日 期：2026-02-27 修 回 日 期：2026-04-08 接 受 日 期：2026-04-12
第一作者简介：苏国新（1976），男，副高级工程师，主要研究方向为火力发电技术，12109621@ceic.com。
通信作者简介：巨重阳（1998），男，硕士，主要研究方向为深度学习、火力发电技术，2226974720@qq.com。

[Methods] First, a multi-source feature space is reconstructed based on combustion and thermodynamic mechanisms. Twenty-seven key physical variables are selected from four categories (load commands, fuel and air-gas system parameters, thermodynamic properties of the working fluid, and historical feedback indicators) to form a physically meaningful input representation. Second, a multi-task deep learning framework combining convolutional neural networks (CNN) and Transformer is developed. The front-end temporal convolutional networks (TCN) extract local transient features and perform dimensionality reduction, while the back-end Transformer introduces a self-attention mechanism with a global receptive field to capture long-term dependencies and large-delay characteristics inherent in boiler dynamics. Finally, a hard parameter-sharing strategy is employed to enable simultaneous prediction of NO_x, SO_2, and CO emissions, as well as main and reheat steam temperatures, within a unified modeling framework. [Results] Experimental results demonstrate that the proposed end-to-end model achieves high prediction accuracy, with root mean square error of 2.895, mean absolute error of 1.654, mean absolute percentage error of 0.565%, and coefficient of determination R^2 of 0.957 on the test set. The prediction errors of all five industrial key indicators remained below 3%, significantly outperforming traditional regression and time-series models such as SVR and LSTM, as well as single-task deep learning baselines. Ablation studies further validate the effectiveness of incorporating physical priors, the complementary roles of the TCN-Transformer hybrid architecture, and the performance gains brought by the multi-task learning mechanism. These components jointly enhance model accuracy, stability, and robustness under complex operating conditions, including load fluctuations, fuel variations, and frequent operational adjustments. [Conclusion] The PID-CTNet model effectively integrates physical mechanisms with deep learning, achieving a favorable balance between prediction accuracy and computational efficiency. By providing lightweight and reliable collaborative prediction of emissions and steam temperature parameters, the proposed method offers a practical and deployable solution for intelligent operation, emission optimization, and safety control of coal-fired power units. The results demonstrate that physics-informed multi-task learning is a promising direction for improving the interpretability, engineering applicability, and long-term maintainability of data-driven models in modern power plant operation and optimization.

Key words: coal-fired power unit; CNN-Transformer; physical mechanism; multi-task learning; collaborative prediction

在当前全球能源结构转型与“双碳”目标的驱动下，燃煤发电机组正逐步由基荷电源向调节型电源转变。为了消纳风电、光伏等新能源的波动性，火电机组频繁参与深度调峰与快速变负荷运行已成为常态[1-3]。然而，这种宽负荷、变工况的运行模式使得锅炉燃烧系统的非线性、大迟延及多参数耦合特性更加显著，极易导致燃烧不稳定、主再热蒸汽汽温超限以及 NO_x 等污染物排放量剧烈波动[4]。

传统的依赖物理传感器的监测手段往往存在测量滞后、维护成本高及恶劣环境下易失效等问题，难以满足实时精细化控制的需求[5]。因此，基于运行数据的软测量技术，即利用易获取的运行参数对难以测量的关键状态进行实时建模与预测[6]，对于提升机组的运行安全性与环保经济性具有重要意义[7]。

近年来，随着工业互联网与人工智能技术的深度融合，基于数据驱动的建模方法已成为电站锅炉状态监测的重要手段。尽管此类方法应用广泛，但现有研究多集中于 NO_x 等单一环保指标的预测，对多参数协同监测的关注不足。在研究初期，支持向量回归（support vector regression，SVR）、随机森林（random forest，RF）等浅层机器学习算法是主流选择[8-9]。谭增强等[10]以机组超低排放与安全稳定运行为目标，应用 SVR 模型对火电厂氨逃逸进行建模，结果表明该方法在处理小样本数据时具有较好的泛化性能。王伟同等[11]为了实现对锅炉出口 NO_x 排放量的精准预测，引入 Spearman 秩相关系数进行特征优选，有效降低了输入维度，并利用 RF 模型实现了快速推理。此外，为了克服静态模型的局限性，Zhang 等人[12]提出了一种改进的 SVR 算法，通过基于样本相似度的在线增量学习机制，使模型能够适应过程特性的动态漂移，从而在一定程度上提高了预测精度。尽管上述传统算法模型结构简单且易于部署，但它们在处理高维时序数据时存在天然缺陷：一是难以捕捉燃烧过程普遍存在的长距离时序依赖与大惯性延迟；二是作为浅层模型，其特征提取能力有限，难以表征多变量间复杂的非线性耦合关系，限制了预测精度的进一步提升。

针对传统浅层模型在时序特征提取上的不足，以循环神经网络（recurrent neural network，RNN）及其变体长短时记忆（long short-term memory，LSTM）神经网络、门控循环单元（gated recurrent unit，GRU）为代表的深度学习方法逐渐成为研究主流[13-16]。众多学者利用此类模型对 NO_x 排放、锅炉

效率及关键参数进行建模，取得了显著优于传统方法的预测精度。赵旭东等[17]提出一种融合改进 LSTM 神经网络与注意力机制的时间序列预测框架，通过双路径特征编码及自注意力融合解码策略，实现了 NO_x 的实时精准预测。Wan 等人[18]同样结合 LSTM 神经网络与注意力机制预测小时级电力需求，优化了火电厂的主动负荷调度策略，从而减少了非必要的机组启停与能源损耗。此外，Gotzman 等人[19]针对汽轮机温度预测问题，利用改进的 GRU 算法结合注意力机制，通过大规模数据集增强模型的鲁棒性，有效降低了因温度应力导致的金属疲劳裂纹萌生风险，其泛化性能优于传统机器学习方法。尽管基于 RNN 的变体模型在一定程度上解决了时序依赖问题，但其固有的串行计算架构导致训练效率较低，且在处理超长历史序列时仍面临长期记忆丢失或梯度消失的风险[20]，难以捕捉长跨度时间相关性。

上述研究大多仅从纯数据驱动的角度出发，将锅炉系统视为“黑箱”，忽略了燃烧过程中客观存在的热力学机理与物理约束，导致模型在工况剧烈波动时鲁棒性不足[21]。同时，现有方法多侧重于单一变量的独立预测，缺乏对 NO_x、SO_2、CO 及蒸汽温度等多目标变量间强物理耦合关系的协同建模，难以实现对锅炉运行状态的全面感知。对此，本文提出一种基于物理信息驱动的卷积神经网络（convolutional neural network，CNN）-Transformer 协同网络（physics-informed dual-branch convolutional transformer network，PID-CTNet），用于火电厂锅炉关键状态参数的实时监测。

1 数据处理

1.1 研究对象与数据来源

本文以某 600 MW 超临界燃煤机组为研究对象。数据采集自电厂分散控制系统（distributed control system，DCS）历史数据库，采样间隔 10 min，样本量超过 5 万条，覆盖 300~600 MW 全负荷运行区间，包含稳态、升降负荷及磨煤机启停等典型工况，不同工况通过机组负荷变化率、主蒸汽流量变化率以及磨煤机电流与启停信号等关键监测量，且均自动识别完成。原始数据包含 50 余维测点，存在噪声、缺失及冗余信息，因此需进行特征筛选与预处理以构建高质量输入数据。

1.2 基于物理机理的特征构建

依据燃烧学与热力学机理[22-25]，从负荷指令、燃料与风烟系统、工质热力参数以及关键指标历史反馈 4 类变量中筛选得到 27 维关键物理量作为模型输入。该特征集能够覆盖锅炉在负荷扰动、煤质变化及空气分配调整过程中的主要动态响应，关键输入特征分类及物理作用如表 1 所示。

1.2.1 输入变量筛选

1）负荷指令　选取机组毛负荷与净负荷。负荷指令是机组运行的全局基准信号，直接决定了给煤量、给水量及总风量的基础设定值。它表征锅炉整体的能量输出水平，决定炉膛基础温度场的分布，是预测排放与蒸汽温度最重要的“宏观背景变量”。

2）燃料与风烟系统　选取给煤量、入炉煤热量、省煤器出口氧量、空预器出/入口氧量、炉膛压力、排烟温度、烟气流速及 CO_2 体积分数。这些参数描述了燃烧侧的能量输入与反应状态，其中，给煤量与入炉煤热量共同决定了炉膛的总热输入；各点氧量与 CO_2 体积分数反映了过量空气系数及燃烧的充分程度，是影响 NO_x 生成与 CO 排放的直接决定因素；而炉膛压力与烟气流速则表征了烟气侧的动力场状态。

3）工质热力参数　选取主蒸汽、再热蒸汽及给水三大系统的流量、压力、温度等参数，这些变量反映了水蒸气侧对燃烧放热的吸热响应。特别地，研究基于 IAPWS-IF97[22]水蒸气状态方程，显式引入主蒸汽与再热蒸汽的焓值与熵值作为物理先验特征，并将其直接输入深度神经网络的输入层。由于工质吸热过程遵循非线性热力学状态方程，引入焓、熵能够更准确地表征工质在不同压力与温度下的真实能量状态，弥补单一温度参数在刻画热惯性方面的不足，同时，汽轮机侧的热力状态还携带了管网传输迟延带来的反馈信息，使模型在输入阶段即可获得明确的热力学约束，从而增强对工况边界变化的物理感知能力。

4）关键指标历史反馈　选取 NO_x、SO_2、CO 排放量及主/再热蒸汽温度的历史时刻观测值。鉴于锅炉系统的大迟延与强惯性特性，关键状态参数的当前值与历史趋势高度相关。将这 5 项预测目标的历史值作为输入特征，能为模型提供当前的“状态起点”与“变化惯性”，显著提升预测的动态跟踪能力。

表 1 关键输入特征分类及物理作用

Tab.1 Classification and physical effects of key input features

变量类别	变量名称	对应代码	符号	单位	属性
1.负荷指令	机组毛负荷	load_gross	P_{gross}	MW	输入特征
	机组净负荷	load_net	P_{net}	MW	输入特征
2.燃料与风烟系统	给煤量	flow_coal	M_{coal}	t/h	输入特征
	入炉煤热值	coal_hhv	Q_{hhv}	kJ/kg	输入特征
	省煤器出口氧量	o2_eco	$O_{2,\mathrm{eco}}$	%	输入特征
	空预器入口氧量	o2_aph_in	$O_{2,\mathrm{in}}$	%	输入特征
	空预器出口氧量	o2_aph_out	$O_{2,\mathrm{out}}$	%	输入特征
	炉膛压力	pres_furnace	p_{fur}	Pa	输入特征
	烟气流速	vel_flue_gas	V_{gas}	m/s	输入特征
	排烟温度	temp_flue_gas	T_{gas}	℃	输入特征
	CO_2 体积分数	conc_co2	$C_{\mathrm{CO_2}}$	%	输入特征
3.工质热力参数	主蒸汽流量	flow_ms	D_{ms}	t/h	输入特征
	主蒸汽压力	pres_ms	p_{ms}	MPa	输入特征
	主蒸汽温度(汽轮机侧)	temp_ms_turb	$T_{\mathrm{ms,tur}}$	℃	输入特征
	再热蒸汽温度(汽轮机侧)	temp_rh_turb	$T_{\mathrm{rh,tur}}$	℃	输入特征
	冷再热蒸汽温度	temp_crh	$T_{\mathrm{rh,c}}$	℃	输入特征
	冷再热蒸汽压力	pres_crh	$p_{\mathrm{rh,c}}$	MPa	输入特征
	参考温度	temp_ref	T_{ref}	℃	输入特征
	给水流量	flow_fw	D_{fw}	t/h	输入特征
	给水温度	temp_fw	T_{fw}	℃	输入特征
	主蒸汽入口焓值	enth_ms_in	H_{ms}	kJ/kg	输入特征
	主蒸汽入口熵值	entr_ms_in	S_{ms}	kJ/(kg·℃)	输入特征
4.关键指标历史反馈	NO_x 质量浓度	conc_nox	$y_{\mathrm{NO_x}}$	mg/m^3	输入(历史)&输出(预测)
	SO_2 质量浓度	conc_so2	$y_{\mathrm{SO_2}}$	mg/m^3	输入(历史)&输出(预测)
	CO 质量浓度	conc_co	y_{CO}	mg/m^3	输入(历史)&输出(预测)
	主蒸汽温度（锅炉侧）	target_temp_ms	$y_{T_{\mathrm{ms}}}$	℃	输入(历史)&输出(预测)
	再热蒸汽温度（锅炉侧）	target_temp_rh	$y_{T_{\mathrm{rh}}}$	℃	输入(历史)&输出(预测)

1.2.2 协同预测目标构建

为了实现对锅炉运行状态的全面监测，选取五维关键指标作为多任务协同预测的目标变量 $\boldsymbol{Y}$，即：

$$\boldsymbol{Y}=[y_{\mathrm{NO}_x},y_{\mathrm{SO}_2},y_{\mathrm{CO}},y_{T_{\mathrm{ms}}},y_{T_{\mathrm{rh}}}] \tag{1}$$

式中：y_{NO_x}、y_{SO_2}、y_{CO} 均为环保排放水平；$y_{T_{\mathrm{ms}}}$、$y_{T_{\mathrm{rh}}}$ 分别为锅炉运行的热经济性与安全性。另外，考虑到锅炉效率指标通常依赖离线飞灰含碳量，不具备实时预测的条件，因此未将其纳入预测目标。同理，烟尘排放受除尘器运行状态影响显著，且其在线监测数据普遍存在采样频率低、滞后大及缺失较多等问题，因此亦未将烟尘作为预测对象。

1.3 相关性分析

为了验证所选物理先验特征的有效性，采用 Pearson 相关系数[23]分析输入特征与预测目标的耦合关系，结果如图 1 所示。

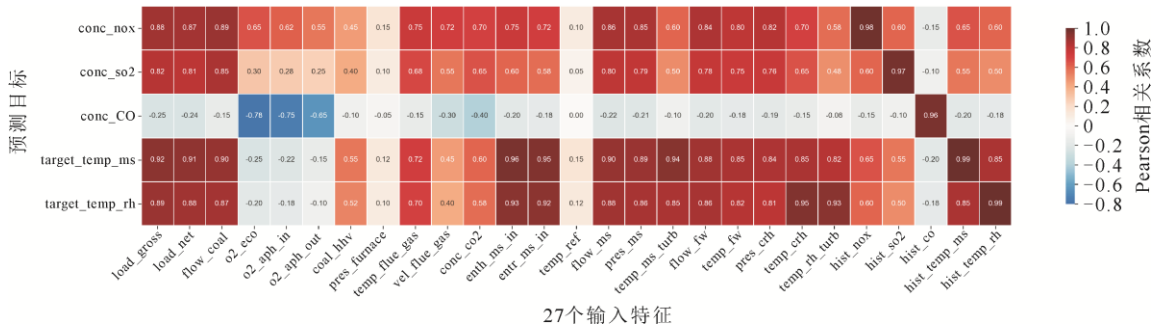


图 1 基于燃烧机理的关键特征先验相关性分析

Fig.1 Prior correlation analysis based on key characteristics of combustion mechanism

由图 1 可见：负荷、给煤量与 NO_x、蒸汽温度呈显著正相关；氧量与 NO_x 和 CO 呈现典型“剪刀差”特性；主蒸汽焓值与蒸汽温度相关系数接近 0.95，验证了物理先验特征在表征热力状态方面的有效性。

2 模型

2.1 模型架构

锅炉燃烧系统运行数据具有非线性、大迟延及局部突变特征，单一的卷积神经网络难以捕捉长距离依赖，而纯 Transformer 结构在处理连续数值的时序局部趋势时存在归纳偏置不足的问题。为此，本文提出一种融合物理先验特征的 PID-CTNet 协同预测模型，其模型架构如图 2 所示。

该模型采用串行融合架构，模型首先将 27 维关键物理特征序列 $X_t \in \mathbb{R}^{B\times L\times D}$ 输入时间卷积神经网络（temporal convolutional networks，TCN）局部卷积模块，利用一维卷积与残差连接对物理特征进行浅层特征提取与维度映射；随后，将提取的特征序列送入 Transformer 时序编码器，利用多头自注意力机制捕捉燃烧过程的全局动态关联；最后，提取当前时刻的隐层状态，通过参数硬共享机制分发至 5 个独立任务头，实现对关键状态的协同预测。

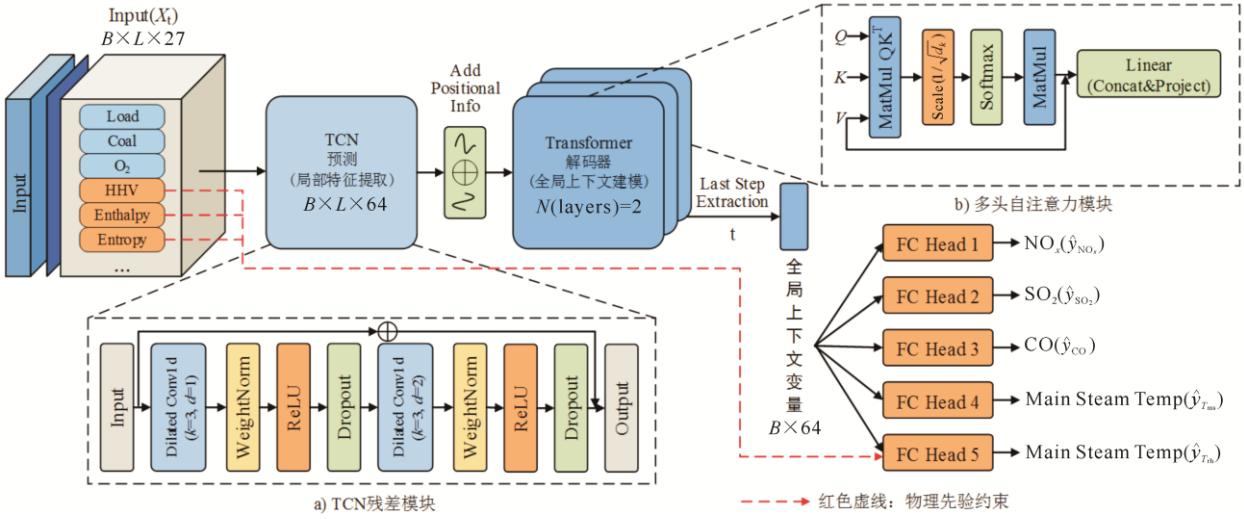


图 2 融合物理机理的 PID-CTNet 协同预测模型架构

Fig.2 Architecture of the PID-CTNet collaborative prediction model integrating physical mechanisms

2.2 TCN 局部特征提取模块

锅炉运行数据中包含大量的局部瞬态特征（如磨煤机启动瞬间的电流脉冲、风压的快速波动）。为了增强模型对局部趋势的感知能力，并在进入 Transformer 前对高维物理特征进行降维与平滑，设计了基于残差结构的时域卷积模块（TCN Block）。该模块由 2 个串联的一维卷积层（Conv1d）与 1 个残差连接（residual connection）组成，对于输入序列 X，计算过程为：

$$H_{\text{conv}} = \sigma(W_2 * \sigma(W_1 * X + b_1) + b_2) \tag{2}$$

$$Y_{\text{tcn}} = H_{\text{conv}} + W_{\text{res}} * X \tag{3}$$

式中：$*$表示一维卷积操作；σ为 ReLU 激活函数；W_1、W_2 均为卷积核权重；W_{res} 为残差映射权重（用于匹配维度）；b_1 与 b_2 均为线性映射的偏置。该模块的作用主要体现在两方面。

1）特征映射　将 27 维关键物理量映射至高维隐层空间（Hidden Dim=64），增加特征表达能力。

2）局部平滑与去噪　利用卷积核的局部感受野，滤除传感器数据中的高频随机噪声，提取具有物理意义的局部变化趋势。

2.3 Transformer 编码器

经过 TCN 提取的特征序列虽已实现局部瞬态特征的有效表征，但受限于其有限的感受野范围，难以捕捉锅炉燃烧系统中普遍存在的大迟延长时序依赖关系，例如给煤量扰动需经过燃烧、传热、汽水循环等多环节耦合传递，其对主/再热蒸汽温度的影响常滞后数分钟至十多分钟。为克服这一局限，引入 Transformer 编码器作为后端时序建模单元，利用其自注意力机制的全局感受野特性，实现对跨时间尺度因果关联的精准捕获。

2.3.1 多头自注意力机制

Transformer 的核心在于自注意力机制，它能够

动态计算历史时刻 t–k 对当前时刻 t 的重要性权重。对于 TCN 输出的序列 $Z=Y_{\text{tcn}}$，通过 3 个线性变换生成查询矩阵 $\boldsymbol{Q}$、键矩阵 $\boldsymbol{K}$ 和值矩阵 $\boldsymbol{V}$，即：

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\text{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{4}$$

采用多头机制（Heads=4），允许模型从“燃料–排放”“风烟–温度”等多个不同的物理子空间并行关注序列的动态关联，从而有效解决大迟延系统的信息滞后问题。

2.3.2 前馈网络与层归一化

每个注意力层后接一个全连接前馈网络（feed-forward network，FFN），与 2.2 节中的卷积层处理不同，FFN 旨在对注意力机制提取的全局特征进行逐位置（position-wise）的非线性增强，其结构由 2 层线性层及中间的激活函数构成，即：

$$\text{FFN}(x)=\sigma(zW_1+b_1)W_2+b_2 \tag{5}$$

式中：z 为注意力机制的输出；σ为激活函数；b_1 和 b_2 分别为对应线性变换的偏置项（bias）。紧接着，模型引入层归一化（layernorm，LN）与残差连接（residual connection），以确保深层网络的训练稳定性，该子层的最终输出 Output 为：

$$\text{Output}=\text{LN}(z+\text{FFN}(z)) \tag{6}$$

通过这种设计，模型能有效防止深层网络中的梯度消失问题。经过多层堆叠处理后，模型提取序列在最后一个时间步 t 的输出量 h_t，作为包含了全历史信息的上下文变量，输入至多任务预测头，即：

$$h_t=\text{Transformer_Encoder}(X)_t \tag{7}$$

2.3.3 多任务协同预测头

为了利用锅炉各状态变量间的物理耦合性，如炉膛温度场同时决定 NO_x 生成与主蒸汽温度，模型后端设计了 5 个并行的任务预测头。上下文变量 h_t 首先经过一个公共的投影层进行特征融合，随后分发至 5 个独立的子网络，如式(8)所示：

$$\hat{y}_t^m=\text{Linear}^m(\text{ReLU}(\text{Linear}_{\text{proj}}(h_t))) \tag{8}$$

式中：$m\in\{NO_x, SO_2, CO, T_{\text{ms}}, T_{\text{rh}}\}$。这种“主干共享+独立分支”的结构构成了多任务学习（multi-task learning）框架。在反向传播过程中，预测蒸汽温度产生的梯度会更新前端共享的 TCN 与 Transformer 参数，这种机制迫使模型学习到更鲁棒的、能够同时解释燃烧与换热过程的通用物理表征，起到正则化的作用，有效降低单任务预测的过拟合风险。

2.4 损失函数

为了平衡不同任务的学习难度与量纲差异，总损失函数 L_{total} 定义为各子任务均方误差的加权和：

$$L_{\text{total}}=\sum_{m=1}^{M}\lambda_m\cdot\frac{1}{N}\sum_{i=1}^{N}(y_i^m-\hat{y}_i^m)^2 \tag{9}$$

式中：M=5 为任务总数；λ_m 为第 m 个任务的权重系数，各任务权重均设为 1.0，是为了避免人为设定权重带来的偏置，并保持各任务在训练过程中的梯度贡献处于同一量级。

采用均等权重策略，一方面，各子任务均采用均方误差作为损失形式，量纲一致、数值范围相近，能够在同一优化尺度下稳定训练；另一方面，环保排放指标与运行安全指标在工程应用中同等重要，因此在任务重要性上不存在需要额外强调的优先级。在预实验中尝试过对排放类任务给予更高权重，但整体性能提升不明显，且会导致训练不稳定，因此最终采用均等权重作为最稳健的配置，以同等关注环保指标与运行安全指标。模型采用 Adam 优化器进行端到端训练，通过反向传播算法同步更新共享参数θ_{shared}与各子任务参数θ^m。

2.5 评价指标

选取均方根误差δ_{RMSE}（root mean square error，RMSE）、平均绝对误差δ_{MAE}（mean absolute error，MAE）、平均绝对百分比误差δ_{MAPE}（mean absolute percentage error，MAPE）和决定系数 R^2（coefficient of determination）为模型预测评估指标[24]，对预测结果进行合理评估。其中，均方根误差能够反映预测误差的整体波动性，对较大偏差更为敏感；平均绝对误差衡量预测值与真实值之间的平均绝对偏差，具有更直观的物理意义；平均绝对百分比误差通过归一化处理刻画相对误差大小，便于不同量纲变量之间的横向比较；而 R^2 用于衡量模型对真实数据的解释能力，其值越接近 1，说明模型拟合效果越好。各计算公式为：

$$\delta_{\text{RMSE}}==\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2} \tag{10}$$

$$\delta_{\text{MAE}}=\frac{1}{N}\sum_{i=1}^{N}|y_i-\hat{y}_i| \tag{11}$$

$$\delta_{\text{MAPE}}=\frac{1}{N}\left(\sum_{i=1}^{N}\frac{|y_i-\hat{y}_i|}{y_i}\right)\times100\% \tag{12}$$

$$R^2=\sum_{i=1}^{N}(y_i-\hat{y}_i)^2\Big/\sum_{i=1}^{N}(y_i-\bar{y})^2 \tag{13}$$

2.6 模型流程

PID-CTNet 模型训练流程如图 3 所示。

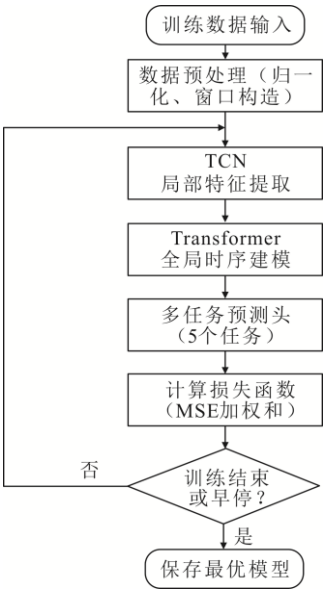


图 3 PID-CTNet 模型训练流程

Fig.3 Flowchart of PID-CTNet model training

3 实验

3.1 实验环境及数据集划分

实验数据来源与特征构建过程已在数据处理一节中详细说明。为验证所提 PID-CTNet 模型的有效性，将处理后的多变量时间序列数据按 0.7:0.15:0.15 的比例划分为训练集、验证集与测试集，以保证模型在不同工况下的泛化能力。

在模型训练阶段，滑动窗口长度设为 L=6（对应 1 h 历史数据），批量大小 Batch=64，最大迭代次数 Epochs=100，并采用早停机制避免过拟合。优化器采用 Adam，初始学习率设为 10^{-4}，并使用余弦退火策略进行学习率衰减，以提升训练稳定性，PID-CTNet 编码器卷积核大小为 3，Transformer 多头注意力头数为 4，隐藏维度为 64。该结构在预实验中表现出较好的收敛速度与泛化能力。实际训练过程中，PID-CTNet 模型展现出优异的学习能力，通常在 50 次迭代内即可触发早停逻辑完成收敛。

为了确保不同模型在相同条件下进行公平对比，所有基准模型（SVR、LSTM、GRU）均使用相同的数据预处理流程、窗口构造方式及训练集划分策略。所有实验均在 Windows 10 环境下完成，处理器为 Intel(R) Core(TM) i5-6300HQ CPU@2.30 GHz，内存容量 32 GB，实验开发工具采用 PyCharm，编程语言为 Python。

3.2 对比实验分析

为了验证所提 PID-CTNet 模型的预测性能，本节研究选取了 3 类具有代表性的基准模型进行对比实验。使用传统机器学习模型（SVR）作为回归问题的经典基准，验证浅层模型在处理高维非线性锅炉数据时的局限性；采用循环神经网络变体（LSTM、GRU）作为处理时序数据的工业标准方法，用于对比提出的模型在捕捉长程依赖及特征提取方面的性能提升。各模型在测试集上的预测性能对比结果如表 2 所示。

表 2 不同模型的预测性能对比

Tab.2 Comparison of prediction performance of different models

模型	δ_{RMSE}	δ_{MAE}	δ_{MAPE}	R^2	Params/kb
SVR	7.126	4.053	1.452	0.782	N/A
LSTM	5.542	3.103	1.054	0.841	57 413
GRU	5.685	3.189	1.088	0.837	43 141
PID-CTNet	2.895	1.654	0.565	0.957	580 229

通过对表 2 中的数据进行分析，可以看出提出模型表现出卓越的性能。首先是精度的提升，即便是与表现最好的基准模型 LSTM 相比，提出模型也在各项误差指标上实现了大幅度下降，均方根误差从 5.542 降至 2.895，下降了约 47.8%，平均绝对百分比误差指标也大幅度压缩，表明预测误差进一步降低，结果表明通过 TCN 与 Transformer 的串并行结合，模型能够更精准地捕捉到燃烧过程中的瞬态变化与长期趋势。其次是拟合优度的突破，在判定系数 R^2 方面，基准模型（SVR、LSTM、GRU）均处于 0.85 以下，难以应对复杂工况下的波动。而提出模型的 R^2 达到了 0.957，这意味着该模型能够解释锅炉系统中 95.7%的变量波动，验证了其在多变量预测任务中的高度可靠性。但是，也可以看到提出模型的参数量（580 kb）明显大于 LSTM 与 GRU，这样增加了计算开销，但是带来的精度收益是巨大的，在锅炉燃烧这种对安全性与节能要求极高的场景下，采用更深、更复杂的模型架构（Transformer）以换取近乎翻倍的预测精度是完全必要且值得的。

3.3 消融实验

3.3.1 模块消融

为了进一步探究 PID-CTNet 模型中各核心组件对预测性能的具体贡献，本节进行模块消融实验。消融实验的基本思路是：通过“控制变量法”，在保持其他网络结构和超参数不变的前提下，有针对性地移除或替换模型中的某个核心模块，观察模型性

能指标的变化。设置了以下 2 个变体模型：

1）w/o TCN　移除前端的时间卷积网络（TCN），直接将数据输入 Transformer，此实验旨在验证 TCN 在提取局部动态特征及多尺度感知方面的作用。

2）w/o Transformer　移除后端的 Transformer 模块，仅依靠 TCN 进行特征建模，此实验旨在验证自注意力机制在捕捉长距离时序依赖及全局信息整合中的必要性。模块消融实验的定量分析结果如表 3 所示。

表 3 模块消融实验结果
Tab.3 Results of the module ablation experiment

模型	δ_{RMSE}	δ_{MAE}	δ_{MAPE}	R^2	Params/kb
w/o TCN	5.854	3.351	1.154	0.825	564 421
w/o Transformer	4.356	2.568	0.941	0.903	17 925
PID-CTNet	2.895	1.654	0.565	0.957	580 229

通过对表 3 的数据进行分析，可以看出核心模块的协同效用，单一使用 TCN 与 Transformer 模块的结果各项指标明显上升，证明了“局部特征提取”与“全局依赖建模”两者缺一不可，双分支架构的串并行结合产生了显著的性能增益。TCN 对局部波动捕捉的重要性体现在 w/o TCN 变体中，模型的 R^2 出现了最为明显的下降，从 0.957 降至 0.825，且均方根误差增幅最大。这说明在锅炉燃烧这种具有高频波动和瞬态响应特性的系统中，缺乏 TCN 的扩张因果卷积支撑，模型将难以准确捕捉燃烧工况的细微突变，导致整体拟合能力大幅削弱。Transformer 对长效趋势的建模能力在 w/o Transformer 情况下，虽然模型保持了较小的参数量（仅 17.9 kb），但其平均绝对百分比误差指标显著升高至 0.941%。这表明尽管 TCN 擅长处理局部特征，但在缺乏 Transformer 全局自注意力机制的情况下，模型难以有效整合长历史窗口内的时序信息，增加 Transformer 模块虽然显著提升了参数量，但换来的是模型对复杂非线性耦合关系更深层次的理解。

3.3.2 参数消融

在验证了模块协同的有效性后，本节进一步开展参数层面的消融实验，旨在定量评估引入的“物理先验知识”对模型性能的贡献，实验通过调整输入特征集来观察模型表现，设置以下 2 个变体模型：

1）w/o PID-CTNet（No Enthalpy）　在保持模型架构与多任务优化策略不变的前提下，移除基于 IAPWS-IF97 构建的热力学先验特征（主蒸汽与再热蒸汽的焓值、熵值），仅保留原始 DCS 测点数据作为输入，该变体用于评估热力学物理先验对模型性能的增益。

2）w/o PID-CTNet（Full）　代表基础的纯数据驱动模型，不包含任何物理先验增强，也不采用多任务协同优化，仅使用原始监测量直接训练单任务预测模型，该模型用于确立性能基线，以对比完整模型在物理先验与多任务机制共同作用下的性能提升。具体的参数消融实验结果如表 4 所示。

表 4 参数消融实验结果
Tab.4 Results of parameter ablation experiments

模型	δ_{RMSE}	δ_{MAE}	δ_{MAPE}	R^2	Params/kb
w/o PID-CTNet (No Enthalpy)	3.561	2.012	0.751	0.929	579 845
w/o PID-CTNet (Full)	4.187	2.254	0.723	0.914	585 413
PID-CTNet	2.895	1.654	0.565	0.957	580 229

结合表 4 的数据结果从物理机理的角度进行分析，可以看到物理先验特征的显著增益，对比 No Enthalpy 版本与提出的完整模型可以发现，引入物理特征后，均方根误差指标从 3.561 显著下降至 2.895（优化幅度达 18.7%）。这有力地证明了焓、熵等热力学参数不仅是数据的简单补充，更是对燃烧过程内在能量转换规律的有效提取。物理特征通过将复杂的非线性关系进行预先“解耦”，降低了深度学习模型从原始高噪声数据中学习机理的难度。“物理-数据”双驱动具有优越性，与基础的纯数据驱动模型 w/o PID-CTNet（Full）相比，提出模型的性能提升更大，均方根误差从 4.187 降低至 2.895，降幅高达 30.8%，同时，R^2 指标从 0.914 提升至 0.957，这说明单纯增加模型复杂度或单纯依赖数据规模，其提升效果远不如将“物理机理”作为归纳偏置引入模型，物理先验知识起到了导航的作用，引导模型在更符合物理逻辑的方向上进行参数寻优。

3.3.3 任务消融

本节进行任务消融实验，探究多任务学习（multi-task learning）架构对各预测目标的协同增强作用，提出模型 PID-CTNet（All）是多任务预测模型，同时预测 NO_x、SO_2 等 5 个目标。为了验证多任务协同预测相较于独立任务预测的优越性，构建 5 个仅针对单一目标进行优化的单任务对比模型，这些模型在网络深度、超参数设置及物理特征输入上与完整模型保持一致。具体的消融实验结果如表 5 所示。

表 5 任务消融实验结果
Tab.5 Results of the task ablation experiment

模型	δ_{RMSE}	δ_{MAE}	δ_{MAPE}	R^2	Params/kb
PID-CTNet (NO_x)	3.521	2.102	0.684	0.932	515 703
PID-CTNet (SO_2)	2.957	1.759	0.591	0.951	515 703
PID-CTNet (CO)	5.854	3.476	1.153	0.865	515 703
PID-CTNet (T_{ms})	1.154	0.694	0.213	0.985	515 703
PID-CTNet (T_{rh})	1.357	0.789	0.256	0.978	515 703
PID-CTNet (All)	2.895	1.654	0.565	0.957	580 229

通过对表 5 的数据进行对比，可以看到任务间的协同增强效应，对比单任务模型与多任务模型可以发现，多任务架构在处理复杂耦合变量时综合表现更优，以 NO_x 为例，单任务预测时的均方根误差为 3.521，而在多任务协同模式下，整体平均预测性能提升至 2.895。这表明通过硬参数共享机制，模型捕捉到了不同排放物与运行参数之间的内在物理联系，实现了信息互补。针对难预测指标的性能修复，CO 的预测在单任务模式下误差极大，均方根误差为 5.854，平均绝对百分比误差为 1.153%，这是由于 CO 排放具有极强的非线性和瞬态波动性。但是在多任务框架下，得益于蒸汽温度等稳定信号的辅助监督，模型对燃烧状态的整体感知能力增强，从而有效抑制了高频噪声对 CO 预测的干扰。参数量上综合较优，所提出的模型在更小的计算成本及存储空间的同时，获得了更高的预测精度，实现了效率与性能的深度权衡。

进一步分析可以发现，主蒸汽温度（T_{ms}）与再热蒸汽温度（T_{rh}）在独立任务预测中表现优异，而排放类指标（NO_x、SO_2、CO）在单任务模式下显著退化，这一“趋势相反”的现象与不同物理量的过程特性密切相关。首先，T_{ms} 与 T_{rh} 属于典型的缓变、强约束、强可控量，其动态主要由燃烧放热与受热面吸热过程决定，并处于完善的闭环控制系统中，时序变化平稳、信噪比较高，物理因果链条清晰。因此，即便在单任务模式下，模型也能依托稳定的热力学关系学习到较准确的映射，从而保持较低的预测误差。相比之下，NO_x、SO_2、CO 等排放指标具有更强的非线性、突变性与工况敏感性，不仅受燃料特性、空气分配、炉膛温度场影响，还受到喷氨脱硝、配风策略、负荷切换等多因素扰动。在单任务模式下，模型难以从高噪声、弱可控的排放数据中提取稳定规律，容易出现过拟合或对局部工况的偏置。多任务框架通过共享主干网络，使排放预测能够借助蒸汽温度等平稳目标的“辅助监督”，在特征空间中形成更具物理一致性的表示，从而显著提升排放类指标的泛化能力。

所以 T_{ms} 与 T_{rh} 在单任务下表现良好，而排放指标更依赖多任务协同，这从物理机理层面验证了多任务学习在强耦合锅炉系统中的必要性与有效性。

3.4 消融实验结果分析

通过上述对比实验与消融实验的图表分析，已经验证了 PID-CTNet 在各项指标上的优越性。为了更直观地展示所提模型在实际工业动态工况下的实时追踪能力，本节使用不同的模型对测试集中随机连续的 200 个采样点进行预测，预测结果进行了可视化展示，如图 4 所示。需要说明的是，图 4 中相邻采样点的实际时间间隔为 10 min，与 DCS 数据的采样周期一致，因此 200 个连续采样点约对应 33 h 的实际运行过程。

图 4 中自上而下分别展示了 NO_x、SO_2、CO 排放质量浓度以及主蒸汽温度（T_{ms}）、再热蒸汽温度（T_{rh}）的预测曲线对比，分析可知：

1）排放指标（NO_x、SO_2、CO 排放质量浓度）的拟合表现出瞬态追踪能力，在 NO_x 和 SO_2 曲线中，工况波动较为频繁，可以看到红色实线（PID-CTNet）几乎与黑色实线（实际值）完全重合，尤其是在波峰与波谷的极值点处，所提模型表现出了比 LSTM 和 GRU 更快的响应速度，有效克服了传统时序模型的“滞后”现象。还有高频噪声抑制能力，CO 质量浓度受燃烧瞬时状态影响，波动剧烈且规律性较差。对比模型（SVR 和 w/o Transformer）在此处出现了明显的震荡和偏离，而 PID-CTNet 凭借 TCN 的局部特征提取能力和多任务协同机制，在保持趋势准确的同时，过滤了不必要的随机噪声。

2）汽温指标（T_{ms}、T_{rh}）的预测表现出稳定性，适合大惯性系统建模，主、再热蒸汽温度属于大滞后、大惯性参数，波动幅度虽小但对精度要求极高。从最后 2 幅子图中可以看出，模型在该指标上也保持了极高的预测精度，所提模型在 565~568 ℃这一极窄的波动范围内，依然保持了极高的拟合优度，相比之下，消融模型（w/o TCN）在温度转折点的捕捉上略显迟钝，这再次印证了 TCN 与 Transformer 串并行架构在处理复杂动力学系统时的鲁棒性。

通过观察全局 5 个任务的拟合情况，PID-CTNet（红色）始终是最贴近真实值曲线的模型，对比模型（SVR、LSTM、GRU）往往只能捕捉大致趋势，

在细节拟合上与提出模型存在明显差距，消融模型（w/o TCN、w/o Transformer）在面对复杂突变工况时容易产生相位偏差，说明缺失任何一个核心模块都会削弱模型对锅炉非线性特性的表达能力。

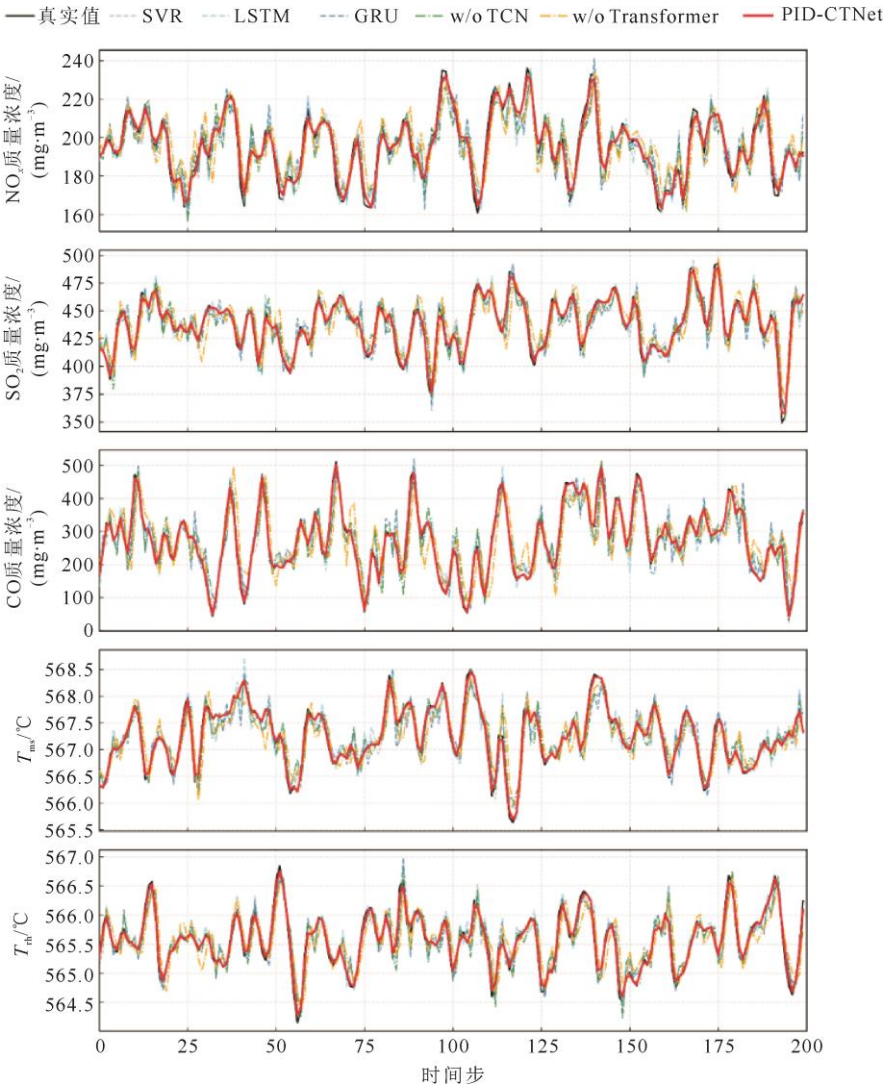


图 4 真实数据预测效果对比

Fig.4 Comparison chart of real data prediction results

4 结 论

针对锅炉燃烧多变量监测，提出融合物理信息的 PID-CTNet 模型，并通过多组实验验证了其有效性，主要结论如下：

1）模型预测性能显著提升。PID-CTNet 在测试集上取得均方根误差为 2.895、平均绝对误差为 1.654、平均绝对百分比误差为 0.565%、R^2=0.957，整体性能明显优于传统 SVR、LSTM 等基准模型。

2）物理先验与模块协同有效增强模型鲁棒性。TCN-Transformer 结构实现局部高频与全局依赖的协同建模；引入焓、熵等热力学先验后，均方根误差进一步降低 18.7%，在复杂工况下表现更稳定。

3）多任务学习实现高效的物理耦合建模。多任务架构通过硬参数共享捕捉排放与蒸汽温度等变量间物理联系，在仅使用单任务总参数量约 23%的情况下完成五指标同步预测，兼顾精度与计算效率。

未来将进一步探索在线学习、模型轻量化及将物理方程嵌入损失函数，以提升工业部署能力。

[参 考 文 献]

[1] 高达, 王峰, 刘伟华, 等. 新型电力系统下熔盐储热技术的研究进展[J]. 动力工程学报, 2026, 46(1): 93.
Gao D, Wang F, Liu W H, et al. Research progress of molten salt thermal energy storage technology under new power system[J]. Journal of Chinese Society of Power Engineering, 2026, 46(1): 93.

[2] 周念成, 徐镇宇, 池源, 等. 计及快速变负荷下煤耗特性及分摊成本的多耦合系统优化调度[J]. 电力系统保

护与控制, 2025, 53(23): 163.
Zhou N C, Xu Z Y, Chi Y, et al. Optimal dispatch of multi-coupling systems considering coal consumption characteristics and cost allocation under rapid load changes[J]. Power System Protection and Control, 2025, 53(23): 163.

[3] Hong F, Zhao Y Z, Ji W M, et al. A feature-state observer and suppression control for generation-side low-frequency oscillation of thermal power units[J]. Applied Energy, 2024, 354: 122179.

[4] 刘振兴, 张锡乾, 杨宏伟, 等. 不同负荷下燃准东煤锅炉的 NO_x 排放特性和炉内结渣特性研究[J]. 热能动力工程, 2025, 40(7): 120.
Liu Z X, Zhang X Q, Yang H W, et al. Study on NO_x emission and slagging characteristics of a Zhundong coal-fired boiler at different loads[J]. Journal of Engineering for Thermal Energy and Power, 2025, 40(7): 120.

[5] 钟天东, 喻小伟, 邹宜金, 等. SCR 脱硝系统运行过程中存在的问题分析[J]. 清洗世界, 2021, 37(8): 59.
Zhong T D, Yu X W, Zou Y J, et al. Analysis of problems existing in the operation process of SCR denitrification system[J]. Cleaning World, 2021, 37(8): 59.

[6] Li G, Wei Y, Li J, et al. Machine learning-driven combustion pressure reconstruction and prediction for free-piston linear generators using kinematic parameters[J]. Energy, 2025, 335: 138140.

[7] 陈亚平, 王焕明, 孙胡彬, 等. 基于煤仓分层和软测量的入炉煤实时监测[J]. 动力工程学报, 2025, 45(2): 190.
Chen Y P, Wang H M, Sun H B, et al. Real-time monitoring of furnace feed coal based on layering model and soft measurement of coal bunker[J]. Journal of Chinese Society of Power Engineering, 2025, 45(2): 190.

[8] 姚顺春, 李龙千, 卢志民, 等. 机器学习驱动锅炉燃烧优化技术的现状与展望[J]. 洁净煤技术, 2024, 30(2): 228.
Yao S C, Li L Q, Lu Z M, et al. Current situation and prospect of machine learning-driven boiler combustion optimization technology[J]. Clean Coal Technology, 2024, 30(2): 228.

[9] Guillermo R, Angel-Iván G. A machine learning-based approach for flames classification in industrial heavy oil-fire boilers[J]. Expert Systems with Applications, 2024, 238: 122188.

[10] 谭增强, 牛拥军, 李元昊, 等. 火电厂氨逃逸的预测研究[J]. 动力工程学报, 2023, 43(7): 917.
Tan Z Q, Niu Y J, Li Y H, et al. Study on prediction of ammonia escape from thermal power plants[J]. Journal of Chinese Society of Power Engineering, 2023, 43(7): 917.

[11] 王伟同, 范海东, 梁成思, 等. 基于随机森林算法的对冲锅炉出口 NO_x 排放量预测模型研究[J]. 热力发电, 2022, 51(4): 96.
Wang W T, Fan H D, Liang C S, et al. Predictive modeling of NO_x outlet of hedged boiler based on random forest[J]. Thermal Power Generation, 2022, 51(4): 96.

[12] Zhang J, Niu K, Nie Q, et al. Prediction of NO_x emission based on online-SVR model adapts to ultra-supercritical coal-fired electric utility boilers[J]. International Journal of Energy Research, 2025, 2025(1): 5597462.

[13] 孙安良, 武利斌, 湛戌, 等. SCR 脱硝系统 NO_x 浓度预测模型与应用[J]. 重庆理工大学学报(自然科学), 2023, 37(10): 312.
Sun A L, Wu L B, Zhan X, et al. Prediction model and application of NO_x emission of SCR denitrification system[J]. Journal of Chongqing University of Technology (Natural Science), 2023, 37(10): 312.

[14] 董渊博, 茅大钧, 章明明. 基于CNN-LSTM的燃气轮机 NO_x 排放预测研究[J]. 热能动力工程, 2021, 36(9): 132.
Dong Y B, Mao D J, Zhang M M. Research on NO_x emission prediction of gas turbine based on CNN-LSTM[J]. Journal of Engineering for Thermal Energy and Power, 2021, 36(9): 132.

[15] Wang Y, Chen S, Kong Q, et al. High-precision concentration detection of CO_2 in flue gas based on BO-LSTM and variational mode decomposition[J]. Measurement Science and Technology, 2024, 35(9): 095202.

[16] Zhou J, Zheng K, Sun Y. Deep reinforcement learning (DRL)-driven intelligent scheduling of virtual power plants[J]. Energies, 2025, 18(23): 6341.

[17] 赵旭东, 李军, 王林, 等. 基于改进 LSTM 与注意力机制的 SCR 入口 NO_x 生成浓度预测[J]. 动力工程学报, 2025, 45(12): 2044.
Zhao X D, Li J, Wang L, et al. Prediction of NO_x concentration at the SCR inlet based on improved LSTM with attention mechanism[J]. Journal of Chinese Society of Power Engineering, 2025, 45(12): 2044.

[18] Wan A, Du C, Bukhaiti A K, et al. Optimizing combined-cycle power plant operations using an LSTM-attention hybrid model for load forecasting[J]. Journal of Mechanical Science and Technology, 2025, 39(10): 1.

[19] Gotzman K S, Bzymek G, Kania K. Resolving steam turbine casing thermal management challenges with a dual attentive Bi-GRU soft sensor for transient operation[J]. Materials, 2025, 18(22): 5213.

[20] 房方, 吴兴堂, 程龙, 等. 电力智算系统构建理论与关键技术[J]. 中国科学:技术科学, 2026, 56(1): 201.
Fang F, Wu X T, Cheng L, et al. Theoretical framework and key technologies of electric power intelligent computing systems[J]. Scientia Sinica Technologica, 2026, 56(1): 201.

[21] 李苏辉, 张归华, 吴玉新. 面向未来燃气轮机的先进燃烧技术综述[J]. 清华大学学报(自然科学版), 2021, 61(12): 1423.
Li S H, Zhang G H, Wu Y X. Advanced combustion technologies for future gas turbines[J]. Journal of Tsinghua University (Science and Technology), 2021, 61(12): 1423.

[22] Gao Y, Guo Y, Xie Y, et al. Kinetic mechanism and NO_x emission characteristics of the ammonia/alcohol co-combustion explored by reactive molecular dynamics calculation and kinetic numerical simulation[J]. Combustion and Flame, 2025, 273: 113953.

[23] 何江永, 张洁, 邹冲, 等. 热解温度对低阶煤热解半焦动态燃烧机理演变的影响[J]. 煤炭转化, 2025, 48(1): 27.
He J Y, Zhang J, Zou C, et al. Effect of pyrolysis temperature on dynamic combustion mechanism evolution of semi-coke from low-rank coal pyrolysis[J]. Coal Conversion, 2025, 48(1): 27.

[24] Sun J H, Yang Q, Zhao N B, et al. Numerically study of CH_4/NH_3 combustion characteristics in an industrial gas turbine combustor based on a reduced mechanism[J]. Fuel, 2022, 327(1): 137349.

[25] Yang M, Zhang Z, Wu H, et al. A high-accuracy prediction approach for NO_x concentration in coal-fired boilers based on optimal subset generation algorithm and bidirectional gated recurrent unit[J]. Fuel, 2026, 407: 137543.

（责任编辑　杜亚勤）