

DOI: 10.19666/j.rlfed.202607055

基于 IHRO 优化 PGRC-LSTM 的燃煤火电厂碳排放软测量模型研究

许鹏成^{1,2,4}、叶健诚^{1,2}、王传博^{1,2}、徐石明^{1,2}、陈哲^{3,5}、王旭⁴

(1. 国网电力科学研究院有限公司, 江苏省 南京市 211106; 2. 国电南瑞科技股份有限公司, 江苏省 南京市 211106; 3. 国网江苏省电力有限公司电力科学研究院, 江苏省 南京市 211100; 4. 华北电力大学, 北京市 102206; 5. 南京大学, 江苏省南京市 210008)

[摘要] 【目的】针对燃煤火电厂小时级碳排放软测量中数据驱动模型物理一致性不足、残差偏差及关键参数难以协同选择的问题, 构建兼顾精度与机理约束的软测量方法。【方法】基于某燃煤机组 2025 年 8760 个小时样本, 完成时间校核、异常筛查和标准化, 利用烟气氧含量、流速、温度、湿度及压力构造氧量稀释、标准干烟气流量和燃烧强度代理特征。以长短期记忆网络(long short-term memory, LSTM)提取当前及历史运行信息, 将流量、氧量和平滑正则嵌入训练, 经岭回归校正校准集残差, 并采用改进心率优化算法(improved heart rate optimizer, IHRO)在统一 90 次评价预算和 20 次搜索下联合选择预训练物理权重模型及残差校正强度。【结果】12 月 744 个小时测试中, 物理引导残差校正长短期记忆网络(physics-guided residual correction-long short-term memory, PGRC-LSTM)平均 MAPE 为 5.5528%, 较 LSTM 降低 1.4305 个百分点。IHRO-PGRC-LSTM 的 MAE、RMSE、MAPE 和 R^2 分别为 34176.5604 kg、46830.4620 kg、4.4631% 和 0.9511, 四项平均指标均优于表 5 所列模型, MAPE 较 BO 和 HRO 分别降低 0.8342 和 0.3197 个百分点。【结论】所提方法可在现有烟气输入同步可得条件下提高碳排放软测量的平均精度, 并为弱机理信息与数据驱动模型融合提供参考, 其跨机组与跨季节适用性仍需进一步验证。

[关键词] 燃煤火电厂; 碳排放软测量; 改进心率优化算法(IHRO); 物理引导残差校正; LSTM 模型

[引用本文格式] (本段作者勿动) 作者姓名. 中文标题[J]. 热力发电, 2026, 55(1): 起始页码-终止页码. 作者姓名. 英文标题[J]. Thermal Power Generation, 2026, 55(1): 起始页码-终止页码.

Research on Carbon Emission Soft-Sensing Model for Coal-Fired Thermal Power Plants Based on IHRO-Optimized PGRC-LSTM

Xu Pengcheng^{1,2,4}、Ye Jiancheng^{1,2}、Wang Chuanbo^{1,2}、Xu Shiming^{1,2}、Chen zhe^{3,5}、Wang Xu⁴

(1.State Grid Electric Power Research Institute Co., Ltd., Nanjing, Jiangsu Province 211106, China; 2.State Grid NARI Technology Co., Ltd., Nanjing, Jiangsu Province 211106, China; 3.State Grid Jiangsu Electric Power Company Limited - Electric Power Research Institute, Nanjing, Jiangsu Province 211100, China; 4.North China Electric Power University, Beijing 102206, China; 5. Nanjing University, Nanjing, Jiangsu Province 210008, China)

Abstract: [Objective] Carbon emissions from coal-fired power plants vary with combustion organization, flue-gas conditions, and operating-load disturbances. Conventional inventory methods cannot represent these hourly variations in a timely manner, whereas purely data-driven soft sensors may produce estimates that conflict with available process knowledge. This study aims to improve the accuracy and physical consistency of hourly carbon-emission soft sensing and to coordinate the selection of physical regularization and residual-correction parameters under a limited search budget. [Methods] The dataset contained 8,760 hourly records from a large coal-fired unit in

收稿日期: XXXX-XX-XX 修回日期: XXXX-XX-XX 接受日期: XXXX-XX-XX 网络首发日期: XXXX-XX-XX
基金项目: 国网科技项目《面向碳双控的电力碳核算与碳足迹关键技术研究及应用》(521001250037-257-YYGG)

Supported by: State Grid Science and Technology Project《Research and Application on Key Technologies of Carbon Accounting and Carbon Footprint of Electric Power for Carbon Dual Control》(521001250037-257-YYGG)

第一作者简介: 许鹏成, 1987.04, 男, 硕士, 高级工程师, 电力碳核算技术研究, xupc09@163.com.

通信作者简介: 叶健诚, 1981.11, 男, 硕士, 高级工程师, 电碳协同技术研究, 479157663@qq.com.

eastern coastal China, covering 1 January to 31 December 2025. The records underwent temporal-consistency checks, screening of missing and physically invalid values, standardization using training-period statistics, and sliding-window construction. Data from January to October were used for training, November data were used for calibration, and 744 observations from December were reserved for testing. Flue-gas oxygen content, velocity, temperature, humidity, and pressure were used to construct proxy features for oxygen dilution, standard dry flue-gas flow, and combustion intensity. A long short-term memory (LSTM) network extracted current and historical operating information. Flow-direction, oxygen-direction, and temporal-smoothness regularization were incorporated into network training. Ridge regression then learned a mapping of the calibration residuals from 15 physics-related features, forming a physics-guided residual correction LSTM (PGRC-LSTM). An improved heart rate optimizer (IHRO) coupled heart-rate amplitude with fitness feedback, reused known fitness values, and performed a final boundary check in the favorable search interval. IHRO jointly selected one of three pretrained physical-weight models and the ridge regularization coefficient. All compared optimizers used the same candidate set, forward temporal-validation objective, search range, and budget of 90 evaluations. Twenty search seeds were used. Performance was evaluated using mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). [Results] Under the fixed configurations used for basic-model screening, LSTM achieved the lowest mean MAPE of 6.9833% among the 11 compared models. Incorporating physical regularization reduced the mean MAPE to 6.7923%, and residual correction further reduced it to 5.5528%. The PGRC-LSTM therefore lowered MAPE by 1.4305 and 1.2395 percentage points relative to LSTM and the physics-guided regularized LSTM, respectively. In the conditional direction-consistency analysis, its average combined violation rate was 22.33%, compared with 24.91% for LSTM; this statistic was interpreted together with prediction error rather than as independent proof of engineering improvement. After optimization, IHRO-PGRC-LSTM obtained an MAE of 34,176.5604 kg, an RMSE of 46,830.4620 kg, a MAPE of 4.4631%, and an R^2 of 0.9511 on the December test set. All four mean metrics were the best among the models listed in Table 5. The mean MAPE was 1.0897, 0.8342, and 0.3197 percentage points lower than those of unoptimized PGRC-LSTM, Bayesian optimization, and the baseline heart rate optimizer, respectively. Its MAPE standard deviation was higher than that of Bayesian optimization, indicating that the observed advantage primarily concerned mean accuracy rather than the lowest search variability. [Conclusion] The proposed method combines weak physical priors, residual correction, and constrained parameter search to improve hourly carbon-emission soft sensing when the required flue-gas inputs are synchronously available. Because online coal quality, total fuel and air flow, and flue-gas carbon dioxide concentration were unavailable, the imposed relationships are directional constraints rather than a strict carbon balance. The findings are limited to the historical unit data and search space examined. Independent-period, cross-season, and cross-unit validation is required before broader deployment.

Key words: Coal-fired power plant; Carbon emission soft sensing; Improved Heart Rate Optimizer (IHRO); Physics-guided residual correction; LSTM model

燃煤电厂承担电力系统调峰保供和稳定支撑任务,也是电力行业碳排放的重要来源^[1,2]。碳排放核算、市场履约和低碳运行需要连续监测与在线评估,但燃烧组织、烟气状态、负荷波动和环境扰动使排放过程呈现非线性、时变性与工况耦合。传统排放因子法和离线核算难以反映动态变化,直接监测又受设备维护、采样滞后和成本限制。因此,利用运行数据构建碳排放软测量模型,可辅助动态感知排放状态和优化锅炉运行。

时序建模是碳排放软测量的基础,当前排放量受即时烟气状态及历史燃烧过程共同影响。长短期记忆网络(long short-term memory, LSTM)和门控循环单元(gated recurrent unit, GRU)通过门控结构缓解梯度衰减^[3,4],注意力机制强化关键历史状态表达^[5,6]。Bai 等^[7]指出时序卷积网络(temporal

convolutional network, TCN)兼具并行计算和长记忆优势, WaveNet 验证了因果膨胀卷积的序列建模能力^[8]。双层校正 LSTM 已用于光伏功率预测^[9], Temporal Fusion Transformer、Informer、PatchTST 和 iTransformer 进一步发展了深度时序方法^[10-13]。本文选取 LSTM 作为软测量主干,并在其上融合物理正则、残差校正和参数寻优。

物理正则权重和残差校正参数会影响组合模型性能。序贯模型优化、随机搜索和贝叶斯优化为昂贵黑箱问题提供了常用调参方法^[14-17],粒子群、差分进化、灰狼和鲸鱼等群智能算法也广泛用于连续寻优^[18-22]。心率优化算法(heart rate optimizer, HRO)模拟心率变化,调节全局探索与局部开发^[23]。本文将心率幅值与适应度反馈耦合,引入已知适应度复用和末次区间边缘检查,形成 IHRO,并在统

一多折时间验证目标下联合搜索 PGRC-LSTM 的物理权重组合和残差校正强度。

机理特征和物理约束有助于提高软测量的可解释性。物理信息神经网络将控制方程或物理残差嵌入训练^[24,25], 理论引导数据科学以及物理引导、混合建模和约束学习进一步促进科学知识机器学习融合^[26,27]。相关方法已用于光伏功率和锅炉 NO_x 排放建模^[28,29], 部分依赖性分析和个体条件期望图可检验输入输出响应^[30,31]。受煤质、负荷及多组分烟气浓度缺失限制, 本文难以建立严格碳元素守恒约束, 因而利用烟气氧含量、流速、温度、湿度和压力构造标准干烟气流量、氧量稀释及燃烧强度代理特征, 并将局部方向约束嵌入损失函数。

据此, 本文提出 IHRO-PGRC-LSTM 碳排放软测量模型, 将烟气机理特征、LSTM、物理正则、残差校正及 IHRO 寻优结合, 并通过基础模型、物理机制和优化算法对比评价其性能。

1 方法概述

1.1 技术路线

技术路线如图 1 所示。原始烟气数据经时间对齐、缺失与异常处理及标准化后, 构造氧量稀释、干烟气状态修正、标准干烟气流量和燃烧强度代理特征。LSTM 提取历史状态, 训练中加入流量、氧量和平滑正则, 并以校准集残差拟合校正模型。IHRO 联合选择预训练物理权重模型和残差校正强度, 最终在 12 月测试集评价精度。

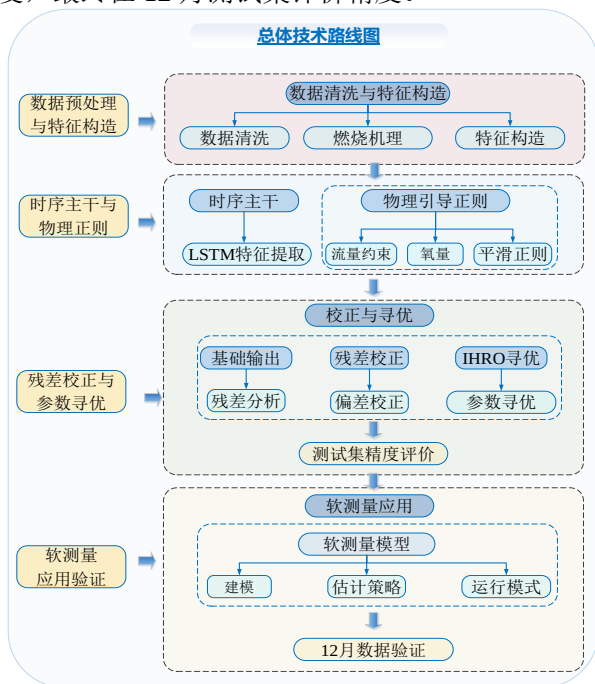


图 1 总体技术路线图

Fig. 1 The technological roadmap

1.2 数据预处理与机理特征构造

异常、漂移和量纲差异会影响工业软测量模型的泛化与稳定性^[32,33]。本文依次进行时间校核、缺失与异常筛查、标准化及滑动窗口构造, 流程见图 2。

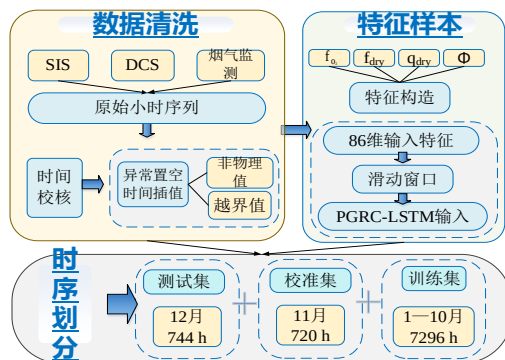


图 2 数据预处理与样本构造流程图

Fig. 2 Flowchart of data preprocessing and sample construction

标准化参数仅由训练段拟合, 校准与测试阶段复用, 再按滑动窗口构造序列。参照碳排放和烟气氧含量软测量研究中的变量筛选与机理特征方法^[34,35], 令流速、温度、湿度、压力和氧含量为 v 、 T 、 H 、 P 和 O_2 , 构造

$$f_{O_2,t} = 1 - \frac{O_{2,t}}{21} \quad (1)$$

$$f_{dry,t} = \frac{P_t + 101.325}{T_t + 273.15} \cdot \left(1 - \frac{H_t}{100}\right) \quad (2)$$

$$q_{dry,t} = v_t f_{dry,t} \quad (3)$$

进一步构造燃烧强度代理量:

$$\phi_t = q_{dry,t} f_{O_2,t} \quad (4)$$

式中, $f_{O_2,t}$ 、 $f_{dry,t}$ 、 $q_{dry,t}$ 和 ϕ_t 分别为氧量稀释、干烟气状态修正、标准干烟气流量及燃烧强度代理量。上述变量由现有烟气监测值计算, 用于表征烟气状态与碳排放强度关系, 不替代严格碳排放核算。

1.3 改进心率优化算法 (IHRO)

1.3.1 心率优化算法 (HRO)

HRO 模拟人体心率调节开展群智能搜索^[23]。基准采用取得的 HROA 实现, 保留均匀随机与反向学习初始化、分组高斯探索、Lévy 扰动、档案引导和逐维择优更新, 并经 MATLAB 运行及 Python 移植核验。

基准 HRO 以相邻轮次最优适应度变化调节步长。本文加入心率幅值控制、已知适应度复用和末次优势区间边缘检查, 改善有限预算下的寻优效果。

1.3.2 改进心率优化算法 (IHRO)

IHRO 将心率幅值与适应度反馈共同用于步长控制, 复用局部候选适应度, 并在末次评价检查优势区域的可行边界。随机与反向学习初始化、Lévy 扰动沿用基准实现, 各算法共用多折验证目标。

设第 i 个搜索个体对应的候选参数向量为:

$$X_i = [b_i]_i, i = \log_{10} \alpha_i \quad (5)$$

式中, b_i 为三个预训练物理权重模型的编号, $\ell_i = \log_{10} \alpha_i$ 为岭系数的对数坐标。模型编号取整并截断至 $\{0,1,2\}$, $\ell_i \in [-4,2]$ 。初始位置均匀随机采样, 并逐对比较反向位置。令 $T=8$ 为迭代轮数上限, 心率幅值定义为

$$h_t = \left| \sin \left(\frac{2\pi t}{T} \right) \right| \quad (6)$$

其中, t 为迭代编号。以本轮开始时与上一轮开始时的最优适应度变化率 Δ_t 构成反馈, 并令心率幅值参与步长系数 a_t 的调节

$$\Delta_t = \frac{|F_{best,t-2} - F_{best,t-1}|}{|F_{best,t-2}| + \varepsilon}$$

$$a_t = \max \left\{ \min \left[1, \max (0, \Delta_t) \right], h_t \right\} \quad (7)$$

式中, $F_{best,t-1}$ 和 $F_{best,t-2}$ 分别为本轮及上一轮开始时记录的最优适应度, ε 为防止分母为零的机器精度, 首轮令 $\Delta_t=0$ 。通常采用下式产生候选位置

$$Y_i = X_i + a_t c_i d \odot z \quad (8)$$

式中, $d=ub-lb$, z 为标准正态随机向量, $\odot$ 为逐元素乘积。前半种群的 $c_i=2$, 后半种群为 0.5。以 0.2 的概率将高斯增量替换为 $0.01L \odot (X_i - X_{best})$, 其中 L 为 $\beta=1.5$ 的 Lévy 随机步长。另以 0.3 的概率进行档案混合

$$Y_i \leftarrow \frac{Y_i + A + 0.1z'}{2} (p=0.3) \quad (9)$$

其中, A 为当前最优档案, z' 为标准正态随机向量。候选越界时截断至合法范围, 随后沿用逐维替换与贪婪保留, 复用本轮已知的相同候选适应度。第 90 次评价若发现当前最优岭系数位于其模型已探索区间边缘且尚未到达该侧可行端点, 则检查该端点, 否则执行原待评价候选。全部算法采用相同三折验证综合误差作为适应度函数。

$$F(X_i) = \text{RMSE}_{\text{val}}(X_i) + 35500 \text{MAPE}_{\text{val}}(X_i) + 0.20 \sigma_{\text{RMSE}} + 1500 \sigma_{\text{MAPE}} \quad (10)$$

1.4 IHRO-PGRC-LSTM 模型

IHRO-PGRC-LSTM 由时序网络、物理引导残

差校正及参数寻优组成。LSTM 提取小时级依赖关系, PGRC 引入流量、氧量和平滑机制, IHRO 在共同候选集合内选择模型编号与残差正则化系数。

1.4.1 基于 LSTM 的碳排放软测量基础网络

LSTM 以输入、遗忘及输出门控制细胞状态, 缓解长序列梯度衰减^[3]。对窗口 $X_{t-L+1:t}$ 逐时更新隐状态和细胞状态, 再将末时刻隐状态映射为同一时刻的小时级 CO_2 排放量, 结构见图 3。

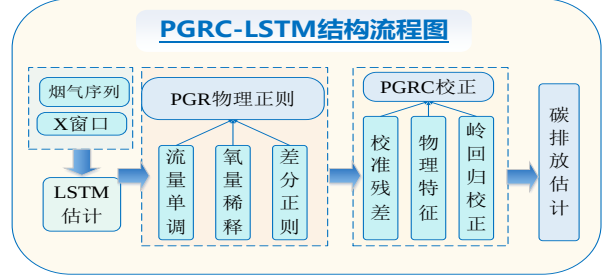


图 3 PGRC-LSTM 软测量模型结构流程图

Fig. 3 Structure flowchart of the PGRC-LSTM soft-sensing model

$$i_t = \sigma(W_i[x_t, h_{t-1}] + b_i),$$

$$f_t = \sigma(W_f[x_t, h_{t-1}] + b_f), \quad (11)$$

$$o_t = \sigma(W_o[x_t, h_{t-1}] + b_o)$$

$$\tilde{C}_t = \tanh(W_c[x_t, h_{t-1}] + b_c),$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t, \quad (12)$$

$$h_t = o_t \odot \tanh(C_t), y_t' = f_o(h_t)$$

式中: i_t 、 f_t 和 o_t 分别为输入门、遗忘门和输出门; C_t 为细胞状态; h_t 为隐状态; $f_o(\cdot)$ 为输出映射函数。

1.4.2 物理机制建模

受现场数据限制, 本研究缺少煤质在线分析、实时入炉煤元素组成及烟气 CO_2 体积分数, 无法建立严格碳平衡。参照物理信息网络及锅炉 NO_x 单调性建模^[24,28,29], 基于可观测烟气量建立图 4 所示弱物理约束。

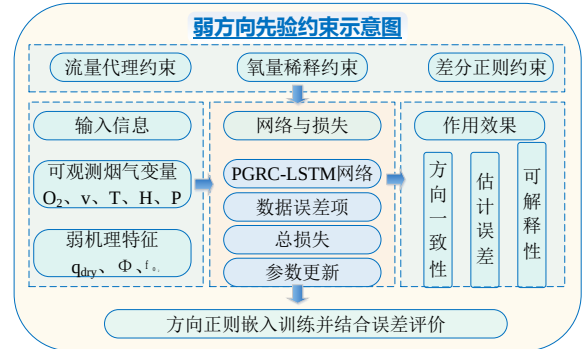


图 4 物理方向先验约束示意图

Fig. 4 Schematic diagram of directional physical priors

在煤质和燃烧组织相对稳定的短时区间, 标准干烟气流量代理量 q_{dry} 增大通常对应单位时间烟气排放通量增加, 据此设定

$$\frac{\partial y'}{\partial q_{\text{dry}}} > 0 \quad (13)$$

式中, q_{dry} 为标准干烟气流量代理量; y' 为模型预测的小时级 CO_2 排放量。

在负荷与烟气流量相近时, 氧含量升高通常对应过量空气增加及 CO_2 浓度稀释, 据此设置方向约束

$$\frac{\partial y'}{\partial \text{O}_2} < 0 \quad (14)$$

式中, O_2 为烟气氧含量。由于本文预测目标为小时级 CO_2 排放量而非烟气 CO_2 浓度, 上式不作为严格热力学等式, 而作为相近运行状态下的方向性物理约束。

为抑制噪声引起的高频波动, 加入相邻时刻平滑约束。总损失由数据误差、流量、氧量和平滑项组成

$$L = L_{\text{data}} + \lambda_1 L_{\text{flow}} + \lambda_2 L_{\text{O}_2} + \lambda_3 L_{\text{smooth}} \quad (15)$$

其中, 软测量误差损失采用均方误差形式:

$$L_{\text{data}} = \frac{1}{n} \sum_{t=1}^n (y_t - y'_t)^2 \quad (16)$$

标准干烟气流量正则根据批内样本按流量代理量排序后的相邻差分计算, 用于惩罚输出与流量代理量的反向变化, 其损失函数为

$$L_{\text{flow}} = \frac{1}{n-1} \sum_{j=2}^n \max(0, -\Delta_{\pi q} q \cdot \Delta_{\pi q} y'_j) \quad (17)$$

氧量正则根据批内样本按氧含量排序后的相邻差分计算, 用于惩罚输出与氧含量的同向变化, 其损失函数为

$$L_{\text{O}_2} = \frac{1}{n} \sum_{j=2}^n \max(0, \Delta_{\pi \text{O}_2} \text{O}_2 \cdot \Delta_{\pi \text{O}_2} y'_j) \quad (18)$$

平滑项按当前顺序批次内相邻窗口输出的差分计算, 其损失函数为

$$L_{\text{smooth}} = \frac{1}{n-1} \sum_{j=2}^n (y_j - y'_j)^2 \quad (19)$$

式中, n 为批次样本数, $\Delta_{\pi q}$ 和 $\Delta_{\pi \text{O}_2}$ 为按流量、氧含量排序后的相邻差分, y'_j 按批次顺序排列, λ_1 、 λ_2 、 λ_3 为正则权重。其物理解释限于给定背景状态下的方向一致性。

未观测煤质、负荷扰动和变量耦合可能造成系统残差, 故在物理正则训练后, 以 PGRC 进一步校

正。校准集残差为

$$r_t = y_t - y'_t \quad (20)$$

以标准干烟气流量代理量、氧含量、燃烧强度代理量、氧量稀释因子、干烟气状态修正因子及其交互项构成残差校正输入向量 $\mathbf{Z}_t$, 采用岭回归学习物理残差映射 $g(\cdot)$:

$$\min_{\theta_r} \sum_{t=1}^m (r_t - g(\mathbf{Z}_t; \theta_r))^2 + \alpha \|\theta_r\|_2^2, \quad y''_t = y'_t + g(\mathbf{Z}_t; \theta_r) \quad (21)$$

式中: r_t 为校准集残差; $\mathbf{Z}_t$ 为物理残差校正输入向量; θ_r 为残差校正模型参数; α 为岭正则化系数; y''_t 为 PGRC 修正后的最终预测值。

1.4.3 IHRO 优化模型的具体流程和参数

IHRO 优化三个预训练物理权重模型的编号 b_i 与岭系数 α_i , 按 11 月验证目标选参后评价 12 月指标, 测试集不用于选参。由于 12 月数据已被查看, 本研究属于历史数据回顾性复核。流程见图 5。

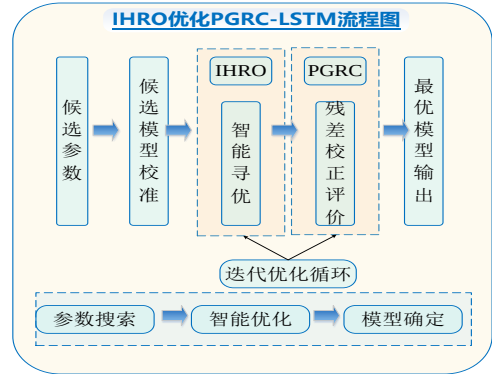


图 5 IHRO 优化 PGRC-LSTM 模型流程图

Fig. 5 Flowchart of the IHRO-optimized PGRC-LSTM model

首先使用训练段得到的三个 PGR 候选模型。按 11 月输出顺序构造三个前向折, 分别以前 50%、60%、70%拟合, 以随后 50%—70%、60%—82%、70%—100%验证。各折独立拟合仿射校准、标准化和岭回归, 以式(10)的三折均值及总体标准差评价候选, MAPE 按百分数代入。全部算法共用候选集、搜索范围和目标, 单次评价预算为 90 次, 使用 9001+101r ($r=0, \dots, 19$) 共 20 个搜索种子。选参后以完整 11 月校准输出重新拟合并评价 12 月测试集, 参数见表 1。

表 1 IHRO 基础参数设置

Table 1 Basic parameter settings of IHRO

参数	设置
优化对象	模型编号 b 、 $\log_{10}\alpha$
物理权重候选组合	PGR-LSTM-w005、w010、w020
岭正则化系数搜索范围	$[10^{-4}, 102]$

参数	设置
种群规模	10
迭代上限 / 评价预算	8 轮 / 90 次评价
初始化方式	均匀随机与反向学习
收敛控制	心率幅值与适应度反馈
评价策略	已知适应度复用及末次边缘检查
适应度函数	多折时间校准稳健综合误差

上述步骤构成 IHRO-PGRC-LSTM 碳排放软测量模型。

1.5 评价指标

采用 MAE、RMSE、MAPE 及 R^2 评价模型 [10,12,34], 分别反映平均绝对偏差、大误差敏感性、相对误差和目标变化解释能力。

$$MAE=\frac{1}{n}\sum_{i=1}^n\left|y_i-y_i'\right| \tag{22}$$

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^n\left(y_i-y_i'\right)^2} \tag{23}$$

$$MAPE=\frac{100\%}{n}\sum_{i=1}^n\left|\frac{y_i-y_i'}{y_t}\right| \tag{24}$$

$$R^2=1-\frac{\sum_{i=1}^n\left(y_i-y_i'\right)^2}{\sum_{i=1}^n\left(y_i-y_{\text{mean}}\right)^2} \tag{25}$$

式中, n 为待预测样本量; y_i 为第 i 个样本的实际碳排放量; y_i' 为第 i 个样本的预测碳排放量; y_{mean} 为测试集实际值均值。MAE、RMSE 和 MAPE 越小, 说明模型软测量误差越低; R^2 越接近 1, 说明模型解释能力越强。

表 3 和表 4 分别给出基础软测量模型与物理机制模型五次训练的指标均值, 图 9 和图 12 的误差棒表示样本标准差。表 4 的 MAPE 差异检验同时对五个共同种子与 24 h 循环时间块进行配对重采样, 共 5000 次, 每次各模型共用种子和时刻索引, 报告 95% 区间, 并对六项与 LSTM 的比较作 Holm 校正。表 5 中各优化算法的指标为共享预训练候选集合下 20 次搜索的均值。

2 数据来源与预处理分析

2.1 变量选取与数据采集

数据来自我国东部沿海某大型燃煤机组烟气连续监测历史数据库, 覆盖 2025 年 1 月 1 日至 12 月 31 日, 间隔 1 h。以烟气运行量估计小时级 CO_2 排放量, 变量见表 2。

本研究面向已配置烟气运行监测系统的燃煤机组, 利用烟气氧含量、流速、温度、湿度和压力

等既有测量信息估计同一时刻的小时级碳排放量, 为运行状态分析提供辅助量化信息。本文的应用价值体现在利用既有监测信息实现碳排放的间接估计。

表 2 碳排放软测量模型相关变量

Table 2 Primary variables for carbon emission soft-sensing model			
标签	变量	描述	单位
x1	O_2	烟气氧含量	%
x2	v	烟气流速	m/s
x3	T	烟气温度	$^{\circ}\text{C}$
x4	H	烟气湿度	%
x5	P	烟气压力	kPa
y	CO_2	小时级二氧化碳排放量	kg

图 6 展示 2025 年 CO_2 排放量及主要烟气变量的变化, 灰线为小时序列, 彩线为 24 h 滑动平均, 虚线标示校准集和测试集边界。 CO_2 排放量与烟气流速总体变化相近, 氧含量在部分时段与排放量呈反向变化。预处理后获得 8760 个小时样本, 1—10 月用于训练、11 月校准、12 月测试。模型比较统一采用 12 月 744 个小时, 其中 12 月 16 日 11 时的无效排放值经时间插值修正, 并在剔除该时刻后的 743 点上敏感性核对。

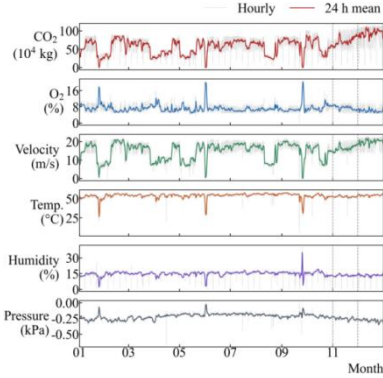


图 6 2025 年全年烟气运行变量与 CO_2 排放量波动特征

Fig. 6 Annual fluctuation characteristics of flue gas variables and CO_2 emissions in 2025

2.2 数据处理过程及效果分析

烟气变量量纲不同, 直接输入会使数值较大的变量主导梯度更新。本文先校核小时序列时间及物理有效性, 再标准化并构造窗口。缺失与非物理、越界值的处理见图 7。

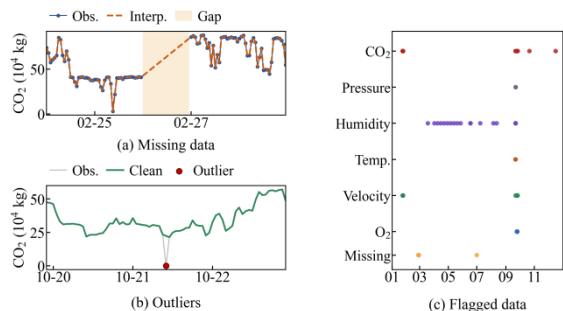


图 7 缺失值与异常值处理结果可视化

Fig. 7 Visualization of missing-value and outlier processing results

图 7 显示,缺失主要集中于 2025 年 2 月 26 日,异常主要涉及烟气湿度、CO₂ 排放量和流速。将非物理或越界值置空后,以离线时间插值补全,蓝色实线和橙色虚线分别表示观测与插值段。在线应用时应改用仅依赖当时可得数据的填补方法。样本充足时可剔除集中异常区段及受影响的输入窗口和滚动特征,避免跨缺口构窗。随后构造机理代理量与多尺度滚动特征,标准化并滑窗采样,变量分布见图 8。

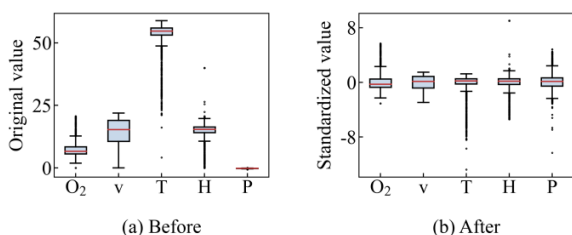


图 8 数据处理前后变量分布对比

Fig. 8 Variable distribution before and after data preprocessing

图 8 中烟气温度约为 50 °C,压力取值较小。训练段 z-score 标准化后,输入均值约 0、标准差约 1,可分布于零两侧,校准和测试段复用训练段参数。

3 结果分析与讨论

3.1 基础软测量模型筛选

比较多层感知机(multilayer perceptron, MLP)、多尺度时序卷积网络(multi-scale temporal convolutional network, MSTCN)、GRU、LSTM、Transformer、TCN、TCN-Attention、TCN-LSTM、TCN-LSTM-Attention、LSTM-Attention 及 CNN-BiLSTM-Attention 等 11 种基础模型。各模型均使用第 2 节的数据划分和 86 维标准化输入,采用 AdamW、批量 128、权重衰减 0.0001 及梯度范数上限 1.0。固定网络结构和超参数,以 101、202、303、404、505 为共同种子独立训练 5 次,并拟合 11 月

仿射校准,表 3 统计 12 月四项指标的均值。

LSTM、GRU、Transformer、TCN、TCN-Attention、TCN-LSTM、CNN-BiLSTM-Attention、MLP 和 MSTCN 采用 48 h 窗长、0.1 的 dropout、0.001 的初始学习率及 90 轮训练。表 3 评价各模型在固定配置下五次独立训练的平均表现。

表 3 基础软测量模型对比结果

Table 3 Comparison results of basic soft-sensing models

模型	MAE	RMSE	MAPE/%	R ²
LSTM	52372.1900	65274.7137	6.9833	0.9158
Transformer	73012.5834	88323.7768	8.0340	0.8455
TCN	77008.4653	91241.4213	8.3039	0.8352
CNN-BiLSTM-Attention	61359.2943	72020.5790	7.2067	0.8973
GRU	59298.9664	74178.9623	8.0995	0.8911
TCN-LSTM	72212.1162	86234.8357	8.5573	0.8527
TCN-LSTM-Attention	90476.1669	105304.4698	10.4306	0.7794
LSTM-Attention	71402.9281	90542.1037	10.2842	0.8372
TCN-Attention	91086.1770	106650.7520	10.0610	0.7748
MLP	76326.0803	97564.2419	10.8549	0.8084
MSTCN	137647.0118	163666.2815	18.6667	0.4668

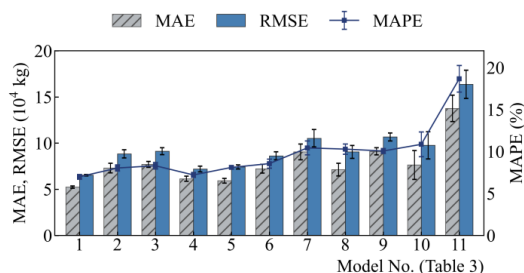


图 9 基础软测量模型误差对比图

Fig. 9 Error comparison of basic soft-sensing models

由表 3 可知, LSTM 的平均 MAE、RMSE 和 MAPE 分别为 52372.1900 kg、65274.7137 kg 和 6.9833%, 均为所比较模型中的最低值, 平均 R² 为 0.9158, 为最高值。CNN-BiLSTM-Attention 的平均 MAPE 为 7.2067%, MSTCN 为 18.6667%, 后者平均 R² 为 0.4668。MSTCN 表现较弱, 可能与多分支结构在当前配置下对本数据时序特征的适配不足有关。图 9 给出重复训练的波动, 其中 LSTM 的 MAPE 和 R² 样本标准差分别为 0.2805 个百分点和 0.0019。综合该组固定配置下的平均精度及与物理引导模块的衔接, 本文选用 LSTM 作为后续研究的基础网络。

3.2 物理机制模型与单调性对比

以 LSTM 为基线, 比较 PI-LSTM、SD-PI-LSTM、PC-LSTM、PGR-LSTM、MPI-PI-LSTM 和 PGRC-LSTM 六种机制。各模型分别引入烟气速度、氧量

与平滑约束, 标准干烟气流代理量, 物理一致性代理项, 流量、氧量与平滑正则, 以及多类物理约束。PGRC-LSTM 在 PGR-LSTM 基础上校正 11 月残差。各模型统一采用 48 h 窗长、两层 96 维 LSTM、

0.1 dropout、0.001 初始学习率、90 轮训练和批量 128, 并使用五个共同种子。PGRC 残差特征为 15 维, 岭系数为 10。表 4 统计 12 月 744 个小时的五次平均误差。

表 4 物理机制对比结果

Table 4 Comparison results of different physical mechanisms					
模型	物理机制	MAE	RMSE	MAPE/%	R ²
PGRC-LSTM	物理引导残差校正	38802.6422	51602.4714	5.5528	0.9473
PGR-LSTM	物理引导正则	51528.9031	64380.1979	6.7923	0.9180
LSTM	基础模型	52372.1900	65274.7137	6.9833	0.9158
SD-PI-LSTM	标准干烟气流量约束	53298.5077	65873.4319	7.1443	0.9140
MPI-PI-LSTM	多物理机制约束	53043.2608	66476.1298	7.2690	0.9123
PC-LSTM	物理一致性约束	53496.9744	66902.4550	7.2599	0.9110
PI-LSTM	原始烟气速度-氧量-平滑约束	56135.3518	68900.5415	7.2268	0.9059

对 12 月 744 个时刻逐一分析输入响应, 保持其余当前输入与历史序列不变, 在训练期 5%—95% 分位内改变氧量, 同步更新派生特征。按 LSTM 反向变化幅度, 选取其随氧量上升而 PGRC 递减且为正的四个不同日期片段, 以各片段流量背景值构造共同二维网格。图 10 蓝色虚线为固定流量下的氧量剖面。

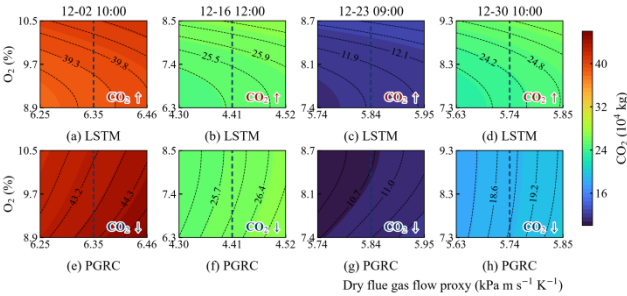


图 10 典型违背片段下的局部响应等高线对比图

Fig. 10 Local response contours at selected violation cases

图 10(a)—(d)对应 12 月 2 日 10 时、16 日 12 时、23 日 9 时和 30 日 10 时。在各片段的流量背景值处, LSTM 估计值随氧量增加分别上升 17597.6、16965.1、9690.0 和 20657.3 kg, 与相近工况下的氧量稀释关系不一致。对应 PGRC 响应(e)—(h)分别下降 11098.8、3915.2、5511.2 和 1841.5 kg, 且估计值均为正, 表明其在这些片段中的氧量响应更符合该物理关系。

以 72 h 窗口、24 h 步长评价方向一致性。先按另一变量的四分位条件分箱, 再按考察变量六分位分箱, 统计相邻箱均值方向。容差取 1800 kg 与窗口估计绝对值中位数的 0.2%中较大者, 将流量增加而输出下降、氧量增加而输出上升计为违背, 两项取平均。图 11 中 LSTM、SD-PI-LSTM、PGR-LSTM、

PGRC-LSTM 平均综合违背率为 24.91%、22.07%、22.67%、22.33%, 后两者较 LSTM 降低 2.24 和 2.58 个百分点。本文结合方向一致性统计与估计误差, 评价物理引导机制的作用。

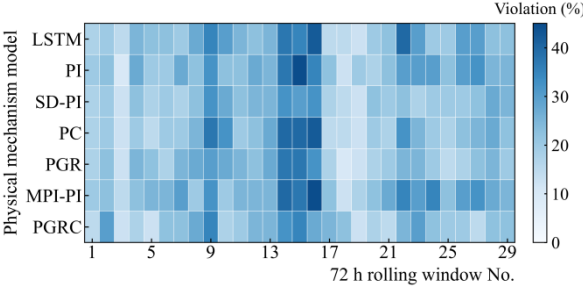


图 11 不同物理机制模型的单调性违背率对比图

Fig. 11 Comparison of monotonicity violation rates for different physical mechanisms

表 4 与图 12 显示, LSTM、PGR-LSTM、PGRC-LSTM 的平均 MAPE 依次为 6.9833%、6.7923%、5.5528%, MAE、RMSE 递减, R2 递增。PGRC 平均 MAE、RMSE 和 R2 为 38802.6422 kg、51602.4714 kg 和 0.9473, 四项指标最优, 按表 4 列示均值计算, MAPE 较 LSTM 和 PGR 分别降低 1.4305 和 1.2395 个百分点, 五个种子均改善, 因此作为后续寻优的基础模型。其余物理机制模型平均 MAPE 为 7.1443%—7.2690%, 高于 LSTM。

按表 4 列示均值计算, SD-PI 相对 LSTM 的平均 MAPE 差为 0.1610 个百分点, 配对重采样 95% 区间为[-0.4182,0.7119], PGR 为-0.1910 个百分点、[-0.5440,0.1756]。两区间均跨 0, Holm 校正 p 均为 1.0000, PGR 仅在五个种子中的四个降低 MAPE。PGRC 相对 LSTM 的表列均值之差为-1.4305 个百分点, 配对重采样 95%区间[-2.0177,-0.8554]、校正

p 为 0.0012, 支持其在本组重复训练及 12 月序列上的误差改善。

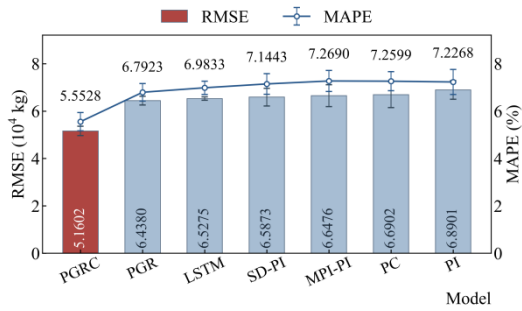


图 12 不同物理机制误差对比图

Fig. 12 Error comparison of different physical mechanisms

3.3 IHRO-PGRC-LSTM 模型性能分析

BO、WOA、PSO、ISA、GWO、HRO 和 IHRO 共用三个预训练 PGR 模型、岭系数范围及验证规则，单次预算为 90 次，种群规模、迭代上限和搜索范围见表 1，20 个搜索种子见第 1.4.3 节。基础 BO 采用高斯过程与 LCB 准则、3 次随机初始化和 2848 个预生成全局候选，并屏蔽已评价配置。表 5 统计相同 12 月 744 点的搜索指标均值，PGRC 参考项采用物理权重 0.010 和岭系数 10。

表 5 优化算法对比结果

Table 5 Comparison results of optimization algorithms

算法	MAE/kg	RMSE/kg	MAPE/% (均值±标准差)	R ²
PGRC 未优化	38802.6422	51602.4714	5.5528±0.3877	0.9473
BO	37450.0725	50630.2516	5.2973±0.5034	0.9491
WOA	38636.3080	51836.6013	5.4848±0.4916	0.9466
PSO	38551.2276	51608.8732	5.4727±0.4105	0.9471
ISA	38658.1344	51745.2294	5.4865±0.4222	0.9469
GWO	38729.8797	51848.6655	5.4932±0.3881	0.9467
HRO	36235.8892	48501.0578	4.7828±0.4916	0.9505
IHRO	34176.5604	46830.4620	4.4631±0.5614	0.9511

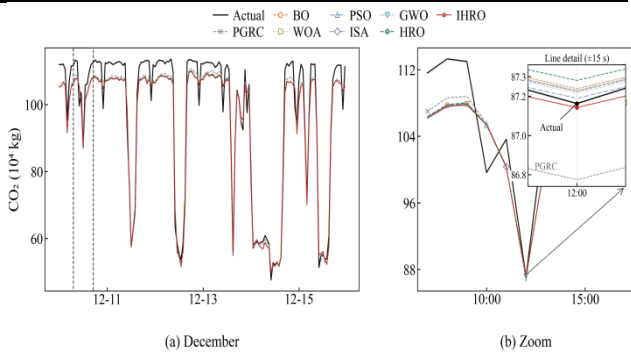


图 13 不同优化算法 PGRC-LSTM 模型估计曲线及局部放大图

Fig. 13 Estimation curves and local zoom of PGRC-LSTM models optimized by different algorithms

表 5 中各优化算法的平均 MAPE 和 MAE 均低于未优化 PGRC，但 WOA、PSO、ISA 和 GWO 的 RMSE 较高且 R^2 较低。IHRO 的平均 MAE、RMSE、MAPE 和 R^2 为 34176.5604 kg、46830.4620 kg、4.4631% 和 0.9511，四项均为表中最优，MAPE 较 PGRC、BO 和 HRO 分别降低 1.0897、0.8342 和 0.3197 个百分点。其 MAPE 标准差为 0.5614 个百分点，高于 BO 的 0.5034 个百分点，表明 IHRO 的优势主要体现在平均精度。图 13(a) 给出 12 月 10 日至 15 日的估计曲线，图 13(b) 放大 12 月 10 日 07:00—17:00 区段及 12:00 附近细部。局部图用于说明估计差异，总体比较以表 5 为准。

4 结论

本文以烟气机理特征构建 LSTM 软测量网络，加入流量、氧量及平滑正则，以校准残差建立 PGRC，并用 IHRO 联合选择物理权重候选模型与残差校正强度，得到以下结论。

(1) 在表 3 固定模型配置、数据处理及评价条件下，LSTM 五次训练的平均 MAE、RMSE、MAPE 最低， R^2 最高，支持其作为本研究的基础网络。

(2) 相同网络与训练配置下，LSTM、PGR-LSTM、PGRC-LSTM 五次平均 MAPE 为 6.9833%、6.7923%、5.5528%，MAE、RMSE 递减， R^2 递增。PGRC 五个种子的 MAPE 均改善，支持物理正则与残差校正组合的作用。

(3) 共用候选集、验证规则和 90 次评价预算时，表 5 中 IHRO 平均 MAPE 最低，为 4.4631%，BO 和 HRO 为 5.2973% 和 4.7828%，分别降低 0.8342 和 0.3197 个百分点，体现其在该联合搜索任务中的平均精度优势。

(4) 表 5 所列 IHRO-PGRC-LSTM 在 12 月的 MAE、RMSE、MAPE、 R^2 为 34176.5604 kg、46830.4620 kg、4.4631%、0.9511，可为输入同步可得时的碳排放软测量提供参考。后续可引入负荷、总煤量、总风量及燃料低位发热量，建立更严格的碳守恒约束，检验跨机组与季节迁移能力。

后续将以独立月份检验输入分布变化下的参数迁移，并结合在线预处理及统一计算环境评价运行效率。

赞助基金：国网科技项目《面向碳双控的电力碳核算与碳足迹关键技术研究及应用》

(项目编码: 521001250037-257-YYGG)

[参考文献]

- [1] Liu Z, Deng Z, He G, et al. Challenges and opportunities for carbon neutrality in China[J]. Nature Reviews Earth & Environment, 2022, 3(2): 141-155.
- [2] 国家能源局. 国家能源局发布 2024 年全国电力工业统计数据[Z]. (2025-01).
National Energy Administration. National Energy Administration releases 2024 national electric power industry statistics. [Z]. (2025-01).
- [3] 周秋阳, 高辉, 李炜卓, 等. 基于 GAN 数据增强与改进 Bi-LSTM 的充电桩故障预测方法[J]. 电力工程技术, 2025, 44(6): 49-61.
ZHOU Q Y, GAO H, LI W Z, et al. A fault prediction method for charging piles based on GAN data augmentation and improved Bi-LSTM[J]. Electric Power Engineering Technology, 2025, 44(6): 49-61. DOI: 10.12158/j.2096-3203.2025.06.005.
- [4] Cho K, Merriënboer B van, Gulcehre C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. 1724-1734, 2014.
- [5] Bahdanau D, Cho K, Bengio Y. Neural machine translation by jointly learning to align and translate[C]//International Conference on Learning Representations, 2015.
- [6] 何海飞, 周良松, 陈俊鹏, 等. 基于 CNN-BiLSTM-Attention 的燃煤电厂二氧化碳排放预测模型[J]. 热力发电, 2026, 55(8): 33-44.
HE H F, ZHOU L S, CHEN J P, et al. Carbon dioxide emission prediction model for coal-fired power plants based on CNN-BiLSTM-Attention[J]. Thermal Power Generation, 2026, 55(8): 33-44. DOI: 10.19666/j.rlfd.202602032.
- [7] Bai S, Kolter JZ, Koltun V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[J]. arXiv preprint arXiv:1803.01271, 2018.
- [8] Oord A van den, Dieleman S, Zen H, et al. WaveNet: A generative model for raw audio[J]. arXiv preprint arXiv:1609.03499, 2016.
- [9] 石宇龙, 彭乔, 刘天琪, 等. 基于相似日和双层校正 LSTM 的光伏功率短期预测方法[J]. 电力工程技术, 2026, 45(3): 85-94.
- [10] Shi YL, Peng Q, Liu TQ, et al. Photovoltaic power short-term forecasting method based on similar days and b_i -layer correction LSTM model[J]. Electric Power Engineering Technology, 2026, 45(3): 85-94.
- [11] Lim B, Arik SO, Loeff N, et al. Temporal fusion transformers for interpretable multi-horizon time series forecasting[J]. International Journal of Forecasting, 2021, 37(4): 1748-1764.
- [12] 詹莹, 王潇添, 王旭, 等. 基于综合距离相似的 CEEMDAN-Informer 光伏功率组合预测模型研究[J]. 太阳能学报, 2025, 46(8): 315-326.
ZHAN Y, WANG X T, WANG X, et al. Combined photovoltaic power forecasting model based on comprehensive distance similarity and CEEMDAN-Informer[J]. Acta Energiae Solaris Sinica, 2025, 46(8): 315-326.
- [13] Nie Y, Nguyen NH, Sinthong P, et al. A time series is worth 64 words: Long-term forecasting with transformers[C]//International Conference on Learning Representations, 2023.
- [14] Liu Y, Hu T, Zhang H, et al. iTransformer: Inverted transformers are effective for time series forecasting[C]//International Conference on Learning Representations, 2024.
- [15] Bergstra J, Bardenet R, Bengio Y, et al. Algorithms for Hyper-Parameter Optimization[C]//Advances in Neural Information Processing Systems: 卷 24. 2546-2554, 2011.
- [16] Bergstra J, Bengio Y. Random Search for Hyper-Parameter Optimization[J]. Journal of Machine Learning Research, 2012, 13(10): 281-305.
- [17] Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms[C]//Advances in Neural Information Processing Systems: 卷 25, 2012.
- [18] Shahriari B, Swersky K, Wang Z, et al. Taking the Human Out of the Loop: A Review of Bayesian Optimization[J]. Proceedings of the IEEE, 2016, 104(1): 148-175.
- [19] 周秋阳, 高辉, 李炜卓, 等. 基于 GAN 数据增强与改进 Bi-LSTM 的充电桩故障预测方法[J]. 电力工程技术, 2025, 44(6): 49-61.
Zhou QY, Gao H, Li WZ, et al. A fault prediction method

- for charging pile based on GAN data enhancement and improved Bi-LSTM[J]. Electric Power Engineering Technology, 2025, 44(6): 49-61.
- [19] Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of ICNN'95 - International Conference on Neural Networks: 卷 4. 1942-1948, 1995.
- [20] Storn R, Price K. Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces[J]. Journal of Global Optimization, 1997, 11(4): 341-359.
- [21] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.
- [22] Mirjalili S, Lewis A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95: 51-67.
- [23] Hosney ME, Emam MM, Saad MR, et al. Heart rate optimizer: a novel bio-inspired metaheuristic algorithm[J]. Scientific Reports, 2026, 16(1): 15985.
- [24] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.
- [25] Karniadakis GE, Kevrekidis IG, Lu L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.
- [26] 赵征, 刘思远. 基于等效动态热值与 SSA-NM 辨识的垃圾焚烧炉温灰盒建模[J/OL]. 热力发电, 2026. DOI: 10.19666/j.rlfd.202605094
- ZHAO Z, LIU S Y. Grey-box modeling of furnace temperature in municipal solid waste incinerator based on equivalent dynamic heating value and SSA-NM identification[J/OL]. Thermal Power Generation, 2026. DOI: 10.19666/j.rlfd.202605094.
- [27] Willard J, Jia X, Xu S, et al. Integrating scientific knowledge with machine learning for engineering and environmental systems[J]. ACM Computing Surveys, 2022, 55(4): 1-37.
- [28] Zhan Y, Wang X, Xu Y, et al. A Hybrid TCN-LSTM-Attention Framework for Multi-Scenario Short-Term Photovoltaic Power Forecasting Incorporating Physics-Informed Neural Network Strategy[J]. Energy, 2026, 344: 139968.
- [29] 谢锐彪, 李新利, 王英男, 等. 基于粒子群优化的双向门控循环神经网络燃煤电厂 NO_x 排放预测[J]. 热力发电, 2021, 50(10): 87-94.
- XIE R B, LI X L, WANG Y N, et al. NO_x emission prediction for coal-fired power plants based on bidirectional gated recurrent neural network optimized by particle swarm optimization[J]. Thermal Power Generation, 2021, 50(10): 87-94.
- [30] Friedman JH. Greedy function approximation: A gradient boosting machine[J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [31] Goldstein A, Kapelner A, Bleich J, et al. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation[J]. Journal of Computational and Graphical Statistics, 2015, 24(1): 44-65.
- [32] 陈凡, 吴凌霄, 王曼, 等. 基于 PPO 算法的 CIES 低碳优化调度方法[J]. 电力工程技术, 2024, 43(6): 88-99. DOI: 10.12158/j.2096-3203.2024.06.009.
- CHEN F, WU L X, WANG M, et al. A low-carbon optimal scheduling method for CIES based on PPO algorithm[J]. Electric Power Engineering Technology, 2024, 43(6): 88-99. DOI: 10.12158/j.2096-3203.2024.06.009.
- [33] Liu Y, Xie M. Rebooting data-driven soft-sensors in process industries: A review of kernel methods[J]. Journal of Process Control, 2020, 89: 58-73.
- [34] Chen J, Zheng L, Che W, et al. A method for measuring carbon emissions from power plants using a CNN-LSTM-Attention model with Bayesian optimization[J]. Case Studies in Thermal Engineering, 2024, 63: 105334.
- [35] Pan H, Su T, Huang X, et al. LSTM-based soft sensor design for oxygen content of flue gas in coal-fired power plant[J]. Transactions of the Institute of Measurement and Control, 2021, 43(1): 78-87.