自动化学报
ACTA AUTOMATICA SINICA

虚实数据混合驱动的MSWI过程CO_2排放浓度预测

汤健 王子 余文 乔俊飞

Prediction of CO_2 Emission Concentration of MSWI Process Driven by Mixed Virtual-real Data

TANG Jian, WANG Zi, YU Wen, QIAO Jun-Fei

在线阅读 View online:
http://www.aas.net.cn/article/current/df7e075e35cdc80e6ec00c022b575db049ab4ede43bb9986770dea36ed27b923

您可能感兴趣的其他文章

基于宽度混合森林回归的城市固废焚烧过程二英排放软测量

Soft Sensing Method of Dioxin Emission in Municipal Solid Waste Incineration Process Based on Broad Hybrid Forest Regression

自动化学报. 2023, 49(2): 343−365 https://doi.org/10.16383/j.aas.c220012

基于仿真机理和改进回归决策树的二英排放建模

Dioxin Emission Concentration Modeling Based on Simulation Mechanism and Improved Linear Regression Decision Tree

自动化学报. 2024, 50(8): 1601−1619 https://doi.org/10.16383/j.aas.c230625

城市固废焚烧过程数据驱动建模与自组织控制

Data-driven Modeling and Self-organizing Control of Municipal Solid Waste Incineration Process

自动化学报. 2023, 49(3): 550−566 https://doi.org/10.16383/j.aas.c220570

城市固废焚烧过程神经网络控制研究综述

A Review of Neural Network Control in Municipal Solid Waste Incineration Processes

自动化学报. 2025, 51(9): 1951−1973 https://doi.org/10.16383/j.aas.c240604

基于潜在特征选择性集成建模的二英排放浓度软测量

Dioxin Emission Concentration Soft Measurement Based on Multi-source Latent Feature Selective Ensemble Modeling for Municipal Solid Waste Incineration Process

自动化学报. 2022, 48(1): 223−238 https://doi.org/10.16383/j.aas.c190254

基于IT2FBLS强化学习PID的MSWI过程炉膛温度控制

Furnace Temperature Control Using IT2FBLS-based Reinforcement Learning PID for MSWI Process

自动化学报. 2025, 51(7): 1626−1641 https://doi.org/10.16383/j.aas.c240074

虚实数据混合驱动的 MSWI 过程 CO_2 排放浓度预测

汤 健[1,2] 王 子[1] 余 文[3] 乔俊飞[1]

摘 要 CO_2 排放浓度动态预测是助力实现双碳战略的关键环节之一, 但城市固废焚烧 (MSWI) 过程受限于燃料与工况波动及数据噪声等多种因素, 传统数据驱动或机理驱动难以实现高精度建模. 针对该问题, 提出虚实数据混合驱动的 CO_2 排放浓度预测方法. 首先, 依据 CO_2 的产生机理特性通过耦合多软件构建数值仿真模型并结合双重正交实验获取虚拟机理数据; 然后, 在采用虚拟机理数据构建基于改进决策树的可解释主模型的基础上, 以真实过程数据为输入得到主模型残差数据后再采用区间三型模糊宽度学习系统构建残差补偿模型, 得到虚实驱动主补机制模型; 接着, 为捕获 CO_2 排放的动态特性, 基于长短期神经网络构建真实数据驱动时序特性模型; 最后, 通过改进岭回归实现上述虚实模型输出的动态融合, 得到 CO_2 排放浓度的预测结果. 基于北京某 MSWI 电厂实际数据验证了模型的有效性和合理性.

关键词 城市固废焚烧; CO_2 排放; 虚拟机理数据; 双重正交实验; 虚实驱动主补建模; 动态融合

引用格式 汤健, 王子, 余文, 乔俊飞. 虚实数据混合驱动的 MSWI 过程 CO_2 排放浓度预测. 自动化学报, xxxx, **xx**(x): x−xx

DOI 10.16383/j.aas.c260092 **CSTR** 32138.14.j.aas.c260092

Prediction of CO_2 Emission Concentration of MSWI Process Driven by Mixed Virtual-real Data

TANG Jian[1,2] WANG Zi[1] YU Wen[3] QIAO Jun-Fei[1]

Abstract The dynamic prediction of CO_2 emission concentration is one of the key links to help realize the dual carbon strategy. However, the municipal solid waste incineration (MSWI) process is limited by various factors such as fuel and working condition fluctuations and data noise, and it is difficult to achieve high-precision modeling by traditional data-driven or mechanism-driven. To solve this problem, this paper proposes a prediction method of CO_2 emission concentration driven by mixed virtual-real data. Firstly, according to the characteristics of CO_2 production mechanism, a numerical simulation model was constructed by coupling multi-software and the virtual mechanism data were obtained by double orthogonal experiment; Then, on the basis of the interpretable improved decision tree-based main model constructed by using the virtual mechanism data, the residual data of the main model was obtained by using real process data as input, and the residual compensation model was constructed by using interval type-3 fuzzy broad learning system to obtain the virtual-real driven main-complement mechanism model; Next, in order to capture the dynamic characteristics of CO_2 emission, a real data driven time series characteristic model was constructed based on long short-term neural network; Finally, the dynamic fusion of the above virtual-real model outputs was realized by improved ridge regression, and the prediction results of CO_2 emission concentration were obtained. Based on the actual data of a MSWI power plant in Beijing, the validity and rationality of the model are verified.

Keywords municipal solid waste incineration; CO_2 emission; virtual mechanism data; double orthogonal experiment; virtual-real driven main-complement modeling; dynamic fusion

Citation Tang Jian, Wang Zi, Yu Wen, Qiao Jun-Fei. Prediction of CO_2 emission concentration of MSWI process driven by mixed virtual-real data. *Acta Automatica Sinica*, xxxx, **xx**(x): x−xx

城市固废焚烧 (municipal solid waste incineration, MSWI) 因具有无害化、减量化和资源化等优势[1], 已成为世界范围内广泛采用的可实现废物到能源转变的核心产业[2]. 由于城市固废 (municipal solid waste, MSW) 分类等诸多不确定因素的影响以及居民生活习惯差异显著, 导致入炉燃料组分存在着较大波动性[3]. 这种组分波动进一步引发焚烧

收稿日期 2026-01-31 录用日期 2026-06-04
Manuscript received January 31, 2026; accepted June 4, 2026
国家自然科学基金 (62573011, 62373017) 资助
Supported by National Natural Science Foundation of China (62573011, 62373017)
本文责任编委 王卓
Recommended by Associate Editor WANG Zhuo
1. 北京工业大学信息科学技术学院 北京 100124 中国 2. 烟台科技学院人工智能学院 烟台 265600 中国 3. 墨西哥国立理工大学高级研究中心 (CINVESTAV-IPN) 墨西哥 07360 墨西哥
1. School of Information Science and Technology, Beijing University of Technology, Beijing 100124, China 2. College of Artificial Intelligence, Yantai Institute of Science and Technology, Yantai 265600, China 3. Centro de Investigation de Estudios Avanzados (CINVESTAV-IPN), National Polytechnic Institute México, Mexico D.F. 07360, Mexico

工况频繁变化, 使得 MSWI 过程稳定控制变得尤为困难[4], 进而影响烟气排放特性. 二氧化碳 (carbon dioxide, CO_2) 作为 MSWI 过程的副产物, 是引起全球变暖或温室效应的关键气体之一[5]. 近年来, 我国已成为全球最大的 CO_2 排放国, 其排放量约占全球总量的 30%[6]. 为应对这一挑战, 我国政府创新性地提出“双碳”战略目标, 即力争 2030 年前实现碳达峰和 2060 年前达成碳中和[7]. 在此背景下, 开展 MSWI 过程 CO_2 排放浓度的精准检测与预测具有双重意义: 一方面能通过历史数据为评估 MSWI 环境绩效提供量化依据; 另一方面能通过实时数据反馈和预测数据反馈优化燃烧控制, 最终实现降碳与能源回收的协同增效.

对 CO_2 排放浓度进行预测的本质是一类时间序列建模问题. 部分学者利用数据驱动 (data-driven, DD) 模型快速建立输入特征与输出目标之间的映射关系. 例如, Taheri 等[8] 比较六种先进的学习算法用于预测室内的 CO_2 浓度, 包括支持向量机 (support vector machine, SVM)、自适应提升 (adaptive boosting, AdaBoost)、随机森林 (random forest, RF)、梯度增强 (gradient boosting, GB)、逻辑回归 (logistic regression, LR) 以及多层感知器 (multilayer perceptron, MLP); 结果表明, MLP 效果最优. Pan 等[9] 为优化 CO_2 捕集的动态运行, 开发基于 Transformer 架构的数据驱动模型预测烟气流量和 CO_2 浓度. 为克服单一模型的局限并进一步提升预测性能, 当前研究集中于构建混合或集成模型. 例如, Kumar[10] 提出混合机器学习模型用于预测 CO_2 排放; Zhang 等[11] 为捕获时序特征, 引入结合 RF、极限梯度提升 (extreme gradient boosting, XGBoost) 和长短期记忆 (long short-term memory, LSTM) 网络的混合机器学习框架. 综上所述, CO_2 排放浓度研究在交通、大气等领域已有涉及. 然而, 目前国内外研究多集中于 MSWI 过程中二噁英、NO_x、SO_2 等污染物的建模与控制, 尚未见针对 MSWI 炉内或烟气中 CO_2 排放浓度的精准预测模型.

近年来, 机器学习作为一种强大的数据驱动技术, 得到飞速发展[12]. 数据驱动模型主要依赖于对历史数据的分析, 挖掘输入与输出的复杂关系[13]. 目前已有诸多学者在不同领域, 通过构建 DD 模型对 CO_2 的排放浓度进行预测建模. 例如, Zhang 等[11] 分别构建基于 RF、决策树、XGBoost 和 LSTM 等算法的预测模型; Zhou 等[14] 应用 Catboost 算法针对轻型汽车的 CO_2 排放进行预测建模, 结果表明预测精度超越了支持向量回归和岭回归等其他模型; Wen 等[15] 针对碳排放数据固有的线性和非线性关系, 开发综合自回归积分滑动平均 (autoregressive integrated moving average, ARIMA) 与 LSTM 的混合模型. 尽管上述 DD 模型能够从数据中学习规律, 但其性能高度依赖历史数据的质量与规模.

在实际工业环境中, 由于传感器噪声、数据缺失以及工况多变等因素, 纯 DD 模型在未知工况下的泛化能力往往不足, 且模型可解释性较差[16]. 与之相对, 机理驱动 (mechanism-driven, MD) 模型, 也被称为白盒模型, 基于过程内部的物理化学原理, 通过质量、动量、能量守恒及反应动力学方程等描述系统行为[17−19]. 对于工业领域, Dang 等[19] 采用非等温热重分析法, 揭示了 CO_2 气化反应中催化作用的机理; Shafeeyan 等[20] 为优化设计, 建立了固定床柱数学 MD 模型来预测 CO_2 吸附分离性能, 其全面考虑了物质、能量和动量平衡以及传质速率和平衡方程; Wang 等[21] 开发瞬态反应扩散模型以用于研究不同强度焦炭在 CO_2 或 H_2O 反应环境中的气化机理, 以此降低炼铁过程中的碳排放. 然而, 尽管 MD 模型具有明确的物理基础和强解释性, 在 MSWI 这类复杂反应系统中, 其精确构建仍面临炉内反应复杂、耦合显著、精确建模困难、计算成本高昂等诸多挑战; 此外, MD 模型通常基于理想化的假设, 与实际工业工况存在着一定偏差, 这限制了其落地应用的预测精度. 综上所述, 单一依赖 MD 或 DD 的模型, 均难以在保证物理合理性的同时, 实现对波动工况下 CO_2 排放浓度的高精度、强鲁棒预测. 因此, 探索一种有效融合机理知识与现场数据的混合建模方法, 已成为提升 MSWI 过程碳排放预测能力的研究热点.

目前, 研究人员已提出多种实现 DD 和 MD 模型的融合策略. Cen 等[22] 通过 MD 模型生成重构特征, 并将重构特征与原始特征同时输入基于随机森林、支持向量回归等四种机器学习算法构建的模型中, 以获得混合驱动模型. 类似地, Zhang 等[23] 先构建 MD 模型, 再将历史真实数据转换为含有机理知识的重构数据, 并与真实历史数据加权混合. 但是上述方法的 MD 模型基于原始特征通过物理化学方程计算得出, 重构特征与原始特征之间存在着高度的共线性和信息重叠会导致特征的冗余. 为避免信息重叠, 部分学者将 MD 与 DD 模型进行严格划分. 例如, Shen 等[24] 用 MD 模型生成虚拟数据后, 应用该数据对 DD 模型进行预训练, 再输入真实数据; Liang 等[25] 利用 MD 模型对真实数据进行检查, 选择符合机理的数据进行后续的 DD 建模. 上述方法虽然均避免了机理数据与真实数据的重叠问题, 但是 MD 模型本身是现实世界的简化, 必然存在建模误差. 因此, 上述方法很大程度上受限于 MD 模

型的准确度. 如果 MD 模型过于粗糙, 将导致生成的虚拟数据或者筛选后的数据与真实数据差距过大, 可能会错误地训练 DD 模型. 为避免上述情况发生, 部分学者考虑采用残差学习方式. 例如, Wang 等[26] 将 MD 模型的普适性与 DD 模型的简单性和实用性相结合, 提出通过神经网络补偿 MD 模型的预测残差的混合驱动方法; Yan 等[27] 首先构建 MD 模型, 再构建 DD 模型以用于 MD 模型的输出补偿, 最后将 MD 和 DD 模型相加, 得到最终结果; Zhu 等[28] 同样先构建水文 MD 模型, 再利用 LSTM 算法构建 DD 模型以修正 MD 模型的残差.

上述混合驱动模型局限于机理与数据的二元融合问题, 即仅将 MD 模型的输出与一个 DD 模型的输出进行结合. 这种方式虽然可以从真实数据中挖掘出机理特性, 但是未能捕获真实工业过程中固有的时序动态特性. 为突破这种局限, 本文进一步引入 LSTM, 构建纯 DD 的时序预测模型以捕获 CO_2 排放浓度在真实工况下的时序依赖与动态波动特性. 该模型直接以历史 CO_2 排放浓度数据为输入, 通过其门控机制与状态记忆结构, 自适应学习不同时刻下 CO_2 浓度之间的时序映射关系, 从而有效捕捉因燃烧波动等实际因素引起的动态演变规律. 此外, 鉴于前文的融合方法采用的简单线性加权或静态集成策略难以适应 MSWI 时变特性的问题, 本文进一步提出基于自适应加权的岭回归方法以构建动态融合机制, 通过自适应地调整各个模块输出的权重贡献得到 CO_2 排放浓度预测值.

综上, 本文创新性地提出基于虚实数据混合驱动的 MSWI 过程 CO_2 排放浓度预测方法. 该方法面向 MSWI 过程, 提出由虚拟机理数据生成、虚实驱动主补机制建模、真实数据驱动时序特性建模和虚实模型输出动态融合模块组成的 CO_2 排放浓度预测框架, 弥补了该领域现有预测模型的短板. 为便于查阅, 本文的符号说明详见附录 A; 缩写词及其含义详见附录 B.

1　面向 CO_2 浓度的 MSWI 过程描述

本研究以北京某日处理量为 800 t 的炉排炉型 MSWI 工艺为例, 从 CO_2 生成机理视角, 划分的工艺流程如图 1 所示.

由图 1 可知, 将 MSWI 过程分为五个区域: 固相燃烧区、气相燃烧区、换热区、烟气净化区和烟气排放区. 其中与 CO_2 排放直接相关的环节如下: 固相燃烧区中垃圾经干燥、燃烧后生成 CO_2、C_mH_n 以及 CO 等中间气体; 气相燃烧区中 C_mH_n、CO 进一步氧化为 CO_2; 换热区仅涉及热量回收, 不改变烟气中的 CO_2 含量; 烟气净化区主要通过氢氧化钙来吸收 CO_2; 烟气排放区对 CO_2 无去除作用.

我国作为推动全球环保事业的重要力量, 于 2020 年 9 月制定了“CO_2 排放力争 2030 年前达到峰值, 努力争取 2060 年前实现碳中和”的双碳战略[29]. 目前, 针对特定真实 MSWI 电厂进行碳排放的预测研究较少[30], 该研究对实施“减污降碳”颇为关键.

此外, 燃烧效率 (combustion efficiency, CE) 反映了燃烧过程的充分性, 其是指定量燃料在燃烧室内燃烧时实际可用于加热燃烧产物的热量与该燃料在绝热条件下实现完全燃烧时所释出的低位发热量之比[1], 其作为能够定量评价燃烧状态的指标[31], 计算公式如下:

$$\xi_{\mathrm{CE}}=\frac{\gamma_{\mathrm{CO_2}}}{(\gamma_{\mathrm{CO}}+\gamma_{\mathrm{CO_2}})}\times 100\% \tag{1}$$

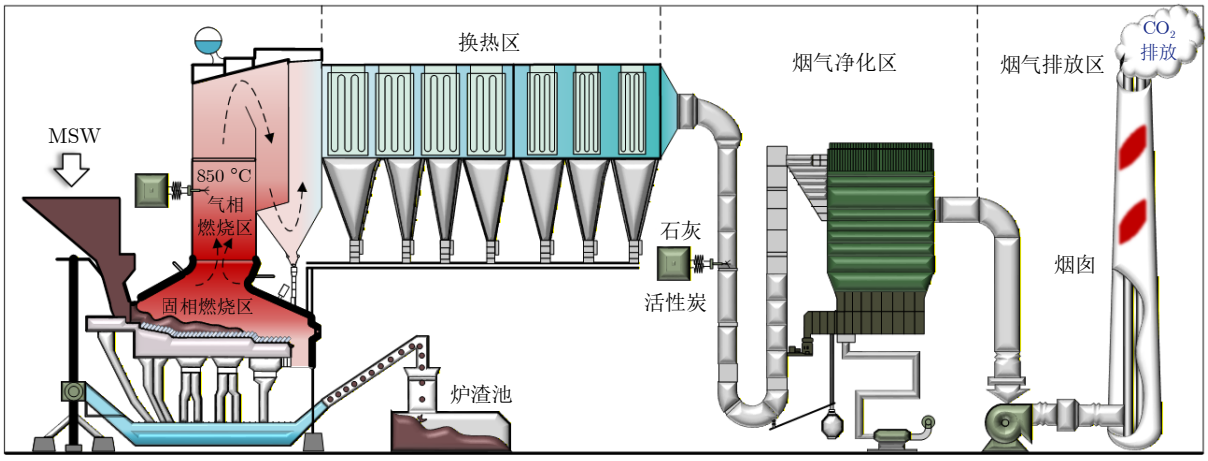


图 1　面向 CO_2 浓度的 MSWI 工艺流程

Fig. 1　MSWI process flow for CO_2 concentration

其中, ξ_{CE} 表示燃烧效率; γ_{CO} 表示 CO 的浓度; γ_{CO_2} 表示 CO_2 的浓度.

由上式可知, 全面和准确地预测 CO_2 的排放可以提前预测 CE 的变化, 从而有助于降低 MSWI 的运行成本, 这也是制定适当地减缓排放策略的先决条件.

2 建模策略

本文提出的建模策略如图 2 所示, 图 2 中四个模块的功能为:

1) 虚拟机理数据生成模块. 首先通过耦合 FLIC 和 Aspen Plus 仿真软件构建表述 CO_2 全生命周期

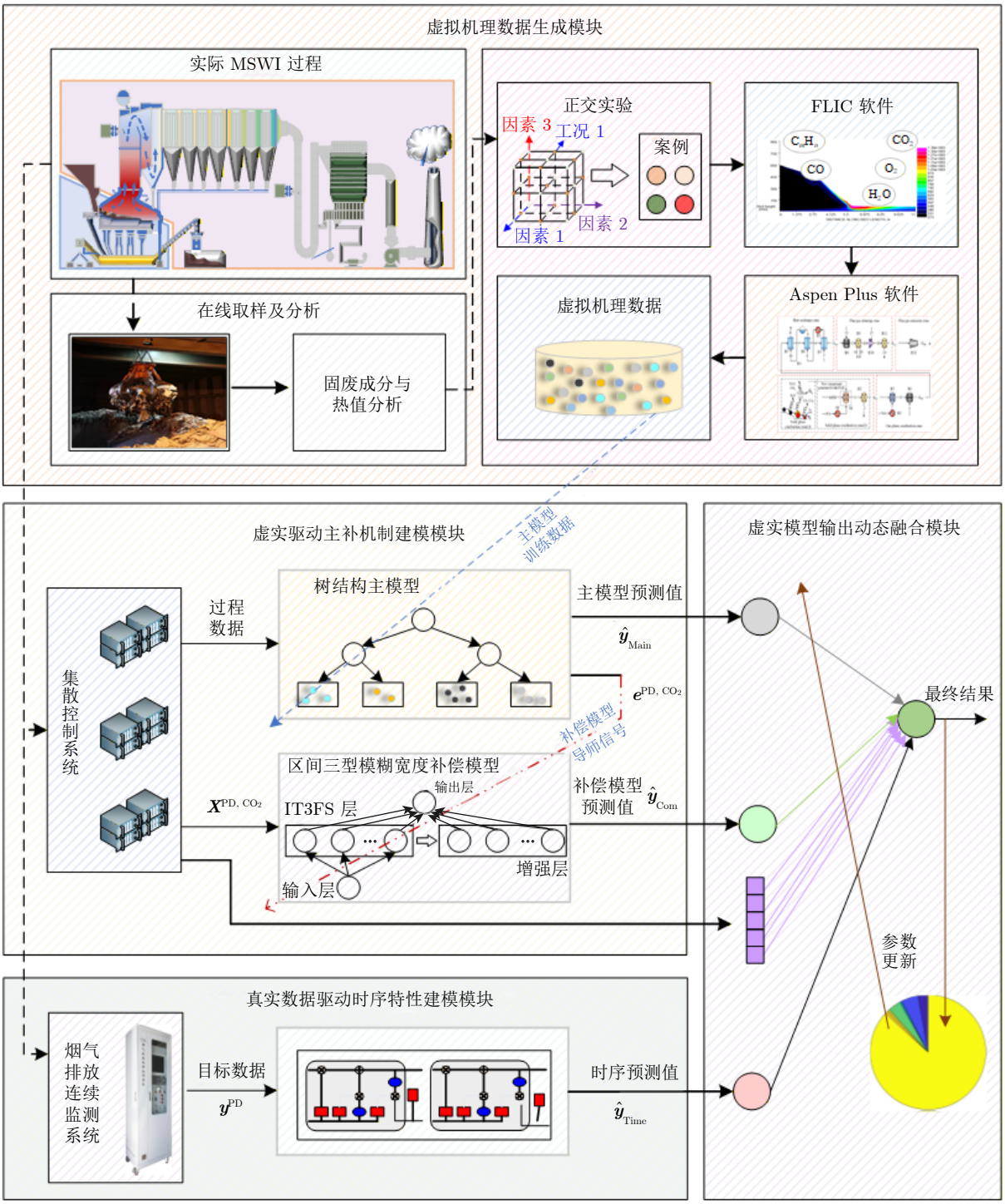


图 2 建模策略

Fig. 2 Modeling strategy

机理的基准机理数值仿真模型, 接着采用双重正交实验设计获取多种运行工况下的多组操纵变量和非操纵变量值, 最后将其输入上述基准工况模型获取多工况机理数据.

2) 虚实驱动主补机制建模模块. 包含虚拟 DD 可解释的树结构主模型子模块和真实 DD 区间三型模糊宽度误差补偿模型子模块两个部分, 其中: a) 基于改进决策树结构的主模型: 首先, 输入虚拟机理数据构建主模型, 进而稳健地捕捉输入与输出之间的复杂模式; 接着, 输入真实过程数据得到主模型的预测值. 该模型基于树结构分支规则与特征重要性排序实现内在可解释性, 能够清晰揭示各工艺变量对 CO_2 排放浓度的贡献程度与影响逻辑, 从根本上避免信息重叠与特征冗余问题. b) 基于区间三型模糊宽度误差补偿模型: 通过构建 IT3FBLS 学习主模型的预测残差, 进而修正主模型无法捕捉的复杂性、不确定性和未知扰动.

3) 真实数据驱动时序特性建模模块. 由于前文构建的主模型、补偿模型已充分挖掘了各工艺特征变量对 CO_2 排放浓度真值的影响规律. 为避免特征冗余并聚焦时序特征挖掘, 在 LSTM 模型建模过程中, 本文利用从烟气排放连续监测系统 (continuous emission monitoring system, CEMS) 得到的目标数据 (CO_2 浓度值) 构建 LSTM 时序模型, 用于学习真实 CO_2 浓度时间序列本身的内在演化规律, 进而弥补主补模型的不足.

4) 虚实模型输出动态融合模块. 采用改进岭回归对虚实驱动主补机制建模和真实数据驱动时序特性建模模块结果以及真实过程数据进行有机融合, 实现优势互补.

3　策略实现

3.1　虚拟机理数据生成模块

3.1.1　MSWI 过程 CO_2 的全生命周期机理数值仿真模型构建子模块

3.1.1.1　基于 FLIC 的固相燃烧模型

FLIC 软件用于模拟 MSW 在炉排上的燃烧过程, 其可将实际三维建模问题转化为二维问题, 便于求解计算. MSW 在炉排上的固相燃烧要满足质量、动量和能量守恒方程. 此外, 固相燃烧子过程主要分为如图 3 所示四个过程: 水分蒸发、挥发分释放、挥发分燃烧以及焦炭燃烧.

上述理论已有诸多文献进行了研究, 在此不再详细介绍[32]. 最终固相燃烧模型的产物为 N_2、O_2、CO_2、H_2O、C_mH_n、CO、H_2 和 NO. 为了便于分析, 本文以 CH_4 代替 C_mH_n 进行模拟.

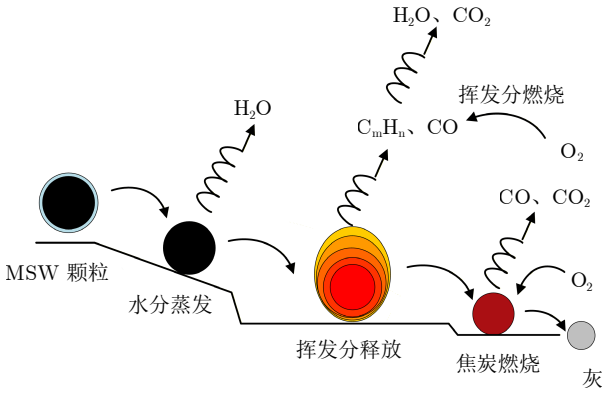


图 3　固相燃烧子过程示意图

Fig. 3　Diagram of solid phase combustion subprocess

3.1.1.2　基于 Aspen Plus 的气相燃烧模型

气相燃烧模型包括固相燃烧区、气相燃烧区、换热区、烟气净化区以及烟气排放区, 如图 4 所示. 其与 FLIC 的耦合表现在: 从 FLIC 获得固相燃烧的烟气组分, 将其输入 B1.

在图 4 中, MSW 流股表示 MSW 进料; PA 和 SA 分别表示一次风和二次风; G_{SC}、G_{GC}、G_{HT}、G_{CL} 和 G_{GE} 分别表示固相燃烧区、气相燃烧区、换热区、烟气净化区和烟气排放区的出口烟气. 各个区域的主要过程如下:

1) 固相燃烧区

一次风 PA 经加热器 (H1, Heater) 预热后, 与 MSW 流股一同进入 B1, 采用 RYield 反应器模块模拟 MSW 在炉排上的固相燃烧. 其中, 相关气体组分由固相燃烧区 a) 仿真获得. 同时, 产物经过分离器 (B2, Sep) 过滤掉灰分、残渣等固体 (流股 2), 得到固相燃烧区 b) 的出口烟气 G_{SC}.

2) 气相燃烧区

流股 G_{SC} 和经 H2 预热后的二次风流股 SA 进入 B3, 采用 RGibbs 反应器模拟 G_{SC} 在炉膛内的气相燃烧.

3) 换热区

流股 G_{GC} 进入换热区, 先后经过过热器 (B5)、蒸发器 (B6) 和省煤器 (B7) 实现烟气降温, 得到换热区的出口烟气 G_{HT}. 其中, 水 (流股 7) 通过 B7 和 H3 进行预热, 然后进入 B6 蒸发, 得到的产物再进入汽包 (Drum) 实现汽水分离, 最后再通过 B5 进一步加热. 在该阶段, 主要是利用燃烧产生的高温烟气通过热交换器将热量传递给其他介质, 从而实现能量的逐级转化, 因此该阶段不涉及 CO_2 的相关反应.

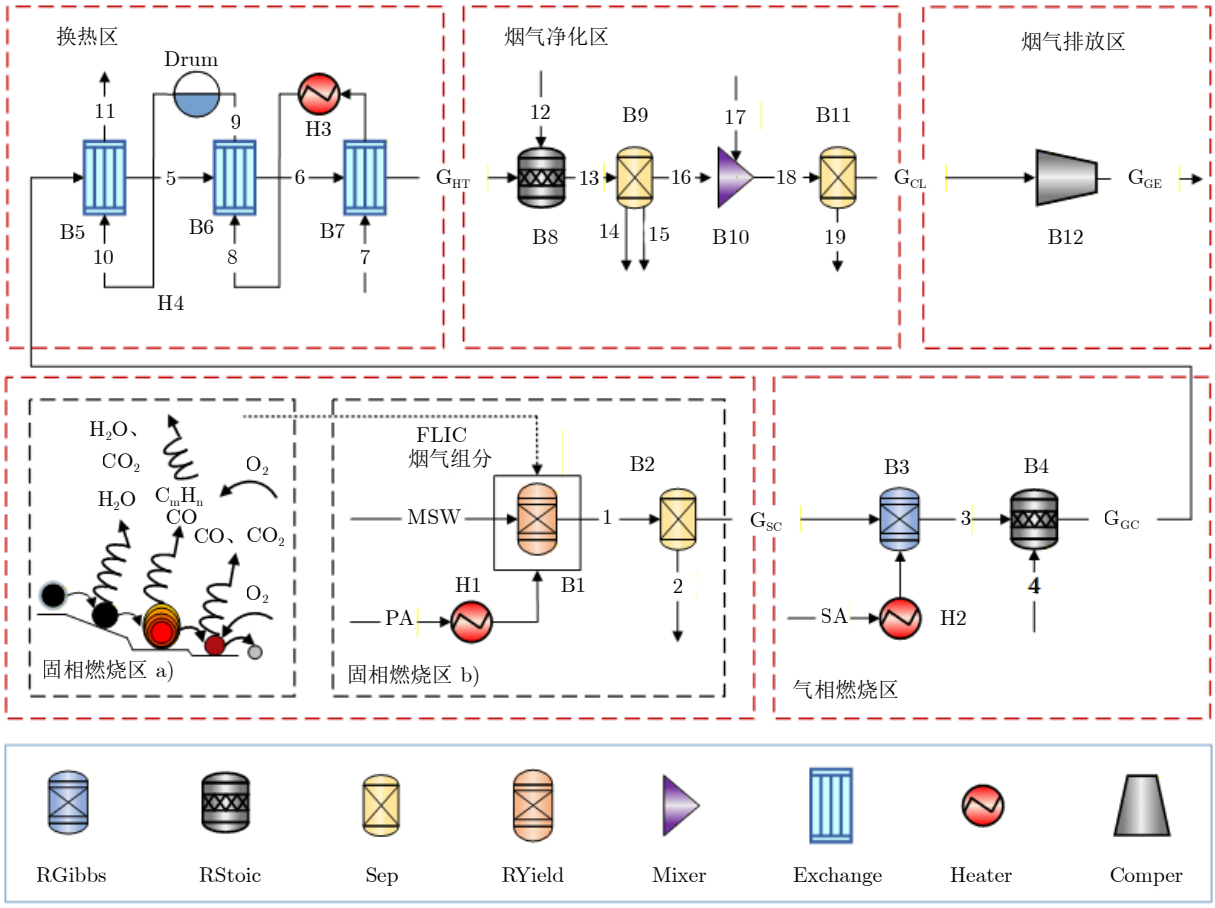


图 4 基于 Aspen Plus 的气相燃烧模型

Fig. 4 Gas phase combustion model based on Aspen Plus

4) 烟气净化区

流股 G_{HT} 进入脱酸反应器 (B8, RStoic) 中, 同时加入消石灰 (流股 12) 实现酸性气体的消除[33].

5) 烟气排放区

净化后的烟气经烟囱 (B12, Comper) 排放到大气中.

3.1.2 基于双重正交实验的多工况虚拟机理数据生成子模块

本文采用双重正交实验设计获取多种运行工况下的多组操纵变量和非操纵变量值, 其过程为: 1) 首先确定实验因子 n 及其水平数 m, 此时可得到正交实验表 $L_B(m^n)$, 进一步可得到 B 个实验案例; 2) 接着依据第 1 次正交实验得到的 B 个案例, 重新确定实验因子 n_1 及其水平数 m_1, 得到 L 个正交实验表 $L_L(m_1^{n_1})$; 3) 最终可以得到 $B \times L$ 个实验案例, 即工况次数为 $B \times L$.

接着, 将上述得到的 $B \times L$ 个实验案例依次输入多软件耦合的 CO_2 全生命周期的机理数值仿真模型 $f_{\text{model}}(\cdot)$ 中, 得到机理模型计算的 CO_2 浓度.

至此, 已获得 $B \times L$ 种工况下的包括 CO_2 全生命周期虚拟机理数据, 将进行处理后的数据记为 $\boldsymbol{D}^{\text{MD, CO}_2}$.

3.2 虚实驱动主补机制建模模块

为了避免由于 MD 模型过于粗糙, 导致生成的虚拟机理数据与真实过程数据差距过大, 本文提出了虚实驱动主补机制建模方法. 通过采用树结构作为机理驱动模型, 从根本上避免了信息重叠以及特征冗余问题. 在补偿模型上, 引入先进的区间三型模糊宽度学习系统, 专门处理 MD 模型残差中的高不确定性与非线性.

3.2.1 虚拟数据驱动可解释主模型子模块

首先, 定义 CO_2 仿真机理数据集为 $\boldsymbol{D}^{\text{MD, CO}_2} = \{\boldsymbol{X}^{\text{MD}}, \boldsymbol{y}^{\text{MD}}\} = \{\boldsymbol{x}_j, y_j\}_{j=1}^{J} \in \mathbf{R}^{J\times(M+1)}$. 其中, $j = 1, \cdots, J$, J 表示 CO_2 虚拟机理数据样本数; $m = 1, \cdots, M$ 表示特征数索引, M 表示 CO_2 虚拟机理数据特征数; $\boldsymbol{x}_j = [x_{j,1}, \cdots, x_{j,m}, \cdots, x_{j,M}]$, $x_{j,m}$ 表示第 j 个样本中的第 m 个特征, y_j 表示第 j 个样本中的目标值; $\boldsymbol{X}^{\text{MD}}$ 表示仿真机理数据集的特征;

$\boldsymbol{y}^{\mathrm{MD}}$ 表示仿真机理数据集的目标值.

然后, 循环遍历数据集 $\boldsymbol{D}^{\mathrm{MD,\,CO_2}}$ 中的每个特征值和可能的分割点 s, 将数据集划分为左、右两个子空间, 表示如下:

$$\begin{cases} R_k^{\mathrm{Left}}(s)=\left\{\boldsymbol{X}_k^{\mathrm{Left}},\ \boldsymbol{y}_k^{\mathrm{Left}}\right\}=\left\{\boldsymbol{x}_{k,\,j}^{\mathrm{Left}},\ y_{k,\,j}^{\mathrm{Left}}\right\}_{j=1}^{N_k^{\mathrm{Left}}}, \\ \qquad\qquad\qquad\qquad x_{k,\,j,\,m}^{\mathrm{Left}}\le s \\ R_k^{\mathrm{Right}}(s)=\left\{\boldsymbol{X}_k^{\mathrm{Right}},\ \boldsymbol{y}_k^{\mathrm{Right}}\right\}=\left\{\boldsymbol{x}_{k,\,j}^{\mathrm{Right}},\ y_{k,\,j}^{\mathrm{Right}}\right\}_{j=1}^{N_k^{\mathrm{Right}}}, \\ \qquad\qquad\qquad\qquad x_{k,\,j,\,m}^{\mathrm{Right}}> s \end{cases} \tag{2}$$

其中, $R_k^{\mathrm{Left}}(s)$ 和 $R_k^{\mathrm{Right}}(s)$ 分别表示在第 k 个节点上, 使用分割点 s 形成的左子树和右子树; $\boldsymbol{X}_k^{\mathrm{Left}}$ 和 $\boldsymbol{X}_k^{\mathrm{Right}}$ 分别表示在第 k 个节点上左、右子树的特征; $\boldsymbol{y}_k^{\mathrm{Left}}$ 和 $\boldsymbol{y}_k^{\mathrm{Right}}$ 分别表示在第 k 个节点上左、右子树的目标值; $\boldsymbol{x}_{k,\,j}^{\mathrm{Left}}$ 和 $\boldsymbol{x}_{k,\,j}^{\mathrm{Right}}$ 分别表示在第 k 个节点上左、右子树的第 j 个特征; N_k^{Left} 和 N_k^{Right} 分别表示在第 k 个节点上划分得到的左、右子空间样本数; $x_{k,\,j,\,m}^{\mathrm{Left}}$ 和 $x_{k,\,j,\,m}^{\mathrm{Right}}$ 分别表示在第 k 个节点上左、右子树中第 j 个样本中的第 m 个特征; $y_{k,\,j}^{\mathrm{Left}}$ 和 $y_{k,\,j}^{\mathrm{Right}}$ 分别表示在第 k 个节点上左、右子树中第 j 个样本的目标值.

然后, 计算两个子空间的均方误差 (mean square error, MSE), 以最小化 MSE 确定分割点. 其计算过程如下:

$$\Omega(s)=\frac{1}{N_k^{\mathrm{Left}}}\sum_{j=1}^{N_k^{\mathrm{Left}}}\left(y_{k,\,j}^{\mathrm{Left}}-\mu_k^{\mathrm{Left}}\right)^2+\frac{1}{N_k^{\mathrm{Right}}}\sum_{j=1}^{N_k^{\mathrm{Right}}}\left(y_{k,\,j}^{\mathrm{Right}}-\mu_k^{\mathrm{Right}}\right)^2 \tag{3}$$

其中, $\Omega(s)$ 表示使用分割点 s 划分后的损失函数值; μ_k^{Left} 和 μ_k^{Right} 分别表示在第 k 个节点上左、右两个子空间的目标值均值, 公式如下:

$$\mu_k^{\mathrm{Left}}=\frac{1}{N_k^{\mathrm{Left}}}\sum_{j=1}^{N_k^{\mathrm{Left}}}y_{k,\,j}^{\mathrm{Left}} \tag{4}$$

$$\mu_k^{\mathrm{Right}}=\frac{1}{N_k^{\mathrm{Right}}}\sum_{j=1}^{N_k^{\mathrm{Right}}}y_{k,\,j}^{\mathrm{Right}} \tag{5}$$

接着, 选择使得总均方误差最小的分割点来结束本次空间划分:

$$s^*=\arg\min_s\Omega(s) \tag{6}$$

重复上述过程, 直到叶节点的最小样本数达到设定的阈值, 得到了 K 个不相交的子空间 R_1, R_2, $\cdots$, R_K:

$$R=\bigcup_{k=1}^{K}R_k \tag{7}$$

传统的决策树算法对于每个子空间 R_k 仅采用该子空间内部样本标签的均值作为输出, 忽略了相似样本特征中存在的非线性关系. 为了克服这一问题, 本文采用基于拟牛顿优化的策略来改进. 以第 k 个子空间为例, 定义目标函数为:

$$\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right)=\frac{1}{N_k}\sum_{j=1}^{N_k}\left(y_{k,\,j}-\left(\boldsymbol{w}_k^{(t)}\right)^{\mathrm{T}}\boldsymbol{x}_{k,\,j}\right)^2+\lambda\left\|\boldsymbol{w}_k^{(t)}\right\|^2 \tag{8}$$

其中, N_k 表示第 k 个子空间内的总样本数; $y_{k,\,j}$ 表示第 k 个子空间的第 j 个样本的目标值; $\boldsymbol{w}_k^{(t)}$ 为第 t 次迭代下第 k 个子空间内样本权重; $\boldsymbol{x}_{k,\,j}$ 表示第 k 个子空间的第 j 个样本的特征; λ 为正则项系数.

此时, 子空间样本权重更新公式可表示为:

$$\boldsymbol{w}_k^{(t+1)}=\boldsymbol{w}_k^{(t)}-\alpha\left(\boldsymbol{H}_k^{(t)}\right)^{-1}\nabla\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right) \tag{9}$$

其中, $\boldsymbol{H}_k^{(t)}$ 表示海森矩阵; α 表示收敛容忍度; $\nabla\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right)$ 为目标函数梯度, 可表示为:

$$\nabla\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right)=-\frac{2}{N_k}\sum_{j=1}^{N_k}\left(y_{k,\,j}-\left(\boldsymbol{w}_k^{(t)}\right)^{\mathrm{T}}\boldsymbol{x}_{k,\,j}\right)\boldsymbol{x}_{k,\,j}+2\lambda\boldsymbol{w}_k^{(t)} \tag{10}$$

在实际应用中直接求解海森矩阵的逆会消耗大量的计算资源. 对此, 使用拟牛顿法通过逐步更新海森矩阵的逆以实现近似, 进而能够显著降低计算的复杂性, 其规则可表示为:

$$\left(\boldsymbol{H}_k^{(t+1)}\right)^{-1}=\left(\boldsymbol{H}_k^{(t)}\right)^{-1}+\frac{\Delta\boldsymbol{w}_k^{(t)}\left(\Delta\boldsymbol{w}_k^{(t)}\right)^{\mathrm{T}}}{\left(\Delta\boldsymbol{w}_k^{(t)}\right)^{\mathrm{T}}\Delta\boldsymbol{G}_k^{(t)}}-\frac{\left(\boldsymbol{H}_k^{(t)}\right)^{-1}\Delta\boldsymbol{G}_k^{(t)}\left(\Delta\boldsymbol{G}_k^{(t)}\right)^{\mathrm{T}}\left(\boldsymbol{H}_k^{(t)}\right)^{-1}}{\left(\Delta\boldsymbol{G}_k^{(t)}\right)^{\mathrm{T}}\left(\boldsymbol{H}_k^{(t)}\right)^{-1}\Delta\boldsymbol{G}_k^{(t)}} \tag{11}$$

其中, $\Delta\boldsymbol{w}_k^{(t)}$ 和 $\Delta\boldsymbol{G}_k^{(t)}$ 分别为权重变化量和梯度变化量.

$$\Delta\boldsymbol{w}_k^{(t)}=\boldsymbol{w}_k^{(t+1)}-\boldsymbol{w}_k^{(t)} \tag{12}$$

$$\Delta \boldsymbol{G}_k^{(t)} = \nabla \mathcal{L}_k^{(t+1)}\left(\boldsymbol{w}_k^{(t+1)}\right) - \nabla \mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right) \tag{13}$$

当满足式 (14) 或达到最大迭代次数 $N_{\text{iter}}^{\text{Main}}$ 时, 则停止优化.

$$\left\|\nabla \mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right)\right\| < N_{\text{tol}}^{\text{Main}} \tag{14}$$

其中, $N_{\text{tol}}^{\text{Main}}$ 表示收敛容忍度.

通过上述方式, 避免了昂贵的矩阵求逆运算. 这种方法不仅提高了计算效率, 也加速了优化过程的收敛速度. 最终, 得到子空间样本的优化权重 $\boldsymbol{w}_k$. 此时, 对于给定的输入样本 $\boldsymbol{x}_k$, 主模型算法输出可表示为:

$$\hat{\boldsymbol{y}}_{\text{Main}} = \sum_{k=1}^{K}\left(\boldsymbol{w}_k^{\text{T}}\boldsymbol{x}_k \times I\left(\boldsymbol{x}_k \in R_k\right)\right) \tag{15}$$

其中, $I(\cdot)$ 为指示函数, 当 $\boldsymbol{x}_k \in R_k$ 存在时为1, 否则为0.

3.2.2 真实数据驱动误差补偿模型子模块

前文已基于 CO_2 虚拟数据建立了可解释主模型. 然而, 由于实际工业过程的物料波动、设备异常、环境变化等因素导致所构建的 MD 模型与真实工业过程存在一定的偏差, 进而显著影响主模型的精度. 为了解决这一问题, 本文提出通过引入神经网络的拟合能力以补偿主模型输出, 采用 IT3FBLS 提取强不确定性关系. 此处的强不确定性关系特指工业现场因 MSWI 过程的燃烧成分强波动、设备烧蚀损坏等因素引起的运行工况频繁波动而导致的主模型与真实过程之间难以以可解释方式表征的具有强非线性和不确定性的未建模动态.

首先, 将真实数据输入树结构主模型中以得到残差 $\boldsymbol{e}^{\text{PD, CO}_2}$. 定义补偿模型的训练数据为 $\boldsymbol{D}^{\text{PD, CO}_2} = \left[\boldsymbol{X}^{\text{PD, CO}_2},\ \boldsymbol{e}^{\text{PD, CO}_2}\right]$, 其中 $\boldsymbol{X}^{\text{PD, CO}_2}$ 为 CO_2 真实数据的输入特征, $\boldsymbol{e}^{\text{PD, CO}_2}$ 为主模型输出与真实数据真值间的残差.

在输入层, $\boldsymbol{X}^{\text{PD, CO}_2}$ 可表示为:

$$\boldsymbol{X}^{\text{PD, CO}_2} = \left[\boldsymbol{x}_1^{\text{PD, CO}_2},\ \boldsymbol{x}_2^{\text{PD, CO}_2},\ \cdots,\ \boldsymbol{x}_{N_{\text{PD}}}^{\text{PD, CO}_2}\right]^{\text{T}} \tag{16}$$

其中, N_{PD} 表示 $\boldsymbol{X}^{\text{PD, CO}_2}$ 的样本数量; $\boldsymbol{x}_n^{\text{PD, CO}_2} = \left[x_{n,\,1}^{\text{PD}},\ x_{n,\,2}^{\text{PD}},\ \cdots,\ x_{n,\,M}^{\text{PD}}\right]$, $x_{n,\,M}^{\text{PD}}$ 表示 $\boldsymbol{x}_n^{\text{PD, CO}_2}$ 的第 M 个特征.

在 IT3FS 层, 该层包含 N_{IT3FS} 个 IT3FS 子系统, 以第 n_{IT3FS} 个 IT3FS 子系统为例, 具体描述如下:

$\boldsymbol{X}^{\text{PD, CO}_2}$ 经输入层神经元直接传递至后续网络, 在隶属函数层中针对输入特征计算模糊隶属函数 $\tilde{A}_m^l\ (l = 1,\ 2,\ \cdots,\ L,\ m = 1,\ 2,\ \cdots,\ M)$ 在不同水平切片 $\mu_g = \alpha_g\ (g = 1,\ 2,\ \cdots,\ G)$ 下的隶属度值如下所示:

$$\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right) = m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) + \sqrt{\ln\left(\frac{1}{\overline{\alpha}_g}\right)\sigma_{V_m}^2} \tag{17}$$

$$\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right) = m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) - \sqrt{\ln\left(\frac{1}{\overline{\alpha}_g}\right)\sigma_{V_m}^2} \tag{18}$$

$$\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right) = m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) + \sqrt{\ln\left(\frac{1}{\underline{\alpha}_g}\right)\sigma_{V_m}^2} \tag{19}$$

$$\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right) = m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) - \sqrt{\ln\left(\frac{1}{\underline{\alpha}_g}\right)\sigma_{V_m}^2} \tag{20}$$

其中, $\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$ 和 $\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$ 以及 $\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$ 和 $\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$ 表示第 m 个输入特征在水平切片等级为 $\overline{\alpha}_g/\underline{\alpha}_g$ 时模糊隶属函数 $\tilde{A}_m^l$ 的上界和下界; $m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right)$ 表示在基准切片 α_0 下, 模糊集合对样本 $x_{n,\,m}^{PD}$ 的隶属度; δ 表示不确定扰动变量; $\left[\underline{\alpha}_g,\ \overline{\alpha}_g\right]$ 表示二级隶属度的不确定性, 可表示为:

$$\underline{\alpha}_g = \alpha_g^{\frac{1}{\Delta}} \tag{21}$$

$$\overline{\alpha}_g = \alpha_g^{\Delta} \tag{22}$$

其中, $\Delta > 1$ 是一个常数, 表示高斯垂直切片的标准差的不确定性.

在 IT3FS 中, 使用下界计算标准差以保证模糊逻辑系统能够在处理不确定性和模糊性时, 从而保持更高的稳定性和可靠性.

$$\sigma_{V_m}^2\left(x_{n,\,m}^{\text{PD}}\right) = \frac{\left(\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) - m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right)\right)^2}{\ln\left(\frac{1}{\varepsilon}\right)} \tag{23}$$

其中,

$$m_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) = \frac{\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right) + \overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\text{PD}}\right)}{2} \tag{24}$$

$$\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\mathrm{PD}}\right)=\exp\left(-\frac{\left(x_{n,\,m}^{\mathrm{PD}}-C_{\tilde{A}_m^l}\right)^2}{\underline{\sigma}_{V_m}^2}\right) \tag{25}$$

$$\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\alpha_0}\left(x_{n,\,m}^{\mathrm{PD}}\right)=\exp\left(-\frac{\left(x_{n,\,m}^{\mathrm{PD}}-C_{\tilde{A}_m^l}\right)^2}{\overline{\sigma}_{V_m}^2}\right) \tag{26}$$

其中, $C_{\tilde{A}_m^l}$、$\underline{\sigma}_{V_m}$ 和 $\overline{\sigma}_{V_m}$ 分别表示在水平切片等级 $\mu_\delta=\alpha_0$ 下模糊隶属函数 $\tilde{A}_m^l$ 的中心、下宽度和上宽度; ε 表示收敛阈值.

然后, 在规则推理层中计算模糊规则激活强度从而确定输出变量的范围. 第 l 条模糊规则的激活强度如下:

$$\tau_l=\mu_{\tilde{A}_1^l}\left(x_{n,\,1}^{PD}\right)\wedge\mu_{\tilde{A}_2^l}\left(x_{n,\,2}^{PD}\right)\wedge\cdots\wedge\mu_{\tilde{A}_M^l}\left(x_{n,\,M}^{PD}\right) \tag{27}$$

其中, τ_l 为第 l 条模糊规则的激活; $\mu_{\tilde{A}_M^l}\left(x_{n,\,M}^{PD}\right)$ 为输入 x 隶属于第 l 条模糊规则的激活强度. 进而对应的第 l 条规则的后件输出区间为 $[\underline{w}_l,\,\overline{w}_l]$. $\underline{w}_l$ 和 $\overline{w}_l$ 表示第 l 条规则的后件参数; 此时, 输入特征变量的范围被划分为多个区间, 进而基于不同区间隶属度值确定每条模糊规则的激活强度, 计算公式如下:

$$\overline{z}_{\mu_\delta=\overline{\alpha}_g}^l=\prod_{g=1}^{G}\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g} \tag{28}$$

$$\underline{z}_{\mu_\delta=\overline{\alpha}_g}^l=\prod_{g=1}^{G}\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\overline{\alpha}_g} \tag{29}$$

$$\overline{z}_{\mu_\delta=\underline{\alpha}_g}^l=\prod_{g=1}^{G}\overline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g} \tag{30}$$

$$\underline{z}_{\mu_\delta=\underline{\alpha}_g}^l=\prod_{g=1}^{G}\underline{\mu}_{\tilde{A}_m^l|\mu_\delta=\underline{\alpha}_g} \tag{31}$$

进而, 第 n_{IT3FS} 个 IT3FS 子系统的输出可表示为:

$$z_{n\mathrm{IT3FS}}=\frac{\sum\limits_{g=1}^{G}\left(\overline{\alpha}_g\overline{f}_g+\underline{\alpha}_g\underline{f}_g\right)}{\sum\limits_{g=1}^{G}\left(\overline{\alpha}_g+\underline{\alpha}_g\right)} \tag{32}$$

其中, $\overline{f}_g$ 和 $\underline{f}_g$ 分别表示模糊系统在切片 $\overline{\alpha}_g$ 和 $\underline{\alpha}_g$ 下对应的加权输出, 计算如下:

$$\overline{f}_g=\frac{\sum\limits_{l=1}^{L}\left(\overline{z}_{\mu_\delta=\overline{\alpha}_g}^l\overline{w}_l+\underline{z}_{\mu_\delta=\overline{\alpha}_g}^l\overline{w}_l\right)}{\sum\limits_{l=1}^{L}\left(\overline{z}_{\mu_\delta=\overline{\alpha}_g}^l+\underline{z}_{\mu_\delta=\overline{\alpha}_g}^l\right)} \tag{33}$$

$$\underline{f}_g=\frac{\sum\limits_{l=1}^{L}\left(\overline{z}_{\mu_\delta=\underline{\alpha}_g}^l\underline{w}_l+\underline{z}_{\mu_\delta=\underline{\alpha}_g}^l\underline{w}_l\right)}{\sum\limits_{l=1}^{L}\left(\overline{z}_{\mu_\delta=\underline{\alpha}_g}^l+\underline{z}_{\mu_\delta=\underline{\alpha}_g}^l\right)} \tag{34}$$

最终, 将 N_{IT3FS} 个 IT3FS 子系统输出进行融合, 得到 IT3FS 层输出:

$$\boldsymbol{Z}=[\boldsymbol{z}_1,\,\boldsymbol{z}_2,\,\cdots,\,\boldsymbol{z}_{N_{\mathrm{IT3FS}}}] \tag{35}$$

在增强层, 该层包含 N_{Enhan} 个增强节点对 IT3FS 层输出进行非线性变换, 可表示为:

$$\boldsymbol{H}=[\boldsymbol{h}_1,\,\boldsymbol{h}_2,\,\cdots,\,\boldsymbol{h}_{N_{\mathrm{Enhan}}}] \tag{36}$$

$$\boldsymbol{h}_{N_{\mathrm{Enhan}}}=\xi_{n^{\mathrm{Enhan}}}\left(\boldsymbol{Z}\boldsymbol{W}_{n^{\mathrm{Enhan}}}+\boldsymbol{\beta}_{n^{\mathrm{Enhan}}}\right) \tag{37}$$

其中, $\boldsymbol{W}_{n^{\mathrm{Enhan}}}$ 和 $\boldsymbol{\beta}_{n^{\mathrm{Enhan}}}$ 分别表示 IT3FS 层与增强层中第 n^{Enhan} 个增强节点连接权值与偏置; $\xi_{n^{\mathrm{Enhan}}}$ 表示第 n^{Enhan} 个增强节点的激活函数.

输出层包含 1 个节点, 表示误差补偿模型的输出, 可表示为:

$$\hat{\boldsymbol{y}}_{\mathrm{Com}}=\boldsymbol{Z}\boldsymbol{W}_f+\boldsymbol{H}\boldsymbol{W}_e=[\boldsymbol{Z},\,\boldsymbol{H}]\begin{bmatrix}\boldsymbol{W}_f\\ \boldsymbol{W}_e\end{bmatrix}=\boldsymbol{A}\boldsymbol{W}_{\mathrm{Com}}^{\mathrm{IT3FBLS}} \tag{38}$$

其中, $\boldsymbol{W}_f$ 和 $\boldsymbol{W}_e$ 分别表示 IT3FS 层和增强层与输出层之间的连接权值; $\boldsymbol{A}$ 为 $\boldsymbol{Z}$ 和 $\boldsymbol{H}$ 的联合矩阵.

此时, 网络连接权重 $\boldsymbol{W}_{\mathrm{Com}}^{\mathrm{IT3FBLS}}$ 可表示为:

$$\boldsymbol{W}_{\mathrm{Com}}^{\mathrm{IT3FBLS}}=\boldsymbol{A}^{+}\boldsymbol{E} \tag{39}$$

其中, $\boldsymbol{A}^{+}$ 表示广义逆矩阵; $\boldsymbol{E}$ 表示模型真值.

采用 L^2 正则化将广义逆矩阵求取问题转换为优化问题, 如下:

$$\arg\min_{\boldsymbol{W}}\|\boldsymbol{A}\boldsymbol{W}-\boldsymbol{E}\|_2^2+\lambda\|\boldsymbol{W}\|_2^2 \tag{40}$$

进而, 采用岭回归估计, 得到 Moore-Penrose 逆矩阵 $\boldsymbol{A}^{+}$ 为:

$$\boldsymbol{A}^{+}=\lim_{\lambda\to0}\left(\lambda\boldsymbol{I}+\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)^{-1}\boldsymbol{A}^{\mathrm{T}} \tag{41}$$

其中, $\boldsymbol{I}$ 为单位矩阵.

最终得到的网络连接权值如下:

$$\boldsymbol{W}_{\mathrm{Com}}^{\mathrm{IT3FBLS}}=\left(\lambda\boldsymbol{I}+\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}\right)^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{E} \tag{42}$$

至此, 得到误差补偿模型的预测值.

3.3　真实数据驱动时序特性建模模块

由于主补模型无法捕获真实工业过程中固有的时序动态特性, 因此本节单独构建基于 LSTM 的时序预测模型. 该模型对真实历史数据进行学习以捕

获和预测 CO_2 排放浓度自身的变化规律与动态模式.

图 5 所示为 LSTM 单元的示例图.

该子模块计算过程如下:

本文的 LSTM 模型的损失函数为均方误差 MSE, 其公式如下:

$$E_t = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_{\text{Time},\,t} - y_t^{\text{PD}}\right)^2 \tag{43}$$

其中, E_t 为 t 时刻的损失; $\hat{y}_{\text{Time},\,t}$ 为 t 时刻的预测值; y_t^{PD} 为 t 时刻的真值.

首先, 将真实数据的目标值 $\boldsymbol{y}^{\text{PD}}$ 作为输入数据, 依据式 (44) 对其进行标准化以得到 $\boldsymbol{y}^{\text{PD, str}}$:

$$\boldsymbol{y}^{\text{PD, str}} = \frac{\boldsymbol{y}^{\text{PD}} - f_{\min}\left(\boldsymbol{y}^{\text{PD}}\right)}{f_{\max}\left(\boldsymbol{y}^{\text{PD}}\right) - f_{\min}\left(\boldsymbol{y}^{\text{PD}}\right)} \times (b-a) + a \tag{44}$$

其中, $f_{\min}(\cdot)$ 和 $f_{\max}(\cdot)$ 分别表示计算最小值和最大值; a 和 b 分别表示标准化的下限和上限.

接着, 计算 t 时刻遗忘门的过程值 $\boldsymbol{f}_t$ 和输出值 $\boldsymbol{g}_t$. 遗忘门通过作用于之前的记忆值可使神经元选择性地记忆或者忘记之前的信息. 进一步, 通过输入 t 时刻的 $y_t^{\text{PD, str}}$ 和 $t-1$ 时刻的 $\boldsymbol{h}_{t-1}$ 和 $\boldsymbol{c}_{t-1}$, 依据式 (45) 和式 (46) 可得 t 时刻遗忘门的过程值 $\boldsymbol{f}_t$ 以及遗忘门的输出值 $\boldsymbol{g}_t$:

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{z}_t^f\right) = \sigma\left(\boldsymbol{w}^{fx} y_t^{\text{PD, str}} + \boldsymbol{W}^{fh}\boldsymbol{h}_{t-1} + \boldsymbol{b}^f\right) \tag{45}$$

$$\boldsymbol{g}_t = \boldsymbol{c}_{t-1} \odot \boldsymbol{f}_t \tag{46}$$

其中, σ 为 Sigmoid 函数; $\boldsymbol{w}^{fx}$、$\boldsymbol{W}^{fh}$ 和 $\boldsymbol{b}^f$ 是遗忘门对应的权重和偏置; $\odot$ 为哈达玛积.

然后, 计算 t 时刻输入门的过程值 $\boldsymbol{i}_t$ 和输出值 $\boldsymbol{c}_t$ 以及单元状态更新值 $\hat{\boldsymbol{c}}_t$: 输入门作用于当前时刻的输入值, 在输入 $y_t^{\text{PD, str}}$ 和 $\boldsymbol{h}_{t-1}$ 后, 依据式 (47) ~ (49) 得到 $\boldsymbol{i}_t$、$\hat{\boldsymbol{c}}_t$ 及 $\boldsymbol{c}_t$:

$$\boldsymbol{i}_t = \sigma\left(\boldsymbol{z}_t^i\right) = \sigma\left(\boldsymbol{w}^{ix} y_t^{\text{PD, str}} + \boldsymbol{W}^{ih}\boldsymbol{h}_{t-1} + \boldsymbol{b}^i\right) \tag{47}$$

$$\hat{\boldsymbol{c}}_t = \text{Tanh}\left(\boldsymbol{z}_t^c\right) = \text{Tanh}\Big(\boldsymbol{w}^{cx} y_t^{\text{PD, str}} + \boldsymbol{W}^{ch}\boldsymbol{h}_{t-1} + \boldsymbol{b}^c\Big) \tag{48}$$

$$\boldsymbol{c}_t = \boldsymbol{i}_t \odot \hat{\boldsymbol{c}}_t + \boldsymbol{g}_t \tag{49}$$

其中, Tanh 为正切函数; $\boldsymbol{w}^{ix}$、$\boldsymbol{W}^{ih}$ 和 $\boldsymbol{b}^i$ 是输入门对应的权重和偏置; $\boldsymbol{w}^{cx}$、$\boldsymbol{W}^{ch}$ 和 $\boldsymbol{b}^c$ 是单元更新状态对应的权重和偏置.

之后, 计算 t 时刻输出门的过程值 $\boldsymbol{o}_t$ 和输出值 $\boldsymbol{h}_t$: 输入 $\boldsymbol{h}_{t-1}$、$y_t^{\text{PD, str}}$ 和 $\boldsymbol{c}_t$ 后, 依据式 (50) 和式 (51) 可得 $\boldsymbol{o}_t$ 以及输出值 $\boldsymbol{h}_t$:

$$\boldsymbol{o}_t = \sigma\left(\boldsymbol{z}_t^o\right) = \sigma\left(\boldsymbol{w}^{ox} y_t^{\text{PD, str}} + \boldsymbol{W}^{oh}\boldsymbol{h}_{t-1} + \boldsymbol{b}^o\right) \tag{50}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \text{Tanh}\left(\boldsymbol{c}_t\right) \tag{51}$$

其中, $\boldsymbol{w}^{ox}$、$\boldsymbol{W}^{oh}$ 和 $\boldsymbol{b}^o$ 是输出门对应的权重和偏置.

最后, 针对输出层, 在输入 $\boldsymbol{h}_t$ 后, 依据式 (52) 得到时序模型输出 $\hat{y}_{\text{Time},\,t+1}$:

$$\hat{y}_{\text{Time},\,t+1} = \sigma\left(\boldsymbol{v}^{\text{T}}\boldsymbol{h}_t + b_v\right) \tag{52}$$

其中, $\boldsymbol{v}$、b_v 为输出层的权重、偏置.

基于上述已完成的前向部分, 再采用梯度下降算法对 LSTM 模型中的 $\boldsymbol{w}^{fx}$、$\boldsymbol{W}^{fh}$、$\boldsymbol{w}^{ix}$、$\boldsymbol{W}^{ih}$、

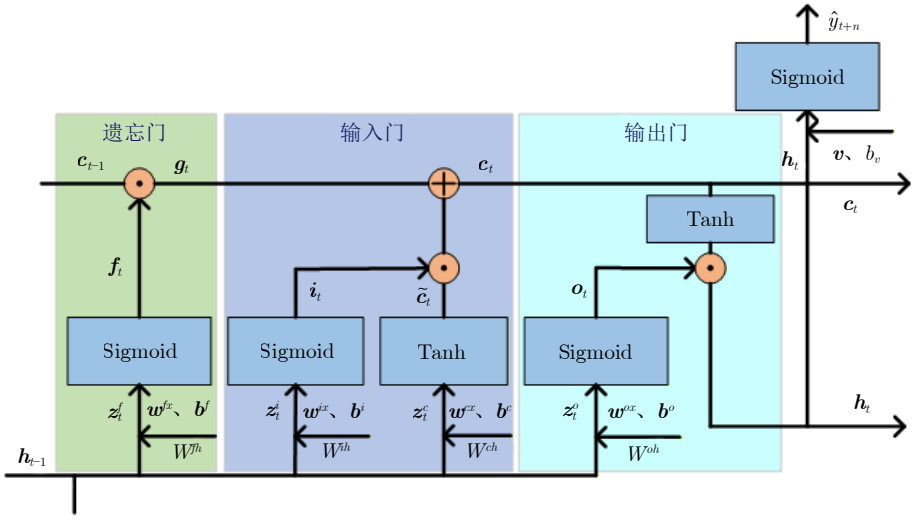


图 5 LSTM 单元图

Fig. 5 LSTM cell diagram

$\boldsymbol{w}^{cx}$、$\boldsymbol{W}^{ch}$、$\boldsymbol{w}^{ox}$、$\boldsymbol{W}^{oh}$、$\boldsymbol{b}^{f}$、$\boldsymbol{b}^{i}$、$\boldsymbol{b}^{c}$ 和 $\boldsymbol{b}^{o}$ 进行参数更新.

至此, 已完成 LSTM 网络的第一次预测及参数更新. 重复上述步骤, 直至预测完整序列, 最终得到预测值 $\hat{\boldsymbol{y}}_{\text{Time}}$.

3.4 虚实模型输出动态融合模块

前文已分别得到虚实驱动的主补机制建模模块和真实数据驱动时序特性建模模块的结果, 此时需进一步将上述结果进行融合, 得到最终预测结果.

模型融合输出模块的输入如下:

$$\boldsymbol{X}^{\text{Final}} = \left[\hat{\boldsymbol{y}}_{\text{Main}},\ \hat{\boldsymbol{y}}_{\text{Com}},\ \boldsymbol{X}^{\text{MD}},\ \hat{\boldsymbol{y}}_{\text{Time}}\right] = \left\{\boldsymbol{x}_1^{\text{Final}},\ \boldsymbol{x}_2^{\text{Final}},\ \cdots,\ \boldsymbol{x}_{N_{\text{len}}}^{\text{Final}}\right\} \tag{53}$$

其中, N_{len} 表示 $\boldsymbol{X}^{\text{Final}}$ 的样本数. 为简化建模复杂度, 此处采用改进岭回归以得到连接权重:

$$\boldsymbol{W} = \left(\lambda \boldsymbol{I} + \boldsymbol{X}^{\text{Final}}\left(\boldsymbol{X}^{\text{Final}}\right)^{\text{T}}\right)^{-1}\left(\boldsymbol{X}^{\text{Final}}\right)^{\text{T}}\boldsymbol{y}^{\text{MD}} \tag{54}$$

此时, 所构建的模型输出可表示为:

$$\hat{\boldsymbol{y}} = \boldsymbol{X}^{\text{Final}}\boldsymbol{W} \tag{55}$$

为进一步提升模型精度, 此处采用 variable step size least mean square (VSS-LMS) 基于误差实现权重在线更新, 如下:

$$\boldsymbol{W}^{(t+1)} = \boldsymbol{W}^{(t)} + \eta^{(t)} e^{(t)}\left(\boldsymbol{X}^{\text{Final}}\right)^{(t)} \tag{56}$$

其中, $\eta^{(t)}$ 为动态更新学习率; $e^{(t)}$ 为预测误差.

为使得在误差较小的情况下, 更新步长较大, 帮助更快速地收敛; 在误差较大的时候, 减小学习率, 避免过度更新, 采用下式为:

$$\eta^{(t+1)} = \frac{\eta^{(t)}}{1+\zeta\left(e^{(t)}\right)^2} \tag{57}$$

其中, ζ 为正则化系数.

3.5 算法伪代码

本文所提方法的伪代码如下所示.

算法 1. 算法伪代码

输入. 虚拟机理数据 $\boldsymbol{D}^{\text{MD, CO}_2}$, 真实过程数据 $\boldsymbol{X}^{\text{PD, CO}_2}$, 叶节点最小样本数 $N_{\text{samples}}^{\text{Main}}$, 主模型收敛容忍度 $N_{\text{tol}}^{\text{Main}}$, 主模型学习率 $N_{\text{learn}}^{\text{Main}}$, 补偿模型模糊规则数 $N_{\text{rule}}^{\text{Com}}$, 子系统数 N_{IT3FS}, 补偿模型增强节点数 $N_{\text{enh}}^{\text{Com}}$, LSTM 隐藏单元数 $N_{\text{unit}}^{\text{LSTM}}$, LSTM 训练轮次 $N_{\text{epo}}^{\text{LSTM}}$, LSTM 学习率 $N_{\text{learn}}^{\text{LSTM}}$, LSTM 的 Dropout 参数 $N_{\text{drop}}^{\text{LSTM}}$, 真实过程数据的目标值 $\boldsymbol{y}^{\text{PD}}$.

输出. 模型最终预测值 $\hat{\boldsymbol{y}}$.

1) While(true)
  For $j = 1$ to J
    For $m = 1$ to M
      由循环分割点 $s = (j,\ m)$ 得到左/右空间 $R_k^{\text{Left}}(s)/R_k^{\text{Right}}(s)$;
      计算该切分方式下的均方误差 $\Omega(s)$;
    end For
  end For
  If $N_k^{\text{Left}} < N_{\text{samples}}^{\text{Main}}$ and $N_k^{\text{Right}} < N_{\text{samples}}^{\text{Main}}$
    break;
  else
    continue;
  end If
  end While
2) 至此, 得到 K 个不相交的子空间 $R_1,\ R_2,\ \cdots,\ R_K$;
3) For $k = 1$ to K
  While(true)
    定义目标函数 $\mathcal{L}_k^{(t)}$;
    计算子空间样本权重 $\boldsymbol{w}_k^{(t)}$;
    计算目标函数梯度 $\nabla\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right)$;
    If $\nabla\mathcal{L}_k^{(t)}\left(\boldsymbol{w}_k^{(t)}\right) < N_{\text{tol}}^{\text{Main}}$
      break;
    else
      continue;
    end If
  end While
  end For
4) 计算主模型预测值 $\hat{\boldsymbol{y}}_{\text{Main}}$;
5) 获得补偿模型的输入数据 $\boldsymbol{D}^{\text{PD}}$;
6) For $n_{\text{IT3FS}} = 1$ to N_{IT3FS}
  For $l = 1$ to $N_{\text{rule}}^{\text{Com}}$
    计算在不同水平切片下的隶属度值;
    计算不同水平切片下隶属函数中心 $C_{\tilde{A}_m^l}$、下宽度 $\underline{\sigma}_{V_m}$ 和上宽度 $\bar{\sigma}_{V_m}$;
    在规则推理层中计算模糊规则激活强度;
  end For
    获得第 n_{IT3FS} 个 IT3FS 子系统的输出 $z_{n_{\text{IT3FS}}}$;
  end For
7) 融合 N_{IT3FS} 个 IT3FS 子系统输出得到 IT3FS 层输出 $\boldsymbol{Z}$;
8) For $n^{\text{Enhan}} = 1$ to $N_{\text{enh}}^{\text{Com}}$
    对 IT3FS 层输出进行非线性变换, 得到 IT3FS 层与增强层中第 n^{Enhan} 个增强节点的连接权值与偏置;
  end For
9) 计算补偿模型的输出值 $\hat{\boldsymbol{y}}_{\text{Com}}$;
10) 定义 LSTM 时序模型的损失函数;

11) 将真实过程数据的目标值 $\boldsymbol{y}^{\mathrm{PD}}$ 作为输入数据, 进行标准化得到 $\boldsymbol{y}^{\mathrm{PD,\,str}}$;

12) For $t=1$ to $N_{\mathrm{len}}^{\mathrm{time}}$ ($N_{\mathrm{len}}^{\mathrm{time}}$ 表示 $\boldsymbol{y}^{\mathrm{PD}}$ 的长度)

输入 $\boldsymbol{y}^{\mathrm{PD,\,str}}$ 以计算得到 t 时刻遗忘门的过程值 $\boldsymbol{f}_t$ 和输出值 $\boldsymbol{g}_t$;

计算 t 时刻输入门的过程值 $\boldsymbol{i}_t$、输出值 $\boldsymbol{c}_t$ 和单元状态更新值 $\hat{\boldsymbol{c}}_t$;

计算 t 时刻输出门的过程值 $\boldsymbol{o}_t$ 和输出值 $\boldsymbol{h}_t$;

计算 $t+1$ 时刻的预测值 $\hat{\boldsymbol{y}}_{\mathrm{Main},\,t+1}$;

采用梯度下降方法对时序模型进行参数更新;

end For

13) 得到时序模型的预测值 $\hat{\boldsymbol{y}}_{\mathrm{Time}}$;

14) 采用岭回归算法获得动态融合权重 $\boldsymbol{W}$;

15) 获得模型预测值 $\hat{\boldsymbol{y}}$;

4 实验验证

4.1 数据描述

本文验证对象是北京某 MSWI 电厂, 其采用的是处理量为 800 t/d 的往复式炉排炉, 炉排总长 11 m, 宽 12.9 m. 一次风总量为 6.54×10^4 $\mathrm{Nm^3/h}$, 从炉排下通过沿炉排长度的 4 个独立部分进入, 每部分的流量为总一次风量的 24.31%、43.35%、19.27% 和 13.07%, 二次风总量为 7.50×10^3 $\mathrm{Nm^3/h}$. 该 MSWI 电厂基准运行工况如表 1 所示.

表 1 焚烧炉的规格和基准工况运行参数表

Table 1 Specification and baseline operating condition parameter table of incinerators

参数	值	单位
额定处理量	800	t/d
实际处理量	624	t/d
炉排类型	往复式逆推	—
炉排长 × 宽	11.0 × 12.9	m
炉排速度	8	m/h
一次风量	65 400	$\mathrm{Nm^3/h}$
二次风量	7 500	$\mathrm{Nm^3/h}$
一次风温度	200	℃
一次风分布	24.31/43.35/19.27/13.07	—

MSW 的取样来自实际的 MSWI 厂, 其在实验室进行化验后的成分分析结果如表 2 所示, 数值仿真模型的输入参数如表 3 所示.

真实数据源自该 MSWI 厂 2021 年 3 月某日实际检测的数据. 采用 10 s 均值处理后, 共计 4 138 条样本. 所对应的输入特征为低位热值、炉排速度、进料量、干燥段风量、燃烧 1 段风量、燃烧 2 段风量、燃烬段风量和一次风量. 所提方法需将真实数据与虚拟数据进行单位统一, 具体处理如下:

表 2 MSW 成分分析

Table 2 MSW component analysis

	项目	值	单位
工业分析	水分 M	49.70	%
	固定碳 FC	7.82	%
	挥发分 V	32.22	%
	灰分 ASH	10.26	%
元素分析	C	65.62	%
	H	8.09	%
	N	1.12	%
	Cl	0.09	%
	S	0.11	%
	O	24.93	%

表 3 数值仿真模型输入参数

Table 3 Numerical simulation model input parameters

软件	参数	值	单位
FLIC	炉排速度	7.5	m/h
	炉排长 × 宽	11.0 × 12.9	m
	MSW 成分	见表 2	—
FLIC 和 Aspen Plus	进料量	624	t/d
	一次风量	65 400	$\mathrm{Nm^3/h}$
	一次风温度	200	℃
Aspen Plus	FLIC 输出烟气组分	见表 4	—
	二次风量	7 500	$\mathrm{Nm^3/h}$
	二次风温度	25	℃
	活性炭	17	kg/h
	氢氧化钙	16	kmol/h
	氨气	500	kg/h

1) MSW 低位热值 (lower heat value, LHV) 计算公式如下[34]:

$$M_{\mathrm{LHV}}=38.2M_{\mathrm{C}}+84.9\left(M_{\mathrm{H}}-\frac{M_{\mathrm{O}}}{8.0}\right)-0.5 \tag{58}$$

其中, M_{LHV} 为 MSW 低位热值; M_{C}、M_{H} 和 M_{O} 分别表示 MSW 中 C、H 和 O 的质量分数; 38.2、84.9、8.0 和 0.5 均为经验值.

2) 基于炉排长度 (11 m) 得到与虚拟数据集中单位一致的真实数据集的炉排速度, 其转换公式如下:

$$V_{\mathrm{Process}}^{\mathrm{Trans}}=V_{\mathrm{Process}}^{\mathrm{Ori}}\times\frac{11}{2\times\bar{V}_{\mathrm{Process}}^{\mathrm{Ori}}} \tag{59}$$

其中, $\bar{V}_{\mathrm{Process}}^{\mathrm{Ori}}$ 和 $V_{\mathrm{Process}}^{\mathrm{Ori}}$ 分别为真实数据集中炉排速度均值和样本实际值, 其单位为 %; $V_{\mathrm{Process}}^{\mathrm{Trans}}$ 为单位

统一转换后的炉排速度, 其单位为 m/h.

3) 由于进料量难以实时获取, 此处基于基准工况实际产能, 在取小时均值后加入均值为 0、方差为 0.1 的高斯噪声以得到不同样本.

4.2　评估指标

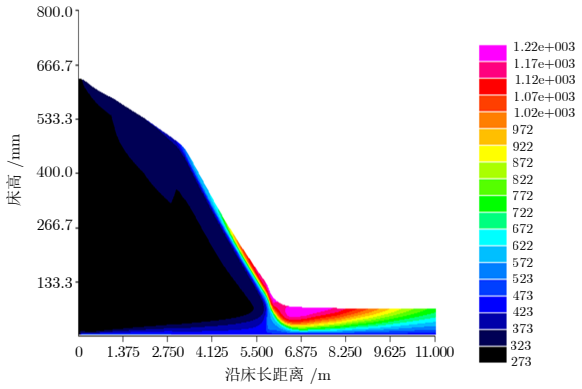
为评估预测模型效果, 选取平均绝对误差 MAE、均方根误差 RMSE 和决定系数 R^2 作为模型的评价指标.

4.3　实验设置及结果

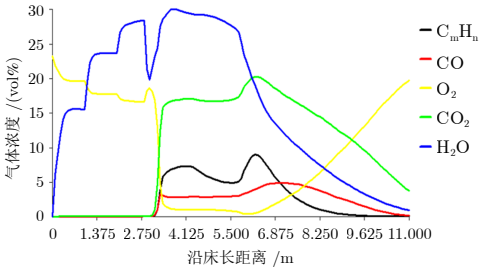
4.3.1　虚拟机理数据生成模块结果

4.3.1.1　全生命周期机理数值仿真模型构建子模块

由 FLIC 仿真的炉排固相燃烧过程结果如图 6 所示.



(a) 固体温度
(a) Solid temperature



(b) 气体浓度
(b) Gas concentration

图 6　炉排固相燃烧过程结果
Fig. 6　Grate solid phase combustion process results

由图 6 可知:

1) 在干燥炉排上, 温度大约为 350 K, MSW 内部的水分大量蒸发. 在 2.7 m 处, 水蒸气浓度达到最高, 但由于温度较低, MSW 未被点燃; 截至炉排床层 2.9 m 处, 炉排上的主要气体为 O_2 和 MSW 蒸发出的水蒸气.

2) 在 2.9 m 至 6.8 m 处, 固体温度开始上升, 挥发分开始释放, 并且 O_2 浓度骤降, 这表征 MSW 开始燃烧, 释放的主要产物有 C_mH_n、CO 和 CO_2.

3) 伴随着焦炭燃烧的结束, 固体温度从最高 1 220 K 逐渐降至 580 K. 由于燃烧结束, 因此 O_2 含量也迅速上升.

至此, 固相燃烧仿真完成, 床层上方混合气体的气体组分如表 4 所示, 其即为 Aspen Plus 中 B1 的输入.

表 4　输入 Aspen Plus 的气体组分
Table 4　Gas components input to Aspen Plus

气体组分	值 (mol%)	气体组分	值 (mol%)
CH_4	0.021 96	CO	0.021 04
CO_2	0.078 78	O_2	0.061 32
N_2	0.564 90	H_2O	0.355 80
H_2	0.022 04	NO	0.032 32

接着, 进行气相燃烧的模拟. 全流程五个区域采用 Aspen Plus 进行模拟, 五个区域的出口 CO_2 浓度以及质量流量如图 7 所示.

由图 7 可知:

1) 将 FLIC 软件得到的气体组分输入 Aspen Plus 后, 在固相燃烧区排放了 3 011 kg/h 的 CO_2. 由于气相燃烧区具有更高的温度以及充分的氧气供应, 因此在该区域能够实现更完全的燃烧过程. 此外, 固体废物在固相燃烧区首先会分解为可燃气体如 C_mH_n 和 CO, 这些气体在气相燃烧区进一步燃烧, 转化为 CO_2. 这些都导致气相燃烧区的 CO_2 质量流量高于固相燃烧区.

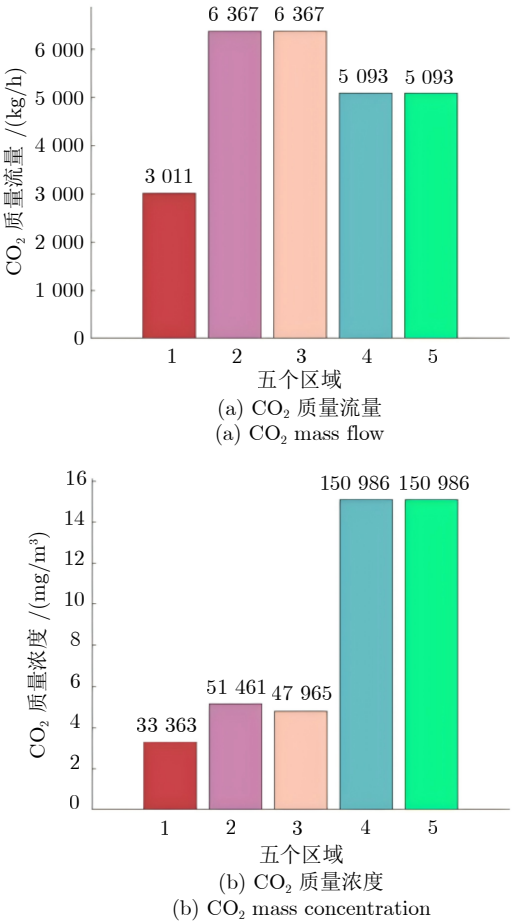
2) 由于换热区不涉及 CO_2 的相关反应, 因此 CO_2 的质量流量不变. 但由于在换热区是通过加入大量水来对高温烟气进行降温, 进而实现将烟气温度快速降至 180 ℃, 这导致烟气中出现部分水蒸气, 进而使烟气的体积增大, 因此换热区的 CO_2 质量浓度相比气相燃烧区有一定程度的减小. 气体质量浓度 C_Q 计算如式 (60) 所示:

$$C_Q = \frac{m_{\text{gas}}}{V_{\text{total}}} \tag{60}$$

其中, m_{gas} 为 CO_2 的质量流量; V_{total} 为混合烟气的总体积流量.

3) 在烟气净化区的脱酸反应器中, 氢氧化钙溶液会吸收 CO_2 形成碳酸钙固体, 导致 CO_2 的质量流量减小; 同时, 烟气净化区的氢氧化钙会与水蒸气发生吸湿反应, 生成氢氧化钙水合物, 导致烟气中水蒸气较少, 烟气的体积增大, 进而 CO_2 的质量浓度快速增大.

4) 在烟气排放区, 不涉及 CO_2 的化学反应以



(a) CO_2 质量流量
(a) CO_2 mass flow

(b) CO_2 质量浓度
(b) CO_2 mass concentration

图 7 不同位置 CO_2 质量浓度以及质量流量

Fig. 7 CO_2 mass concentration and mass flow at different locations

及温度变化, 因此 CO_2 的质量流量与质量浓度无显著差异.

为了进一步判断本文构建的 CO_2 全流程数值机理仿真模型的有效性, 需要与真实数据进行对比. 由于目前仅有烟气 G_{HT} 和 G_{GE} 处的化验结果, 因此只针对 G_{HT} 和 G_{GE} 处进行 CO_2 浓度仿真值和真实值的比较.

本文选取真实值将每分钟内的数据作均值处理, 共有 720 组, 每组包含了 194 个过程变量. 对该 720 组的 CO_2 浓度作均值处理, 得到 G_{GE} 处 CO_2 的真实排放浓度为 8.21%; 同时, 相关检测报告显示 G_{HT} 处 CO_2 的真实排放浓度为 2.58%. 此处的化验结果为 CO_2 体积浓度, 因此需要对数值仿真值进行换算, 如下[35]:

$$C_V = C_Q \times \frac{22.4}{44 \times 10} \tag{61}$$

其中, C_V 为 CO_2 的体积分数; 44 为 CO_2 摩尔质量; 22.4 为标准状况下气体摩尔体积.

经过计算, 本文所构建的全流程机理数值仿真模型中烟气 G_{HT} 和 G_{GE} 处 CO_2 的体积浓度分别为 2.40% 和 7.70%, 与 G_{HT} 和 G_{GE} 处真实值的相对误差分别为 6.90% 和 6.20%. 上述结果表明, 本文构建的基准工况机理数值仿真模型具有一定的可信度和有效性.

4.3.1.2 基于双重正交实验的多工况机理数据生成子模块

按照多因素多水平的双重正交实验设计得到的因素水平表如表 5 所示. "—"表示该参数未参与第二次正交实验.

按照表 5 的设计实施仿真, 其中第一次正交实验为 10 因素 5 水平 (49 例), 第二次对操纵变量进一步设计 6 因素 5 水平 (42 例), 共获得 2 058 (49 × 42) 条工况数据, 经人工处理得到 1 960 组虚拟机理数据. 虚拟样本数量减小的原因在于: 数值仿真软件在部分极端工况下会出现数值不收敛或结果物理意义异常的情况, 导致输出结果为无效值 (如 NaN);

表 5 正交实验参数信息

Table 5 Orthogonal experimental parameter information

	10 因素 5 水平		6 因素 5 水平	
参数	因素	单位	水平-1	水平-2
操纵变量	炉排速度	m/h	7.0, 7.5, 8.0, 8.5, 9.0	−0.1
	给料量	t/h	24.2, 24.7, 25.2, 25.7, 26.2	+0.1
	干燥段风量	Nm^3/h	16 080, 16 440, 16 800, 17 160, 17 520	+1.8
	燃烧 1 段风量	Nm^3/h	28 620, 29 280, 29 940, 30 600, 31 260	+3.2
	燃烧 2 段风量	Nm^3/h	12 660, 12 960, 13 260, 13 560, 13 860	+1.4
	燃烬段风量	Nm^3/h	8 640, 8 820, 9 000, 9 180, 9 360	+1.0
微观参数	颗粒大小	mm	15, 20, 25, 30, 35	—
	颗粒混合系数	无	$2e^{-6}$, $3e^{-6}$, $4e^{-6}$, $5e^{-6}$, $6e^{-6}$	—
组分参数	水分含量	%	48.00, 49.75, 51.50, 53.25, 55.00	—
	C : H : O 比率	%	(58.0 : 7.5 : 33.0), (59.0 : 7.5 : 32.0), (60.0 : 7.5 : 31.0), (61.0 : 7.5 : 30.0), (62.0 : 7.5 : 29.0)	—

此处对无效值样本进行了删除操作. 由于烟气净化区和烟气排放区的 CO_2 浓度相同, 因此 5 个区域出口处的 CO_2 仿真浓度数据的直方图如图 8 所示.

由图 8 可知, G_{SC}、G_{GC}、G_{HT} 和 G_{CL} (G_{GE}) 处的 CO_2 浓度分别在 3.20×10^4 ~ 3.50×10^4, 4.40×10^4 ~ 4.60×10^4, 4.00×10^4 ~ 4.20×10^4 和1.30×10^5 ~ 1.32×10^5 之间出现最大频率. 同时, 各个区域的 CO_2 仿真浓度直方图均呈现正态分布的形状.

需要提出的是, 由于基于 FLIC 和 Aspen Plus 软件的数值仿真是建立在众多假设的基础上的, 仿真所产生的数据也只是保障了该特定焚烧厂在特定工况下的虚拟机理数据的可靠性, 其普适性还有待研究; 另外, 炉内高温烟气的成分在目前阶段是无法实时检测的, 这也需要我们后续考虑获得炉内烟气成分化验分析结果以提升虚拟机理数据的准确性.

4.3.2　虚实驱动主补机制建模模块结果

本文所提的虚实驱动主补机制建模模块的超参数设置如表 6 所示.

通过上述超参数构建本文的虚实驱动主补机制模型, 重复 10 次实验, 其评价指标的均值统计结果和预测曲线如表 7 和图 9 所示.

针对虚实驱动主补机制建模模块的结果, 具体分析如下:

在预测精度方面, 主补模型在训练集、验证集与测试集上的均方根误差均值分别为 0.402 0、0.409 7 与 0.408 9, 平均绝对误差均值分别为 0.304 1、0.313 7 与 0.314 1. 各项误差指标在三个数据集上表现接近, 且其方差均处于较低数量级, 表明该模型具有较好的稳定性与泛化能力, 未出现明显过拟合现象.

在拟合效果方面, 训练集、验证集与测试集的决定系数均值分别为 0.455 2、0.426 0 与 0.430 2. 该结果表明, 主补模型能够解释目标变量约 42% ~ 46% 的方差, 说明其具备一定的输入/输出关系捕捉能力. 然而, R^2 仍处于中等水平, 反映出模型对系统的不确定性及未知扰动的刻画仍存在局限, 部分关键信息尚未被充分提取.

综合而言, 基于树结构的主模型与基于 IT3F-BLS 的补偿机制协同作用下, 主补机制模型表现出良好的泛化性与稳定性, 为后续融合建模奠定了基

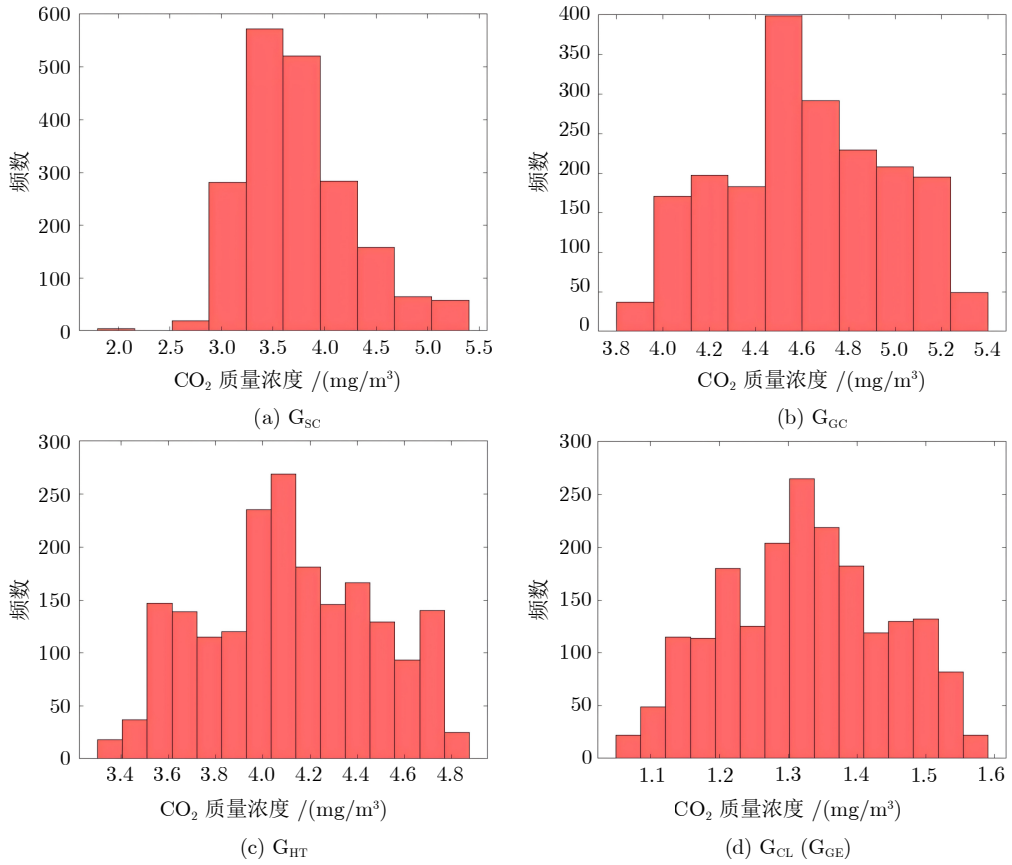


图 8　CO_2 仿真浓度直方图

Fig. 8　CO_2 simulation concentration histogram

表 6 虚实驱动主补机制建模模块超参数设置
Table 6 Virtual-real driven main-complement mechanism modeling module hyperparameter setting

模型超参数	符号	取值
叶节点最小样本数	$N_{\text{samples}}^{\text{Main}}$	400
主模型最大迭代次数	$N_{\text{iter}}^{\text{Main}}$	1 000
主模型收敛容忍度	$N_{\text{tol}}^{\text{Main}}$	0.01
主模型学习率	$N_{\text{learn}}^{\text{Main}}$	0.02
补偿模型模糊规则数	$N_{\text{rule}}^{\text{Com}}$	50
补偿模型子系统数	$N_{\text{IT3FS}}^{\text{Com}}$	50
补偿模型增强节点数	$N_{\text{enh}}^{\text{Com}}$	50

础. 但 R^2 仍有较大提升空间, 也进一步印证了后续引入 LSTM 时序模型以捕捉时间演化规律, 并采用岭回归进行多元输出融合的必要性, 从而进一步提升 CO_2 浓度的预测性能.

4.3.3 真实数据驱动时序特性建模模块结果

本文所提的真实数据驱动时序特性建模模块的超参数设置如表 8 所示.

通过上述超参数构建本文的真实数据驱动时序特性建模模型, 重复 10 次实验, 其评价指标的均值统计结果和预测曲线如表 9 和图 10 所示.

针对真实数据驱动时序特性建模模块的结果, 具体分析如下:

首先, 模型具有出色的预测精度与拟合优度. 在测试集上, 其 RMSE 与 MAE 均值分别低至 0.218 9 与0.156 2, 相较于主补模型降低了约 46.5% 与 50.3%, 预测误差显著减小. 同时, 测试集的 R^2 均值达到 0.7 762, 意味着该模型能够解释目标变量近 78% 的方差, 其拟合能力远优于 R^2 均值仅为 0.430 2 的主补模型. 这充分证明, LSTM 网络能够有效学习并捕捉 CO_2 排放浓度序列中复杂的长期依赖性与非线性时序动态.

其次, 模型表现出高度的稳定性与鲁棒性. 所有评估指标在训练集、验证集与测试集上的数值均非常接近, 其中关键指标 R^2 在三个数据集间的极差仅为 0.013 6. 各指标的标准差均处于 10^{-3} 至 10^{-4} 量级, 波动极小. 这表明模型训练过程收敛稳定, 且未对训练数据产生过拟合, 具备良好的外推泛化性能.

最后, 该结果从时序维度揭示了主补模型的局限性与互补空间. 主补机制模型基于静态的输入–输出映射关系, 难以充分刻画浓度序列的时变特性. LSTM 时序模型仅依赖浓度序列自身的历史值, 便实现了更高的预测精度, 这清晰印证了 CO_2 浓度演化中蕴含的强时序规律. 因此, LSTM 时序模块与主补机制模型形成了机理与 DD、静态关系与动态演化之间的有效互补, 为后续通过岭回归进行多元融合提供了关键支撑与实证依据.

4.3.4 虚实输出动态融合模块结果

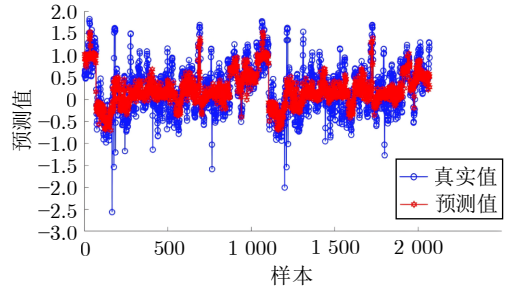
第 4.3.2 节和第 4.3.3 节已完成了虚实驱动主补机制建模与真实数据驱动时序特性建模, 本节采用改进岭回归分别将虚实驱动主补机制建模模块、真实数据以及真实数据驱动时序特性建模结果进行融合, 其评价指标的均值统计结果和预测曲线如表 10 和图 11 所示.

由表 10 和图 11 可知, 在测试集上, 融合模型的 R^2 均值高达 0.932 0, 表明其能够解释目标变量超过 93% 的方差, 模型拟合优度达到了优秀水平. 融合模块之所以能够实现如此大幅的精度提升, 其根本原因在于多个来源信息固有的本质差异性与高度互补性, 其中: 1) 虚实驱动主补机制建模模块能够表征虚拟机理数据中蕴含的稳态机理映射, 进而具有明确的物理可解释性, 其以当前时刻的过程变量为输入能够快速响应工况突变, 但受限于仅利用了同一时刻的截面信息, 未能捕捉 CO_2 浓度自身的历史时序累积效应; 2) 真实数据驱动时序特性建模模块能够表征时间序列的自回归演化规律, 其以历史时刻的 CO_2 浓度序列为输入进而能够充分利用历史信息; 3) 虚实输出动态融合模块通过引入原始变量信息和自适应地调整包括前述模块在内的多信息分量的融合权重, 这使得当工况发生变化时融合模型能够通过赋予高精度模块分量以更高权重而快速跟踪变化, 进而通过多元信息的协同互补而使得融合后的预测精度显著超越任一单一模块.

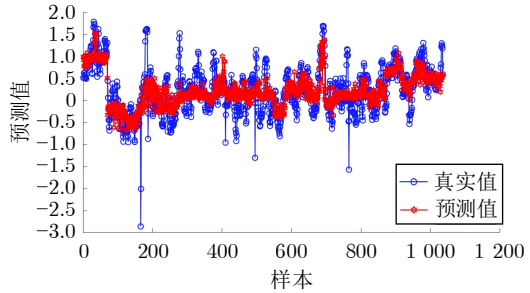
上述结果验证了改进岭回归融合机制不仅提升了预测精度, 还通过正则化约束有效制约了模型复

表 7 虚实驱动主补机制建模模块实验结果
Table 7 Experimental results of virtual-real driven main-complement mechanism modeling module

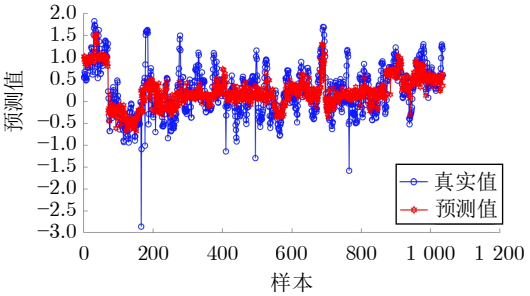
方法	数据集	RMSE		MAE		R^2	
		均值	方差	均值	方差	均值	方差
本文方法	训练集	0.402 0	2.3e−05	0.304 1	2.7e−05	0.455 2	1.9e−04
	验证集	0.409 7	1.2e−04	0.313 7	8.2e−05	0.426 0	9.5e−04
	测试集	0.408 9	8.3e−05	0.314 1	8.3e−05	0.430 2	6.7e−04



(a) 训练集
(a) Training set



(b) 验证集
(b) Validation set



(c) 测试集
(c) Test set

图 9　虚实驱动主补机制建模模块的拟合曲线

Fig. 9　The fitting curve of virtual-real driven main-complement mechanism modeling module

杂度, 避免了过拟合风险, 保障了模型在未知数据上的可靠性与鲁棒性.

为了进一步分析各个分量对最终预测结果的影响, 此处给出了岭回归算法的分配权重, 如表 11 所示:

由表 11 可知: 1) 主模型、补偿模型和时序模型在最终预测结果中均发挥了重要作用, 呈现出“时序主导、机理支撑、偏差修正”的协同互补特征; 2) 时序模型权重最高, 为 0.424 9, 表明 CO_2 排放浓度作为典型的时序变量, 其时间序列内在演化规律对最终预测精度的影响最为显著; 3) 主模型预测值权重为 0.186 2, 仅次于时序模型, 充分体现了机理驱动模型的核心价值; 4) 补偿模型预测值权重为 0.152 7, 与主模型权重相近, 验证了不确定性补偿机制的有效性. 补偿模型通过 IT3FBLS 学习主模型的预测残差, 以修正主模型在真实工况下因物料波动、设备异常、环境变化等强不确定性因素所导致的偏差, 进一步提升预测结果的准确性与抗干扰能力; 5) 一次风量等变量的权重均低于三个预测子模型, 整体对预测结果的贡献较小. 其中, 热值是全部变量中权重 (0.148 3) 最高的, 但仍低于主模型和补偿模型, 原因在于: 主模型基于已嵌入工艺规律虚拟机理数据构建, 补偿模型与时序模型进一步拟合真实数据特征, 三个模型已较为充分地表征了 CO_2 产生工艺过程中的有效信息, 进而导致作为单独输入的工艺变量对预测结果的贡献较小.

综上所述, 基于改进岭回归的虚实输出动态融合模块充分发挥了各模块的特长, 显著提升了模型的综合性能.

表 8　真实数据驱动时序特性建模模块超参数设置

Table 8　Hyperparameter settings of real data driven time series characteristics modeling module

模型超参数	符号	取值
隐藏层单元数	$N_{\text{unit}}^{\text{LSTM}}$	64
训练轮次	$N_{\text{epo}}^{\text{LSTM}}$	300
学习率	$N_{\text{learn}}^{\text{LSTM}}$	0.01
Dropout 系数	$N_{\text{drop}}^{\text{LSTM}}$	0.2

4.4　方法比较

为验证本文所提方法的有效性, 本文采用反向传播神经网络 (back propagation neural network, BPNN)、LSTM、RF、模糊决策森林 (fuzzy decision forest, FDF) 和 Transformer 模型与本文所提模型进行对比, 相应模型的参数设置为: 1) BPNN, 训练轮次为 500, 隐藏层单元数为 50, 学习率为 0.01,

表 9　真实数据驱动时序特性建模模块实验结果

Table 9　Experimental results of real data driven time series characteristics modeling module

方法	数据集	RMSE		MAE		R^2	
		均值	方差	均值	方差	均值	方差
本文方法	训练集	0.213 2	4.5e−06	0.155 4	2.9e−06	0.784 7	1.7e−05
	验证集	0.221 5	4.8e−06	0.156 0	3.4e−06	0.771 1	2.0e−05
	测试集	0.218 9	4.8e−06	0.156 2	2.9e−06	0.776 2	2.0e−05

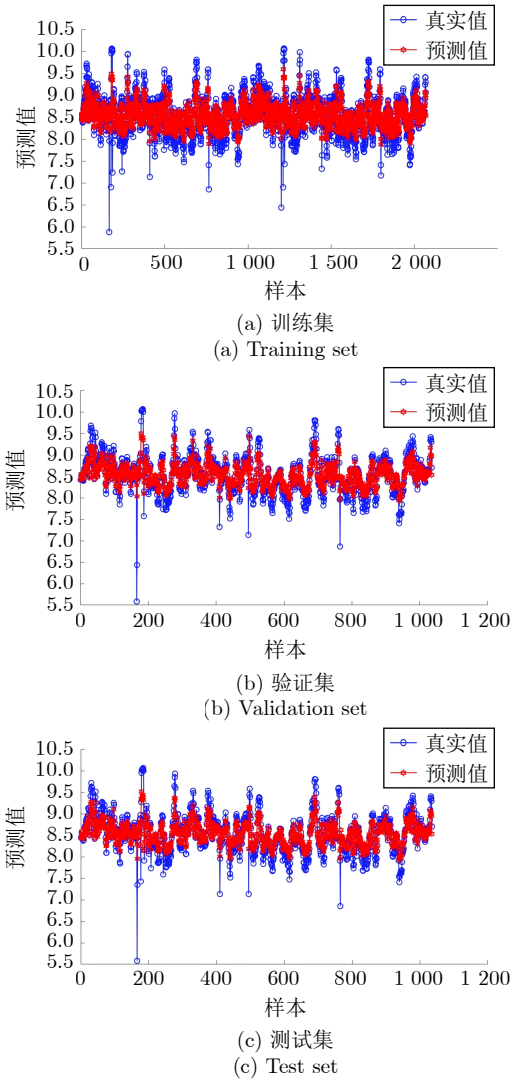


图 10 真实数据驱动时序特性建模模块的拟合曲线

Fig. 10 The fitting curve of real data driven time series characteristic modeling module

收敛误差为 0.01; 2) LSTM, 训练轮次为 300, 隐藏层单元数为 64, 学习率为 0.01, 惩罚因子为 0.2, Dropout 参数为 0.8; 3) RF, 叶节点最小样本个数为 500, 特征选择数量为 7, 决策树数量为 50; 4) FDF, 输出参数为 2.4, 叶节点最小样本个数为 500, 特征选择数量为 7, 决策树数量为 50; 5) Transformer, 隐藏层维度为 64, 注意力头数为 4, 编码器层数为 2, 学习率为 0.001.

上述方法均重复 10 次实验, 其评价指标的均值统计结果和预测曲线如表 12 和图 12 所示. 需要特别说明的是, 在采用对比方法建立单一模型时, 为保证模型训练的公平性, 将虚拟数据集与划分用于训练的真实数据合并后作为对比方法的训练数据.

针对对比实验结果, 具体分析如下: 在测试集上, 本文方法的 RMSE 均值为 0.120 5、MAE 均值为 0.059 9、R^2 均值达到 0.932 0, 相比性能次优的 LSTM 方法, RMSE 均值降低了 58.4%, R^2 均值提升了 53.5%, 充分证明了本文方法的优越性. 从泛化能力来看, 本文方法在训练集、验证集和测试集上的 R^2 均值分别为 0.943 5、0.927 3 和 0.932 0, 三者高度一致, 差异不超过 1.6%, 远小于其他对比方法 (如 LSTM 的验证集与训练集 R^2 均值差异达 12.0%), 表明本文方法具有良好的泛化性能和稳定性. 此外, 本文方法的 MAE 方差在各数据集上均低于 1.060×10^{-5}, 远小于 LSTM的 1.840×10^{-4}, 体现了其出色的鲁棒性. 相比之下, 传统机器学习方法 BPNN 和 RF 的 R^2 均值均低于 0.300 0, 难以有效捕捉 MSWI 过程的非线性动态特性; LSTM 虽优于传统方法, 但其 R^2 均值仅为 0.607 1, 仍有较大提升空间; 标准 Transformer 受限于数据规模, 表现不及 LSTM; FDF 方法因线性假设限制, R^2 为负值, 无法有效拟合数据分布.

综上所述, 本文方法在 MSWI 过程 CO_2 浓度预测任务中取得了显著优于现有方法的预测精度、泛化能力和鲁棒性.

4.5 消融实验

为验证本文所提方法中各个模块的有效性, 分别设计了无主补机制、无时序机制以及无岭回归三种消融实验, 其中后者是采用均值融合的方式获得模型输出. 上述方法均重复 10 次实验, 其评价指标的均值统计结果和预测曲线如表 13 和图 13 所示.

针对消融实验结果的具体分析如下.

1) 无虚实驱动主补机制建模模块在测试集上 RMSE 均值为 0.275 4, MAE 均值为 0.194 6, R^2 均值为 0.645 8, 性能大幅下降. 其中, RMSE 相较于完整方法上升了 128.5%, 这表明虚实驱动主补机

表 10 虚实输出动态融合模块实验结果

Table 10 Experimental results of virtual-real output dynamic fusion module

方法	数据集	RMSE		MAE		R^2	
		均值	方差	均值	方差	均值	方差
本文方法	训练集	0.109 1	1.760e−05	0.057 5	8.940e−06	0.943 5	1.840e−05
	验证集	0.124 7	2.230e−05	0.061 1	1.062e−05	0.927 3	2.990e−05
	测试集	0.120 5	1.823e−05	0.059 9	8.930e−06	0.932 0	2.250e−05

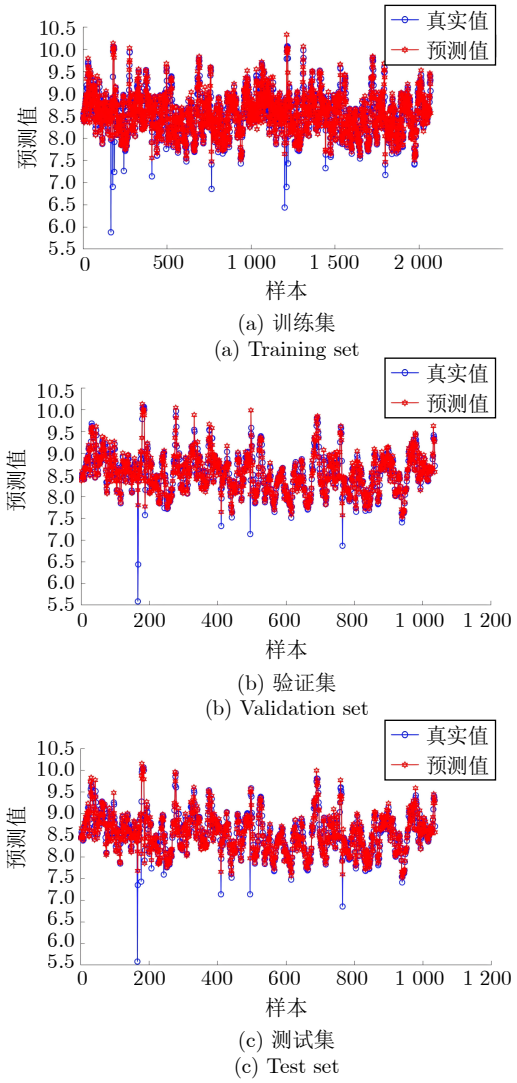


图 11　虚实输出动态融合模块的拟合曲线

Fig.11　The fitting curve of virtual-real output dynamic fusion module

制在捕捉 CO_2 浓度与操纵变量之间的复杂非线性关系方面起着一定的作用. 纯时序模型虽能学习序列演化规律, 但无法充分利用过程机理信息, 导致泛化能力有限.

2) 无真实数据驱动时序特性建模模块在测试集上 RMSE 均值为 0.407 6, MAE 均值为 0.315 7, R^2 均值为 0.225 9, 性能最差. RMSE 相较于完整方法上升了 238.3%, 这突显了时序建模的关键作用. CO_2 排放浓度具有显著的时间依赖性和动态特性, LSTM 模型能够有效捕捉序列中的长期依赖关系, 弥补了虚实驱动主补机制建模模型在时序预测上的不足.

3) 无虚实输出动态融合模块在测试集上 RMSE 均值为 0.273 5, MAE 均值为 0.202 6, R^2 均值为 0.650 6, 相较于完整方法 RMSE 增加了 127.0%. 这一结果验证了岭回归融合策略的有效性, 原因在于均值融合为各模型分配固定权重, 无法自适应调整不同工况下各模型的贡献度; 改进岭回归通过学习最优权重系数, 实现了预测结果的优势互补, 显著提升了最终预测精度.

表 11　岭回归融合算法的分配权重

Table 11　Assigned weights for ridge regression fusion algorithm

分量	权重
主模型预测值	0.186 2
补偿模型预测值	0.152 7
炉排速度	0.012 3
热值	0.148 3
进料量	0.014 5
干燥段风量	−0.013 6
燃烧 1 段风量	0.042 1
燃烧 2 段风量	0.009 6
燃烬段风量	−0.008 4
一次风量	0.006 8
时序模型	0.424 9

综上, 消融实验结果表明, 任一模块的缺失都会导致性能显著下降, 验证了本文所提框架的完整性和必要性.

4.6　超参数分析

为了进一步对本文提出的模型进行性能评估, 本节对叶节点最小样本数、主模型最大迭代次数、主模型收敛容忍度、主模型学习率、补偿模型模糊规则数、补偿模型子系统数、补偿模型增强节点数、隐藏层单元数、训练轮次、学习率和 Dropout 系数进行敏感性分析, 其设置区间如表 14 所示.

实验中采用单变量分析策略, 每次只改变一个参数, 每个参数重复十次实验, 其最终测试集的统计结果如图 14 所示.

1) 叶节点最小样本数 $N_{\text{samples}}^{\text{Main}}$. 由图 14(a) 可见, 所有参数配置下的决定系数均维持在 0.910 0 以上的高水平, 证实了所提模型对于叶节点最小样本数的强适应能力. 当叶节点最小样本数设置为 350 时, 模型达到最优性能, 中位数 R^2 为 0.921 1 且变异程度较小. 当样本数增加至 650 时, 平均性能最佳, 表现出良好的集中趋势. 需注意的是, 模型性能对参数变化相对不敏感: 从较小样本数 (100 ~ 200) 到较大样本数 (700 ~ 800), R^2 值的波动范围有限, 最大差异仅约 0.003 6, 且各配置下实验结果的分布均较为集中. 这一现象表明, 在本研究场景下, 所提模型对叶节点最小样本数的选择具有较强的容错性.

2) 主模型最大迭代次数 $N_{\text{iter}}^{\text{Main}}$. 由图 14(b) 可见, 当迭代次数处于 100 ~ 1 500 区间时, 决定系数

表 12 对比实验统计结果

Table 12 Statistical results of the comparative experiments

方法	数据集	RMSE		MAE		R^2	
		均值	方差	均值	方差	均值	方差
BPNN	训练集	0.386 9	7.560e−06	0.293 9	1.600e−05	0.291 4	1.000e−04
	验证集	0.397 2	4.420e−06	0.298 6	2.840e−06	0.263 7	5.950e−05
	测试集	0.400 9	4.070e−05	0.302 3	3.090e−06	0.249 5	5.770e−04
LSTM	训练集	0.262 5	6.390e−05	0.183 7	4.770e−05	0.673 3	3.920e−04
	验证集	0.294 8	5.790e−04	0.212 3	4.670e−05	0.592 2	4.580e−03
	测试集	0.289 7	2.510e−05	0.205 8	1.840e−04	0.607 1	0.188e−03
RF	训练集	0.421 0	6.250e−06	0.321 4	2.590e−06	0.160 9	9.700e−05
	验证集	0.426 1	3.130e−06	0.322 7	1.340e−06	0.153 1	4.990e−05
	测试集	0.428 5	3.970e−06	0.324 2	2.740e−06	0.149 7	1.080e−04
FDF	训练集	0.467 6	2.850e−04	0.361 9	1.740e−04	−0.036 1	5.620e−03
	验证集	0.469 8	2.720e−04	0.361 3	1.720e−04	−0.030 6	5.250e−03
	测试集	0.473 2	2.650e−04	0.364 6	2.310e−04	−0.045 9	5.120e−03
Transformer	训练集	0.343 5	1.800e−07	0.258 8	8.000e−08	0.441 4	1.280e−06
	验证集	0.362 8	7.200e−06	0.269 3	2.650e−06	0.385 9	8.450e−05
	测试集	0.354 9	3.920e−06	0.269 2	8.000e−08	0.412 1	4.050e−05
本文方法	训练集	0.109 1	1.760e−05	0.057 5	8.940e−06	0.943 5	1.840e−05
	验证集	0.124 7	2.230e−05	0.061 1	1.062e−05	0.927 3	2.990e−05
	测试集	0.120 5	1.823e−05	0.059 9	8.930e−06	0.932 0	2.250e−05

表 13 消融实验结果

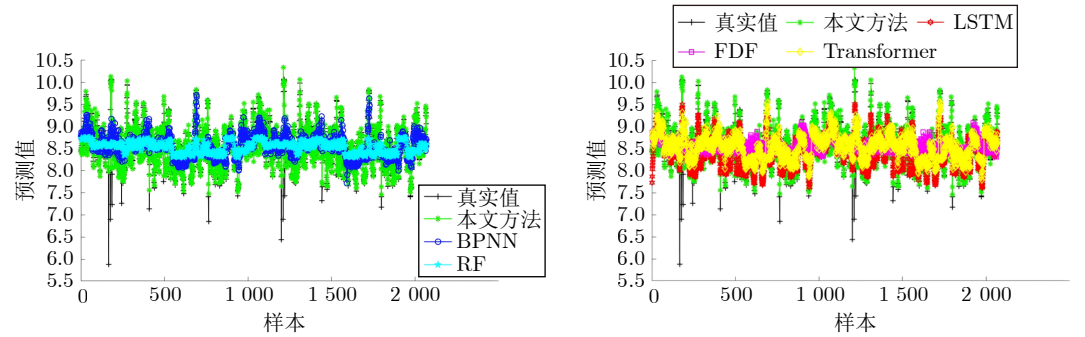
Table 13 Results of ablation experiments

方法	数据集	RMSE		MAE		R^2	
		均值	方差	均值	方差	均值	方差
无虚实驱动主补机制建模模块	训练集	0.199 7	1.420e−05	0.135 3	9.260e−06	0.811 1	5.010e−05
	验证集	0.264 8	8.160e−05	0.187 0	9.150e−05	0.672 4	5.110e−04
	测试集	0.275 4	1.850e−05	0.194 6	8.450e−05	0.645 8	1.220e−04
无真实数据驱动时序特性建模模块	训练集	0.385 3	3.830e−06	0.294 3	6.500e−06	0.297 0	5.100e−05
	验证集	0.408 1	2.380e−05	0.316 9	1.650e−05	0.223 0	3.440e−04
	测试集	0.407 6	4.440e−05	0.315 7	4.650e−05	0.225 9	5.970e−04
无真实输出动态融合模块	训练集	0.263 7	1.460e−05	0.196 7	4.700e−06	0.670 6	9.070e−05
	验证集	0.274 8	1.640e−05	0.202 6	1.090e−05	0.648 4	8.430e−05
	测试集	0.273 5	3.420e−05	0.202 6	2.520e−05	0.650 6	2.240e−04
本文方法	训练集	0.109 1	1.760e−05	0.057 5	8.940e−06	0.943 5	1.840e−05
	验证集	0.124 7	2.230e−05	0.061 1	1.062e−05	0.927 3	2.990e−05
	测试集	0.120 5	1.823e−05	0.059 9	8.930e−06	0.932 0	2.250e−05

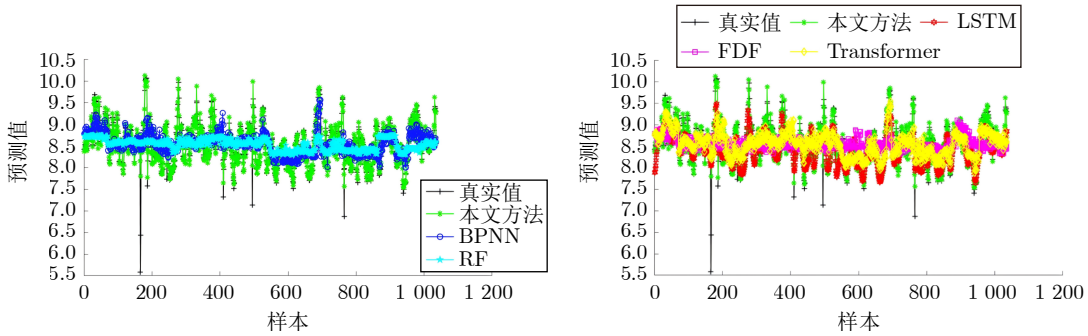
整体保持在 0.912 0 以上的较高水平, 表明模型在该超参数范围内具有较好的拟合稳定性. 其中, 迭代次数为 300 时, 决定系数达到局部峰值; 迭代次数为 700 时, 决定系数达到最低值. 同时, 伴随着该参数 $N_{\text{iter}}^{\text{Main}}$ 的增加, 曲线出现较大浮动, 说明超参数的增大并非单调提升模型拟合效果, 存在对拟合性能的非线性影响. 这一结果提示, 在超参数调优中, 可优先选择 300 次左右的迭代次数以平衡拟合效果与计算成本.

3) 主模型收敛容忍度 $N_{\text{tol}}^{\text{Main}}$. 由图 14(c) 可见, 在收敛容忍度 0.01 ~ 0.10 的调节范围内, 模型决定系数整体维持在 0.914 0 以上的较优区间, 说明该超参数对模型拟合效果的敏感性较低, 整体鲁棒性较强. 这一结果表明, 在该模型的超参数调优中, 收敛容忍度的选择无需过度精细化.

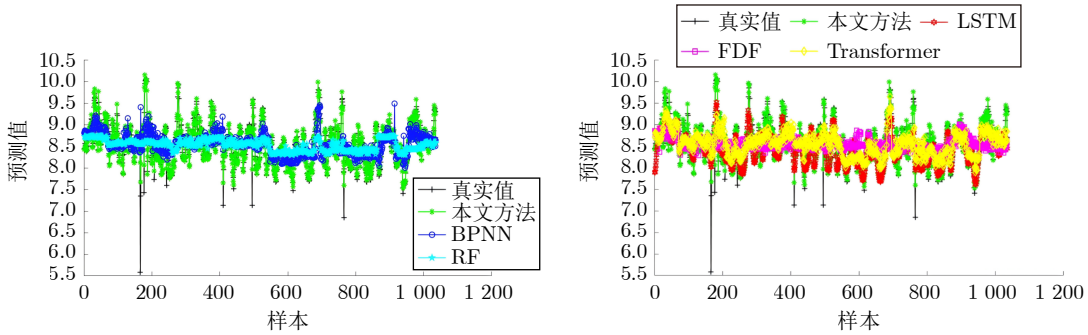
4) 主模型学习率 $N_{\text{learn}}^{\text{Main}}$. 由图 14(d) 可见, 在学



(a) 训练集
(a) Training set



(b) 验证集
(b) Validation set



(c) 测试集
(c) Test set

图 12　对比实验的拟合曲线

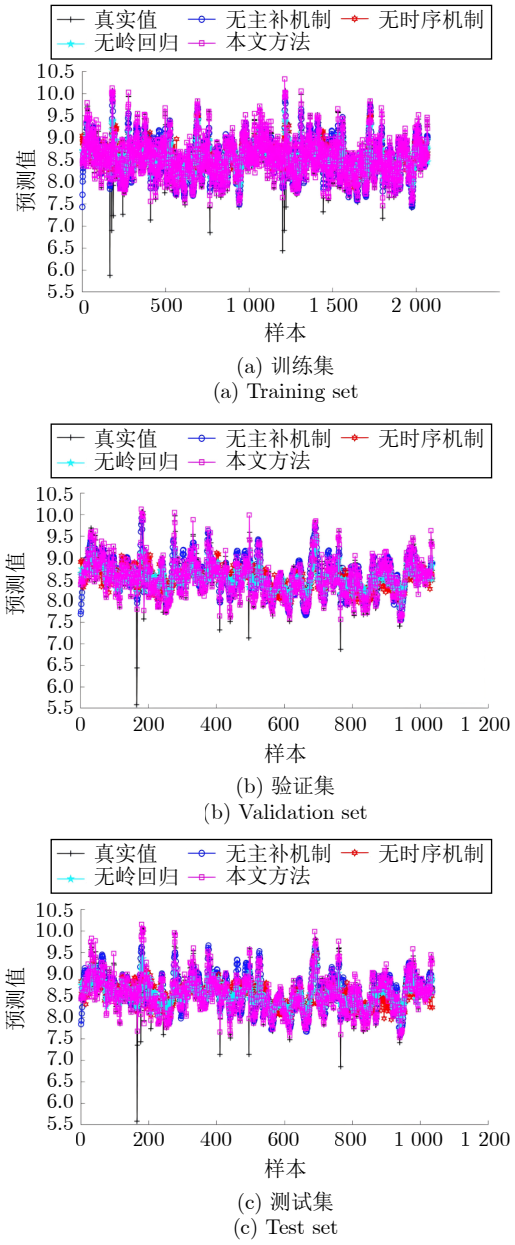
Fig. 12　Fitting curves of the comparative experiments

习率处于 0.010 ~ 0.100 的调节区间内, 决定系数整体处于 0.910 0 ~ 0.924 0 的范围. 但伴随着学习率的增大, 模型性能的波动幅度呈现出增大的趋势. 这一现象表明, 较大的学习率会导致拟合效果呈现出较大的波动. 因此, 可优先选取如 0.015 的学习率, 可以同时保证模型的高性能以及稳定性.

5) 补偿模型模糊规则数 $N_{\text{rule}}^{\text{Com}}$. 由图 14(e) 可见, 在规则数处于 10 ~ 100 的区间内, 决定系数整体处于 0.916 0 ~ 0.926 0 范围, 且随规则数增加呈现出非线性波动特征. 当规则数为 100 时, 决定系数达到峰值, 而在 50、80 等取值下则出现局部波动. 这表明模糊规则数的调整对模型拟合效果存在一定影响, 且规则数的增加并非持续提升性能, 但高规则数更易获得较优拟合结果. 因此, 在超参数调优中, 可优先选择 100 左右的模糊规则数以提升模型拟合能力.

6) 补偿模型子系统数 $N_{\text{IT3FS}}^{\text{Com}}$. 由图 14(f) 可见, 在子系统数处于 10 ~ 100 的区间内, 决定系数整体维持在 0.916 0 ~ 0.924 0 的较优范围, 未出现显著的性能滑坡. 这表明模型对该参数的敏感性较低, 子系统数在较宽区间内均可保障拟合效果, 无需过度精细化调优. 在实践中, 可结合计算资源成本, 在 10 ~ 100 的子系统数区间内灵活选择.

7) 补偿模型增强节点数 $N_{\text{enh}}^{\text{Com}}$. 由图 14(g) 可



(a) 训练集
(a) Training set

(b) 验证集
(b) Validation set

(c) 测试集
(c) Test set

图 13 消融实验的拟合曲线

Fig. 13 Fitted curves for ablation experiments

见, 在增强节点数处于 10 ~ 100 的区间内, 决定系数整体处于 0.912 0 ~ 0.924 0 的较优范围, 未出现明显的性能衰减, 但呈现出上升的趋势. 这表明模型对该参数有一定的正相关, 但仍需考虑模型计算成本, 灵活选择超参数.

8) 隐藏层单元数 $N_{\text{unit}}^{\text{LSTM}}$. 由图 14(h) 可见, 当单元数处于 10 ~ 30 区间时, 决定系数维持在 0.935 0 左右的较高水平; 随着单元数增加, 决定系数呈现出逐步下降的趋势, 取值 100 时降至 0.905 0 附近. 这表明所提模型对隐藏层单元数的敏感性较高, 单元数并非越多性能越优, 过多的单元数反而会导致拟合效果衰减. 因此, 本文所提模型的隐藏层单元数的最优区间 10 ~ 30, 此区间既能保障模型的效果, 又能避免过拟合与复杂度冗余的问题.

表 14 超参数分析区间设置

Table 14 Setting of hyperparameter analysis range

模型超参数	符号	范围
叶节点最小样本数	$N_{\text{samples}}^{\text{Main}}$	[100, 800]
主模型最大迭代次数	$N_{\text{iter}}^{\text{Main}}$	[100, 1500]
主模型收敛容忍度	$N_{\text{tol}}^{\text{Main}}$	[0.01, 0.10]
主模型学习率	$N_{\text{learn}}^{\text{Main}}$	[0.001, 0.100]
补偿模型模糊规则数	$N_{\text{rule}}^{\text{Com}}$	[10, 100]
补偿模型子系统数	$N_{\text{IT3FS}}^{\text{Com}}$	[10, 100]
补偿模型增强节点数	$N_{\text{enh}}^{\text{Com}}$	[10, 100]
隐藏层单元数	$N_{\text{unit}}^{\text{LSTM}}$	[10, 100]
训练轮次	$N_{\text{epo}}^{\text{LSTM}}$	[50, 500]
学习率	$N_{\text{learn}}^{\text{LSTM}}$	[0.001, 0.100]
Dropout 系数	$N_{\text{drop}}^{\text{LSTM}}$	[0.1, 1.0]

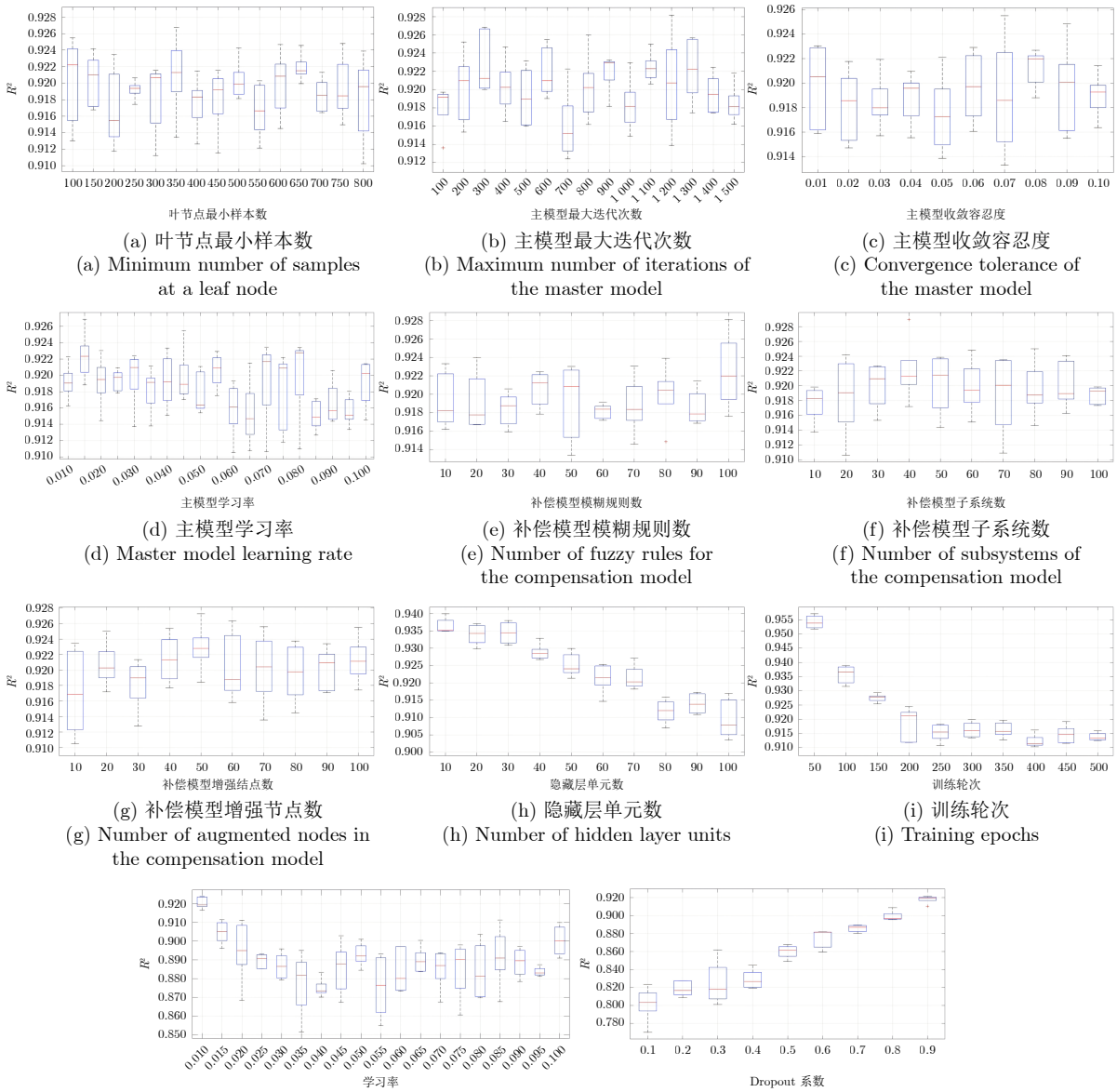
9) 训练轮次 $N_{\text{epo}}^{\text{LSTM}}$. 由图 14(i) 可见, 决定系数随轮次增加呈明显下降趋势, 50 轮时达到 0.955 0 的峰值, 500 轮时降至 0.915 0 左右. 这表明模型在训练初期快速学习了数据的有效模式, 但轮次的增加已开始出现拟合效果的衰减. 原因很可能是, 训练轮次过多, 模型开始学习噪声而非真实规律. 因此, 需严格控制训练轮次以平衡拟合效果与泛化能力.

10) 学习率 $N_{\text{learn}}^{\text{LSTM}}$. 由图 14(j) 可见, 当学习率为 0.010 时, 决定系数达到 0.920 0 的峰值, 是所有取值中拟合效果最优的. 随着学习率增至 0.035, 决定系数显著下滑至 0.875 0 左右, 这表明过大的学习率会导致参数更新步长失控, 模型难以稳定收敛. 当学习率超过 0.045 0 后, 决定系数虽有小幅回升, 但整体维持在 0.880 0 ~ 0.900 0 的较低水平且波动明显, 说明高学习率下模型收敛状态不稳定. 这一结果表明, 该模型对学习率的敏感性较高, 最优学习率为 0.010.

11) Dropout 系数 $N_{\text{drop}}^{\text{LSTM}}$. 由图 14(k) 可见, 当 Dropout 系数从 0.10 增至 0.90 时, 决定系数整体呈上升趋势, 其中系数 0.90 时达到 0.92 的峰值, 而系数为 0.10 时仅为 0.80 左右. 这表明该模型在训练过程中过拟合风险较高, 适当增大 Dropout 系数可有效提升模型的泛化能力, 进而提高决定系数. 因此, 本文所提模型的最优 Dropout 系数应选择 0.90 左右, 以平衡正则化效果与模型的特征学习能力.

5 结束语

针对 MSWI 过程 CO_2 排放浓度预测建模的局限, 本文提出虚实驱动新型混合建模方法, 经北京



(a) 叶节点最小样本数
(a) Minimum number of samples at a leaf node

(b) 主模型最大迭代次数
(b) Maximum number of iterations of the master model

(c) 主模型收敛容忍度
(c) Convergence tolerance of the master model

(d) 主模型学习率
(d) Master model learning rate

(e) 补偿模型模糊规则数
(e) Number of fuzzy rules for the compensation model

(f) 补偿模型子系统数
(f) Number of subsystems of the compensation model

(g) 补偿模型增强节点数
(g) Number of augmented nodes in the compensation model

(h) 隐藏层单元数
(h) Number of hidden layer units

(i) 训练轮次
(i) Training epochs

(j) 学习率
(j) Learning rate

(k) Dropout 系数
(k) Dropout coefficient

图 14　超参数敏感性分析曲线

Fig. 14　Sensitivity analysis curves of hyperparameter

某 MSWI 电厂实际数据验证了其有效性与合理性. 主要贡献如下: 1) 耦合 FLIC 与 Aspen Plus 构建 CO_2 全流程机理数值仿真模型, 结合双重正交实验设计生成多工况虚拟机理数据, 为混合建模提供基础; 2) 采用可解释树结构作为机理主模型以规避信息重叠与特征冗余, 引入 IT3FBLS 构建残差补偿模型以提升模型对复杂扰动与不确定性的适应能力; 3) 构建基于 LSTM 的真实数据驱动时序模型以挖掘 CO_2 排放动态演化规律; 4) 通过改进岭回归动态融合各模块输出以实现多信息源优势互补. 该方法为 MSWI 过程碳排放精准预测与优化控制提供了有力支撑, 也为复杂工业过程多源信息融合建模提供了可借鉴的一体化思路, 可推广应用于具有相似特征的复杂工业建模场景. 未来将进一步探索该模型在在线学习、多目标优化等方面的应用潜力.

附录 A 符号说明

表 A1　符号说明

Table A1　Symbolic description

符号	含义
ξ_{GE}	燃烧效率
γ_{CO}	CO 的浓度
γ_{CO_2}	CO_2 的浓度

表 A1 符号说明 (续表)
Table A1 Symbolic description (continued table)

符号	含义
ϕ	床层空隙率
V_g	气体速度
V_s	固体颗粒速度
V_B	移动边界的速度
S_{sg}	固体转化为气体的转化速率
S_s	质量源项
ρ_g	气体的密度
ρ_{sb}	多孔介质的密度
ρ_s	MSW 的固体密度
p_g	气体压力
$F(v)$	固体对流体流动的阻力
σ	床内的法向应力张量
τ	床内的切向应力张量
A	随机扰动
H_g	气体焓
H_s	固相焓
T_s	固体温度
T_g	气体温度
Q_h	气体的热损益
Q_{sh}	固相热源项
q_r	辐射热通量
S_a	颗粒表面积
$\dot{h}_s$	对流换热系数
λ_s	有效导热系数
λ_g	热分散系数
λ_0	有效热扩散系数
d_p	颗粒直径
C_{pg}	混合气体的比热容
R_{evp}	水分蒸发反应速率
H_{evp}	固体水分蒸发热
$\dot{h}$	固体和气体之间对流传质系数
$C_{w,\,s}$	固相中的水分浓度
$C_{w,\,g}$	气相中的水分浓度
Q_{cr}	对流换热和辐射换热固体吸收的热量
ξ_s	固体辐射率
σ_b	玻尔兹曼辐射常数
T_{env}	环境温度
R_v	挥发分反应速率
v_∞	挥发分的最终产率
v	挥发分在 t 时刻的释放量
R_g	通用气体常数
A_v	挥发分燃烧速率的指前因子

表 A1 符号说明 (续表)
Table A1 Symbolic description (continued table)

符号	含义
E_v	活化能
C_{mix}	经验常数
$C_{(\cdot)}$	气态反应物的质量分数
D_g	气体的质量扩散系数
R_C	焦炭燃烧的反应速率
S	反应中的化学计量系数
$P_{\mathrm{O_2}}$	氧的分压
k_r	化学动力学的速率常数
k_d	扩散引起的速率常数
$\mathrm{G_{SC}}$	固相燃烧区的出口烟气
$\mathrm{G_{GC}}$	气相燃烧区的出口烟气
$\mathrm{G_{HT}}$	换热区的出口烟气
$\mathrm{G_{CL}}$	烟气净化区的出口烟气
$\mathrm{G_{GE}}$	烟气排放区的出口烟气
x_{feed}	给料量
x_{speed}	炉排速度
x_{air1}	干燥段风量
x_{air2}	燃烧 1 段风量
x_{air3}	燃烧 2 段风量
x_{air4}	燃烬段风量
x_{particle}	粒径
x_{mix}	颗粒混合系数
x_{moisture}	水分含量
x_{other}	其他变量
$\boldsymbol{D}^{\mathrm{MD,\ CO_2}}$	虚拟机理数据
$\boldsymbol{X}^{\mathrm{MD}}$	虚拟机理数据特征
$\boldsymbol{y}^{\mathrm{MD}}$	虚拟机理数据目标值
$R_k^{\mathrm{Left}}(s)$	第 k 个节点上使用分割阈值 s 形成的左子树
$R_k^{\mathrm{Right}}(s)$	第 k 个节点上使用分割阈值 s 形成的右子树
N_k^{Left}	第 k 个节点上划分得到的左子树样本数量
N_k^{Right}	第 k 个节点上划分得到的右子树样本数量
$x_{k,\,j,\,m}^{\mathrm{Left}}$	第 k 个节点上左子树中第 j 个样本中的第 m 个特征
$x_{k,\,j,\,m}^{\mathrm{Right}}$	第 k 个节点上右子树中第 j 个样本中的第 m 个特征
$y_{k,\,j}^{\mathrm{Left}}$	第 k 个节点上左子树中第 j 个样本的目标值
$y_{k,\,j}^{\mathrm{Right}}$	第 k 个节点上右子树中第 j 个样本的目标值
$\Omega(s)$	分割阈值 s 划分后的损失函数值
μ_k^{Left}	第 k 个节点上左子树的目标值均值
μ_k^{Right}	第 k 个节点上右子树的目标值均值
R_k	子空间

表 A1　符号说明 (续表)
Table A1　Symbolic description (continued table)

符号	含义
N_k	第 k 个子空间内的总样本数
$\boldsymbol{w}_k^{(t)}$	第 t 次迭代下第 k 个子空间内样本权重
λ	正则项系数
$\boldsymbol{H}_k^{(t)}$	海森矩阵
$\nabla\mathcal{L}_k^{(t)}(\boldsymbol{w}_k^{(t)})$	目标函数梯度
$\Delta\boldsymbol{w}_k^{(t)}$	权重变化量
$\Delta\boldsymbol{G}_k^{(t)}$	梯度变化量
$N_{\text{tol}}^{\text{Main}}$	收敛容忍度
$I(\cdot)$	指示函数
$\boldsymbol{e}^{\text{PD, CO}_2}$	主模型残差
$\boldsymbol{X}^{\text{PD, CO}_2}$	真实过程数据集输入特征
N_{PD}	$\boldsymbol{X}^{\text{PD, CO}_2}$ 的样本数量
$\boldsymbol{x}_{n,\,M}^{\text{PD}}$	$x_n^{\text{PD, CO}_2}$ 的第 M 个特征
N_{IT3FS}	IT3FS 子系统数量
$\overline{\mu}_{\tilde{A}_m^l\|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$	第 m 个输入特征在水平切片等级为 $\overline{\alpha}_g$ 时模糊隶属函数 $\tilde{A}_m^l$ 的上界
$\overline{\mu}_{\tilde{A}_m^l\|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$	第 m 个输入特征在水平切片等级为 $\underline{\alpha}_g$ 时模糊隶属函数 $\tilde{A}_m^l$ 的上界
$\underline{\mu}_{\tilde{A}_m^l\|\mu_\delta=\overline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$	第 m 个输入特征在水平切片等级为 $\overline{\alpha}_g$ 时模糊隶属函数 $\tilde{A}_m^l$ 的下界
$\underline{\mu}_{\tilde{A}_m^l\|\mu_\delta=\underline{\alpha}_g}\left(x_{n,\,m}^{\text{PD}}\right)$	第 m 个输入特征在水平切片等级为 $\underline{\alpha}_g$ 时模糊隶属函数 $\tilde{A}_m^l$ 的下界
$\left[\underline{\alpha}_g,\,\overline{\alpha}_g\right]$	二级隶属度的不确定性
$C_{\tilde{A}_m^l}$	模糊隶属函数 $\tilde{A}_m^l$ 的中心
$\underline{\sigma}_{V_m}$	模糊隶属函数 $\tilde{A}_m^l$ 的下宽度
$\overline{\sigma}_{V_m}$	模糊隶属函数 $\tilde{A}_m^l$ 的上宽度
$\underline{w}_l$、$\overline{w}_l$	第 l 条规则的后件参数
$\boldsymbol{W}_n^{\text{Enhan}}$	IT3FS 层与增强层中第 n^{Enhan} 个增强节点连接权值
$\boldsymbol{\beta}_n^{\text{Enhan}}$	IT3FS 层与增强层中第 n^{Enhan} 个增强节点偏置
$\boldsymbol{W}_f$	IT3FS 层和增强层与输出层之间的连接权值
$\boldsymbol{A}^+$	广义逆矩阵
$\boldsymbol{E}$	模型真值
$\boldsymbol{I}$	单位矩阵
E_t	t 时刻的损失
$\hat{y}_{\text{Time},\,t}$	t 时刻的预测值
y_t^{PD}	t 时刻的真值
$f_{\min}(\cdot)$	最小值
$f_{\max}(\cdot)$	最大值
a	标准化下限
b	标准化上限
$\boldsymbol{f}_t$	t 时刻遗忘门的过程值

表 A1　符号说明 (续表)
Table A1　Symbolic description (continued table)

符号	含义
$\boldsymbol{g}_t$	t 时刻遗忘门的输出值
σ	sigmoid 函数
$\boldsymbol{w}^{fx}$	遗忘门对应的权重
$\boldsymbol{W}^{fh}$	遗忘门对应的权重
$\boldsymbol{b}^f$	遗忘门对应的偏置
$\odot$	哈达玛积
$\boldsymbol{i}_t$	t 时刻输入门的过程值
$\hat{\boldsymbol{c}}_t$	t 时刻单元状态更新值
Tanh	Tanh 函数
$\boldsymbol{w}^{ix}$	输入门对应的权重
$\boldsymbol{W}^{ih}$	输入门对应的权重
$\boldsymbol{b}^i$	输入门对应的偏置
$\boldsymbol{w}^{cx}$	单元更新状态对应的权重
$\boldsymbol{W}^{ch}$	单元更新状态对应的权重
$\boldsymbol{b}^c$	单元更新状态对应的偏置
$\boldsymbol{o}_t$	t 时刻输出门的过程值
$\boldsymbol{w}^{ox}$	输出门对应的权重
$\boldsymbol{W}^{oh}$	输出门对应的权重
$\boldsymbol{b}^o$	输出门对应的偏置
$\boldsymbol{v}$	输出层的权重
b_v	输出层的偏置
$\eta^{(t)}$	动态更新学习率
$e^{(t)}$	预测误差
$N_{\text{learn}}^{\text{Main}}$	主模型学习率
$N_{\text{rule}}^{\text{Com}}$	补偿模型模糊规则数
$N_{\text{enh}}^{\text{Com}}$	补偿模型增强节点数
$N_{\text{unit}}^{\text{LSTM}}$	LSTM 隐藏层单元数
$N_{\text{epo}}^{\text{LSTM}}$	LSTM 训练轮次
$N_{\text{learn}}^{\text{LSTM}}$	LSTM 学习率
$N_{\text{drop}}^{\text{LSTM}}$	LSTM 的 Dropout 参数
$\hat{\boldsymbol{y}}$	模型预测值
M_{LHV}	MSW 低位热值
M_C	MSW 中 C 的质量分数
M_H	MSW 中 H 的质量分数
M_O	MSW 中 O 的质量分数
$\bar{V}_{\text{Process}}^{\text{Ori}}$	炉排速度均值
$V_{\text{Process}}^{\text{Ori}}$	炉排速度样本实际值
$V_{\text{Process}}^{\text{Trans}}$	单位统一转换后的炉排速度
C_Q	CO_2 的质量浓度
m_{gas}	CO_2 的质量流量
V_{total}	混合烟气的总体积流量
C_V	CO_2 的体积分数

附录 B 缩写词及其含义

表 B1 缩写词及其含义

Table B1 Abbreviations and their meanings

缩写词	中文全称	英文全称
MSWI	城市固废焚烧	municipal solid waste incineration
IT3FBLS	区间三型模糊宽度学习系统	interval type-3 fuzzy broad learning system
LSTM	长短期记忆	long short-term memory
MSW	城市固废	municipal solid waste
CO_2	二氧化碳	carbon dioxide
DD	数据驱动	data-driven
XGBoost	极限梯度提升	extreme gradient boosting
ARIMA	自回归积分滑动平均	autoregressive integrated moving average
FLIC	床层固废燃烧模拟软件	fluid dynamic incinerator code
Aspen Plus	过程工程先进系统	advanced system for process engineering plus
Catboost	分类提升	categorical boosting
MD	机理驱动	mechanism-driven
DXN	二噁英	dioxin
CE	燃烧效率	combustion efficiency
GS	炉排速度	grate speed
FC	给料量	feeding capacity
PA	一次风量	primary airflow
LHV	低位热值	lower heat value
BPNN	反向传播神经网络	back propagation neural network
RF	随机森林	random forest
FDF	模糊决策森林	fuzzy decision forest
GBDT	梯度提升决策树	gradient boosting decision tree

参考文献

1 Tang Jian, Wang Tian-Zheng, Sun Ming-Jun, Pi Meng, Qian Yu, Jiang He. Review of AI technology research on end-edge-cloud collaborative safety empowerment for municipal solid waste incineration. *Control Engineering of China*, DOI: 10.14107/j.cnki.kzgc.20250600
(汤健, 王天峥, 孙明俊, 皮猛, 钱雨, 姜河. 城市固废焚烧过程 AI 安全赋能智慧化发展与展望. 控制工程, DOI: 10.14107/j.cnki.kzgc.20250600)

2 Kammen D M, Sunter D A. City-integrated renewable energy for urban sustainability. *Science*, 2016, **352**(6288): 922−928

3 Tang Jian, Xia Heng, Yu Wen, Qiao Jun-Fei. Research status and prospects of intelligent optimization control for municipal solid waste incineration process. *Acta Automatica Sinica*, 2023, **49**(10): 2019−2059
(汤健, 夏恒, 余文, 乔俊飞. 城市固废焚烧过程智能优化控制研究现状与展望. 自动化学报, 2023, **49**(10): 2019−2059)

4 Xia H, Tang J, Aljerf L, Cui C L, Gao B Y, Ukaogo P O. Dioxin emission modeling using feature selection and simplified DFR with residual error fitting for the grate-based MSWI process. *Waste Management*, 2023, **168**: 256−271

5 Gao X D, Zhang K, Ji D Q, Ma H M, Meng Y, Jiang S, et al. A proposed layout of CO_2 capture utilization and storage for coal fired power plants in China. *Fuel*, 2025, **401**: Article No. 135723

6 Cao Y P, Mirzaliev S, Hamza F, Gulzar F, Ming J. Financing sustainable energy: Cost-benefit analysis of CO_2 reduction in China. *Energy Strategy Reviews*, 2025, **58**: Article No. 101667

7 Chai Tian-You, Qian Feng, Guan Xiao-Hong, Ding Jin-Liang, Du Wei, Xu Zhan-Bo, et al. Theory and technology of automation and intelligence for dual carbon target. *Bulletin of National Natural Science Foundation of China*, 2024, **38**(4): 560−570
(柴天佑, 钱锋, 管晓宏, 丁进良, 堵威, 徐占伯, 等. 面向双碳目标的自动化和智能化理论与技术. 中国科学基金, 2024, **38**(4): 560−570)

8 Taheri S, Razban A. Learning-based CO_2 concentration prediction: Application to indoor air quality control using demand-controlled ventilation. *Building and Environment*, 2021, **205**: Article No. 108164

9 Pan S Y, Shi X D, Dong B B, Skvaril J, Zhang H R, Liang Y T, et al. Multivariate time series prediction for CO_2 concentration and flowrate of flue gas from biomass-fired power plants. *Fuel*, 2024, **359**: Article No. 130344

10 Kumar S. A novel hybrid machine learning model for prediction of CO_2 using socio-economic and energy attributes for climate change monitoring and mitigation policies. *Ecological Informatics*, 2023, **77**: Article No. 102253

11 Zhang Y T, Pakrashi A, Dev S. Multi-regional CO_2 emission forecasting using advanced machine learning and temporal feature engineering. *Environmental Modelling & Software*, 2026, **195**: Article No. 106728

12 Ge Y D, Zhang R, Wang Z, Yan B B, Tao J Y, Song Y J, et al. Optimization of Gompertz model with machine learning towards applicable and accurate simulation of anaerobic digestion. *Energy*, 2025, **333**: Article No. 137395

13 Yan M, Hui S C, Jiang N, Li N. A review on data-driven prognostics and health management for wind turbine systems. *Engineering Applications of Artificial Intelligence*, 2025, **159**: Article No. 111484

14 Zhou G L, Mao L N, Bao T W, Zhuang F P. Machine learning-driven CO_2 emission forecasting for light-duty vehicles in China. *Transportation Research Part D: Transport and Environment*, 2024, **137**: Article No. 104502

15 Wen T X, Liu Y Z, Bai Y H, Liu H Y. Modeling and forecasting CO_2 emissions in China and its regions using a novel AR-

IMA-LSTM model. *Heliyon*, 2023, **9**(11): Article No. e21241

16 Li Y, Shi K, Zhu M Y, Li H Y, Guo Y L, Miao S, et al. Data-driven models for forecasting algal biomass in a large and deep reservoir. *Water Research*, 2025, **270**: Article No. 122832

17 Wyper P F, Antiochos S K, DeVore C R. A universal model for solar eruptions. *Nature*, 2017, **544**(7651): 452−455

18 Yang Chun-Hua, Sun Bei, Li Yong-Gang, Huang Ke-Ke, Gui Wei-Hua. Cooperative optimization and intelligent control of complex production processes. *Acta Automatica Sinica*, 2023, **49**(3): 528−539
(阳春华, 孙备, 李勇刚, 黄科科, 桂卫华. 复杂生产流程协同优化与智能控制. 自动化学报, 2023, **49**(3): 528−539)

19 Dang H, Xu R S, Zhang J L, Wang M Y, Zhang J Y. Interaction mechanism and kinetic modeling of anthracite and power plant biomass waste during CO_2 co-gasification process. *Energy*, 2025, **325**: Article No. 136115

20 Shafeeyan M S, Daud W M A W, Shamiri A. A review of mathematical modeling of fixed-bed columns for carbon dioxide adsorption. *Chemical Engineering Research and Design*, 2014, **92**(5): 961−988

21 Wang A, Rish S K, Jenkins D R, Tahmasebi A. Mechanism of metallurgical coke gasification in CO_2 and H_2O using a reaction-diffusion model. *Chemical Engineering Journal*, 2025, **512**: Article No. 162471

22 Cen X, Hong B Y, Yang Y X, Wu Z, Zhao Z H, Yang J. A high-precision and interpretable prediction method of the gas-solid erosion rate in elbows based on hybrid mechanism-data-driven models. *Measurement*, 2025, **251**: Article No. 117251

23 Zhang Y A, Yan G W, Xiao S Y, Wang F, Zhao G J, Ma S X. Mechanism- and data-driven based dynamic hybrid modeling for multi-condition processes. *Chemometrics and Intelligent Laboratory Systems*, 2025, **260**: Article No. 105353

24 Shen W H, Hu X J, Liu J L, Li S J, Wang H D. A pre-trained multi-step prediction informer for ship motion prediction with a mechanism-data dual-driven framework. *Engineering Applications of Artificial Intelligence*, 2025, **139**: Article No. 109523

25 Liang M Y, Zheng Y, Li S Y. Towards hybrid modeling with mechanistic and real-time data embed iterative co-optimization for industrial processes. *Journal of Process Control*, 2025, **155**: Article No. 103567

26 Wang X L, Cao J F, Cao Y. A hybrid mechanism modeling and data-driven method for energy prediction and optimization for a class of industrial robots. *Measurement*, 2026, **258**: Article No. 119227

27 Yan Z Q, Chu Z J, Hui J Z, Donald H, Lv J X, Xu Z G. Prediction, assessment and optimization of energy consumption in selective laser melting using hybrid mechanism-based and data-driven approaches. *Measurement*, 2026, **257**: Article No. 118895

28 Zhu F L, Zhu O, Han M Y, Liu W F, Guo X N, Hou T T, et al. A hybrid process-data driven framework for real-time hydrological forecasting with interpretable deep learning. *Journal of Hydrology*, 2025, **662**: Article No. 134082

29 Pu J C, Chun W D, Yang H. Low carbon strategy and coordination for a two-echelon automotive supply chain considering dual credit policy. *Computers & Industrial Engineering*, 2024, **192**: Article No. 110229

30 Liu Y, Wang J L. Spatiotemporal patterns and drivers of carbon emissions from municipal solid waste treatment in China. *Waste Management*, 2023, **168**: 1−13

31 Zhang S R, Jiang X G, Lv G J, Liu B X, Jin Y Q, Yan J H. SO_2, NO_x, HF, HCl and PCDD/Fs emissions during co-combustion of bituminous coal and pickling sludge in a drop tube furnace. *Fuel*, 2016, **186**: 91−99

32 Chen J K, Tang J, Xia H, Yu W, Qiao J F. Modelling the furnace temperature field of a municipal solid waste incinerator using the numerical simulation and the deep forest regression algorithm. *Fuel*, 2023, **347**: Article No. 128511

33 Wehrung Q, Pastero L, Bernasconi D, Cotellucci A, Bruno M, Cavagna S, et al. Impact of operational parameters on the CO_2 absorption rate in $Ca(OH)_2$ aqueous carbonation-implications for process efficiency. *Energy & Fuels*, 2024, **38**(17): 16678−16691

34 Hosokai S, Matsuoka K, Kuramoto K, Suzuki Y. Modification of Dulong′s formula to estimate heating value of gas, liquid and solid fuels. *Fuel Processing Technology*, 2016, **152**: 399−405

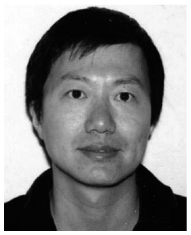
35 China Association of Environmental Protection Industry. Specifications of continuous emission monitoring system for CO_2 in flue gas emitted from stationary sources, T/CAEPI 47-2022, 2022.
(中国环境保护产业协会. 固定污染源二氧化碳排放连续监测系统技术要求, T/CAEPI 47-2022, 2022.)



汤 健 北京工业大学信息科学技术学院教授. 主要研究方向为小样本数据建模, 城市固废处理过程智能控制. 本文通信作者.
E-mail: freeflytang@bjut.edu.cn
(**TANG Jian** Professor at the School of Information Science and Technology, Beijing University of Technology. His research interests include small sample data modeling and intelligent control of municipal solid waste treatment process. Corresponding author of this paper.)



王 子 北京工业大学信息科学技术学院硕士研究生. 主要研究方向为城市固废焚烧过程碳排放预测建模.
E-mail: wangzibjut@163.com
(**Wang Zi** Master student at the School of Information Science and Technology, Beijing University of Technology. His main research interest is carbon emission prediction modeling of municipal solid waste incineration process.)



余 文 墨西哥国立理工大学高级研究中心 (CINVESTAV-IPN) 教授. 主要研究方向为复杂工业过程建模与控制, 机器学习.
E-mail: yuw@ctrl.cinvestav.mx
(**YU Wen** Professor at the Centro de Investigacion de Estudios Avanzados (CINVESTAV-IPN), National Polytechnic Institute México. His research interests include modeling and control of complex industrial process, and machine learning.)



乔俊飞 北京工业大学信息科学技术学院教授. 主要研究方向为污水处理过程智能控制, 神经网络结构设计与优化.
E-mail: junfeiq@bjut.edu.cn
(**QIAO Jun-Fei** Professor at the School of Information Science and Technology, Beijing University of Technology. His research interests include intelligent control of waste water treatment process, structure design and optimization of neural networks.)