DOI: 10.19666/j.rlfd.202512048

基于堆叠集成学习的微混燃烧室 NO_x 排放预测

李松波[1,2]，刘 月[1,2,3]，张哲巅[1,2]，路 源[1,2]，徐 祥[1,2]
（1. 中国科学院工程热物理研究所能源动力研究中心，北京 100190；
2. 中国科学院大学工程科学学院，北京 100049；
3. 江苏中科能源动力研究中心，江苏 连云港 222069）

［摘 要］【目的】氮氧化物（NO_x）排放量是燃气轮机燃烧室的重要环保指标，NO_x 快速精准预测有助于燃烧优化与低排放控制。针对传统基于数值模拟进行预测时计算成本高、数据驱动单一模型泛化能力有限的问题，提出 NO_x 排放预测的基于异质基学习器与 TabNet 元学习器的堆叠集成学习框架。【方法】该框架采用 XGBoost、支持向量机（support vector machine，SVM）、随机森林（random forest，RF）和多层感知机（multi-layer perceptron，MLP）4 种异质基学习器，从不同学习范式生成初级预测；继而引入专为表格数据设计的 TabNet 模型作为元学习器，利用其顺序注意力机制动态融合基学习器输出。【结果】在经预处理的微混燃烧室实验数据集上，所提堆叠集成模型的预测性能优于所有单一基学习器，其均方根误差和平均绝对误差分别低至 0.784 0 mg/m³和 0.173 0 mg/m³。通过动态训练−稳态测试的跨工况验证，模型表现出良好的泛化能力。基于 SHAP 的可解释性分析表明，一、二级燃料流量是影响 NO_x 预测的最关键特征。【结论】所提出的堆叠集成学习框架能有效提升燃烧室 NO_x 排放的预测精度与泛化能力，其关键特征识别结果与燃料分级燃烧的物理机理一致。

［关 键 词］微混燃烧室；NO_x 排放预测；堆叠集成学习；SHAP 分析

［引用本文格式］李松波，刘月，张哲巅，等. 基于堆叠集成学习的微混燃烧室 NO_x 排放预测[J]. 热力发电，2026，55(7)：149.
Li S B, Liu Y, Zhang Z D, et al. Prediction of NO_x emissions in a micro-mixing combustor based on stacking ensemble learning[J]. Thermal Power Generation, 2026, 55(7): 149.

Prediction of NO_x emissions in a micro-mixing combustor based on stacking ensemble learning

Li Songbo[1,2], Liu Yue[1,2,3], Zhang Zhedian[1,2], Lu Yuan[1,2], Xu Xiang[1,2]
(1. Research Center for Energy and Power, Institute of Engineering Thermophysics, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Engineering Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;
3. Jiangsu Zhongke Research Center for Clean Energy and Power, Lianyungang 222069, China)

Abstract: [Objective] Nitrogen oxide (NO_x) emissions are a crucial environmental indicator for gas turbine combustors, and their accurate and rapid prediction plays a vital role in combustion optimization and low-emission control. To address the high computational cost of traditional numerical simulations and the limited generalization capability of single data-driven models, as well as the respective shortcomings of neural networks and ensemble learning methods in tabular data processing and small-sample generalization found in existing research, this study proposes a stacked ensemble learning framework based on heterogeneous base learners and a TabNet meta-learner, aiming to integrate the advantages of multiple models to enhance prediction accuracy and cross-condition adaptability. [Methods] The framework employs four heterogeneous base learners—XGBoost, Support Vector Machine (SVM), Random Forest (RF), and Multi-Layer Perceptron (MLP)—which represent four distinct learning

收 稿 日 期：2025-12-17 修 回 日 期：2026-02-08 接 受 日 期：2026-03-04
基 金 项 目：连云港市科技项目（RQ2504）
Supported by: Science and Technology Project of Lianyungang (RQ2504)
第一作者简介：李松波（2001），男，硕士研究生，主要研究方向为燃气轮机智能算法建模，lisongbo@iet.cn。
通信作者简介：刘月（1984），男，博士，高级工程师，主要研究方向为航机/燃机控制系统设计建模，liuyue16@iet.cn。

paradigms, namely Boosting, kernel methods, Bagging, and neural networks, to generate initial predictions from perspectives of different bias-variance trade-offs and feature representations. Subsequently, TabNet, a model specifically designed for tabular data, is introduced as the meta-learner to leverage its sequential attention mechanism for instance-wise feature selection, dynamically integrating the base learner outputs**.** During the model training process, a Bayesian optimization strategy is adopted for efficient hyperparameter tuning of the key base learners, utilizing the root mean square error from 5-fold cross-validation as the objective function to avoid local optima. [Results] Evaluated on a preprocessed experimental dataset from a micro-mixing combustor, the proposed stacked ensemble model demonstrates significantly better prediction performance than all individual base learners, achieving a root mean square error (RMSE) and mean absolute error (MAE) as low as 0.784 0 mg/m ³and 0.173 0 mg/m ³, respectively. Cross-condition validation, using dynamic data for training and steady-state data for testing, shows that the model exhibits good generalization capability, with an RMSE of 3.647 1 mg/m ³and an MAE of 1.426 2 mg/m ³, indicating strong generalization performance. Furthermore, interpretability analysis based on SHapley Additive exPlanations (SHAP) indicates that the primary and secondary fuel flow rates are the most critical features affecting NO_x prediction. This finding aligns closely with the physical mechanism wherein the temperature in the primary combustion zone, influenced by staged fuel combustion, governs thermal NO_x formation, thereby enhancing the physical plausibility of the model predictions. [Conclusion] The proposed stacked ensemble learning framework can effectively improve the prediction accuracy and generalization capability for NO_x emissions from combustors. Its architecture, which integrates heterogeneous learners with an attention-based meta-learner, provides a high-precision and interpretable data-driven solution for combustion emission prediction. The identification of key features is consistent with the physical mechanism of staged fuel combustion, demonstrating promising potential for engineering applications.

Key words: micro-mixing combustor; NO_x emission prediction; stacking ensemble learning; SHAP analysis

氮氧化物（NO_x）是典型的大气污染物，可导致酸雨、臭氧层耗竭并引发呼吸道疾病[1-2]。在能源结构低碳化转型的背景下，高效率、清洁的大功率燃气轮机成为发展重点[3-5]。然而，燃烧室高温环境引发 NO_x 生成问题制约了燃气轮机的环保性能。传统基于详细化学反应的数值模拟计算成本高昂，而离线实验测量周期长、成本高，二者均难以满足燃烧室的状态监测与优化需求。因此，开发基于关键操作参数快速、准确预测 NO_x 排放的数据驱动模型，对于控制燃气轮机污染物排放具有重要意义[6-7]。

机器学习相较于传统建模具有独特优势，其无需依赖复杂化学反应机理即可建立高精度预测模型，同时具备良好的实时响应特性，能够有效满足燃烧过程在线监测的时效性需求。当前基于机器学习的 NO_x 排放预测研究主要采用神经网络模型与梯度提升决策树（gradient boosting decision tree，GBDT）等集成算法。在神经网络方面，Su 等人[8]提出一种结合 DBSCAN（density-based spatial clustering of applications with noise）、Savitzky-Golay 滤波器（Savitzky-Golay filter）、多层残差网络与改进 Informer 的 SGM-Resinformer 模型，用于预测燃气轮机中 CO 与 NO_x 的混合排放。Yousif 等人[9]提出了一种结合前馈神经网络与采用 k 最近邻算法（k-nearest neighbors，KNN）优化随机数策略的粒子群优化算法的混合模型，通过邻域成分分析选择特征，预测燃气轮机的 NO_x 排放。Li 等人[10]提出了一种基于改进 Transformer 编码器的轻量级增强模型，通过引入因果卷积模块和正则化策略，实现了对燃气轮机 NO_x 排放的快速预测。姚顺春等[11]结合最大信息系数（maximal information coefficient，MIC）、主成分分析（principal component analysis，PCA）和长短时记忆（long short-term memory，LSTM）神经网络，提出了混合模型，用于预测垃圾焚烧炉选择性催化还原（selective catalytic reduction，SCR）脱硝系统出口的 NO_x 排放。张鹏新等[12]结合鲸鱼优化算法（whale optimization algorithm，WOA）和 LSTM 神经网络，提出了 WOA-LSTM 混合模型，用于预测循环流化床机组深度调峰运行过程中的 NO_x 排放。王勇权等[13]提出了一种结合卷积神经网络、门控循环单元和多头注意力机制的混合模型（CNN-GRU-MHA），实现了循环流化床机组在稳态与变负荷工况下 NO_x 排放的动态预测。在集成算法方面，Hoque 等人[14]通过对比多种传统与集成机器学习算法在燃气轮机 NO_x 排放预测中的性能，发现集成方法尤其是 CatBoost（categorical boosting）在预测精度上显著优于传统方法。Naghibi A[15]比较了 bagging 与 boosting 2 种集成学习方法在工业燃气轮机长期性能预测中的应用，结合 PCA 有效解决了多重共线性问题。Coelho 等人[16]提出了一种基于特征工程与深度森林回归器相结合的建模流程，用于预测燃气轮机中 CO 和 NO_x 的排放。

现有研究多采用神经网络或集成学习等单一模型预测 NO_x 排放。然而，这些方法各有局限：深度神经网络结构复杂、训练成本高，对表格数据特征交互学习效率不高，易在小样本下过拟合[17-18]；另一方面，梯度提升决策树等集成方法在表格数据中表现稳健，但其加法模型结构在拟合复杂非线性模式时表征能力有限。

为了克服单一模型的不足，发挥不同模型的互补优势，本文提出一种堆叠集成模型框架，选取极端梯度提升[19]（extreme gradient boosting, XGBoost）、支持向量机（support vector machine, SVM）[20]以及多层感知机（multi-layer perceptron, MLP）[21]与随机森林（random forest, RF）[22]，分别从 Boosting、核方法、神经网络和 Bagging 等不同机制建模。进一步引入专为表格数据设计的 TabNet 作为元学习器，构建两阶段预测架构，以提升对微混燃烧室 NO_x 排放的预测精度，并通过可解释性分析增强结果在物理机理上的可信度。

1 堆叠集成学习框架

本文提出的 NO_x 排放预测堆叠集成学习框架如图 1 所示，采用两级级联结构。第 1 层由 4 个异质的基学习器（XGBoost、SVM、随机森林、MLP）组成，基于不同的机器学习算法对原始输入特征进行学习并生成初级预测。第 2 层为元学习器，采用专为表格数据设计的 TabNet 模型，它将第 1 层输出的初级预测结果拼接成一个新的特征矩阵作为输入，通过其内在的顺序注意力机制，动态学习各基学习器预测之间的非线性交互关系，并输出最终的 NO_x 排放预测值。

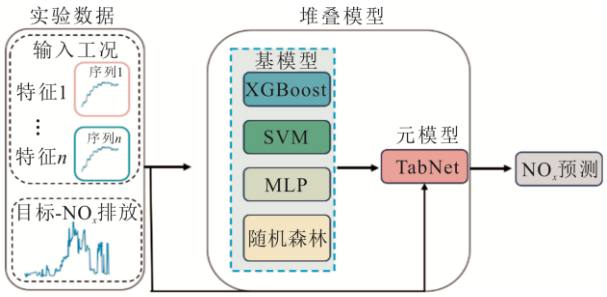


图 1 堆叠集成学习框架

Fig.1 Stacking ensemble learning framework

1.1 基学习器介绍

为提升基学习器多样性，本文选取了 XGBoost、SVM、RF 与 MLP 4 种原理不同的代表性算法。它们分别属于 Boosting、核方法、Bagging 和神经网络范式。Boosting 通过序列化建模降低预测偏差，Bagging 通过并行化聚合控制模型方差，核方法基于结构风险最小化原则保障泛化稳健性，神经网络则具备非线性拟合能力。这些模型在偏差-方差特性、风险控制和特征表达上各有侧重，能为元学习器提供互补的预测信息，从而提升最终预测模型的稳健性与准确性。

XGBoost 作为 Boosting 类型的代表，其模型架构如图 2 所示。具体而言，对于给定的数据集，$D=\{\boldsymbol{x}_i, y_i\}(i=1, 2, \ldots, n, \boldsymbol{x}_i\in \mathrm{R}^m, y_i\in \mathrm{R})$，$\boldsymbol{x}_i$ 为 m 维的特征向量，y_i 为目标值。对于预测值 $\hat{y}_i$ 有：

$$\hat{y}_i=\sum_{k=1}^{M}\phi_k(\boldsymbol{x}_i),\ \phi_k\in\Phi \tag{1}$$

式中：Φ 为回归树的空间，$\Phi=\{\Phi(x)=\omega_{\zeta(x)}\}$。这里 ζ 表示每棵树的结构，它将一个样本映射到对应的叶节点索引。每个 ϕ_k 对应一个独立的树结构 ζ 和叶子权重 ω。

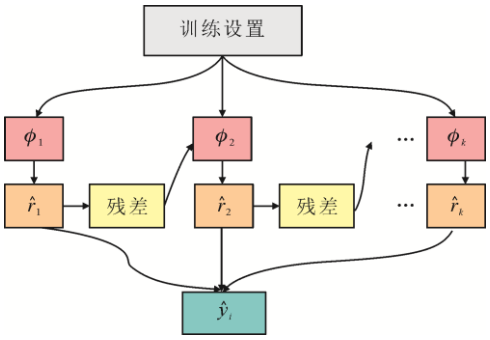


图 2 XGBoost 模型结构示意

Fig.2 Schematic diagram of the XGBoost model structure

为了学习模型中使用的函数集合，需要最小化以下损失函数:

$$L_t=\sum_{i=1}^{n}\left[p_i\phi_t(x_i)+\frac{1}{2}q_i\phi_t^2(\boldsymbol{x}_i)\right]+\Omega(\phi_t) \tag{2}$$

式中：$p_i=\partial_{\hat{y}^{t-1}}l(y_i,\hat{y}^{t-1})$，$q_i=\partial^2_{\hat{y}^{t-1}}l(y_i,\hat{y}^{t-1})$ 分别为损失函数的一阶和二阶梯度，$\Omega(\phi_t)=\alpha U+\frac{1}{2}\beta\sum_{u=1}^{U}\omega_u^2$ 为正则化项。

定义 $j_u=\{i|\zeta(\boldsymbol{x}_i=u)\}$ 为叶子 u 的集合，则最优叶节点权重 ω_u^* 为:

$$\omega_u^*=-\frac{\sum_{i\in j_u}p_i}{\sum_{i\in j_u}q_i+\beta} \tag{3}$$

支持向量机作为核方法的代表，其核心思想是基于结构风险最小化原则，在高维特征空间中寻找

最优回归超平面。对于给定的数据集，SVM 的优化目标可表述为最小化以下损失函数：

$$\min_{w,b}\frac{1}{2}\|\boldsymbol{w}\|^2+C\sum_{i=1}^{n}\xi_i \tag{4}$$

其约束条件为 $y_i-\langle \boldsymbol{w},\Phi_1\rangle-b<\varepsilon+\xi_i$。式中：$\Phi_i$ 为通过核函数实现的非线性映射；ε 为不敏感损失参数；$C>0$ 为惩罚系数；ξ_i 为松弛变量；$\boldsymbol{w}$ 为超平面向量。该模型通过最大化间隔并控制预测误差，在保证拟合精度的同时提升泛化能力。

随机森林采用 Bagging 集成策略，通过 Bootstrap 抽样生成多个训练子集并独立构建决策树，同时在节点分裂时引入特征随机选择机制。这种双重随机化过程有效降低了模型的方差，提升了预测的稳定性和泛化能力。

多层感知机以前馈神经网络架构为基础，通过多层非线性变换建立从输入到输出的复杂映射关系。对于一个包含 L 个隐含层的 MLP，令第 l 层（l=1, …, L）的权重矩阵和偏置向量分别为 $\boldsymbol{W}_l$ 和 $\boldsymbol{b}_l$，并设激活函数为 $\Phi(\cdot)$。则该网络第 l 层的输出 h_l 为：

$$h_l=\sigma(\boldsymbol{W}_l h_{l-1}+\boldsymbol{b}_l) \tag{5}$$

模型通过反向传播算法最小化均方误差损失函数：

$$L=\frac{1}{n}\sum_{i=1}^{n}\|y_i-h_l(\boldsymbol{x}_i)\|^2 \tag{6}$$

逐层更新网络参数 $\{\boldsymbol{W}_l,\boldsymbol{b}_l\}_{l=1}^{L}$，使 MLP 能够自动学习数据中的层次化特征表示。

1.2 TabNet 模型介绍

本文选用 TabNet 作为元学习器，因其专为表格数据设计，能有效捕捉基学习器输出间的复杂非线性关系[23]。其核心的顺序注意力机制可实现实例级特征选择，即针对每一个输入样本，动态地筛选出最相关的一部分基学习器预测结果进行后续处理，而非平等对待所有特征。

TabNet 结构如图 3 所示，TabNet 采用多步的决策过程实现上述功能。在每一步中，模型学习一个稀疏的注意力掩码，该掩码为每个样本选择当前步骤最关键的少量特征。为鼓励多特征探索，引入特征使用先验以惩罚重复选择。被选中的特征子集随后送入一个共享-专用的特征变换器中进行非线性处理。每一步的输出分为两部分：一部分直接贡献于该步的决策，另一部分作为状态信息传递至下一步，以指导后续的特征选择。最终，所有步骤的决策输出被聚合，形成元学习器的最终预测。

这种方式使 TabNet 能够学习基学习器预测之间复杂的非线性交互关系，并根据样本动态调整整合策略，从而更好利用异质信息。

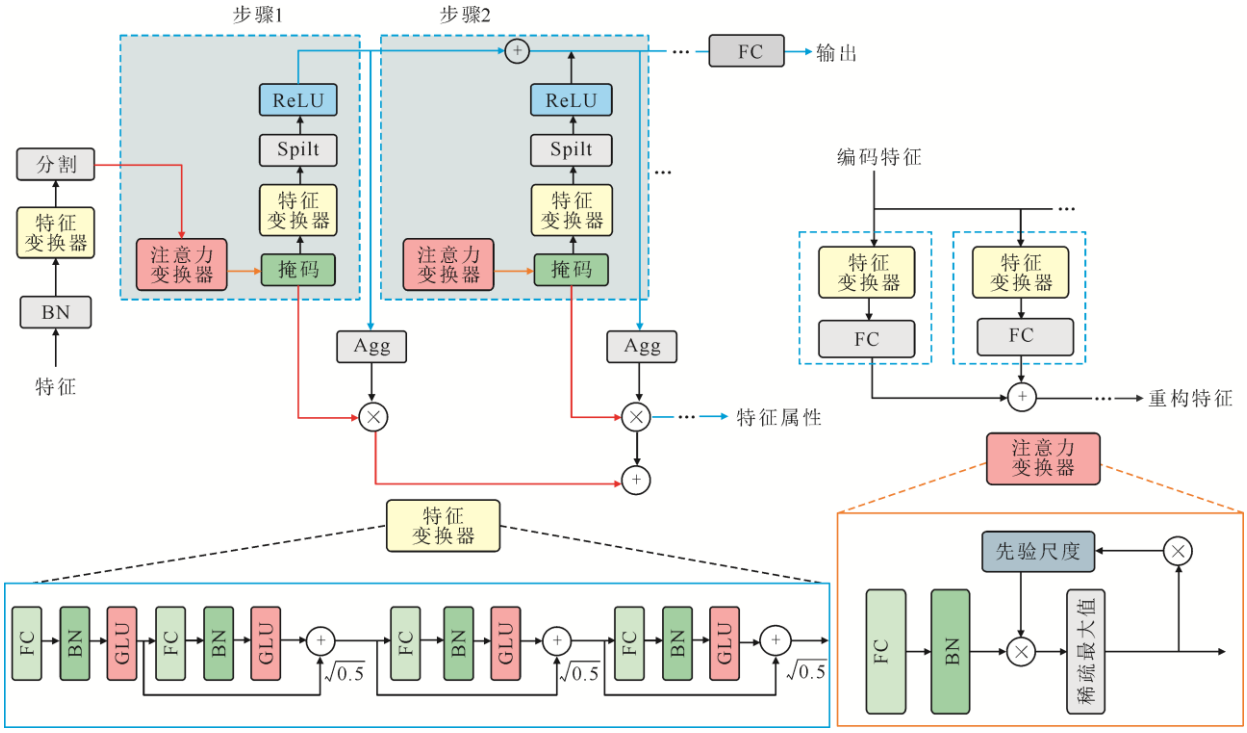


图 3 TabNet 结构示意

Fig.3 Schematic illustration of the TabNet architecture

1.3 超参数优化

对于集成学习算法而言，超参数的配置对模型性能至关重要。采用传统的网格搜索或随机搜索方法，不仅计算成本高昂，且难以高效地找到最优解。因此，本文采用贝叶斯优化作为超参数搜索策略。与前述方法不同，贝叶斯优化通过构建目标函数的概率代理模型，来主动学习超参数与模型性能之间的映射关系[24]。目标函数通过 5 折交叉验证计算均方根误差。针对每个基学习器，独立进行优化迭代共 100 次，以避免陷入局部最优。贝叶斯优化流程如图 4 所示。

考虑到不同模型的特性和计算效率，采用了差异化的超参数优化策略。对于超参数空间较大且对最终性能影响显著的 XGBoost 和随机森林模型，采用贝叶斯优化以自动化寻找其最优配置，贝叶斯超参数空间三维可视化如图 5 所示。

对于 SVM，其超参数较少，采用有限范围的网格搜索即可快速确定稳健取值。对于 MLP 和元学习器 TabNet，由于其神经网络结构导致单次训练耗时显著，若进行需多次模型拟合的贝叶斯优化将带来极高的时间成本。因此，为在可接受的总训练时间内完成调优，对二者采用网格搜索策略。4 种基学习器的最终超参数设置见表 1，元学习器 TabNet 的超参数 n_a、n_d、n_steps、lr 分别设置为 128、128、5、0.002。

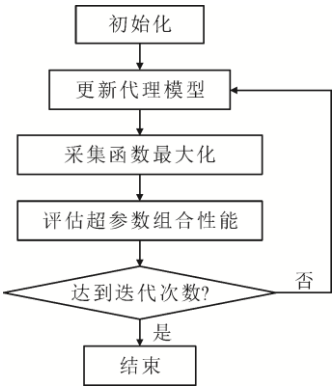


图 4 贝叶斯优化流程

Fig.4 Flowchart of the Bayesian optimization process

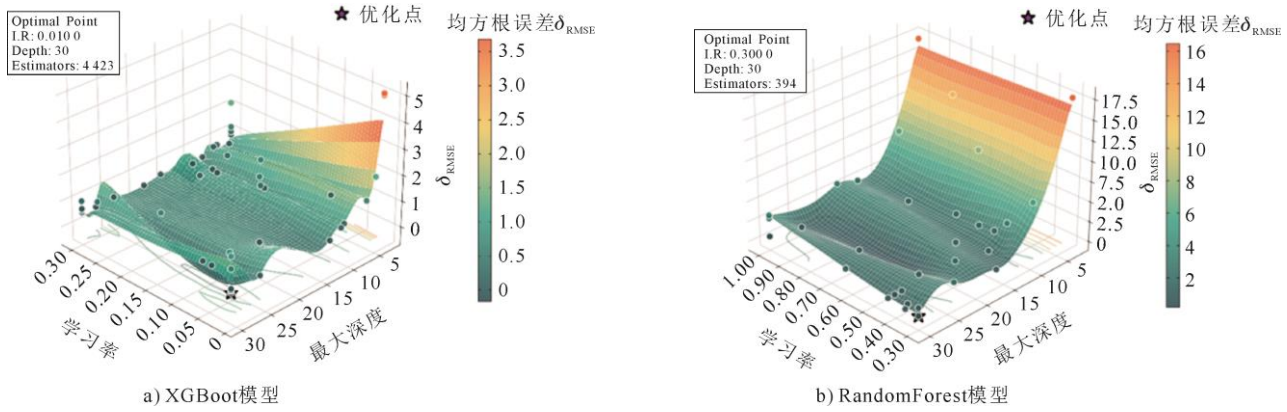


图 5 贝叶斯超参数空间三维可视化

Fig.5 Three-dimensional visualization of the Bayesian hyperparameter optimization space

表 1 基学习器超参数设定

Tab.1 Hyperparameter settings of the base models

模型	超参数	设定值
XGBoost	Estimators	4 423
	Max Depth	30
	Learning Rate	0.01
随机森林	Estimators	394
	Max Depth	30
	Max Feature	0.3
SVM	kernel	rbf
	C	1.0
	epsilon	0.1
MLP	隐含层结构	(100, 50)
	激活函数	ReLU
	Learning Rate	0.01
	epoch	1 000

2 实验数据与处理

2.1 实验系统与数据采集

实验数据来源于某微混燃烧室。其喷嘴由 95 根直径为 8 mm 的微管按特定比例分级排列构成，其中一级微管 28 个、二级 22 个、三级 45 个。三级微管的数量比例决定了燃烧室进口空气流量的分配比例。实验采用市政天然气作为燃料，通过高精度传感器采集系统状态参数。为构建模型，筛选出 10 个关键输入特征，包括进口空气压力、进口空气温度、空气总流量、燃料总流量以及 3 个分级的燃料流量与当量比。各级当量比由对应燃料流量与空

气流量计算得出。目标变量为燃烧室出口的 NO_x 排放质量浓度，由排气段烟气分析仪连续测量。表 2 列出了输入特征选取与取值范围。

表 2 输入特征选取与取值范围

Tab.2 Selection and value ranges of input features

特征	最小值	最大值
进口空气压力/kPa	148.55	1 024.65
进口空气温度/K	337.09	601.73
空气总流量/(kg s^{-1})	0.69	2.62
燃料总流量/(g s^{-1})	7.66	39.14
一级燃料流量/(g s^{-1})	0.002 6	18.38
二级燃料流量/(g s^{-1})	0.001 1	12.37
三级燃料流量/(g s^{-1})	0.000 48	8.38
一级当量比	0.028	0.66
二级当量比	0.032	0.65
三级当量比	0.027	0.15

2.2 数据预处理与数据集划分

为了保证模型训练的稳健性，对数据进行预处理。首先，采用 *z*-score 方法剔除全局离群值。其次，仅使用训练集数据计算均值和标准差，并对所有特征进行 *z*-score 标准化。最后，随机划分训练集与测试集，并固定随机数种子为 42，以保证结果可复现。此流程避免了数据泄露。

为了评估模型对动态过程的预测能力，训练与验证均使用动态数据。原始数据集共 69 000 个样本，清洗后得到 64 945 个有效样本。将其按 80%与 20%的比例随机划分为训练集与测试集。

2.3 评估指标

本文采用均方根误差δ_{RMSE}（root mean squared error，RMSE）、平均绝对误差δ_{MAE}（mean absolute error，MAE）以及决定系数（R^2）3 种评价指标，其数学定义如下：

$$\delta_{\mathrm{RMSE}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{7}$$

$$\delta_{\mathrm{MAE}}=\frac{1}{n}\sum_{i=1}^{n}\left|y_i-\hat{y}_i\right| \tag{8}$$

$$R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \tag{9}$$

式中：y_i 代表真实值；$\hat{y}_i$ 代表预测值；$\bar{y}$ 代表样本平均值；n 代表总样本数。

3 结果与讨论

3.1 堆叠集成模型与基学习器性能对比

为评估堆叠集成模型的动态预测性能，本文以 XGBoost、随机森林、支持向量机和多层感知机作为基学习器，并以 TabNet 为元学习器构建集成模型，得到各模型性能指标对比如图 6 所示。由图 6 可见，在动态测试集上，堆叠集成模型的均方根误差与平均绝对误差均低于所有基学习器，表明其预测精度更高且更稳定。

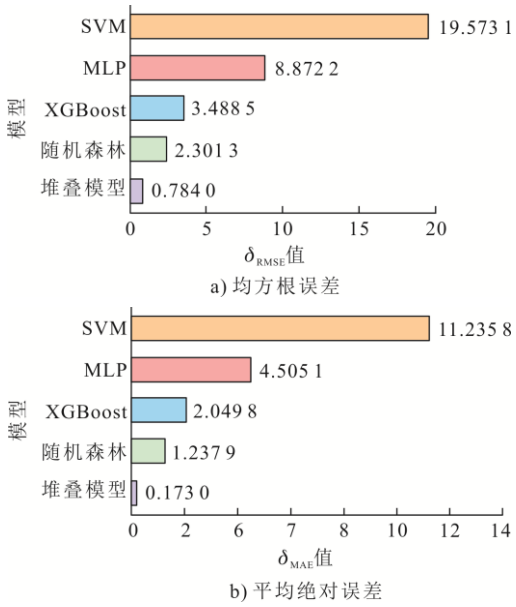


图 6 各模型性能指标对比

Fig.6 Comparison of performance metrics for different models

堆叠集成模型的优势源于其融合架构能综合利用各基学习器的优势。一方面，异质基学习器提供了互补的预测视角。XGBoost、随机森林、支持向量机和多层感知机分别基于 Boosting、Bagging、核方法及神经网络范式，从偏差优化、方差控制、结构风险最小化及非线性映射等不同角度建模，为元学习器提供了多样化的初级预测特征。另一方面，元学习器 TabNet 实现了动态融合。其顺序注意力机制能动态评估各基学习器预测的重要性，并进行非线性整合，生成更精确的预测。

3.2 稳态工况分析

为检验所提模型在训练数据分布之外的泛化能力，进行了跨工况验证：将在动态数据集上训练完成的堆叠集成模型及各基学习器，直接应用于独立的稳态测试集进行预测。

首先采用统一流形逼近与投影（uniform manifold approximation and projection，UMAP）方法[25]对稳态与动态数据的分布进行降维可视化，结果如图 7 所示。稳态数据点（橙色）与动态数据点（蓝色）在降维空间中交织，边缘分布基本重叠。这表明

稳态工况的参数范围未明显偏离动态训练数据所覆盖的特征空间主体，为后续泛化评估提供了依据。

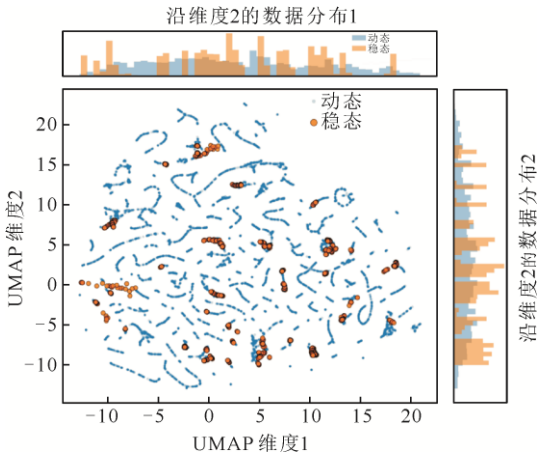
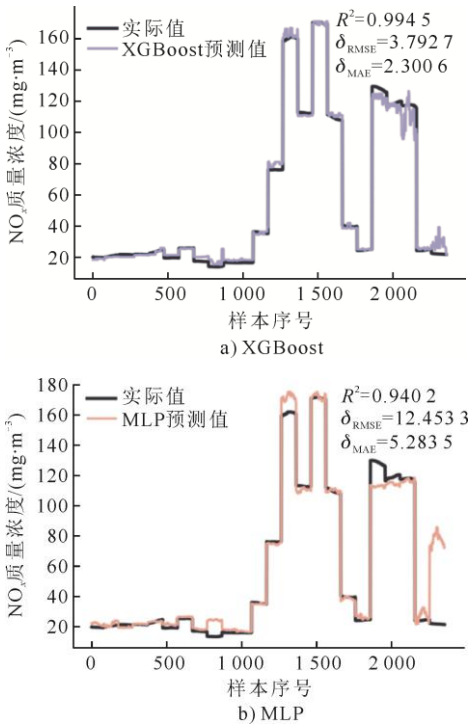


图 7 稳态数据与动态数据 UMAP 可视化

Fig.7 UMAP visualization of the distribution of steady-state and dynamic data

各模型在稳态测试集上的预测结果如图 8 所示。稳态测试集中 NO_x 实际排放质量浓度的统计分析显示，其范围为 13.80~171.37 mg/m³，平均值为 56.99 mg/m³，方差为 2 594.1，表明排放过程具有显著的波动性。从时序变化来看，NO_x 质量浓度呈现阶段性波动：在样本序号 0–1 000 区间内，NO_x 质量浓度维持在 13.80~25.95 mg/m³的低水平，变化平缓；在 1 000–1 500 区间内，NO_x 质量浓度呈阶梯式上升，逐步突破 100 mg/m³并接近峰值；在 1 500–2 000 区间内，NO_x 质量浓度进入高波动状态，后期虽回落至 20 mg/m³ 左右，仍有小幅波动。



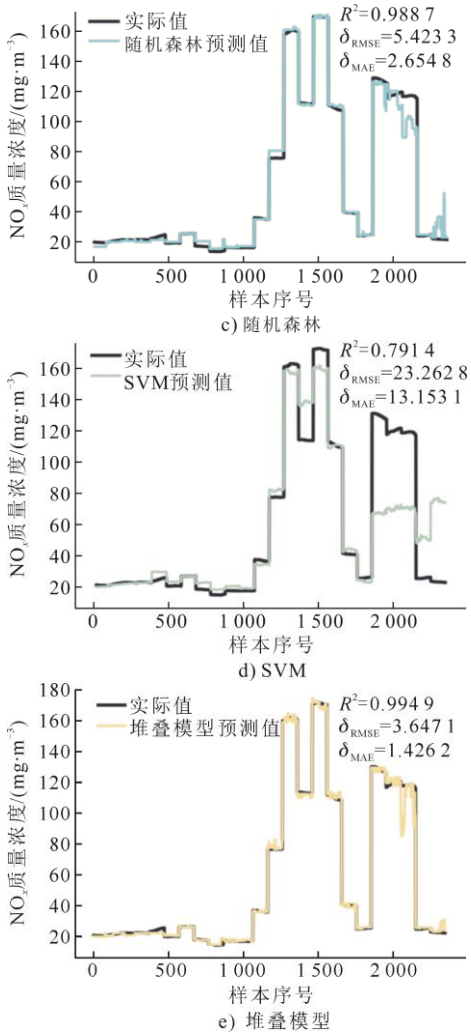


图 8 各模型预测曲线对比

Fig.8 Comparison of prediction curves of different models

对比各模型的预测曲线与实际浓度可见，堆叠集成模型的预测曲线与实际值贴合最为紧密，即使在高波动阶段仍能较好捕捉 NO_x 质量浓度变化趋势。定量结果表明，堆叠集成模型的均方根误差为 3.647 1 mg/m³，平均绝对误差为 1.426 2 mg/m³，在所有模型中最优，进一步验证了其在稳态工况下仍具有优异的泛化能力和预测精度。

3.3 模型可解释性及效率分析

为深入探究各基学习器的决策机制并验证其预测行为与物理机理的一致性，针对 XGBoost 与随机森林 2 种基学习器进行了可解释性分析。通过 SHAP 方法[26]量化了不同输入特征对 NO_x 排放预测的贡献度。图 9 展示了 2 个基学习器的特征重要性排序结果，即其及 SHAP 分析。由图 9 可见，尽管不同模型在特征权重的具体分配上存在细微差异，但呈现出一致的总体趋势。SHAP 分析显示一级燃料流量是对 NO_x 排放预测最具影响力的特征，表明其对

NO_x 排放预测贡献最显著。其次，二级燃料流量与进口空气压力是 NO_x 排放影响次重要因素。

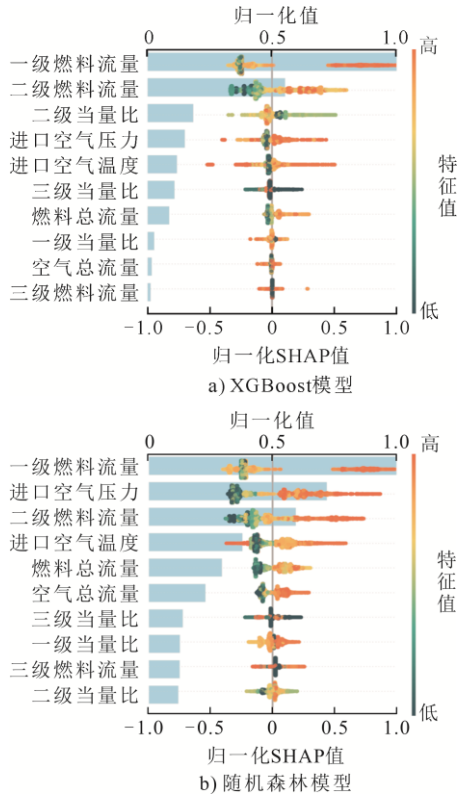


图 9 基学习器 SHAP 分析
Fig.9 SHAP analysis of the base models

为理解关键特征与预测目标之间的关联，图 9 也给出了一级与二级燃料流量与 NO_x 排放的实际分布散点图。结果显示，一级与二级燃料流量与 NO_x 排放呈现出清晰的正相关趋势，这主要是因为作为主燃级燃料，一级与二级燃料流量的增加将直接导致主燃区温度上升，从而促进热力型 NO_x 的生成。

在污染物排放监测与实时控制场景中，预测模型不仅需具备高精度，还应满足工程部署中对推理速度的实时性要求。如前所述，数据驱动模型相较于传统数值模拟的一大优势在于其快速响应能力，可支撑在线监测与实时调控。

为量化分析各模型的推理效率，在相同硬件环境（AMD R7 8845H，32GB RAM，RTX 4070 Laptop GPU）与软件框架下，对训练完成的各模型进行推理耗时测试。测试数据为动态测试集中的 1 个批次，记录每个样本的平均推理时间，结果见表 3。由表 3 可知，在各基学习器中，MLP 推理速度最快，其次为 XGBoost 与随机森林，而 SVM 因核函数计算复杂度较高，推理时间显著增加。堆叠集成模型的单样本推理时间高于所有基学习器。这主要源于其串行架构：在推理阶段，需要先并行执行 4 个基学习器的前向预测，再将结果输入 TabNet 元学习器进行最终决策。

表 3 各模型推理时间
Tab.3 Comparison of inference time per sample for different models

模型	单样本推理时间/ms
XGBoost	0.189 0
随机森林	0.526 4
SVM	1.925 8
MLP	0.071 3
堆叠集成模型	2.125 1

为了进一步分析堆叠模型的耗时构成，对其内部推理过程进行了分解计时，结果显示，基学习器总耗时为 2.000 5 ms，占总推理时间的 94.1%；堆叠集成模型总耗时为 2.125 1 ms，TabNet 元学习器耗时为 0.123 2 ms，仅占约 5.8%，其余时间为数据传递与预处理开销。这表明堆叠模型的时间开销主要来源于基学习器的串联执行。

尽管堆叠集成模型的推理时间较单一基学习器有所增加，但其在动态与稳态工况下均取得了最优的预测精度。在实际工程中，2.125 1 ms 的单样本推理时间仍远低于实时监测系统的延时阈值。因此，在预测精度显著提升的前提下，该模型在精度与效率之间取得了良好平衡，能够满足燃气轮机燃烧过程在线监测与快速优化的实时性需求。

4 结　论

本文针对微混燃烧室 NO_x 排放的快速预测问题，提出了一种基于异质基学习器与 TabNet 元学习器的堆叠集成学习框架，并采用 SHAP 方法对预测结果进行可解释性分析，主要结论如下。

1）提出了一种堆叠集成预测框架。该框架融合了 XGBoost、随机森林、SVM 与 MLP 4 种异质基学习器，并引入具有顺序注意力机制的 TabNet 作为元学习器，结合了不同机器学习范式的优势，为 NO_x 排放预测提供了一种高精度解决方案。

2）模型在动态与稳态工况下均表现出优异的预测性能与泛化能力。在动态测试集上，堆叠集成模型的均方根误差与平均绝对误差分别降低至 0.784 0 mg/m³ 与 0.173 0 mg/m³，优于所有单一基学习器。在未参与训练的稳态测试集上，该模型同样取得最优预测结果，均方根误差为 3.647 1 mg/m³，平均绝对误差为 1.426 2 mg/m³，表明其具有良好的跨工况适应能力。

3）可解释性分析验证了模型的物理一致性。基于 SHAP 的分析表明，一、二级燃料流量是对 NO_x 排放影响最大的特征，这与燃料分级燃烧中主燃区温度影响热力型 NO_x 生成的基本机理相吻合，增强了预测结果的物理可信度与工程应用价值。

[参 考 文 献]

[1] Mollo M, Kolesnikov A, Makgato S. Simultaneous reduction of NO_x emission and SO_x emission aided by improved efficiency of a once-through benson type coal boiler[J]. Energy, 2022, 248: 123551.

[2] Wang C L, Liu Y, Zheng S, et al. Optimizing combustion of coal fired boilers for reducing NO_x emission using Gaussian process[J]. Energy, 2018, 153: 149.

[3] Gülen S C, Curtis M. Gas turbine's role in energy transition[J]. Journal of Engineering for Gas Turbines and Power-Transactions of the Asme, 2024, 146(10): 101008.

[4] Weerakoon A H S, Assadi M. Techno-economic analysis and energy management strategies for micro gas turbines: a state-of-the-art review[J]. Energy Conversion and Management-X, 2024, 24: 100750.

[5] Genova T, Otero M, Morales A, et al. Preheating and premixing effects on NO_x emissions in a high-pressure axially staged combustor[J]. Combustion and Flame, 2022, 235: 111710.

[6] Huang D W, Tang S H, Zhou D J, et al. NO_x emission estimation in gas turbines via interpretable neural network observer with adjustable intermediate layer considering ambient and boundary conditions[J]. Measurement, 2022, 189: 110429.

[7] Zhao Q N, Liu F, Jiao A Y, et al. Prediction model of NO_x emissions in the heavy-duty gas turbine combustor based on MILD combustion[J]. Energy, 2023, 282: 128974.

[8] Su X, Zhang Y M, Zhang Y Y, et al. An enhanced multilayer resinformer with savitzky-golay filter for predicting mixed CO and NO_x emissions in gas turbines[J]. Chemometrics and Intelligent Laboratory Systems, 2025, 261: 105379.

[9] Yousif S T, Ismail F B, Al-Bazi A. A hybrid neural network-based improved PSO algorithm for gas turbine emissions prediction[J]. Advanced Theory and Simulations, 2024, 7(9): 2301222.

[10] Li J J, Han J, Niu D P, et al. Fast and accurate gas turbine emission prediction based on a light and enhanced transformer model[J]. Fuel, 2024, 376: 132750.

[11] 姚顺春, 李龙千, 刘文, 等. 基于 MIC-PCA-LSTM 模型的垃圾焚烧炉 NO_x 排放浓度预测[J]. 华南理工大学学报(自然科学版), 2025, 53(7): 1.
Yao S C, Li L Q, Liu W, et al. Emission concentration prediction of NO_x from waste incinerator based on MIC-PCA-LSTM model[J]. Journal of South China University of Technology (Natural Science Edition), 2025, 53(7): 1.

[12] 张鹏新, 高明明, 解沛然, 等. 基于数据驱动的循环流化床机组深度调峰 NO_x 预测[J]. 发电技术, 2025, 46(3): 627.
Zhang P X, Gao M M, Xie P R, et al. NO_x prediction for deep peaking regulation of circulating fluidized bed units based on data-driven[J]. Power Generation Technology, 2025, 46(3): 627.

[13] 王勇权, 高明明, 王唯铧, 等. 基于 CNN-GRU-MHA 的 CFB 机组污染物排放动态预测[J]. 热力发电, 2025, 54(7): 33.
Wang Y Q, Gao M M, Wang W H, et al. Dynamic prediction of pollutants emission from circulating fluidized bed unit based on CNN-GRU-MHA[J]. Thermal Power Generation, 2025, 54(7): 33.

[14] Hoque K E, Hossain T, Haque A, et al. NO_x emission predictions in gas turbines through integrated data-driven machine learning approaches[J]. Journal of Energy Resources Technology-Transactions of the Asme, 2024, 146(7): 071201.

[15] Naghibi A. Exploring explainable ensemble machine learning methods for long-term performance prediction of industrial gas turbines: a comparative analysis[J]. Engineering Applications of Artificial Intelligence, 2024, 138: 109318.

[16] Coelho L D, Ayala H V H, Mariani V C. CO and NO_x emissions prediction in gas turbine using a novel modeling pipeline based on the combination of deep forest regressor and feature engineering[J]. Fuel, 2024, 355: 129366.

[17] Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need[J]. Information Fusion, 2022, 81: 84.

[18] Borisov V, Leemann T, Sessler K, et al. Deep neural networks and tabular data: a survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(6): 7499.

[19] Chen T Q, Guestrin C, Assoc Comp M. XGBoost: a scalable tree boosting system[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD), San Francisco, CA, F Aug. 13-17, 2016.

[20] Zhou H, Tang Q, Yang L B, et al. Support vector machine based online coal identification through advanced flame monitoring[J]. Fuel, 2014, 117: 944.

[21] Ren T, Modest M F, Fateev A, et al. Machine learning applied to retrieval of temperature and concentration distributions from infrared emission measurements[J]. Applied Energy, 2019, 252: 113448.

[22] Yan P L, Fan W J, Zhang R C. Predicting the NO_x emissions of low heat value gas rich-quench-lean combustor via three integrated learning algorithms with Bayesian optimization[J]. Energy, 2023, 273: 127227.

[23] Arik S O, Pfister T. Assoc Advancement Artificial I. TabNet: attentive interpretable tabular learning[C]// Proceedings of the 35th AAAI Conference on Artificial Intelligence/33rd Conference on Innovative Applications of Artificial Intelligence/11th Symposium on Educational Advances in Artificial Intelligence, Electr Network, F Feb 02-09, 2021.

[24] Victoria A H, Maragatham G. Automatic tuning of hyperparameters using Bayesian optimization[J]. Evolving Systems, 2021, 12(1): 217.

[25] Wang Y F, Huang H Y, Rudin C, et al. Understanding how dimension reduction tools work: an empirical approach to deciphering t-SNE, UMAP, TriMap, and PaCMAP for data visualization[J]. Journal of Machine Learning Research, 2021, 22: 1.

[26] Lundberg S M, LEE S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st Annual Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, F Dec 04-09, 2017.

（责任编辑　杜亚勤）