

DOI: 10.19666/j.rlfed.202302378

基于 mRMR-BO 优化 Stacking 集成模型的 NO_x 浓度动态软测量

金秀章, 乔鹏, 史德金

(华北电力大学控制与计算机工程学院, 河北 保定 071003)

[摘要] 针对火电厂选择性催化还原 (selective catalytic reduction, SCR) 烟气脱硝系统中, 由于影响入口 NO_x 质量浓度因素过多及系统大延迟大惯性, 导致入口 NO_x 质量浓度难以准确及时测量的问题, 提出了利用最大相关-最小冗余算法 (max-relevance and min-redundancy, mRMR) 结合贝叶斯优化算法 (Bayesian optimization, BO) 优化 Stacking 集成模型的 SCR 烟气脱硝系统入口 NO_x 质量浓度动态软测量模型。针对动态 NO_x 生成过程中静态单一模型预测精度降低及辅助变量与入口 NO_x 质量浓度时间异步的问题, 利用 mRMR-BO 结合模型进行辅助变量筛选, Copula 熵 (copula entropy, CE) 确定辅助变量延迟, BO 结合模型确定辅助变量阶次, 将 TCN 及 LASSO 利用 Stacking 法集成, 使用含有延迟时间及阶次信息的辅助变量构建动态 Stacking 集成软测量模型。仿真结果显示: 集成模型较 TCN 及 LASSO 单一网络的均方根误差、平均绝对误差、平均绝对百分比误差最小; 动态集成模型对比静态集成模型, 预测精度更高, 能够实现对入口 NO_x 质量浓度的准确软测量。

[关键词] NO_x 动态建模; 最大相关-最小冗余; 贝叶斯优化; Stacking 集成模型

[引用本文格式]金秀章, 乔鹏, 史德金. 基于 mRMR-BO 优化 Stacking 集成模型的 NO_x 浓度动态软测量[J]. 热力发电, 2023, 52(10): 122-128. JIN Xiuzhang, QIAO Peng, SHI Dejin. Dynamic soft measurement of NO_x concentration based on mRMR-BO Stacking ensemble model[J]. Thermal Power Generation, 2023, 52(10): 122-128.

Dynamic soft measurement of NO_x concentration based on mRMR-BO Stacking ensemble model

JIN Xiuzhang, QIAO Peng, SHI Dejin

(School of Control and Computer Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: Aiming at the problem that it is difficult to accurately and timely measure the inlet NO_x concentration in the denitrification system of selective catalytic reduction (SCR) in thermal power plants, due to the excessive factors affecting the inlet NO_x concentration and the large delay and inertia of the system, the Max-Relevance and Min-Redundancy (mRMR) combined with Bayesian optimization (BO) algorithm is proposed, optimize the dynamic soft measurement model of NO_x concentration at the inlet of the SCR denitration system of the stacking ensemble model. Aiming at the problem of reduced prediction accuracy of static single model and asynchronous timing of auxiliary variables and inlet NO_x concentration in the process of dynamic NO_x generation, the mRMR-BO combined with model was used to screen the auxiliary variables, Copula Entropy (CE) determined the delay of auxiliary variables, the BO combined with model determined the order of auxiliary variables, and TCN and LASSO were integrated by Stacking method. The auxiliary variables containing delay time and order information were used to construct a dynamic stacking ensemble soft measurement model, and the simulation results showed that the root mean square error, average absolute error, and average absolute percentage error of the integrated model compared with TCN and LASSO single networks were the smallest. Compared with the static ensemble model, the dynamic ensemble model has higher prediction accuracy and can achieve accurate soft measurement of the inlet NO_x concentration.

Key words: NO_x dynamic modeling; mRMR; Bayes optimization; Stacking ensemble model

修回日期: 2023-02-11 网络首发日期: 2023-09-19

第一作者简介: 金秀章 (1969), 男, 副教授, 主要研究方向为先进控制策略在大型火电机组的应用、信息融合技术等, jinxsy@163.com。

通信作者简介: 乔鹏 (1995), 男, 硕士研究生, 主要研究方向为先进控制策略在大型火电机组的应用, 220212216043@ncepu.edu.cn。

近年来,我国对于火电厂污染物排放指标要求愈发严格,其中针对排放烟气中的 NO_x 质量浓度,国家要求为排放质量浓度小于 50 mg/m³,必须实时监测^[1-2]。在电厂实际生产中,NO_x 质量浓度控制一般采用 SCR 烟气脱硝反应器进行喷氨中和,由于实际系统中 SCR 烟气脱硝反应器的入口 NO_x 质量浓度测量存在延迟,无法精确控制出口 NO_x 质量浓度,故需对 SCR 入口 NO_x 质量浓度建立精确度较高的软测量模型,以满足实际排放的要求。软测量是基于计算机编程,分析脱硝系统中易于准确测得的辅助变量进而实现对目标变量即入口 NO_x 质量浓度的精确预测,用预测值代替实际测量值作为前馈,精确控制喷氨量,改善脱硝系统 NO_x 质量浓度控制品质^[3]。

随着人工智能算法的发展,针对上述问题已有较多学者建立了相关软测量模型。如 Zhou 等人^[4]采用 GA-SVM 模型预测火电厂 NO_x 排放浓度,文中所采用的输入变量仅有 3 个,不利于模型泛化。Li 等人^[5]采用深度极限学习机构建 NO_x 浓度预测模型,采用 Rajda 标准进行数据滤波,并采用互信息进行输入特征选择,最后选择了 19 个输入变量作为模型的输入,输入变量过多,模型运行成本较高。Sun 等人^[6]提出了采用门控循环网络对 NO_x 浓度进行预测建模,采用手动调参确定模型的最优超参数,耗时较多。Tang 等人^[7]提出了基于差分进化优化最小二乘支持向量机的 NO_x 预估模型,采用偏最小二乘法做特征选择,文中对输入数据仅做了归一化处理,未做滤波处理。Cui 等人^[8]提出了一种基于非线性滑动平均模型来预测 NO_x 浓度,该文考虑了各输入变量与输出变量间的延迟,并用相关系数法确定延迟时间并做时序统一。

上述研究都采用单一模型对 NO_x 质量浓度进行软测量,而 SCR 烟气脱硝系统运行工况复杂,单一模型不能较为充分地挖掘数据特征,满足复杂系统建模要求。同时,建立模型时,未充分考虑数据间的全部信息,研究主要集中在模型超参数及结构调整。

本文采用 Stacking 法将 TCN 模型及 LASSO 模型作为集成模型的第 1 层,线性模型作为集成模型的第 2 层。在模型层方面,通过第 1 层的强预测模型充分挖掘数据间的时序信息及提升处理高维度信息能力,同时为了避免过拟合,第 2 层选用简单模型;在数据层方面,采用 Copula 熵 (Copula entropy, CE)^[9]计算辅助变量与目标变量间的延迟,利用贝叶斯优化算法 (Bayesian optimization, BO)

结合模型确定辅助变量历史数据阶次,筛选后的辅助变量添加延迟和阶次信息后构建动态软测量模型,并利用火电厂厂级监控系统 (SIS) 采集的数据进行验证,结果表明动态建模预测效果优于静态建模,集成模型预测效果优于单一 TCN、LASSO 模型。

1 模型理论基础

1.1 Stacking 集成模型

Stacking 集成法^[10]是将不同的预测模型分层集成的方法,本文设计集成模型为 2 层,第 1 层模型称为基模型,第 2 层模型称为元模型。

Stacking 集成模型在进行训练及预测时,分为 2 个部分。首先将数据集分为训练集 Train 及测试集 Test,在基模型训练时,将训练集利用 5 折交叉验证法分为 5 部分,Train=(T1,T2,T3,T4,T5),T1—T5 依次做验证集,其余用作训练,共分出 5 组不同数据,利用该方法对基模型进行训练,得到基模型的 5 个子模型及对对应验证集在基模型中的预测值 pre_val=(pre_val1,pre_val2,pre_val3,pre_val4,pre_val5);然后将测试集数据分别输入训练出的基模型的 5 个子模型,所得到的 5 个测试集预测结果取均值,作为基模型的最终预测值 pre_test;最后对基模型中的 2 个模型做上述操作,得到 2 个模型对应的验证集预测结果的拼接 (TCN_pre_val、LASSO_pre_val) 及测试集预测结果拼接 (TCN_pre_test、LASSO_pre_test)。

为避免数据重复利用产生过拟合,元模型的训练集为基模型输出的验证集预测结果,元模型的测试集为原数据测试集在基模型的预测结果。元模型层在训练时同样采用 5 折交叉验证法,充分提取数据特性,元模型最终的输出就是集成模型的输出结果。

1.2 基模型选取

Stacking 集成法的优势在于集合不同强预测模型,充分提取数据特征,实现效果最优预测。针对火电厂 SIS 系统中采集的数据为时间序列数据,选取时序卷积网络 (TCN)^[11]充分提取时序信息,其主要模型构成为因果卷积,膨胀卷积^[12],残差连接^[13];针对 SCR 系统运行复杂性、共线性及高维度选取 LASSO 模型^[14],上述 2 种模型组成集成模型的基模型。

1.3 贝叶斯优化算法

BO 最初由 BergstraJ 等人^[15]在 2011 年提出,

是一种先验式的搜索方法,其运算量相较网格搜索或随机搜索的穷举后验式搜索方法减少较多,且能保证搜索精度。贝叶斯优化算法的核心内容为:

1) 定义目标函数 $f(x)$ 及确定 x 的取值范围;

2) 取有限个定义域内的值,求解相应的 $f(x)$ 作为观测值;

3) 根据有限的观测值,采用概率代理模型预估函数的分布;

4) 采用采集函数确定下一个观测值 $f(x)$ 所对应的 x ;

5) 重复 2) 至 4) 步,直至求出符合设定要求的目标函数的最小值。

BO 的核心部分为确定概率代理模型及采样函数。传统贝叶斯优化采用的概率代理模型为高斯过程 (Gaussian process), 本文采用 TPE 过程 (tree parzer estimator process) 作为概率代理模型。常见的采样函数有概率增量, 期望增量, 置信度上界及信息熵等, 本文采用期望增量作为采样函数。

2 构建 Stacking 动态软测量模型

2.1 mRMR-BO 筛选辅助变量

本文使用某 600 MW 电厂 SIS 提供的间隔 10 s 采样的 10 000 组数据, 对应负荷范围为 300~660 MW, 包含波动工况及稳定工况, 确保模型针对不同工况都能实现精确预测, 并从中随机选取连续的 6 000 组数据作为模型的训练集, 1 000 组数据作为测试集。

由 SIS 得到的样本数据较为复杂, SCR 入口 NO_x 质量浓度和燃烧情况密切相关, 直接影响因素包括煤质及配风, 但是所选用的电厂对于煤质工业参数分析次数有限, 实时获取存在难度, 同时煤质变化会影响配风等相关特征参数变化^[16]。综合考虑, 反向利用相关特征代替煤质变化影响, 同时考虑锅炉负荷, 给煤量变化、SCR 反应器相关参量, 风烟系统中烟气流量、温度及烟气管道压力等影响 NO_x 生成因素, 通过机理分析, 初步选取了烟气流量、引风机出口烟气温度、锅炉总风量、一次风量、二次风量、机组发电功率、给煤机总煤量、给煤机瞬时煤块流量、磨煤机入口风量、SCR 反应器入口烟气温度、锅炉出口烟气管道压力、SCR 反应器入口压力、锅炉氧量共计 13 个初始特征作为模型辅助变量组。但此时辅助变量组中各特征与目标变量的互相关性及特征之间的自相关性有待进一步确定, 且存在冗余的辅助变量会降低模型预测精度。

利用最大相关-最小冗余算法 (max-relevance and min-redundancy, mRMR)^[17], 第 1 步依次计算辅助变量 X_i 与目标变量 Y 间的互信息, 用最大互信息 $I(X_i; Y)$ 表征其最大相关性。

$$\max D(S, Y), D = \frac{1}{|S|} \sum_{X_i} I(X_i; Y) \quad (1)$$

式中: S 为辅助变量数据集; $D(\cdot)$ 为筛选与目标变量最大互相关的辅助变量组函数。

经过式(1)得到的辅助变量组满足了与目标变量间最大相关, 但是存在一定的辅助变量间的冗余, 为同时满足最大相关及最小冗余, 第 2 步消除筛选出的辅助变量组的相互冗余。

$$\min R(S), R = \frac{1}{|S|^2} \sum_{X_i, X_j \in S} R(X_i, X_j) \quad (2)$$

式中: $R(\cdot)$ 为消除辅助变量冗余函数。

结合式(1)、式(2), mRMR 算法可表示为:

$$\max \Phi(D, R) = \max_{X_i, X_j \in X - S_{m-1}} [I(X_i; Y) - \frac{1}{m-1} I(X_i; X_j)] \quad (3)$$

式中: m 为辅助变量组中特征个数。

通过 mRMR 对 13 个辅助变量做最大相关-最小冗余计算, 得到 13 个候选顺序辅助变量组 $S_1 \sim S_{13}$, 其中包含的特征为 1 到 13 个, 且满足 $S_1 \subset S_2 \subset \dots \subset S_{13}$ 。

为确定预测效果最好的一组辅助变量, 利用贝叶斯算法结合集成模型, 采用训练集数据, 经过 5 折交叉验证训练模型后的验证集预测结果与对应 SIS 采集的 NO_x 质量浓度实际值的平均绝对百分比误差 (mean absolute percentage error, MAPE) 作为评价指标, 最终确定输入模型的辅助变量, mRMR-BO 筛选辅助变量流程如图 1 所示。

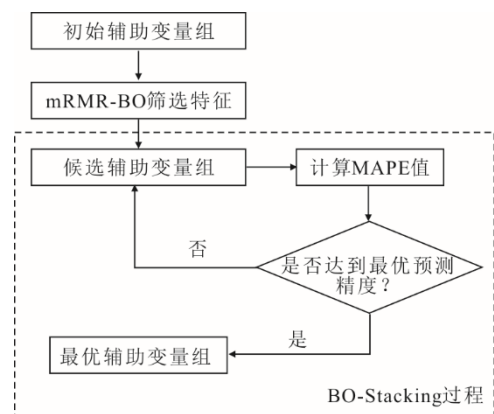


图 1 mRMR-BO 特征选择流程
Fig.1 mRMR-BO feature selection flowchart

为避免输入模型特征过少导致模型欠拟合，故输入模型辅助变量从 3 个递增，逐一计算其预测精度，各候选辅助变量组[S₃, ..., S₁₃]结合 BO-Stacking 模型，预测精度如图 2 所示。从图 2 可以看到：当辅助变量个数为 8 时，即 S₈ 作为最终的模型输入变量组，模型预测精度最高；同时，还显示出 Stacking 集成模型预测精度要高于其基模型中的单一模型。此时输入模型辅助变量分别为给煤机总煤量、给煤机瞬时煤块流量、锅炉氧量、二次风量、一次风量、引风机出口烟气温度的、机组发电功率及锅炉出口烟气管道压力。至此，辅助变量完成降维，从 13 维降至 8 维。

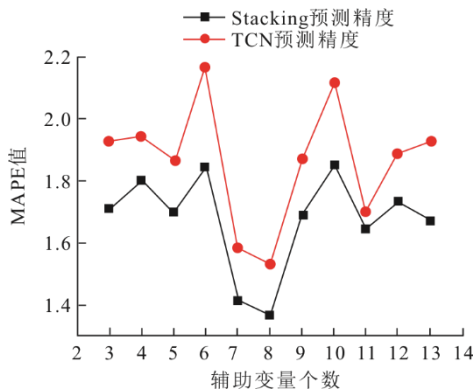


图 2 候选辅助变量组预测精度
Fig.2 Prediction accuracy of candidate auxiliary variable groups

2.2 计算辅助变量延迟及阶次

由于锅炉燃烧系统的大延迟、大惯性特性，对于动态 NO_x 生成过程，t 时刻入口 NO_x 质量浓度可能与[t-k, ..., t-k-d]时刻的辅助变量 X_i 有关，即存在目标变量与辅助变量变化不同步、时间异步的情况，同时存在当前时刻目标变量变化与一段历史时间内的辅助变量相关的情况，具体如图 3 所示。

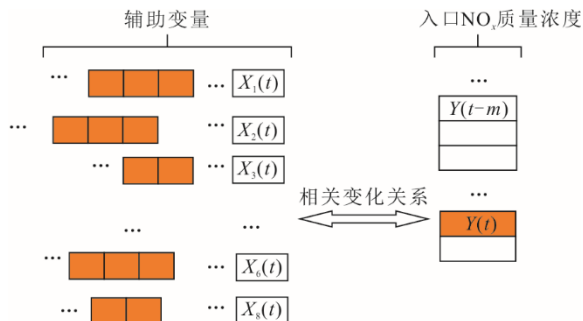


图 3 辅助变量与目标变量时序及阶次关系
Fig.3 Temporal and order relationship between auxiliary variables and target variable

结合广义控制理论的 CARIMA 模型^[16]，可将动态软测量模型表达为：

$$Y(t) = M[X_1(t-k_1, t-k_1-1, t-k_1-d_1), \dots, X_8(t-k_8, t-k_8-1, t-k_8-d_8), Y(t-1), \dots] \quad (4)$$

式中：Y 为入口 NO_x 质量浓度；M(·) 为模型预测过程；X_i 为辅助变量；t 为时间步；k_i 为延迟时间步；d_i 为历史数据阶次。

根据式(4)，入口 NO_x 质量浓度历史数据同样对当前时刻浓度变化有影响，为避免模型过拟合，补充上一时刻入口 NO_x 质量浓度作为模型输入变量。

2.2.1 确定辅助变量延迟

为解决时间异步的问题，利用 Copula 熵计算辅助变量[X₁, ..., X₈]与目标变量 Y 即入口 NO_x 质量浓度之间的时序最大相关。

Copula 熵 (Copula entropy, CE) 本质是一种香浓熵，MA Jian 等^[18]提出 CE 为互信息添加了一种新的数学层面的理解，同时在公式推导中证明了互信息就是负的 Copula 熵。CE 还具备可处理非线性关系、统计独立性度量，计算便捷等优点。

设定最大延迟为 5 min，数据间隔为 10 s，换算为时间步[0,30]，利用 CE 计算辅助变量延迟的具体步骤为：

- 1) 固定 Y 的时间步 t;
- 2) 分别计算 CE[X_i(t-0);Y(t)], ..., CE[X_i(t-30);Y(t)];

3) 第 2 步中 CE 最小值对应的时间步即为辅助变量 X_i 与 Y 的延迟时间步 k_i。

2.2.2 确定辅助变量历史数据阶次

确定各辅助变量延迟时间步 k_i 后，依据式(4)分别确定影响目标变量的辅助变量历史数据阶次 d_i，利用 BO-Stacking，将不同辅助变量的历史数据阶次作为超参数，设定最大历史阶次为 4，加入辅助变量延迟时间步，以训练集数据结合集成模型，利用 5 折交叉验证后验证集预测结果与对应 SIS 采集的 NO_x 浓度实际值的 MAPE 作为评价指标，最终求取 d_i。

表 1 辅助变量延迟和阶次
Tab.1 Auxiliary variables delay and order

项目	延迟时间步	阶次
给煤机总煤量	12	3
给煤机瞬时煤块流量	13	2
锅炉氧量	11	3
二次风量	11	4
一次风量	23	4
引风机出口烟气温度的	9	2
机组发电功率	19	2
锅炉出口烟气管道压力	29	3

2.3 集成模型超参数寻优

在确定辅助变量历史数据阶次的同时,对集成模型中基模型 TCN 网络及 LASSO 网络超参数进行寻优。对于 TCN 网络,主要超参数为训练轮数及学习率,分别设定寻优范围为:训练轮数[100,1 000],寻优步长 50,学习率[0.001,0.01],寻优步长 0.001;对于 LASSO 网络,主要超参数为惩罚因子 α ,寻优范围[1×10^{-4} , 1×10^{-3}],寻优步长 1×10^{-4} 。

利用 BO,将辅助变量历史数据阶次设置为超参数,同时寻优后,TCN 网络最优超参数为:训练轮数 800,学习率 0.006;LASSO 网络最优超参数惩罚因子 $\alpha 4 \times 10^{-4}$ 。

3 实验数据处理

直接从 SIS 系统中获取的数据带有一定的空缺值、异常值及噪声,不利于模型进行精确预测,同时不同特征因计量单位不同,数值差异过大,同样会影响动态建模,需要对所使用的数据做一定预处理。

3.1 处理空缺值、异常值

对于特征列数据中的空缺值,采用该特征列的数据中位数进行填补,对于异常值,利用 3σ 法则判定异常值并剔除,同样采用该特征列数据中位数补充剔除后的空缺值,保证不同特征列之间的时序连续。

3.2 数据降噪处理

SIS 系统所取数值由于设备性能和温度、湿度等影响容易产生噪声^[19]。卡尔曼滤波算法是基于状态变量线性最小方差的递归估计算法^[20],该滤波算法的核心部分为预测与校正 2 部分,在预测阶段,滤波器使用上一状态的滤波器估计值 \hat{X}_{k-1} 预测当前状态的预测值 \hat{X}'_k ;在校正阶段,滤波器利用对当前状态的实际观测值 X_k 修正在预测阶段得到的预测值 \hat{X}'_k ,以获得一个更接近真实的当前时刻滤波器估计值 \hat{X}_k 。其核心计算公式如下:

预测阶段

$$\hat{x}'_k = A\hat{x}_{k-1} + Bu_k \quad (5)$$

$$P'_k = AP_{k-1}A^T + Q \quad (6)$$

校正阶段

$$\hat{z}_k = z_k - H\hat{x}'_k \quad (7)$$

$$K_k = P'_k H^T (HP'_k H^T + R)^{-1} \quad (8)$$

$$\hat{x}_k = \hat{x}'_k + k_k \hat{z}_k \quad (9)$$

更新阶段

$$P_k = (I - k_k H)P'_k \quad (10)$$

式中: A 为状态转移矩阵; u_k 为系统输入向量; B 为输入增益矩阵; Q 为过程噪声协方差矩阵; R 为测量噪声协方差矩阵; H 为测量矩阵; \hat{x}_k 为卡尔曼估计值; P_k 为卡尔曼估计误差协方差矩阵; \hat{x}'_k 为预测值; P'_k 为预测误差协方差矩阵; K_k 为卡尔曼增益; \hat{z}_k 为测量余量; z_k 为处理后的实际观测值,按式(11)计算。

$$z_k = Hx_k + v_k \quad (11)$$

式中: x_k 为实际观测值; v_k 为服从高斯分布的噪声。

所有辅助变量数据均经过卡尔曼滤波,其中省煤器出口烟气流量的数据进行卡尔曼滤波后效果如图 4 所示。

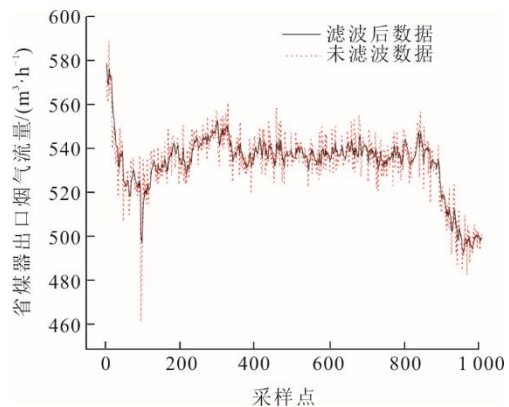


图 4 卡尔曼滤波降噪结果
Fig.4 Noise reduction results of Kalman filter

3.3 数据归一化

各输入变量间存在单位及数值大小差异,为了便于模型进行预测,故对所有输入特征做归一化处理,归一化表达式为:

$$\hat{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (12)$$

4 实验及结果分析

4.1 动、静态建模对于模型输出结果影响

为验证加入辅助变量延迟及历史数据阶次信息的动态建模对比未加入延迟及阶次信息的静态建模对模型预测精度的提升,选用集成模型整体作为预测模型,实验用数据均经过数据预处理的测试集数据,所用模型均经过 BO 寻优,采用模型测试集输出预测值与对应 SIS 采集的 NO_x 质量浓度实际值的均方根误差 (root mean squared error, RMSE) 作为算法的评价标准。测试集[400,600]段预测结果放大展示如图 5 所示,测试集整体预测精度对比见表 2。

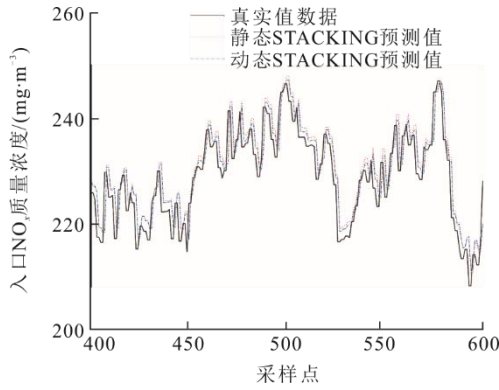


图 5 动、静态建模输出结果对比

Fig.5 Comparison of dynamic and static modeling outputs

表 2 动、静态建模输出结果评价

Tab.2 Evaluation of dynamic and static modeling output results

模型	δ_{RMSE}
静态 Stacking	4.579
动态 Stacking	4.421

分析表 2 中数据可得,对测试集数据来说,Stacking 集成模型动态建模输出结果 δ_{RMSE} 相比静态建模降低了 3.45%。实验结果证明,动态建模能够进一步提升模型预测精度。

4.2 集成模型与单一模型输出结果对比

为证明集成模型能够更充分挖掘脱硝系统数据特性,适应复杂工况建模,将 Stacking 集成模型与其基模型 TCN、LASSO 网络,就动态建模做对比,测试集[0,500]段预测结果放大展示如图 6 所示。采用模型在测试集整体输出预测值与对应 SIS 采集的 NO_x 质量浓度实际值的均方根误差 δ_{RMSE} (root mean squared error, RMSE), 平均绝对值误差 δ_{MAE} (mean absolute error, MAE) 及平均绝对百分比误差 δ_{MAPE} (mean absolute percentage error, MAPE) 作为算法的评价标准,整体预测精度见表 3。

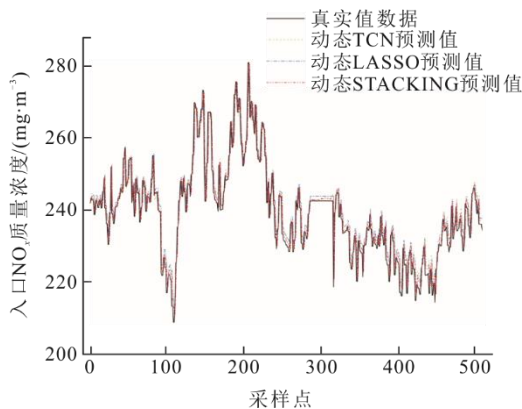


图 6 不同模型预测结果对比

Fig.6 Comparison of prediction results of different models

表 3 3 种模型输出结果评价
Tab.3 Evaluation of output results of three models

模型	δ_{MAPE}	δ_{RMSE}	δ_{MAE}
TCN	1.521	4.793	3.615
LASSO	1.373	4.575	3.265
Stacking 集成模型	1.229	4.421	2.932

由表 3 可见,Stacking 集成模型在 δ_{MAPE} 、 δ_{RMSE} 、 δ_{MAE} 3 项评价指标下,预测效果均优于其基模型中的单一网络模型。集成模型对比单一 TCN 模型, δ_{MAPE} 下降了 19.19%, δ_{RMSE} 下降了 7.76%, δ_{MAE} 下降了 18.89%; 对比单一 LASSO 模型, δ_{MAPE} 下降了 10.49%, δ_{RMSE} 下降了 3.37%, δ_{MAE} 下降了 10.20%。综上所述,集成模型预测效果最优,具备较强的泛化能力,能够实现对 SCR 脱硝系统入口 NO_x 质量浓度的准确预测。

5 结 论

1) 针对运行工况复杂的 SCR 烟气脱硝系统,提出 Stacking 集成模型,将不同强预测模型分层集成,实验结果显示,集成模型对比单一模型,能更好地适应复杂系统,对于入口 NO_x 质量浓度预测效果最优。

2) 针对 NO_x 质量浓度动态生成过程,在筛选出的辅助变量基础上,增加辅助变量延迟及历史数据阶次信息构成动态建模,充分在数据层面分析脱硝系统运行特性,实验结果显示,动态数据建模的模型预测精度优于静态数据建模的模型预测精度。

3) 针对脱硝系统中与入口 NO_x 质量浓度变化相关变量繁多,不利于降低模型运行成本,提升模型预测精度的问题,利用机理分析结合 mRMR-BO-Stacking 集成模型,筛选出与入口 NO_x 质量浓度最大相关且变量间最小冗余的辅助变量组,达到模型最佳预测精度。

[参 考 文 献]

[1] 金秀章,丁续达,赵立慧. 传递熵变量选择的非线性系统时序预测模型[J]. 中国电机工程学报, 2018, 38(增刊 1): 192-200.
JIN Xiuzhang, DING Xuda, ZHAO Lihui. Variables selection for nonlinear system time series prediction model by transfer entropy[J]. Proceedings of the CSEE, 2018, 38(Supply.1): 192-200.

[2] 付忠广,高学伟,李闯,等. 基于改进模糊聚类与 IPSO-SVM 的燃煤电站 NO_x 排放多模型预测[J]. 动力工程学报, 2019, 39(5): 387-393.
FU Zhong guang, GAO Xuwei, LI Chuang, et al. Multi-model NO_x emission prediction based on IGASA-FCM and IPSO-SVM for coal-fired power plants[J]. Journal of

- Chinese Society of Power Engineering, 2019, 39(5): 387-393.
- [3] 赵征, 李悦宁, 袁洪. 燃煤机组 NO_x 生成量动态软测量模型[J]. 动力工程学报, 2020, 40(7): 523-529. ZHAO Zheng, LI Yuening, YUAN Hong. Dynamic model for soft sensing of NO_x generation in coal-fired units[J]. Journal of Chinese Society of Power Engineering, 2020, 40(7): 523-529.
- [4] ZHOU J, LIANG H. Prediction of the NO_x emissions from thermal power plant based on support vector machine optimized by chaos optimization algorithm[C]//IEEE International Conference on Information & Financial Engineering, 2010.
- [5] LI Y, LI F. NO_x prediction method based on deep extreme learning machine[C]//2018 3rd International Conference on Computational Intelligence and Applications (ICCIA), 2018.
- [6] SUN X, JIA X, HOU P, et al. NO_x emission prediction for a coal-fired power process using gated recurrent unit approach[C]//2020 Chinese Automation Congress (CAC), 2020.
- [7] TANG Z, ZHANG H. Modeling NO_x emission of coal-fired boiler with differential evolution optimized least square support vector machine[C]//2018 Chinese Control And Decision Conference (CCDC), 2018.
- [8] CUI J, JIA X. Prediction of NO_x generation process based on a nonlinear ma model[C]//第 31 届中国控制与决策会议, 2019.
- [9] MA J, SUN Z. Mutual information is copula entropy[J]. Tsinghua Science & Technology, 2011, 16(1): 51-54.
- [10] 袁铁江, 郭泽林, 方彤. 基于运行数据时空特征和 Stacking 集成学习的 PEMFC 故障诊断[J/OL]. 中国电机工程学报: 1-10. (2022-07-08)[2022-11-28]. <https://kns.cnki.net/kcms/detail/11.2107.tm.20220707.1344.003.html>. YUAN Tiejia, GUO Zelin, FANG Tong. PEMFC fault diagnosis based on operation data temporal and spatial characteristics and stacking ensemble learning[J/OL]. Proceedings of the CSEE: 1-10. (2022-07-08)[2022-11-28]. <https://kns.cnki.net/kcms/detail/11.2107.tm.20220707.1344.003.html>.
- [11] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[J]. Computation and Language 2018. DOI:10.48550/arXiv.1803.01271.
- [12] 王万良, 胡明志, 张仁贡, 等. 改进 TCN 和 LSTM 的泸水河流域月径流量预测模型[J]. 计算机集成制造系统, 2022, 28(11): 3558-3575. WANG Wanliang, HU Mingzhi, ZHANG Rengong, et al. Monthly runoff prediction model of Lushui river basin based on improved TCN and LSTM[J]. Computer Integrated Manufacturing Systems, 2022, 28(11): 3558-3575.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
- [14] 郑茗友, 王伟, 赵文杰, 等. 基于 PSO-Lasso 算法的电站湿法脱硫出口 SO₂ 浓度预测[J]. 华北电力大学学报(自然科学版), 2022, 49(1): 112-117. ZHENG Mingyou, WANG Wei, ZHAO Wenjie, et al. Prediction of SO₂ concentration at WFGD outlet of power station based on PSO-Lasso algorithm[J]. Journal of North China Electric Power University, 2022, 49(1): 112-117.
- [15] BERGSTRA J, BARDENET R, KÉGL B, et al. Algorithms for hyper-parameter optimization[C]//Advances in Neural Information Processing Systems, 2011.
- [16] 金秀章, 史德金, 乔鹏. 基于沙地猫群优化-最小二乘支持向量机的动态 NO_x 排放预测[J/OL]. 中国电机工程学报: 1-9. (2022-11-04)[2022-11-28]. <https://doi.org/10.13334/j.0258-8013.pcsee.222144>. JIN Xiuzhang, SHI Dejin, QIAO Peng. Dynamic NO_x emission prediction based on sandcat swarm optimization-least squares support vector machine[J/OL]. Proceedings of the CSEE: 1-9. (2022-11-04)[2022-11-28]. <https://doi.org/10.13334/j.0258-8013.pcsee.222144>.
- [17] 胡琰华, 姚圣, 曾成. 面向康复训练的多通道 mRMR-PSO 肌电特征选择算法[J]. 电子测量技术, 2022, 45(11): 72-77. HU Ronghua, YAO Shen, ZENG Cheng. Multi-channel mRMR-PSO sEMG feature selection algorithm for rehabilitation training[J]. Electronic Measurement Technology, 2022, 45(11): 72-77.
- [18] 马剑. Copula 熵: 理论和应用[R/OL]. (2022-10-18)[2022-12-01]. <http://eprint.las.ac.cn/abs/202105.00070>. MA Jian. Copula entropy: theory and applications[R/OL]. (2022-10-18)[2022-12-01]. <http://eprint.las.ac.cn/abs/202105.00070>.
- [19] 金秀章, 于静, 刘岳. 基于人工鱼群-径向神经网络的 NO_x 预测模型[J]. 动力工程学报, 2021, 41(7): 551-557. JIN Xiuzhang, YU Jing, LIU Yue. NO_x prediction model based on artificial fish-radial basis neural network[J]. Journal of Chinese Society of Power Engineering, 2021, 41(7): 551-557.
- [20] 杜鹏, 包晓安, 胡逸飞, 等. 基于卡尔曼滤波的无线传感网时空数据融合算法[J]. 电子科技, 2022, 35(6): 21-27. DU Peng, BAO Xiaolan, HU Yifei, et al. Research on spatio-temporal data fusion algorithm of wireless sensor network based on Kalman filter[J]. Electronic Science and Technology, 2022, 35(6): 21-27.

(责任编辑 邓玲惠)