

DOI: 10.19666/j.rlfed.202408194

# 基于 CNN-GRU-MHA 的 CFB 机组污染物 排放动态预测

王勇权<sup>1,2</sup>, 高明明<sup>1,2</sup>, 王唯铨<sup>1,2</sup>, 张鹏新<sup>1,2</sup>, 成永强<sup>2</sup>

(1. 新能源电力系统全国重点实验室 (华北电力大学), 北京 102206;

2. 控制与计算机工程学院 (华北电力大学), 北京 102206)

**[摘要]** SO<sub>2</sub> 与 NO<sub>x</sub> 排放质量浓度的精准预测可以有效指导污染物排放控制, 对 CFB 机组环保运行具有重要意义。以某 330 MW CFB 机组为研究对象, 采用 Pearson 相关系数实现输入变量筛选, 应用四分位距 (interquartile range, IQR) 方法筛选并替换异常值, 同时进行归一化, 完成数据预处理; 随后, 通过卷积神经网络 (convolutional neural network, CNN) 提取输入变量的特征, 进入门循环单元 (gated recurrent unit, GRU) 处理时间序列特征, 并引入多头自注意力 (multi-head attention, MHA) 机制捕捉特征之间的重要关系, 经训练后反归一化得到模型输出; 最后, 使用平均绝对误差 MAE、平均绝对百分比误差 MAPE 和决定系数 R<sup>2</sup> 评估测试集的结果。结果表明, 该 CNN-GRU-MHA 模型能够较为准确地预测 CFB 机组的污染物排放质量浓度。消融实验与模型对比证明了该模型的优越性能。该 CNN-GRU-MHA 模型可以实现 CFB 机组污染物排放质量浓度的监测与优化指导, 从而使电厂及时调整运行参数, 确保污染物排放达标。

**[关键词]** CFB; 污染物排放预测; 深度学习; 数据驱动

**[引用本文格式]** 王勇权, 高明明, 王唯铨, 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): 33-42. WANG Yongquan, GAO Mingming, WANG Weihua, et al. Dynamic prediction of pollutants emission from circulating fluidized bed unit based on CNN-GRU-MHA[J]. Thermal Power Generation, 2025, 54(7): 33-42.

## Dynamic prediction of pollutants emission from circulating fluidized bed unit based on CNN-GRU-MHA

WANG Yongquan<sup>1,2</sup>, GAO Mingming<sup>1,2</sup>, WANG Weihua<sup>1,2</sup>, ZHANG Pengxin<sup>1,2</sup>, CHENG Yongqiang<sup>2</sup>

(1.State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (North China Electric Power University), Beijing 102206, China;

2.School of Control and Computer Engineering (North China Electric Power University), Beijing 102206, China)

**Abstract:** The accurate prediction of SO<sub>2</sub> and NO<sub>x</sub> emission mass concentrations can effectively guide the control of pollutants emissions, which is of great significance for the environmental protection operation of circulating fluidized bed (CFB) units. A 330 MW CFB unit is taken as the research object, and the Pearson coefficient is used to realize the screening of input variables, and the interquartile range (IQR) method is applied to screen the outliers and replace them with the normalization at the same time, to complete the data preprocessing. Subsequently, the features of input variables are extracted by convolutional neural network (CNN), and by entering into the gate-recurrent unit (GRU) the time-series features are processed. The multi-head self-attention (MHA) mechanism is introduced to capture the important relationships between features, and the model output is obtained after training. Finally, the results of the test set are evaluated using the mean absolute error (MAE), mean

修回日期: 2024-12-24 网络首发日期: 2025-03-06

基金项目: 国家重点研发计划项目 (2022YFB4100304)

Supported by: National Key Research and Development Program (2022YFB4100304)

第一作者简介: 王勇权 (2001), 男, 硕士研究生, 主要研究方向为深度学习在 CFB 机组中的应用, wyq20010501@163.com。

通信作者简介: 高明明 (1979), 男, 教授, 主要研究方向为新能源电力系统人工智能与应用, gmm1@ncepu.edu.cn。

absolute percentage error (MAPE), and the coefficient of determination ( $R^2$ ). The results show that the model is able to predict the pollutants mass concentration in CFBs more accurately and achieve good prediction results, and the superior performance of the model is proved by the comparison of ablation experiments with the model. The proposed CNN-GRU-MHA model can realize the monitoring and optimization guidance of pollutants emissions CFB units, so that the power plant can adjust the operation parameters in time to ensure that the pollutants emissions meet the standards.

**Key words:** circulating fluidized bed; pollutant emission prediction; deep learning; data-driven

随着全球环境保护意识的增强,控制燃煤电厂的  $\text{SO}_2$  与  $\text{NO}_x$  排放成为一项重要任务<sup>[1]</sup>。为响应国家绿色环保精神,2015 年 12 月,我国生态环境部印发了《全面实施燃煤电厂超低排放和节能改造工作方案》<sup>[2]</sup>,鼓励新建 CFB 机组  $\text{NO}_x$  排放质量浓度标准不高于  $50 \text{ mg/m}^3$ ,  $\text{SO}_2$  排放质量浓度不高于  $35 \text{ mg/m}^3$ <sup>[3]</sup>。仅依靠 CFB 锅炉低排放优势很难满足新的排放标准<sup>[4]</sup>。同时,由于 CFB 机组具有大惯性、大迟延特征,且火电机组需频繁进行变负荷以维持电网稳定<sup>[5]</sup>,导致污染物排放控制难以达标。因此,需要对污染物排放进行预测,以实现对其控制的指导<sup>[6]</sup>。

在污染物预测中,传统的机器学习算法如  $K$  近邻算法<sup>[7]</sup>、支持向量机<sup>[8]</sup>、随机森林<sup>[9]</sup>、BP 神经网络<sup>[10]</sup>等已经被广泛应用。这些算法在污染物检测中取得了一定的准确率提升,但仍存在表达复杂函数能力有限和泛化能力不强的问题<sup>[11]</sup>。为了进一步提高污染物预测的性能,近年来,深度学习算法如卷积神经网络(convolutional neural network, CNN)、循环神经网络和深度自编码器等也被引入污染物预测领域,并在实验中表现出了很好的性能。Adams 等人<sup>[12]</sup>开发了一种新型深度神经网络(deep neuronal network, DNN)和最小二乘支持向量机(least square support vector machine, LSSVM)算法,用于预测 CFB 锅炉中的  $\text{SO}_2$  和  $\text{NO}_x$  排放量;夏炫昊<sup>[13]</sup>基于长短时记忆(long short-term memory, LSTM)深度神经网络,构建 CFB 锅炉污染物模型,实现对当前时刻  $\text{NO}_x$ 、 $\text{SO}_2$  的浓度计算;Yu 等人<sup>[14]</sup>提出了一种基于 CNN-LSTM-ATTENTION 的混合神经网络模型预测了 CFB 中的运行参数,其中包括  $\text{SO}_2$  和  $\text{NO}_x$  的排放浓度;Chen 等人<sup>[15]</sup>分析了一阶 Taylor 展开下的微分方程模型与 GRU 模型之间的关系,提出了混合模型,预测了  $\text{SO}_2$  的浓度;Han 等人<sup>[16]</sup>提出了一种建立对抗性去噪自编码器(adversarial denoising autoencoder, ADAE)提取火焰深度特征,然后利用最小支持向量回归(least square support vector regression, LSSVR)分析提取特征的模型以预测  $\text{NO}_x$  排放;Wang 等人<sup>[17]</sup>使用 CEEMDAN 算法

与注意力机制(ATTENTION)和 LSTM 神经网络结合形成的深度学习网络来预测  $\text{NO}_x$  浓度。在前人的研究中,许多学者已经探索了使用深度学习进行污染物预测,取得了较好的效果。

尽管许多研究已经使用深度学习进行 CFB 锅炉的污染物预测,但是仍存在一些问题。例如,一些研究使用单一模型导致预测精度不高,而有些研究使用混合模型但未解释算法模型的计算成本以及实用性,还有些研究未对数据进行预处理导致效果不佳。

针对上述问题,本文提出门控循环神经网络(gated recurrent unit, GRU),结合卷积神经网络与多头注意力机制的混合模型(CNN-GRU-MHA)对 CFB 机组污染物浓度进行预测,并通过四分位距法预处理数据进一步提升污染物排放预测性能。同时分析算法模型的复杂度,增强其在 CFB 锅炉污染物预测方面的可行性。结果表明,该混合模型能够有效实现 CFB 机组污染物排放浓度监测,并为其污染物排放控制提供指导。

## 1 模型介绍

### 1.1 卷积神经网络

卷积神经网络主要由输入层、卷积层、激活函数层、池化层和全连接层等部分组成<sup>[18]</sup>。卷积层通过卷积捕捉输入数据中的局部特征信息,并在卷积层之后加入非线性激活函数以增强模型的表达能力。池化层则用于降低特征维度,提高模型的平移不变性和鲁棒性<sup>[19]</sup>。最后,通过全连接层汇总卷积层提取的局部特征,生成最终输出。本文 CNN 模块用于对输入数据进行特征提取。

### 1.2 门控循环神经网络

相比于传统神经网络,门控循环神经网络通过引入更新门和重置门机制,能够有效地学习和利用时间序列数据的长期依赖关系<sup>[20]</sup>。GRU 结构如图 1 所示。

$$Z_t = \sigma(W_z \cdot [H_{t-1}, x_t]) \quad (1)$$

$$R_t = \sigma(W_r \cdot [H_{t-1}, x_t]) \quad (2)$$

$$\tilde{H}_t = \tanh(W \cdot [R_t \cdot H_{t-1}, x_t]) \quad (3)$$

$$H_t = (1 - Z_t) \cdot H_{t-1} + Z_t \cdot \tilde{H}_t \quad (4)$$

式中： $x_t$ 为输入， $H_t$ 为隐状态， $\tilde{H}_t$ 为候选隐状态， $R_t$ 为重置门， $Z_t$ 为更新门， $W$ 为权重， $\sigma$ 为激活函数。

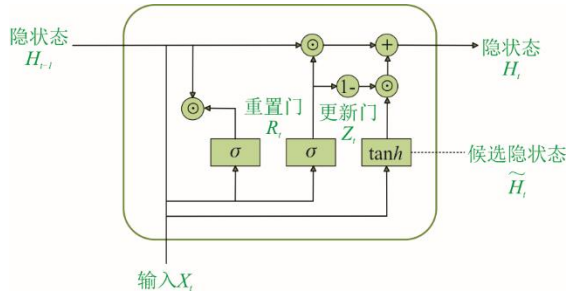


图 1 GRU 结构  
Fig.1 Structural diagram of GRU

GRU 的更新门  $Z_t$  控制着当前时刻的隐藏状态  $H_t$  应该从先前的隐藏状态  $H_{t-1}$  中吸收多少信息<sup>[21]</sup>；而重置门  $R_t$  则控制着当前时刻的隐藏状态  $H_t$  应该从先前的隐藏状态  $H_{t-1}$  中忽略多少信息。GRU 根据当前输入和重置后的先前隐藏状态计算出候选隐藏状态  $\tilde{H}_t$ ，最终的隐藏状态由先前隐藏状态  $H_{t-1}$  和候选隐藏状态  $\tilde{H}_t$  进行加权求和得到。本文中，GRU 可用于建模  $\text{SO}_2$  和  $\text{NO}_x$  浓度随时间的变化趋势，从而捕捉数据中的时间依赖。

### 1.3 多头注意力机制

多头注意力机制可以帮助模型更好地关注输入数据中的关键部分，从而提高模型性能<sup>[22]</sup>。其首先将输入序列进行 3 个不同的线性变换，得到查询（query）、键（key）和值（value）矩阵。然后，对于每个注意力头，计算查询与键的点积，经过 softmax 归一化得到注意力权重。最后，将注意力

权重应用到值矩阵上得到加权输出。本文中，多头注意力机制帮助模型更好地理解输入数据中  $\text{SO}_2$  和  $\text{NO}_x$  浓度变化的时空依赖关系，增强对关键特征的关注，从而提高整体的预测准确性。多头注意力机制结构如图 2 所示。

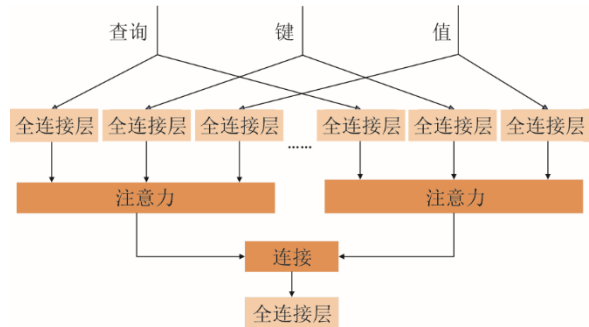


图 2 多头注意力机制结构  
Fig.2 Structural diagram of multi-head attention mechanism

### 1.4 CNN-GRU-MHA 模型设计

本模型的输入为 CFB 机组运行过程中各种传感器数据，如温度、压力、流量等时间序列特征。CNN 部分首先对这些输入数据进行特征提取，利用卷积核在数据上滑动并执行卷积运算，捕捉到局部相关性等特征。将这些经过 CNN 学习的特征进一步输入到 GRU 部分，GRU 通过更新门和重置门，可以准确地刻画  $\text{SO}_2$  和  $\text{NO}_x$  浓度的变化趋势。最后，MHA 部分增强模型对关键时间点或关键运行参数的关注，让模型更好地理解输入数据中  $\text{SO}_2$  和  $\text{NO}_x$  排放的依赖关系。最终，经过全连接层处理，得到对  $\text{SO}_2$  和  $\text{NO}_x$  排放浓度的准确预测。三者结合可以更好地挖掘 CFB 数据中的复杂关系，使预测结果更加精准。其模型结构如图 3 所示。

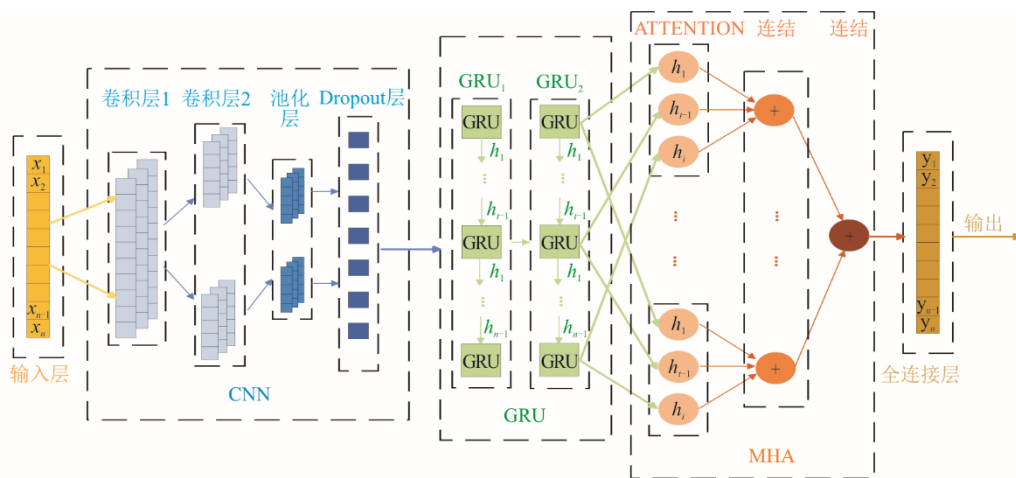


图 3 CNN-GRU-MHA 模型结构  
Fig.3 Structural diagram of the CNN-GRU-MHA model

## 2 案例分析

以某 330 MW CFB 机组作为研究对象,该机组采用炉内炉外相结合的脱硫方式,采用选择性非催化还原(selective non-catalytic reduction, SNCR)脱硝方式,脱硝还原剂为尿素溶液。分别对 SO<sub>2</sub> 和 NO<sub>x</sub> 进行稳态以及变负荷工况下的预测。其中数据集的采样时间为 30 s。

### 2.1 预测流程

选取 CFB 锅炉的污染物 SO<sub>2</sub> 和 NO<sub>x</sub> 作为研究对象,使用 Pearson 相关系数,计算各传感器测点数据与两者的关联度,从而选择相关性高的数据作为输入。然后,对输入数据进行四分位距(interquartile range, IQR)数据预处理,筛选并替换异常值后进行归一化处理,从而提升预测效果。随后,按比例划分训练数据和测试数据,将训练数据送入模型中计算损失函数,并通过 Adam 优化器模块对模型中的参数进行优化。在每一步训练过程中对模型进行测试,计算相应的平均绝对误差(mean absolute error, MAE)、平均绝对百分比误差(mean absolute percentage error, MAPE)、R<sup>2</sup>,保存最佳模型权重以及预测结果,最终反归一化并输出。预测流程如图 4 所示。

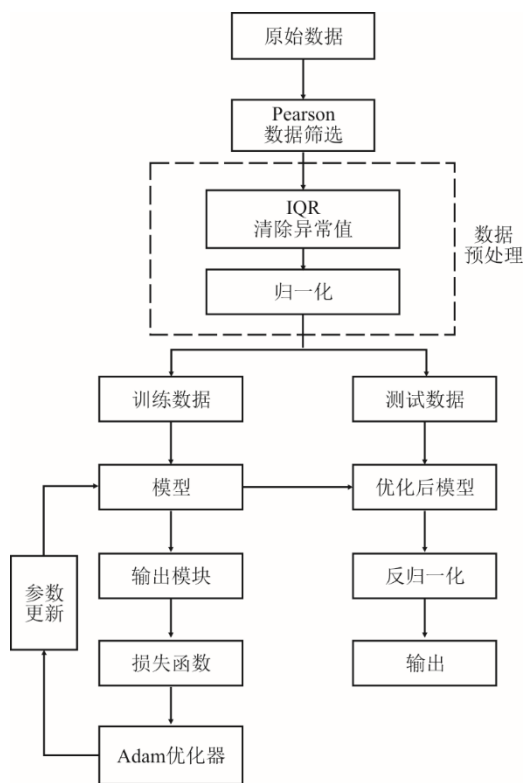


图 4 预测流程

Fig.4 Prediction flow chart

### 2.2 参数选取

Pearson 相关系数可以用来评估 2 个变量之间的线性相关程度<sup>[23]</sup>,计算公式为:

$$\rho_{x,y} = \frac{E(xy) - E(x)E(y)}{\sqrt{E(x^2) - E^2(x)}\sqrt{E(y^2) - E^2(y)}} \quad (5)$$

式中:  $x$ 、 $y$  分别为输入变量和目标变量;  $E(x)$ 、 $E(y)$ 、 $E(xy)$  分别表示  $x$ 、 $y$ 、 $xy$  的期望;  $\rho_{x,y}$  为  $x$ 、 $y$  之间的 Pearson 相关系数。

本文使用 Pearson 相关系数筛选输入变量,可构建更加精准高效的 CFB 锅炉污染物预测模型。图 5 为 SO<sub>2</sub> 和 NO<sub>x</sub> 与其他测点数据之间的 Pearson 相关系数热力图。其中第 1 行是 NO<sub>x</sub> 与其他测点之间的 Pearson 相关系数,最后一行是 SO<sub>2</sub> 与其他测点之间的 Pearson 相关系数。

Pearson 相关系数的取值范围为[-1,1],当 Pearson 相关系数等于 1 时表示完全正相关, Pearson 相关系数等于-1 时表示完全负相关<sup>[24]</sup>。通过 Pearson 相关系数筛选出与 NO<sub>x</sub> 和 SO<sub>2</sub> 具有较强线性相关性的输入特征(即 Pearson 相关系数更接近于-1 或 1 的变量),这些特征包含有价值的预测信息,有助于提高模型的准确性。因此,输入变量选取见表 1。

表 1 输入变量选取  
Tab.1 Selections of input variables

| 输入变量                   | 是否选择            |                 |
|------------------------|-----------------|-----------------|
|                        | NO <sub>x</sub> | SO <sub>2</sub> |
| 原烟气温度                  | 否               | 否               |
| 原烟气压力                  | 否               | 是               |
| 原烟气 SO <sub>2</sub> 浓度 | 否               | 是               |
| 尿素溶液流量                 | 是               | 否               |
| 总风量                    | 是               | 是               |
| 一次风总量                  | 否               | 是               |
| 二次风总量                  | 是               | 是               |
| 尿素输送频率                 | 是               | 否               |
| 风煤比                    | 是               | 是               |
| 负荷                     | 否               | 是               |
| 炉膛出口温度                 | 是               | 是               |
| 床温均值                   | 是               | 是               |
| 净烟气 O <sub>2</sub> 浓度  | 是               | 是               |

### 2.3 四分位距

本文选取的数据段质量有限,因此需要对原始数据进行预处理。首先使用 IQR 法检测并清除一些异常值<sup>[25]</sup>,如图 6 所示。

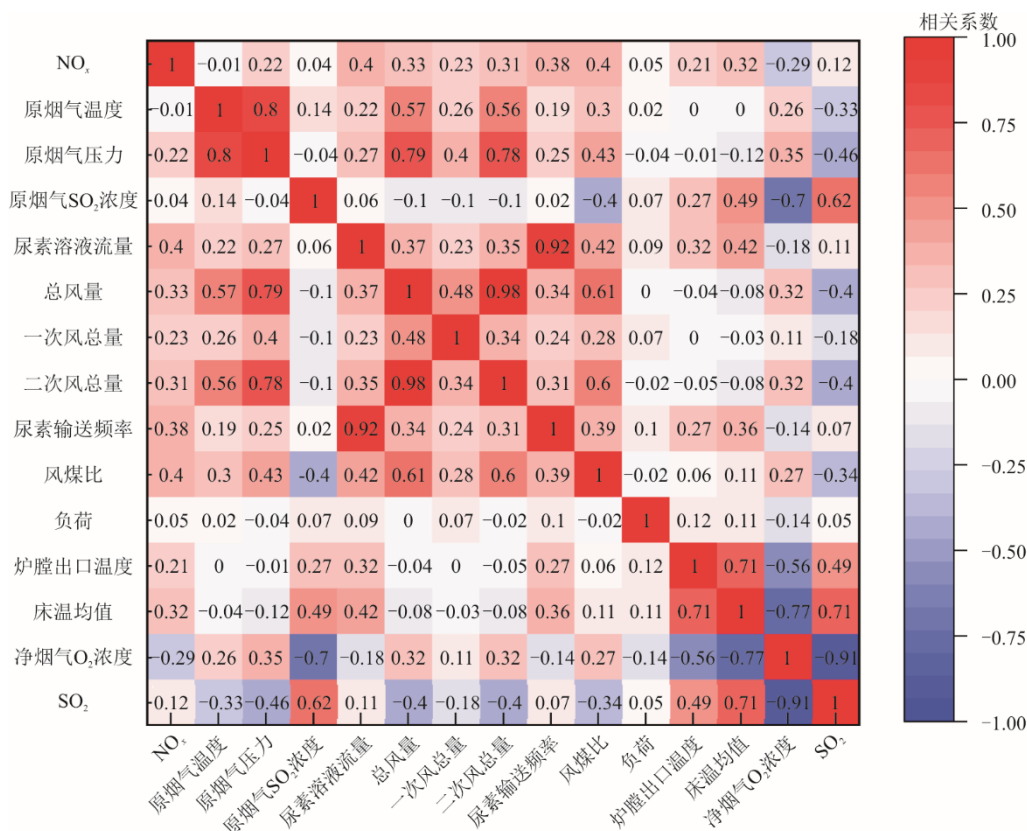


图 5 Pearson 相关系数热力图  
Fig.5 Pearson coefficient heat map

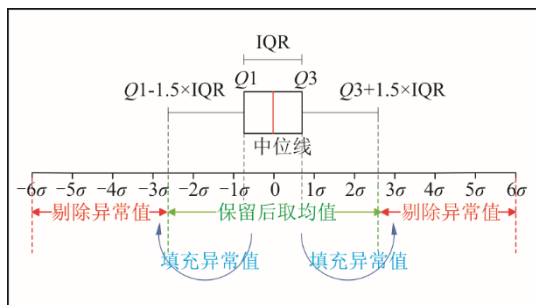


图 6 使用 IQR 法清除异常值  
Fig.6 Clearing the outliers using IQR method

对数据按滑动窗口数进行划分，本文滑动窗口数设置为 21。对每个窗口列表从小到大排序，然后四等分数据段。1/4 位置的点为  $Q_1$ ，同理得到  $Q_2$ 、 $Q_3$ 。其中  $Q_2$  作为中点。令  $Q_3$  和  $Q_1$  之间的差为  $\Delta$ ，有效数据的上限为  $Q_3+1.5\Delta$ ，下限为  $Q_1-1.5\Delta$ ，在  $(Q_1-1.5\Delta, Q_3+1.5\Delta)$  之外的数据则视为异常值。随后将异常值剔除，并对上、下限之内的数据取均值并对异常值位置进行填充。对于 CFB 锅炉的部分测点经常出现异常值等情况，选用该方法能够很好地剔除部分无效数据，进而减少噪声数据对模型训练的干扰，使预测模型不易受极端值的影响。

### 2.4 归一化

由于各参数值量级不同，将影响模型的收敛速度、性能、泛化能力等。因此，在进入模型之前，采用 Min-Max 方法对原始数据进行归一化，将数据值映射在 0 到 1 之间，公式如下：

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (6)$$

式中： $X'$ 为归一化后的值， $X$ 为原始值， $X_{\min}$ 为数据集中的最小值， $X_{\max}$ 数据集中的最大值。

### 2.5 数据集划分

CFB 机组在静态工况和动态工况下运行参数和排放特点存在明显差异。在这 2 种典型工况下都进行污染物预测，可以使预测更全面。同时，在静态工况下可以更好地理解 CFB 机组在稳定负荷下  $SO_2$  和  $NO_x$  的排放特征。而动态工况下，可以分析这些污染物在变负荷时的排放规律，揭示其时间动态特性。数据集划分如图 7 所示。

在众多数据集中选择 4 620 个数据点(数据点时间间隔 30 s)，其负荷变化如图 7 所示。选择前 3 780 个点作为训练集，第 3 781~4 200 个数据点作为变负荷工况下的测试集，此段数据集包含一段连

续降负荷再升负荷数据。随后,第4 201~4 620 个数据点作为静态工况下的测试集,此段数据稳定在300 MW 运行。

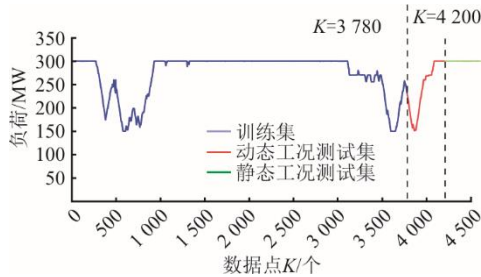


图7 数据集划分  
Fig.7 Data set partitioning

## 2.6 模型网络参数设置

不同工况下  $\text{SO}_2$  和  $\text{NO}_x$  模型网络结构参数的设置需根据实验比较获得最优参数。其中特征维数取值取决于输入变量个数,由 Pearson 相关系数计算后  $\text{SO}_2$  实验取 10 个变量,  $\text{NO}_x$  实验取 8 个变量;由于数据集不大,为避免过拟合问题,通过实验对比,  $\text{SO}_2$  与  $\text{NO}_x$  预测实验中 GRU 层隐藏状态维数均选取 32, 32 维隐含层通常能有效捕捉到所选输入数据的复杂性,平衡模型的复杂度和计算效率。多层 GRU 可以更深入地学习污染物数据中的复杂模式,但层数的选择受到计算资源的限制,GRU 层数均选取 4, 4 层 GRU 通常能够保证在较好的计算效率时捕获序列中较深的时间依赖关系。多头注意力头数过高会增加计算复杂度和内存需求,因此考虑实验中使用的硬件资以及数据集大小,选取头数 3 较为合适。相较于 SGD 等优化器,本文采用 Adam 优化器通常可使用较大的学习率,但学习率较大可能会导致收敛速度过快,损失不稳定,因此学习率设置为 0.001。较大的批尺寸可以加快训练速度,但可能降低收敛速度,通过实验尝试批尺寸设置为 64 能达到较好的训练效果。考虑到可用的计算资源和时间,选取较小的训练轮数 25,可在保证预测效果的同时,降低计算时间。各参数设置见表 2。

表 2 超参数设置  
Tab.2 Hyperparameter settings

| 超参数设置       | $\text{SO}_2$ | $\text{NO}_x$ |
|-------------|---------------|---------------|
| 特征维度        | 10            | 8             |
| GRU 层隐藏状态维度 | 32            | 32            |
| GRU 层数      | 4             | 4             |
| 多头注意力头数     | 3             | 3             |
| 学习率         | 0.001         | 0.001         |
| 批尺寸         | 64            | 64            |
| 训练轮数        | 25            | 35            |

## 2.7 实验环境搭建

操作系统选择 Windows10, GPU 选择 NVIDIA GTX1660Ti, CPU 选择 Intel(R) Core(TM) i5-9300H, 内存 16 GB, 编程语言 Python3.9, 深度学习框架 TensorFlow2.16+Keras3.1。

## 2.8 复杂度分析

在评估算法模型复杂度时,通常使用  $O$  表示法来描述复杂度的增长率,忽略常数和低阶项。例如,  $O(n^2)$  表示比  $O(n)$  增长得更快。本文使用混合模型,其复杂度为各部分复杂度之和。其中时间复杂度以及空间复杂度计算公式如下<sup>[26]</sup>:

$$O_T = O\{(T - K + 1) \times K \times D_1 \times C_{out} + T' \times (3H \times C_{out} + 3H^2) + T''^2 \times D_2\} \quad (7)$$

$$O_S = O\{(T - K + 1) \times C_{out} + H \times C_{out} + 3H^2 + T'' \times D_2\} \quad (8)$$

式中:  $T$  为输入序列长度,  $K$  为卷积核大小,  $D_1$  为输入特征维度,  $C_{out}$  为输出特征维度,  $H$  为 GRU 隐含层维数,  $T'$  为经过 CNN 处理后时间序列长度,  $T''$  为 GRU 输出序列长度  $D_2$  为 MHA 各头的维数总和,  $O_T$  为本文模型的时间复杂度,  $O_S$  为本文模型的空间复杂度。

由式(7)可知,该模型时间复杂度取最高阶结果则为  $O(n^2)$ , 其中包含 2 个平方项  $O(H^2)$  与  $O(T'^2)$ , 由于本文 GRU 隐含层数  $H$  设置为 32, 而 GRU 输出序列长度  $T'$  主要由输入序列长度决定, 因此  $T'$  比  $H$  大, 所以模型的时间复杂度为  $O(T'^2)$ 。由此可见,算法时间复杂度偏大,但是在达到良好预测效果的要求下,所选数据集并不复杂,因此算法运行时间较短,在对 CFB 机组污染物预测的实际案例中较为实用。

由式(8)可知,该模型空间复杂度取最高阶则结果为  $O(n^2)$  (其中为  $n$  为 GRU 隐含层维数  $H$ )。所以模型的时间复杂度为  $O(H^2)$ 。由于 GRU 隐含层维数  $H$  设置为 32, 空间复杂度偏小,因此在整个运行过程中占据的内存空间较小,其运行效率较高,也符合搭建的实验环境。

## 2.9 模型评估参数

在模型误差分析中,使用平均绝对误差 MAE、平均绝对百分比误差 MAPE 和决定系数  $R^2$  作为评判指标。其中, MAE 衡量预测值与真实值之间的绝对差值的平均大小,而 MAPE 是一个无量纲指标,表示预测值相对于真实值的平均百分比偏差,

$R^2$  则反映模型预测值与真实值之间的线性相关程度, 取值范围为[0, 1], 越接近 1 表示模型预测性能越强。三者公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (9)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (10)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^n (\bar{y}_i - \bar{y}_i)^2} \quad (11)$$

式中:  $\hat{y}$  为预测值,  $y_i$  为真实值,  $\bar{y}_i$  为均值,  $n$  为评估值个数。

### 3 结果分析

该实验分为收敛性验证、数据预处理结果对比以及  $NO_x$  和  $SO_2$  在稳态和变负荷工况下的预测。经过消融及对比实验, 得出 CNN-GRU-MHA 预测 CFB 机组的  $SO_2$  和  $NO_x$  排放具有优越性。

#### 3.1 收敛性分析

损失值是评估模型优劣的指标, 由其曲线可看出模型的收敛性。以变负荷工况  $SO_2$  实验中的损失值数据进行展示 (4 组实验结果相似), 结果如图 8 所示。

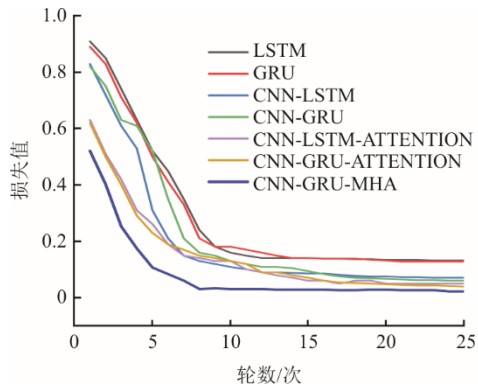


图 8 变负荷工况  $SO_2$  实验损失值对比

Fig.8 Comparison of loss values in  $SO_2$  test under varying load conditions

由图 8 可知, 随着训练轮数的增加, 本文提出的 CNN-GRU-MHA 模型损失值快速下降并趋于稳定状态, 相较于其他模型, 该模型的收敛速度更快且损失值最低。

#### 3.2 数据预处理效果分析

使用 IQR 筛选并替换异常值的预处理方式有

助于提高模型预测精度。使用 CNN-GRU-MHA 作为主体进行建模实验, 含有 IQR 预处理以及不含 IQR 预处理的变负荷  $NO_x$  预测对比结果如图 9 所示。由图 9 可看出, 在 4 000~5 000 s 间, 未使用 IQR 数据预处理的 CNN-GRU-MHA 主体模型的预测效果不佳。检查数据集发现, 原始输入数据中存在部分数据异常等情况。但是经过 IQR 预处理之后, 会消除部分输入数据的异常值, 降低其对预测结果的影响。因此, 在进行 CFB 机组污染物预测的过程中, 数据的预处理十分必要, 而本文提出的 IQR 检测并替换异常值的方法也是可行方法。

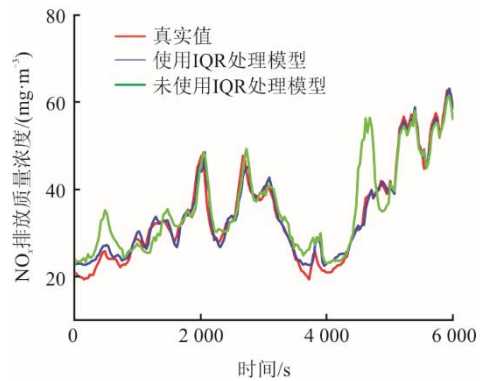


图 9 (不) 含 IQR 预处理的变负荷  $NO_x$  预测结果对比  
Fig.9 The  $NO_x$  prediction results under varying load conditions with and without IQR data preprocessing

#### 3.3 稳态工况分析

对于稳态工况下的  $NO_x$  预测, 此时负荷稳定在 300 MW, 选择 CNN-GRU-MHA 与其他模型对比, 选取的测试集的 200 个点 (数据点时间间隔 30 s) 的结果如图 10 所示。由图 10 可以看出, 在 0~500 s 和 1 500~2 000 s 时, CNN-GRU-MHA 预测效果明显优于 LSTM 与 GRU 单一网络。同时在 4 000~6 000 s 时, CNN-GRU-MHA 的预测结果明显优于其他模型。

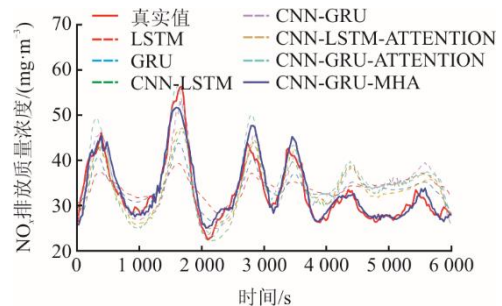


图 10 稳态工况  $NO_x$  预测对比

Fig.10 The  $NO_x$  prediction results under static condition

稳态工况下各模型  $NO_x$  预测结果的 MAE、MAPE、 $R^2$  见表 3。由表 3 可看出: 相较于单一模

型 LSTM、GRU，本文模型的 MAE 等参数均有很大提升；与其他混合模型相比，其评估参数也占优，特别是相较于 CNN-GRU-ATTENTION 模型，本文模型 MAE 降低了  $1.367\ 3\ \text{mg}/\text{m}^3$ ，MAPE 降低了 39.56%， $R^2$  提升了 20.50%。

表 3 稳态工况各模型  $\text{NO}_x$  预测结果  
Tab.3 The  $\text{NO}_x$  prediction results of each model under steady state condition

| 模型                 | MAE/( $\text{mg}\cdot\text{m}^{-3}$ ) | MAPE/% | $R^2$   |
|--------------------|---------------------------------------|--------|---------|
| LSTM               | 4.549 5                               | 11.97  | 0.604 3 |
| GRU                | 3.838 7                               | 9.87   | 0.686 3 |
| CNN-LSTM           | 4.086 6                               | 10.63  | 0.663 3 |
| CNN-GRU            | 3.332 2                               | 9.05   | 0.765 1 |
| CNN-LSTM-ATTENTION | 3.617 3                               | 9.38   | 0.728 0 |
| CNN-GRU-ATTENTION  | 3.757 3                               | 9.91   | 0.702 5 |
| CNN-GRU-MHA        | 2.390 0                               | 5.99   | 0.846 5 |

对于稳态工况下的  $\text{SO}_2$  预测，此时负荷稳定在 300 MW，CNN-GRU-MHA 与其他模型预测结果对比如图 11 所示。由图 11 可看出，在 2 200~2 800 s 时，CNN-GRU-MHA 预测结果明显优于单一模型 LSTM、GRU，同时在整体结果中，CNN-GRU-MHA 的预测结果仍优于其他模型。

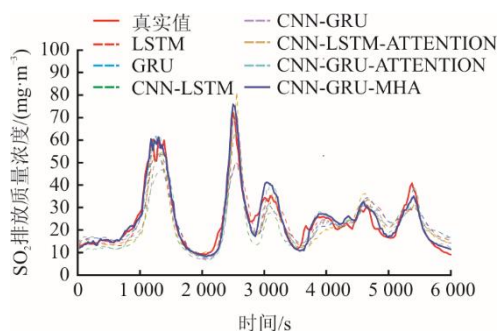


图 11 稳态工况  $\text{SO}_2$  预测对比

Fig.11 The  $\text{SO}_2$  prediction results under static condition

稳态工况下各模型  $\text{SO}_2$  预测结果的 MAE、MAPE、 $R^2$  见表 4。

表 4 稳态工况各模型  $\text{SO}_2$  预测结果  
Tab.4 The  $\text{SO}_2$  prediction results of each model under steady state condition

| 模型                 | MAE/( $\text{mg}\cdot\text{m}^{-3}$ ) | MAPE/% | $R^2$   |
|--------------------|---------------------------------------|--------|---------|
| LSTM               | 4.303 6                               | 20.89  | 0.709 0 |
| GRU                | 4.472 1                               | 22.26  | 0.698 8 |
| CNN-LSTM           | 4.345 0                               | 20.36  | 0.719 8 |
| CNN-GRU            | 4.244 9                               | 19.67  | 0.713 6 |
| CNN-LSTM-ATTENTION | 3.933 8                               | 17.56  | 0.751 3 |
| CNN-GRU-ATTENTION  | 3.084 9                               | 14.59  | 0.859 2 |
| CNN-GRU-MHA        | 2.805 9                               | 13.12  | 0.878 8 |

由表 4 可看出，相较于单一的 LSTM、GRU 模型，本文模型的 MAE 等评估参数均有很大提升，与其他混合模型如 CNN-LSTM 及 CNN-GRU 相比，其评估参数也占优，尤其相较于 CNN-GRU-ATTENTION 模型，本文模型 MAE 降低了  $0.279\ 0\ \text{mg}/\text{m}^3$ ，MAPE 降低了 10.08%， $R^2$  提升了 2.28%。

### 3.4 变负荷工况分析

对于变负荷工况下的  $\text{NO}_x$  预测，选择 CNN-GRU-MHA 与其他模型对比，选取的测试集的 200 个点（数据点时间间隔 30 s）的结果如图 12 所示。在变负荷过程中， $\text{NO}_x$  排放质量浓度波动较大，预测难度同时加大。在 0~3 000 s 间，机组快速降负荷， $\text{NO}_x$  排放质量浓度明显上升，单靠单一模型 LSTM 或 GRU 模型来预测，精度较差，使用其他混合模型精度也不高。但由图 12 可看出，不论是升负荷过程还是降负荷过程，CNN-GRU-MHA 模型的预测都较为准确，相对于其余模型具有较大优势，特别是在 5 000~6 000 s 时，CNN-GRU-MHA 模型能够很好地拟合真实的  $\text{NO}_x$  排放。

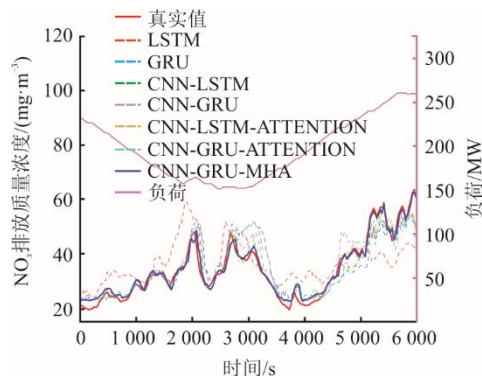


图 12 变负荷工况  $\text{NO}_x$  预测对比

Fig.12 The  $\text{NO}_x$  prediction results under variable load conditions

变负荷工况下各模型  $\text{NO}_x$  预测结果的 MAE、MAPE、 $R^2$  见表 5。由表 5 可看出，LSTM 与 GRU 预测精度较差，加入 CNN 后，精度有所提升，但与 CNN-LSTM-ATTENTION、CNN-GRU-ATTENTION 相比仍有差距，而本文 CNN-GRU-MHA 模型最优。与 CNN-GRU-ATTENTION 模型相比，本文 CNN-GRU-MHA 模型加入多头注意力机制后，表现力更强，其中 MAE 降低了  $1.062\ 1\ \text{mg}/\text{m}^3$ ，MAPE 降低了 37.44%， $R^2$  提升了 18.85%。

对于变负荷工况下的  $\text{SO}_2$  预测，选择 CNN-GRU-MHA 与其他模型对比，选取的测试集的 200 个点（数据点时间间隔 30 s）的结果如图 13 所示。在降

负荷初期, SO<sub>2</sub> 排放质量浓度有波动, 特别是在 0~500 s 时, CNN-GRU-MHA 模型拟合程度高于其他模型。在 3 000~6 000 s 时, 机组快速升负荷, SO<sub>2</sub> 排放质量浓度波动较大, 特别是在 3 500~4 500 s 时, 相较于其他模型, CNN-GRU-MHA 模型能够更好地拟合真实值。

表 5 变负荷工况各模型 NO<sub>x</sub> 预测结果  
Tab.5 The NO<sub>x</sub> prediction results of each model under variable load conditions

| 模型                 | MAE/(mg·m <sup>-3</sup> ) | MAPE/% | R <sup>2</sup> |
|--------------------|---------------------------|--------|----------------|
| LSTM               | 9.300 7                   | 23.48  | 0.343 3        |
| GRU                | 6.299 1                   | 14.71  | 0.502 5        |
| CNN-LSTM           | 5.97 6                    | 13.76  | 0.512 3        |
| CNN-GRU            | 5.671 0                   | 12.48  | 0.517 9        |
| CNN-LSTM-ATTENTION | 4.270 0                   | 8.88   | 0.690 6        |
| CNN-GRU-ATTENTION  | 4.022 7                   | 8.52   | 0.720 5        |
| CNN-GRU-MHA        | 2.960 6                   | 5.33   | 0.856 3        |

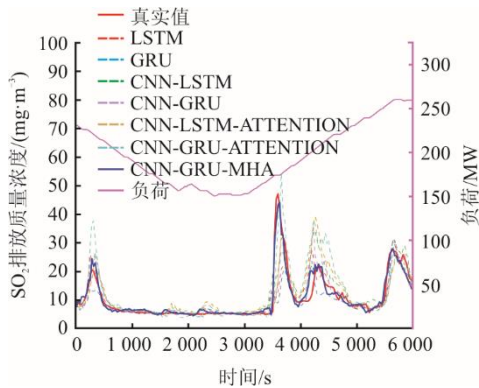


图 13 变负荷工况 SO<sub>2</sub> 预测对比

Fig.13 The SO<sub>2</sub> prediction results under variable load conditions

变负荷工况下各模型 SO<sub>2</sub> 预测结果的 MAE、MAPE、R<sup>2</sup> 见表 6。

表 6 变负荷工况各模型 SO<sub>2</sub> 预测结果  
Tab.6 The SO<sub>2</sub> prediction results of each model under variable load conditions

| 模型                 | MAE/(mg·m <sup>-3</sup> ) | MAPE/% | R <sup>2</sup> |
|--------------------|---------------------------|--------|----------------|
| LSTM               | 6.624 0                   | 29.93  | 0.625 8        |
| GRU                | 5.575 6                   | 27.34  | 0.636 9        |
| CNN-LSTM           | 3.596 8                   | 25.17  | 0.796 6        |
| CNN-GRU            | 2.910 1                   | 17.45  | 0.816 0        |
| CNN-LSTM-ATTENTION | 3.746 7                   | 24.79  | 0.747 0        |
| CNN-GRU-ATTENTION  | 3.407 7                   | 22.57  | 0.798 5        |
| CNN-GRU-MHA        | 2.334 8                   | 13.52  | 0.876 0        |

由表 6 可看出, 相较于单一的 LSTM、GRU 模型, 本文模型的 MAE、MAPE、R<sup>2</sup> 等评估参数均有很大提升, 与其他混合模型相比, 也有较大优势, 特别是与 CNN-GRU-ATTENTION 模型相比, 本文 CNN-GRU-MHA 模型的 MAE 降低了 1.072 9 mg/m<sup>3</sup>, MAPE 降低了 40.10%, R<sup>2</sup> 提升了 9.71%, 预测效果良好。

## 4 结 论

1) 本文构建了 CNN-GRU-MHA CFB 机组污染物排放质量浓度动态预测模型。CNN 提取输入特征的空间相关性模型可以更好地泛化到新的工况条件或环境因素, 提高了应用的广泛性。GRU 模块能够有效地建模污染物排放质量浓度时间序列的长期依赖关系, 为预测模型提供了强大的时序建模能力。MHA 可以学习到更丰富和复杂的注意力权重分布, 从而增强模型的表现力和表达能力。三者结合应用于 CFB 机组的污染物预测, 效果优良, 可为之后电厂的优化与控制工作提供依据。

2) 通过损失值对比、数据预处理效果对比、稳态与变负荷工况下污染物排放预测效果对比, 验证了 CNN-GRU-MHA 模型具有良好的收敛性、预测准确性, 且预处理操作十分必要。

3) 通过从时间和空间两方面分析模型的复杂度, 证明了 CNN-GRU-MHA 模型应用在 CFB 污染物预测方面的可行性。然而, 由于该模型仅分别从单一模块简单介绍了对数据处理的解释性, 整体的解释性不强、透明度不高。此问, 下一阶段将更加关注模型的透明度与解释性。

### [参 考 文 献]

[1] 姚禹歌, 黄中, 张曼, 等. 中国 CFB 燃烧技术的发展与展望[J]. 热力发电, 2021, 50(11): 13-19.  
YAO Yuge, HUANG Zhong, ZHANG Man, et al. Development and prospects of circulating fluidized bed combustion technology in China[J]. Thermal Power Generation, 2021, 50(11): 13-19.

[2] 《全面实施燃煤电厂超低排放和节能改造工作方案》[J]. 节能与环保, 2016(1): 32.  
Work program for the comprehensive implementation of ultra-low emission and energy-saving retrofit of coal-fired power plants[J]. Energy Conservation & Environmental Protection, 2016(1): 32.

[3] 张洪福. CFB 机组负荷响应特性与灵活运行控制研究[D]. 北京: 华北电力大学, 2023: 1.  
ZHANG Hongfu. Study on load response characteristics and flexible operation control of circulating fluidized bed unit[D]. Beijing: North China Electric Power University, 2023: 1.

[4] 高明明, 郭炯楠, 于浩洋, 等. 适应深度调峰的

- CFBNO<sub>x</sub> 排放建模[J]. 洁净煤技术, 2023, 29(6): 24-31.  
GAO Mingming, GUO Jiongnan, YU Haoyang, et al. Modeling NO<sub>x</sub> emissions from circulating fluidized beds adapted to deep peaking[J]. Clean Coal Technology, 2023, 29(6): 24-31.
- [5] 杨敬池, 王菁, 王鹏程, 等. 超临界 350 MW 循环流化床锅炉宽负荷 NO<sub>x</sub> 排放特性及控制技术[J]. 洁净煤技术, 2025, 31(1): 78-85.  
YANG Jingchi, WANG Jing, WANG Pengcheng, et al. Characterization and mitigation of NO<sub>x</sub> emissions across a wide load range in supercritical 350 MW circulating fluidized bed boiler[J]. Clean Coal Technology, 2025, 31(1): 78-85.
- [6] YUE G, CAI R, LU J, et al. From a CFB reactor to a CFB boiler—The review of R&D progress of CFB coal combustion technology in China[J]. Powder Technology, 2017, 316: 18-28.
- [7] KACHAVIMATH A V, NAZARE S V, AKKI S S. Distributed denial of service attack detection using naïve bayes and k-nearest neighbor for network forensics[C]// Proceedings of the 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), F, 2020. IEEE.
- [8] DONG S. Multi class SVM algorithm with active learning for network traffic classification[J]. Expert Systems with Applications, 2021, 176: 114885.
- [9] ELMRABIT N, ZHOU F, LI F, et al. Evaluation of machine learning algorithms for anomaly detection[C]// Proceedings of the 2020 International Conference on Cyber Security and Protection of Digital Services (Cyber Security), F, 2020. IEEE.
- [10] 张砺彦, 麦永强, 张月辉, 等. 基于神经网络的大型燃煤电厂 SO<sub>2</sub> 和 NO<sub>x</sub> 的污染物预测模型[J]. 热力发电, 2008, 37(10): 9-13.  
ZHANG Liyan, MAI Yongqiang, ZHANG Yuehui, et al. Neural network-based pollutant prediction model for SO<sub>2</sub> and NO<sub>x</sub> in large coal-fired power plants[J]. Thermal Power Generation, 2008, 37(10): 9-13.
- [11] LIU L, WANG P, LIN J, et al. Intrusion detection of imbalanced network traffic based on machine learning and deep learning[J]. IEEE Access, 2020, 9: 7550-7563.
- [12] ADAMS D, OH D H, KIM D W, et al. Prediction of SO<sub>x</sub>-NO<sub>x</sub> emission from a coal-fired CFB power plant with machine learning: plant data learned by deep neural network and least square support vector machine[J]. Journal of Cleaner Production, 2020, 270: 122310.
- [13] 夏炫昊. CFB 机组污染物模型研究与优化[D]. 北京: 华北电力大学, 2023: 1.  
XIA Xuanhao. Pollutant modeling and optimization of circulating fluidized bed units[D]. Beijing: North China Electric Power University, 2023: 1.
- [14] YU H, GAO M, ZHANG H, et al. Data-driven optimization of pollutant emission and operational efficiency for circulating fluidized bed unit[J]. Energy, 2023, 281: 128338.
- [15] CHEN J, GAO M, ZHANG H, et al. Dynamic prediction of SO<sub>2</sub> emission based on hybrid modeling method for coal-fired circulating fluidized bed[J]. Fuel, 2023, 346: 128284.
- [16] HAN Z, XIE Y, HOSSAIN M M, et al. A hybrid deep neural network model for NO<sub>x</sub> emission prediction of heavy oil-fired boiler flames[J]. Fuel, 2023, 333: 126419.
- [17] WANG X, LIU W, WANG Y, et al. A hybrid NO<sub>x</sub> emission prediction model based on CEEMDAN and AM-LSTM[J]. Fuel, 2022, 310: 122486.
- [18] SHIN H C, ROTH H R, GAO M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1285-1298.
- [19] 张浩, 刘振娟, 李宏光, 等. 基于关联变量时滞分析卷积神经网络的生产过程时间序列预测方法[J]. 化工学报, 2017, 68(9): 3501-3510.  
ZHANG Hao, LIU Zhenjuan, LI Hongguang. Process time series prediction based on application of correlated process variables to CNN time delayed analyses[J]. CIESC Journal, 2017, 68(9): 3501-3510.
- [20] GRAVES A, MOHAMED A R, HINTON G. Speech recognition with deep recurrent neural networks[C]// Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, F, 2013. IEEE.
- [21] SAJJAD M, KHAN Z A, ULLAH A, et al. A novel CNN-GRU-based hybrid approach for short-term residential load forecasting[J]. IEEE Access, 2020, 8: 143759-143768.
- [22] 韩玉民, 郝晓燕. 基于子词嵌入和相对注意力的材料实体识别[J]. 计算机应用, 2022, 42(6): 1862-1868.  
HAN Yumin, HAO Xiaoyan. Material entity recognition based on subword embedding and relative attention[J]. Journal of Computer Applications, 2022, 42(6): 1862-1868.
- [23] COHEN I, HUANG Y, CHEN J, et al. Pearson correlation coefficient[J]. Noise Reduction in Speech Processing, 2009, 18: 1-4.
- [24] 王雪松, 谷阳阳, 程玉虎. 基于复杂网络的时延基因调控网络构建[J]. 电子学报, 2010, 38(11): 2518-2522.  
WANG Xuesong, GU Yangyang, CHENG Yuhu. Construction of delay gene regulatory network based on complex network[J]. Acta Electronica Sinica, 2010, 38(11): 2518-2522.
- [25] WHALEY III D L. The interquartile range: theory and estimation[D]. East Tennessee State University, 2005: 1.
- [26] 郭喜峰, 王凯泽, 单丹, 等. 多角度基于 CEEMDAN-CNN-BiLSTM 模型的锂离子电池 RUL 预测[J]. 太阳能学报, 2024, 45(7): 181-189.  
GUO Xifeng, WANG Kaize, SHAN Dan, et al. RUL prediction for lithium ion batteries based on CEEMDAN-CNN-BiLSTM model from multiple perspectives[J]. Acta Energetica Solaris Sinica, 2024, 45(7): 181-189.

(责任编辑 李园)