

文章编号:1672-3961(2026)01-0169-10

DOI:10.6040/j.issn.1672-3961.0.2024.194

基于 D-Mamba 模型的超短期火电机组发电负荷预测

王新建¹, 景志滨¹, 孟凡成¹, 石建国¹, 张敏昊¹, 张一帆¹, 王庆华^{2*}, 朱彦恺²

(1. 内蒙古电力(集团)有限责任公司内蒙古电力调度控制分公司, 内蒙古 呼和浩特 010010; 2. 华北电力大学控制与计算机工程学院, 北京 102206)

摘要:随着新型电力系统逐渐完善,火电逐步从基荷电源向调峰电源转变。在新的发电背景下,火电机组面临着越来越严峻的运行考验和各类响应指标考核。超短期机组发电负荷预测需要考虑机组运行状态,以评估机组在未来超短期内的变负荷能力,精准的超短期机组发电负荷预测能够有效表征机组动态性能指标,有利于操作人员做出实时运行调整。针对投入自动发电量控制(auto generation control, AGC)运行方式的机组而言,AGC 指令在发电负荷预测任务中起关键作用。因此,本研究在 Mamba 模型的基础上围绕 AGC 指令构建动态修正模块,实现对超短期机组负荷的精准预测。通过实际机组运行数据验证模型的预测精度。预测结果表明,本研究提出的预测模型能够实现更加精准的负荷预测。

关键词:火电机组;深度学习;负荷预测;状态空间模型;自动发电量控制

中图分类号:TM621.6 **文献标志码:**A

引用格式:王新建,景志滨,孟凡成,等. 基于 D-Mamba 模型的超短期火电机组发电负荷预测[J]. 山东大学学报(工学版), 2026, 56(1): 169-178.

WANG Xinjian, JING Zhibin, MENG Fancheng, et al. Ultra-short-term thermal power unit load forecasting based on D-Mamba model [J]. Journal of Shandong University (Engineering Science), 2026, 56(1):169-178.

Ultra-short-term thermal power unit load forecasting based on D-Mamba model

WANG Xinjian¹, JING Zhibin¹, MENG Fancheng¹, SHI Jianguo¹, ZHANG Minhao¹, ZHANG Yifan¹, WANG Qinghua^{2*}, ZHU Yankai²

(1. Inner Mongolia Electric Power (Group) Co., Ltd., Inner Mongolia Electric Power Dispatching Control Branch, Hohhot 010010, Inner Mongolia, China; 2. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: China progressively promoted the transition toward a new power system, with thermal power shifting from a base-load to a peak-load source. In this evolving power-generation context, thermal power units faced stricter challenges in unit tests and diverse response assessments. Ultra-short-term load forecasting for power units needed to account for the operating state of the unit to accurately assess its near-future load adjustment capability. Precise ultra-short-term load forecasting was critical for revealing a unit's dynamic performance indicators and assisting real-time operational adjustments. For units equipped with Auto-Generation Control (AGC), the AGC command served as a pivotal factor in forecasting power-generation loads. Hence, leveraging the Mamba model, this paper introduced a dynamic correction module centered on the AGC command, thereby ensuring precise ultra-short-term unit load forecasting. The model's performance was validated using actual operational data from a power unit, demonstrating its capability to achieve more accurate load forecasts.

Keywords: thermal power unit; deep learning; load forecasting; state space models; auto generation control

收稿日期:2024-08-04

基金项目:内蒙古电力(集团)有限责任公司科技资助项目(内电科创[2024]5号)

第一作者简介:王新建(1979—),男,内蒙古乌海人,高级工程师,硕士,主要研究方向为电网调度运行与网源协调优化。

E-mail:a18810215030@163.com

*通信作者简介:王庆华(1987—),男,山东临沂人,副研究员,博士,主要研究方向为发电过程建模与控制、先进能源系统分析与优化、智能发电技术。E-mail:a13439827134@163.com

0 引言

随着“双碳”目标的推进,以风能、太阳能为主的可再生能源在全国范围内迅速发展^[1]。随着可再生能源具有波动性、间歇性及随机性的特征,其大规模并网会对电网的安全稳定运行带来挑战。随着可再生能源的发电占比不断增加,火电逐步从基荷电源向调峰电源转变,火电机组因此面临着更为极端的工况场景和更为严格的运行考核。借助超短期预测方法,能够为机组优化运行提供更多的支撑,提高机组运行灵活性。

负荷预测按照预测时间尺度可以分为超短期、短期、中期以及长期。其中,超短期负荷预测由于时间短,实际任务中以预测机组发电负荷为主。超短期机组发电负荷预测时间尺度一般以小时或者更短的时间为基本单位。随着人工智能技术发展,以深度学习模型为主的人工智能技术在火电领域得到迅猛发展^[2-4]。以长短期记忆神经网络^[5]、Elman神经网络^[6]、Transformer模型^[7]等深度学习模型为主的预测方法在火电机组负荷预测任务中占据主要位置。短期负荷预测与超短期机组发电负荷预测由于时间尺度不同,两者之间存在本质区别。短期负荷预测主要用于电网侧进行日前调度分配^[8],该任务更关注区域内电力系统的供需平衡关系^[9]、可再生资源评估^[10]以及人类社会活动规律^[11-12]。超短期机组发电负荷预测更侧重于机组本身的运行状态以及在接收不同自动发电量控制(auto generation control, AGC)指令后的发电行为。由于机组本身运行状态很难进行量化,因此超短期机组发电负荷预测任务很难用传统建模方法实现,通常采用人工智能方法进行实现^[13-15]。文献[16]提出一种基于特征提取和多级深度学习的mRMR-mLSTM预测模型,并成功应用于火电机组AGC响应特性评估。文献[17]采用灰色关联时间卷积网络模型预测火电机组超短期负荷,并结合遗传算法进行机组参数优化和调度优化。文献[18]基于树的管道优化工具TPOT库自动机器学习实现了火电机组超短期负荷预测,并成功将其应用到飞轮-火电模型的构建中。然而,这些研究中并未探究AGC指令对于超短期机组发电负荷预测的影响。

AGC方式是火电机组并网参与电网调节的主要运行方式。电网对火电机组的调度主要通过下

发AGC指令实现。由于电网对于火电机组的考核主要集中于AGC指令的响应情况上,因此AGC指令能够很大程度上决定机组未来的发电行为。机组的运行状态由给煤、给水、风量等一系列机组状态参数来表征^[19]。当机组接收AGC指令时,AGC指令通过指令分解处理生成锅炉主控信号、汽机主控、给水主控等各类主控信号,从而调整风、煤、水等控制量^[20]。AGC指令与当前机组负荷的关系决定机组未来时刻的升/降负荷行为。当AGC指令变化时,各主控信号在原信号的基础上叠加增量信号,这种增量信号具有正负性,其正负性与变负荷方向一致。主控信号的变化导致机组的风、煤、水等控制量也发生变化。因此,对于超短期机组发电负荷预测而言,考虑AGC指令是至关重要的。在火电机组AGC控制运行方式中,AGC指令经过指令分解处理后分发给各个系统主控,通过改变火电机组协调控制系统的控制目标实现机组的变负荷运行。基于这样的运行方式,本研究提出了一种动态修正模块,该模块通过模拟实际机组运行方式,合理地将AGC指令应用于输入变量的修正过程中,有效地提高了模型的预测精度。

为了实现高精度的预测效果,预测模型的结构应该更加贴近机组机理过程。在火力发电领域,状态空间方程常被用于火电机组的机理建模^[21-22]。Mamba模型^[23]通过离散化状态空间方程,实现将深度学习模型与状态空间方程相结合,增强了模型的表达能力。由于超短期机组发电负荷预测需要预估机组发电行为,需要对模型的输入特征进行一定的修正补偿。因此,本研究提出了一种基于D-Mamba模型的超短期火电机组发电负荷预测方法,通过动态修正模型的输入变量,提高模型的预测精度。

1 火电机组超短期发电负荷预测模型

1.1 D-Mamba模型

为了提高火电机组发电行为预测准确性,预测模型需要更加贴近实际机组运行过程。状态空间模型能够有效描述动态系统的行为,被广泛应用于控制系统、时序分析、经济学等领域。随着研究者的不断探索,基于状态空间序列改进的深度学习模型(state space models, SSM)^[24]已经具有相当大的

发展潜力,其中以 SSM 为核心的 Mamba 模型尤其受到关注。Mamba 模型采用零阶保持器离散方法进一步细化状态空间方程的离散形式,并以此构建新的序列预测模型。同时,本研究基于火电机组 AGC 运行方式,在 Mamba 模型的基础上增加动态修正模块,由此提出一种基于 D-Mamba 模型的超

短期火电机组发电负荷预测方法。D-Mamba 模型的结构如图 1 所示。D-Mamba 模型将输入变量划分为一级变量与二级变量。一级变量通过动态修正模块形成补偿值,用于对二级变量进行动态修正。修正后的二级变量通过 Mamba 模型的计算,得到未来时刻的机组发电负荷。

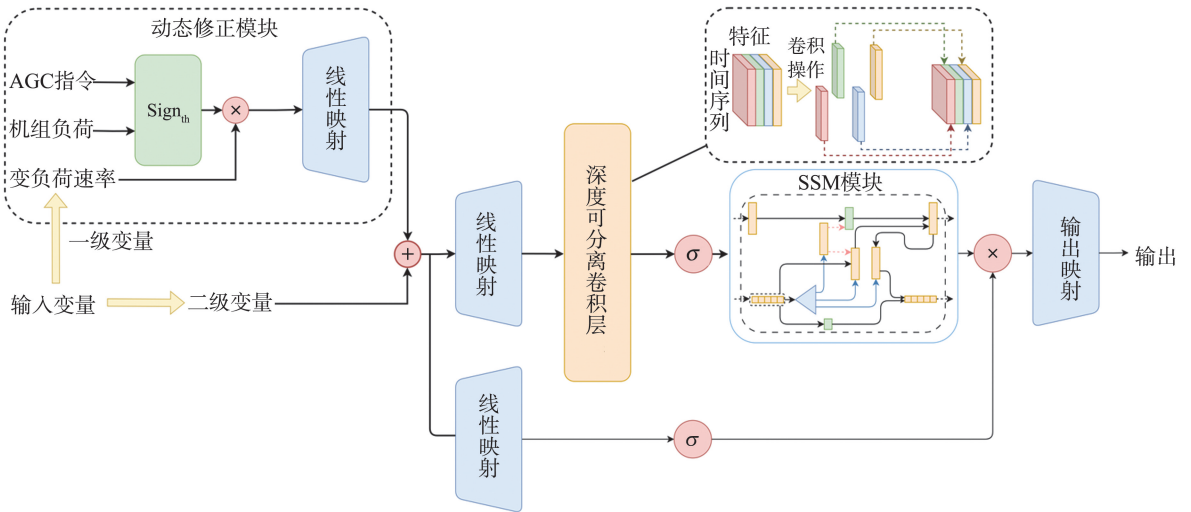


图 1 D-Mamba 模型结构示意图
Fig.1 Structure diagram of D-Mamba model

1.2 基于协调控制方式的动态修正模块

在本研究的预测任务中,模型通过对当前状态参数以及当前时刻获得的 AGC 指令进行计算,得出未来 30 s 的机组实际负荷趋势。在本研究所采用的数据集中,机组全程投入 AGC 运行方式。本研究中对数据及机组运行做出如下假设。假设一:针对每一个输入样本而言,忽略当前时刻到预测时刻之间 AGC 指令变化对预测效果的影响。假设二:控制逻辑及运行方式未发生改变。

本研究所采用的输入参数变量为 AGC 负荷指令、机组负荷、变负荷速率、机组主蒸汽压力、机组主蒸汽压力设定值、机组主蒸汽流量、机组主蒸汽温度、机组再热蒸汽温度、汽包水位、机组给煤量、给水流量、烟气含氧量、总风量。需要注意的是,输入参数变量均为当前时刻及以前时刻的数据,不存在在未来时刻的信息泄露问题。本研究的预测目标为 30 s 后的机组运行负荷。

在输入的变量中,AGC 负荷指令、机组负荷、变负荷速率与负荷预测直接相关,这些变量能够确定机组变化的大致趋势。本研究将这三个变量称为一级变量。其余变量主要用于表征机组的运行状态和控制状态,这些变量与负荷预测之间不存在较为直接的关系,这类变量在本研究中被称作二级变量。

在火电机组的协调运行方式中,一级变量经由指令分解下发到各个主控系统中,形成二级变量的控制目标。由于指令分解处理是线性的,因此可以用一级变量对二级变量进行线性修正,以提高预测精度。在一级变量中,变负荷速率确定主控信号变化的速率。然而,由于变负荷速率设定值在运行操作过程中变化较少,神经网络模型无法对这部分信息进行有效学习。因此,需要对变负荷速率设定值进行特殊处理。由于 AGC 指令与当前机组负荷的关系决定机组在未来一段时间内负荷变化方向,而变化方向信息无法直接参与训练。因此,变化方向信息也需要进行特殊处理。

针对一级变量的处理问题,本研究提出了一种基于动态信息的特征修正神经网络模块。动态信息围绕 AGC 指令展开,同时还结合机组当前发电负荷以及变负荷速率设定值的信息。二级变量经过动态信息的修正后形成 Mamba 模型的输入。基于动态信息的特征修正过程如图 2 所示。其中,电网侧给出的 AGC 指令的刷新时间不固定,而本研究所采用的 AGC 指令是经过电厂 DCS 处理后的数据,该数据的采样时间与机组参数的采样时间一致。同时,由于 AGC 指令主要用于机组变负荷方向判断,该方法与火电机组协调控制系统中的升/降负荷状态判断方法一致,因此,AGC 指令的刷新

时间不会影响本研究的预测方法。

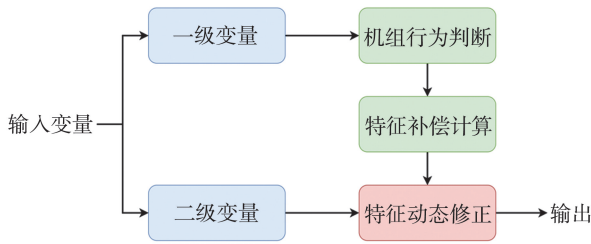


图2 基于动态信息的特征修正过程示意图
Fig.2 Feature correction diagram based on dynamic information

动态修正模块计算公式为

$$s = \text{Sign}(x_{\text{AGC}} - x_{\text{Load}}), \quad (1)$$

$$\text{Sign}(x) = \begin{cases} 1, & x \geq 0.01 \\ 0, & -0.01 < x < 0.01 \\ -1, & x \leq -0.01 \end{cases}, \quad (2)$$

$$X = \text{Linear}(sL_c)W_1 + X_2, \quad (3)$$

式中: s 为趋势信号; x 为单变量输入; Linear 为全连接神经网络层; x_{AGC} 为AGC指令; x_{Load} 为机组负荷; Sign 为带死区的正负判断函数; L_c 为变负荷速率; W_1 为可训练权重; X_2 为二级变量,包含机组主蒸汽压力、机组主蒸汽压力设定值、机组主蒸汽流量、机组主蒸汽温度、机组再热蒸汽温度、汽包水位、机组给煤量、给水流量、烟气含氧量、总风量; X 为经动态信息修正后的特征。

1.3 离散化状态空间模型

为了实现精准预测超短期机组发电负荷,预测模型需要能够更加准确地表征火电机组的运行特性。Mamba模型源自于对状态空间方程的离散化,能更好地表征火电机组的运行特性。

火电机组中经典的三入三出模型^[25]可以表示为如下的状态空间方程

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}, \quad (4)$$

式中: $x(t)$ 为入炉煤量、中间点压力以及中间点焓值; $\dot{x}(t)$ 为 $x(t)$ 的微分值; $u(t)$ 为输入变量,包含给水流量、给煤指令以及汽机调门开度; $y(t)$ 为输出变量,包含机组发电负荷、主蒸汽压力、中间点焓值; A, B, C, D 为权重矩阵。

定义采样间隔为 T ,采样间隔的表达式为

$$T = t_{k+1} - t_k, \quad (5)$$

式中, t_k 为 k 采样时刻对应的连续时间。

对式(4)进行离散化,状态空间方程中的隐藏变量公式 $\dot{x}(t) = Ax(t) + Bu(t)$ 可以离散为

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau)}Bu(\tau)d\tau, \quad (6)$$

式中 e^A 为离散状态下矩阵 A 的表示。

采用零阶保持器离散化方法,假设在时刻 t_k 到 t_{k+1} 之间时, $u(\tau)$ 保持 t_k 时刻的值不变。可以得到如下形式。

$$\begin{cases} x(k) = \bar{A}x(k-1) + \bar{B}u(k) \\ y(k) = \bar{C}x(k) + \bar{D}u(k) \end{cases}, \quad (7)$$

式中, $\bar{A}, \bar{B}, \bar{C}, \bar{D}$ 为权重矩阵 A, B, C, D 修正后的中间变量, $\bar{A} = e^{TA}$, $\bar{B} = (TA)^{-1}(e^{TA} - I)TB$, $\bar{C} = C$, $\bar{D} = D$, $x(k-1)$ 为离散后 $k-1$ 时刻的值。

式(7)为Mamba模型的基础结构。因此,Mamba模型能够有效地表征火电机组的运行特性。

为了在神经网络模型中构建式(7)的表达形式,Mamba模型中采用输入映射的方法对矩阵 B, C, T 进行构建。其中,对采样间隔矩阵 T 进行非线性映射,是为了在离散化状态空间建模中统一不同采样率对动态表达的影响,并确保系统稳定性与可训练性。

$$\begin{cases} B = \text{Linear}_1(u) \\ C = \text{Linear}_2(u) \\ T = \text{SoftPlus}[\text{Linear}_3(u)] \end{cases}, \quad (8)$$

$$\text{SoftPlus}(x) = \log_e(1 + e^x), \quad (9)$$

式中 Linear 为线性映射层。

为了进行简化,Mamba模型中的SSM模块通过 $\bar{A} = e^{TA}$, $\bar{B} = TB$, $\bar{C} = C$, $\bar{D} = D$ 进行离散状态空间模型中系数的计算。其中, A 与 D 为可训练的权重矩阵。在本研究的模型中,计算过程主要通过 $\bar{A}, \bar{B}, \bar{C}, \bar{D}$ 计算得到。其中, T 通过输入 u 映射得到,受到输入 u 的影响。 $\bar{A} = e^{TA}$ 也受到输入 u 的影响,因此 $\bar{A} = e^{TA}$ 可以表示为 $\bar{A} = e^{f(u)}$ 。在将 A 由可训练参数替换成经过输入 u 映射得到的矩阵后, $\bar{A} = e^{f(u)}$ 的形式并未发生变化。在训练过程中, \bar{A} 与输入 u 之间的分布并未发生变化,因此,在经过模型训练后, \bar{A} 的计算并未发生较大变化,对整个模型的预测效果影响不大。为了简化模型,加快模型收敛速度,本研究将 A 设置为可训练参数。可训练参数 D 使得模型呈现出ResNet结构^[26],有利于加快模型训练过程中的收敛速度,提高预测精度。

由于 B, C, T 分别通过输入 u 进行映射,在映射过程中 x 中的时序信息并未被打乱。因此, $\bar{A}, \bar{B}, \bar{C}$ 包含时序性,使得Mamba模型具有表征时变系统的能力。SSM模块在每一个时间步上的计算过程如

图 3 所示。图 3 中,下标 t 表示当前时刻, $t-1$ 表示前一时刻, $\bar{A}x_{t-1}$ 、 $\bar{B}_t u_t$ 、 $\bar{C}_t x_t$ 、 $\bar{D}_t u_t$ 分别代表式(7)中的 $\bar{A}x(k-1)$ 、 $\bar{B}u(k)$ 、 $\bar{C}x(k)$ 、 $\bar{D}u(k)$,橙色部分显示

的张量具有时序信息,在不同的时间步上具有不同的表达;绿色部分不具有时序信息,在不同时间步上的表达相同。

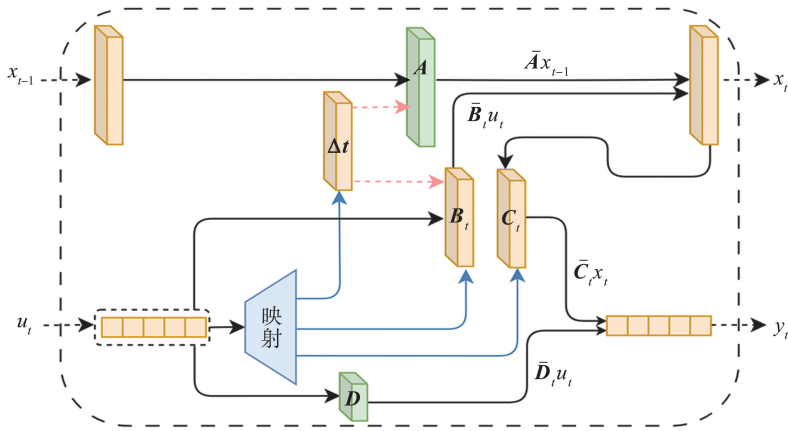


图 3 Mamba 模型的 SSM 模块计算结构图

Fig.3 The structure diagram of SSM module in Mamba

Mamba 模型结构如图 4 所示,图中 σ 为 sigmoid 函数。模型结构采用门控机制的形式进行构建,数据通过映射层形成门控信息及高维映射数据,高维映射数据通过深度可分离卷积层进行时序特征的处理。门控信息对经过卷积处理后的数据进行选择,最终通过输出映射层形成最终的输出。其中,卷积操作能够有效提取数据中的时序特征,过滤数据噪声,而深度可分离卷积针对每一个特征采用单独的卷积核进行处理,能够有效减少模型参数量,提高运算效率。Mamba 模型计算流程为

$$O_1 = \text{Lin}_1(X), \quad (10)$$

$$\text{Res} = \text{SiLU}(\text{Lin}_2(X)), \quad (11)$$

$$X_2 = \text{SiLU}[\text{CNN}(O_1)], \quad (12)$$

$$O_3 = \text{SSM}(O_2), \quad (13)$$

$$Y = \text{Res} \odot O_3, \quad (14)$$

式中: X 为经过动态信息修正后的特征; Y 为模型输出; Lin 表示线性映射层,与式(9)中的 Linear 所表示的结构相同,为了进行区分这里的线性映射层使用 Lin 进行表示; CNN 为深度可分离卷积操作; SSM 为 SSM 模块; \odot 为哈达玛积操作; SiLU 为 SiLU 激活函数

SiLU 激活函数的表达式为

$$\text{SiLU}(x) = x \frac{1}{1+e^{-x}}. \quad (15)$$

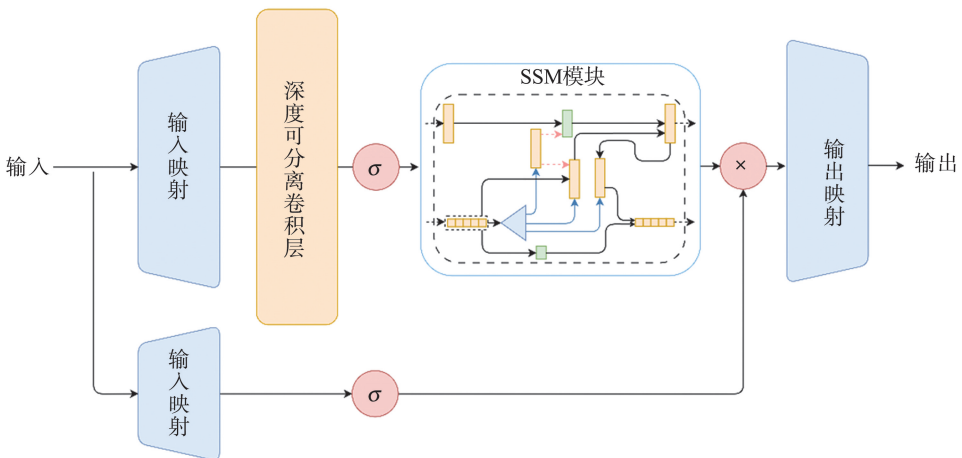


图 4 Mamba 模型结构示意图

Fig.4 Structure diagram of Mamba model

1.4 性能评价指标

为了评估不同试验场景下的预测性能,选择多种科学性能指标预测时间序列。本研究用平均绝

对误差(E_{mae})、平均绝对百分比误差(E_{mape})、均方根误差(E_{rmse})评估不同模型的精度。

平均绝对误差

$$E_{\text{mac}} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i|, \quad (16)$$

平均绝对百分比误差

$$E_{\text{mape}} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y'_i}{y_i} \right|, \quad (17)$$

均方根误差

$$E_{\text{rmse}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2}, \quad (18)$$

式中, y'_i 为预测模型的预测值, n 为预测数据的数量, y_i 为实际值。通常情况下, E_{mac} 、 E_{mape} 以及 E_{rmse} 越低, 表示预测精度越高。

2 数据处理及模型建立

2.1 数据处理

论文数据取自于某 600 MW 亚临界汽包炉机组实际运行数据, 数据采样间隔为 10 秒。数据按照时间顺序划分, 将数据依次划分为训练集、验证集及测试集。数据集大小分别为 68 000 个, 2 000 个, 2 000 个。模型的输入变量在 1.2 节第 2 段中已提及。预测目标为 30 s 后的机组实际运行负荷。其中, 训练集用于训练模型参数, 验证集用于选取模型超参数, 测试集用于展示预测结果。本研究的程序使用 Python 进行编译, 数据处理部分采用 Numpy 和 Scikit-learn 框架进行计算, 算法采用 Pytorch 框架。所有试验均在英特尔酷睿 i9-13900K CPU 和 RTX4090 GPU 机器上进行。算法使用 CUDA 和 cuDNN 进行加速。

首先对收集到的数据进行归一化处理, 归一化公式为

$$X_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \quad (19)$$

式中, X_i 为经过归一化处理之后的数据, x_i 为归一化处理前的数据, x_{\min} 与 x_{\max} 分别为数据中的最小值与最大值。为避免预测过程中新数据超过范围, x_{\min} 与 x_{\max} 应与机组运行过程中的参数上下限值设置相同。

需要注意的是, 机组负荷、AGC 指令预测信号、当前时刻 AGC 指令三个变量采用相同的归一化参数。

2.2 模型超参数选择

预测模型的超参数包含 B 、 C 、 T 的映射维度 n_1 、 n_2 、 n_3 , 深度可分离卷积层的卷积核大小 w , 时间序列长度, 批训练的数据批次大小, 学习率。本研究通过训练集进行模型参数训练, 通过模型在验证

集上的预测性能指标选择超参数, 选择后的超参数设置如表 1 所示。

表 1 模型超参数选择结果

超参数	数值
n_1	16
n_2	16
n_3	32
w	2
时间序列长度	50
数据批次大小	100
学习率	0.001

模型训练过程使用 Early-Stopping 策略, 优化器使用 AdamW^[27]。Early-Stopping 策略通过观察训练误差的变化训练回合数控制, 一旦模型的训练误差发生收敛, 则停止模型训练。

3 模型对比与分析

为了保证试验的有效性, 无论在消融试验还是对比试验中, 参与比较的模型采用与本研究模型相同的输入变量, 其中包含一级变量与二级变量, 输入参数变量均为当前时刻及之前时刻的数据, 不存在未来时刻的信息泄露问题。

3.1 预测模型消融试验

为了验证 D-Mamba 模型结构的有效性, 本节中加入了几种不同结构的模型进行对比。其中, Mamba 模型表示原始的 Mamba 模型; D-Mamba (A) 将 D-Mamba 模型中的矩阵 A 由可训练参数替换成了通过输入 u 映射得到, 即 $A = \text{Linear}_1(u)$; 同理, D-Mamba (D) 将 D-Mamba 模型中的矩阵 D 由可训练参数替换成了通过输入 u 映射得到, 即 $D = \text{Linear}_2(u)$ 为了保证模型结构的一致性, $A = \text{Linear}_1(u)$ 与 $D = \text{Linear}_2(u)$ 中的 $\text{Linear}(u)$ 均进行了维度和矩阵形状的转换, 以保证替换前后的 A 和 D 具有相同的维度和矩阵形状。D-Mamba (A)、D-Mamba (D) 以及 Mamba 模型的超参数设置与本研究模型的设置相同。消融试验预测结果如图 5 所示, 预测性能指标结果如表 2 所示。本研究表格中的加粗数据代表最佳的预测性能指标结果。

表 2 消融试验预测性能指标结果

模型	E_{mac}/MW	$E_{\text{mape}}/\text{MW}$	$E_{\text{rmse}}/\%$
D-Mamba	0.978 4	1.539 7	0.27
D-Mamba (A)	0.991 0	1.561 7	0.28
D-Mamba (D)	1.198 8	1.865 6	0.33
Mamba	1.165 8	1.857 8	0.32

由试验结果可以得出,本研究提出的变量修正模块能够有效提高预测精度。采用特征修正模块后, E_{mac} 指标由 1.165 8 MW 提升至 0.978 4 MW, E_{mapc} 指标由 0.32% 提升至 0.27%, E_{rmse} 指标由 1.857 8 MW 提升至 1.539 7 MW。相较于 D-Mamba 模型,经过结构调整后的 D-Mamba (A) 与

D-Mamba(D) 也出现不同程度的下降。其中,替换可训练参数 A 后的预测效果变化不大。替换可训练参数 D 后的预测效果出现较大变化,主要原因在于替换后的模型破坏了原有的 ResNet 结构,降低了模型参数的收敛能力,导致模型精度降低。

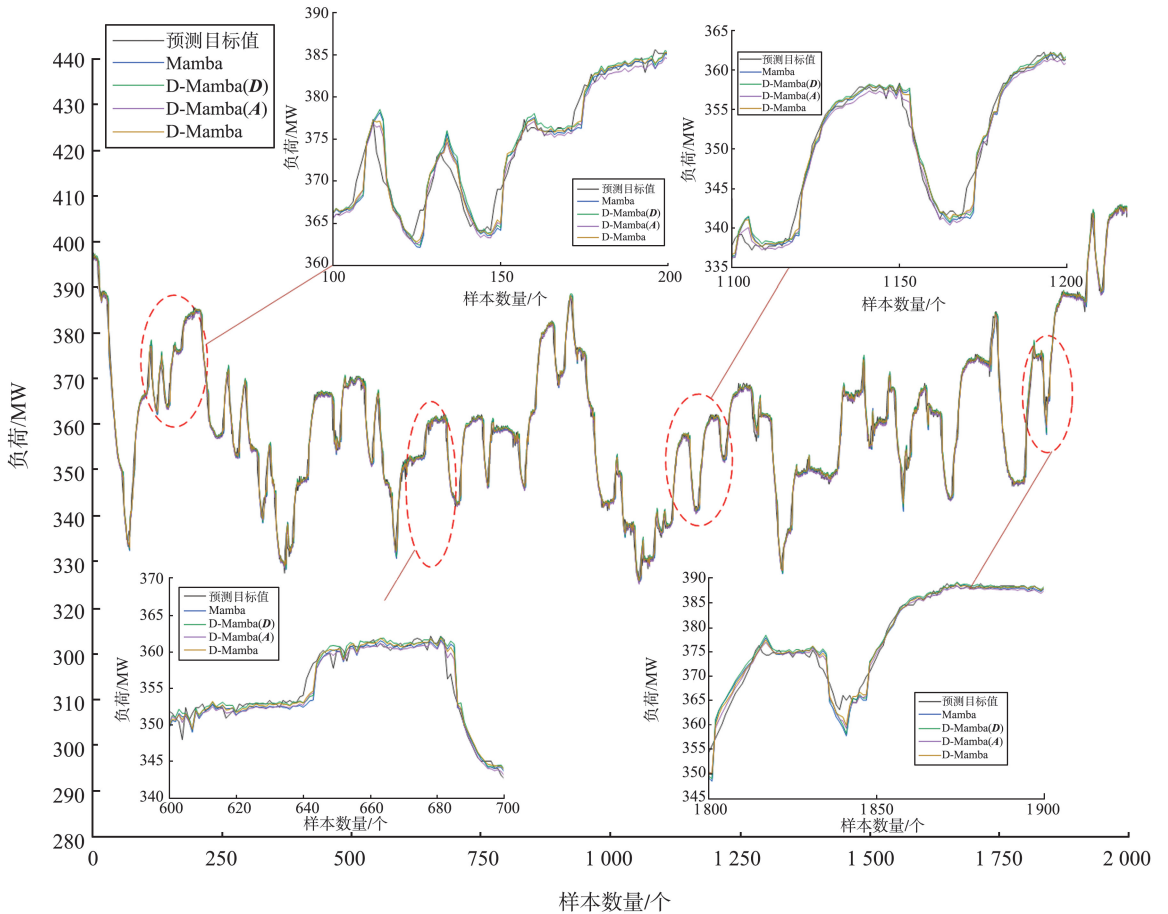


图 5 消融试验预测结果
Fig.5 The predicted results of ablation experiments

3.2 预测模型对比试验

为了凸显模型的有效性,本节采用主流的时序预测模型进行对比试验。对比试验采用了 LSTM 模型^[28]、GRU 模型^[29]以及 Transformer 模型^[7]。同时,为了更加有效地进行对比试验,本节保留了动态信息修正模块,分别在 LSTM 模型、GRU 模型以及 Transformer 模型的基础上增加了动态修正模块,分别命名为 D-LSTM、D-GRU、D-Transformer。对比试验的所有模型均采用与本研究模型相同的输入。另外,对比试验的所有模型均采用相同的方法进行了超参数的选择。对比试验结果如图 6、7 所示,预测性能指标结果如表 3 所示。

表 3 对比试验预测性能指标结果

Table 3 Comparison experiment prediction performance index results

模型	E_{mac} /MW	E_{mapc} /MW	E_{rmse} /%
D-Mamba	0.978 4	1.539 7	0.27
GRU	1.275 5	2.033 0	0.35
LSTM	1.280 1	2.027 9	0.36
Transformer	1.454 6	2.159 5	0.40
D-GRU	1.179 0	1.745 5	0.33
D-LSTM	1.153 9	1.828 2	0.32
D-Transformer	1.258 4	1.923 0	0.35

从表 3 可知,本研究提出的 D-Mamba 模型取得最佳预测精度。而当前应用最广泛的 Transformer 模型在本研究的预测任务中精度最差,主要原因在于 Transformer 模型通过自注意力机制提高模型在跨时间步上的信息提取,这种特

性在自然语言处理任务中极为重要,但该特性在本研究的预测任务中显现出局限性。GRU、LSTM以及本研究的预测模型通过递归的方式实现了时序信息的学习和处理。相较于Transformer模型,这些模型具有更强的时序信息提取能力,但是在跨时间步的信息提取能力上就非常有限了。本研究的负荷预测任务更注重时序信息的提取能力,因此能取得更好的预测性能。D-GRU模型在 E_{mac} 、 E_{mape} 、 E_{rmse} 3个指标上分别达到了

1.179 0 MW、0.33%与1.745 5 MW,D-LSTM模型在 E_{mac} 、 E_{mape} 以及 E_{rmse} 性能指标方面分别达到了1.153 9 MW、0.32%与1.828 2 MW。作为LSTM模型的一类较为成功的变体模型,GRU模型具有更少的门控结构,计算效率更高,但是由于这两类模型具有相同的计算结构,都是通过门控机制丢弃和保留时序信息,因此这两类模型在负荷预测任务中的性能表现较为接近。

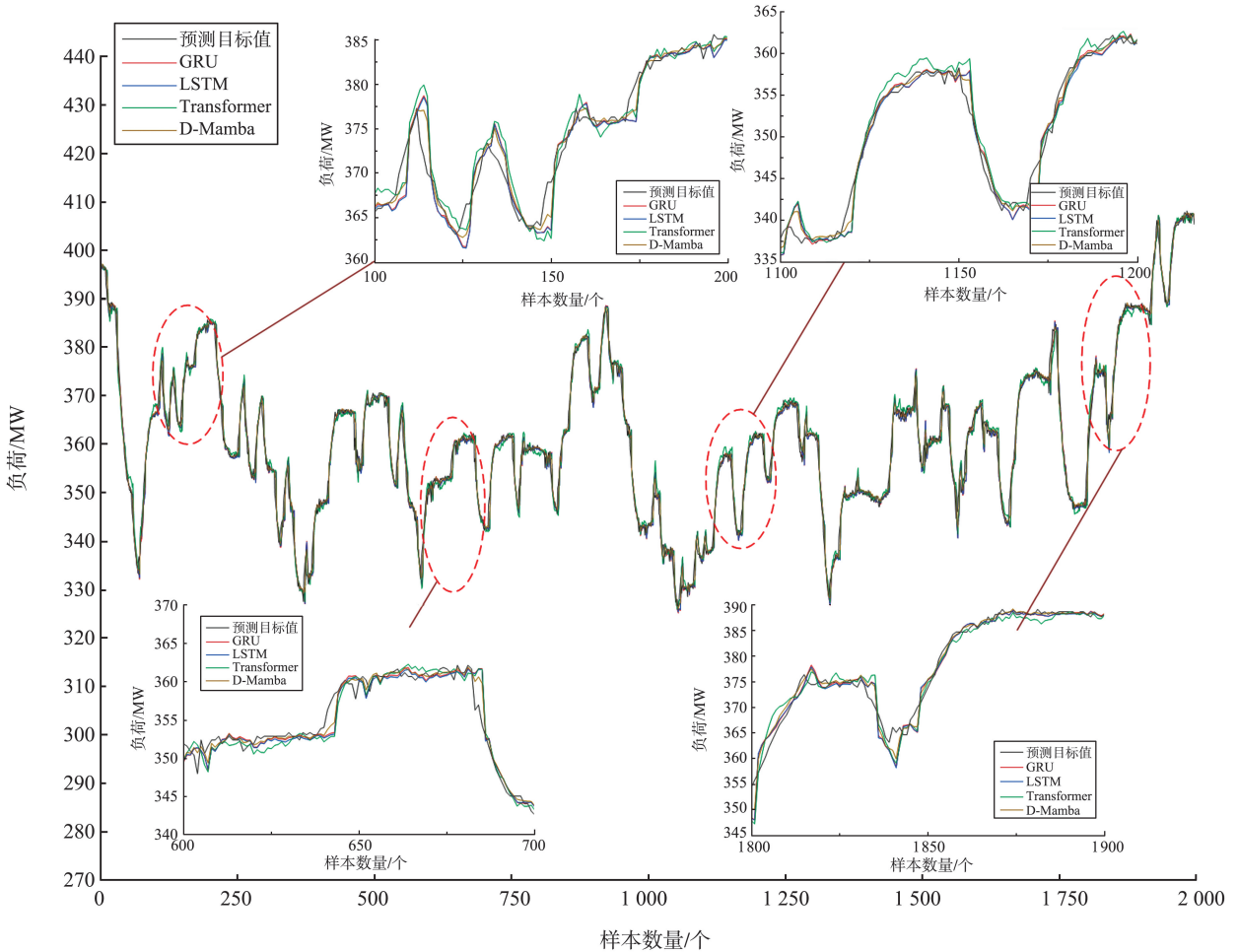


图6 各主流预测模型预测结果

Fig.6 Mainstream prediction model prediction result graph

综合对比来看,动态修正模块能够有效提高模型的预测精度。D-GRU、D-LSTM以及D-Transformer模型均在原来的基础上有所提高。从横向对比来看,采用Mamba模型能够取得更高的预测精度,原因在于Mamba模型具有更强的表征能力。对于Mamba模型而言,若 B,C 的最后一个维度设置为1,并且 $A=-1, B=1$,则SSM模块可以写成如下形式^[23]

$$g_t = \sigma[\text{Linear}(x_t)], \quad (20)$$

$$h_t = (1-g_t)h_{t-1} + g_t x_t, \quad (21)$$

$$\sigma(x) = \frac{1}{1+e^{-x}}, \quad (22)$$

式中, g_t 为门控信号, h_{t-1} 为候选状态信号, h_t 为状态信号。

从化简后的形式可以看出,Mamba模型在形式上可以在特殊情况下转换成GRU与LSTM模型的递归计算形式。GRU模型与LSTM模型的表达形式更接近于一种特殊对象场景下的状态空间方程,具有一定的局限性,因此在精度方面弱于本研究的Mamba模型。

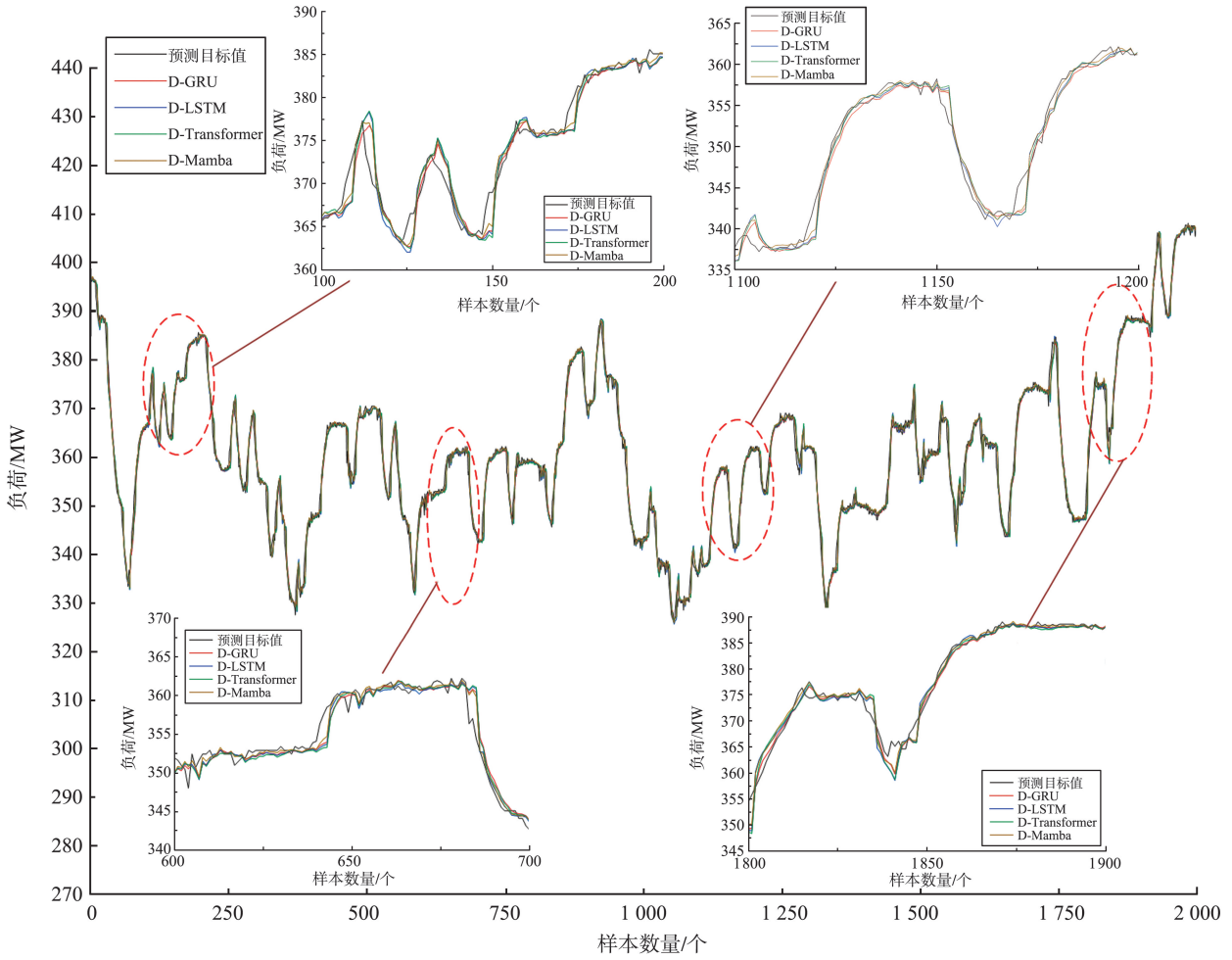


图 7 基于动态修正模块的主流模型预测结果图

Fig.7 Mainstream model prediction result diagram based on dynamic correction module

4 结论

通过分析动态过程中机组运行方式,本研究提出一种基于 D-Mamba 模型的超短期火电机组发电负荷预测方法,通过实际机组运行数据验证了该方法的准确性。本研究所提出的方法为机组动态性能的在线评估奠定研究基础,为新型电力系统中的火电机组智能化灵活性优化改造做铺垫。后续研究将从基于预测结果的机组运行控制优化入手,探究预测信息对机组运行方式的影响关系,分析预测置信度对控制稳定性的扰动问题。

参考文献:

[1] 陈焱, 何山, 谢少华, 等. 基于合作博弈的风-光-电氢微网容量配置 [J]. 太阳能学报, 2024, 45 (2): 395-405.
CHEN Yi, HE Shan, XIE Shaohua, et al. Capacity configuration of wind-photovoltaic-electric hydrogen

microgrid based on cooperative game [J]. Acta Energaie Solaris Sinica, 2024, 45(2): 395-405.
[2] LV M L, ZHAO J P, CAO S X, et al. Prediction of temperature distribution in a furnace using the incremental deep extreme learning machine [J]. PeerJ Computer Science, 2023, 9: e1218.
[3] KHALID S, SONG J, RAOUF I, et al. Advances in fault detection and diagnosis for thermal power plants: a review of intelligent techniques [J]. Mathematics, 2023, 11 (8): 1767.
[4] CUI J H, CHAI T Y, LIU X J. Deep-neural-network-based economic model predictive control for ultrasupercritical power plant [J]. IEEE Transactions on Industrial Informatics, 2020, 16(9): 5905-5913.
[5] WAN A P, CHANG Q, AL-BUKHAITI K, et al. Short-term power load forecasting for combined heat and power using CNN-LSTM enhanced by attention mechanism [J]. Energy, 2023, 282: 128274.
[6] 骆小满, 皇甫成, 阮江军, 等. 基于神经网络的热电联产机组热负荷和电负荷预测 [J]. 热力发电, 2019, 48

- (9): 46-50.
- LUO Xiaoman, HUANGFU Cheng, RUAN Jiangjun, et al. Prediction of heat and electric load of cogeneration unit based on neural network[J]. Thermal Power Generation, 2019, 48(9): 46-50.
- [7] 樊建升, 吴海滨, 刘泽军. 融合时间序列趋势的 Dual-ESN 机组负荷预测模型[J]. 电力系统及其自动化学报, 2023, 35(1): 152-158.
- FAN Jiansheng, WU Haibin, LIU Zejun. Dual-ESN prediction model for unit load fused with time series trend [J]. Proceedings of the CSU-EPSA, 2023, 35(1): 152-158.
- [8] 徐聪, 胡永锋, 张爱平, 等. 基于特征筛选的综合能源系统多元负荷日前-日内预测[J]. 综合智慧能源, 2024, 46(3): 45-53.
- XU Cong, HU Yongfeng, ZHANG Aiping, et al. Multi-load day-ahead and intra-day forecasting for integrated energy systems based on feature screening[J]. Integrated Intelligent Energy, 2024, 46(3): 45-53.
- [9] 刘璐瑶, 陈志刚, 沈欣炜, 等. 基于 EMD-MLP 组合模型的用电负荷日前预测[J]. 南方能源建设, 2024, 11(1): 143-156.
- LIU Luyao, CHEN Zhigang, SHEN Xinwei, et al. Day-ahead forecast of electrical load based on EMD-MLP combination model [J]. Southern Energy Construction, 2024, 11(1): 143-156.
- [10] WEI N, YIN C, YIN L H, et al. Short-term load forecasting based on WM algorithm and transfer learning model[J]. Applied Energy, 2024, 353: 122087.
- [11] WAZIRALI R, YAGHOUBI E, ABUJAZAR M S S, et al. State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques [J]. Electric Power Systems Research, 2023, 225: 109792.
- [12] FAN G F, HAN Y Y, LI J W, et al. A hybrid model for deep learning short-term power load forecasting based on feature extraction statistics techniques [J]. Expert Systems with Applications, 2024, 238: 122012.
- [13] 王欣然. 基于相空间重构和双向长短期记忆网络的火电机组负荷预测建模研究[D]. 长沙: 长沙理工大学, 2022.
- WANG Xinran. Modeling of thermal power unit load prediction based on phase space reconstruction and bidirectional long short-term memory network [D]. Changsha: Changsha University of Science and Technology, 2022.
- [14] 彭维珂, 聂椿明, 陈衡, 等. 基于智能算法的空冷火电机组负荷预测研究[J]. 华电技术, 2021, 43(3): 57-64.
- PENG Weike, NIE Chunming, CHEN Heng, et al. Study on load forecasting for air cooling thermal power units based on intelligent algorithm [J]. Huadian Technology, 2021, 43(3): 57-64.
- [15] 张然然, 刘鑫屏. 火电机组超短期负荷预测[J]. 热力发电, 2018, 47(7): 52-57.
- ZHANG Ranran, LIU Xinping. Ultra-short-term load forecasting for thermal power units[J]. Thermal Power Generation, 2018, 47(7): 52-57.
- [16] 郝晓光, 金飞, 张庆浩, 等. 基于 mRMR 与多级 LSTM 网络的火电机组响应 AGC 调控能力评估[J]. 热能动力工程, 2024, 39(5): 57-64.
- HAO Xiaoguang, JIN Fei, ZHANG Qinghao, et al. Assessment of AGC regulation capability of thermal power units based on mRMR and multi-level LSTM networks[J]. Journal of Engineering for Thermal Energy and Power, 2024, 39(5): 57-64.
- [17] 颜子翼. 基于深度学习的火电机组短期负荷预测[D]. 青岛: 青岛大学, 2023.
- YAN Ziyi. Short-term load forecasting of thermal power units based on deep learning [D]. Qingdao: Qingdao University, 2023.
- [18] 魏乐, 苏少忻, 房方, 等. 基于负荷预测的飞轮-火电系统自动发电控制响应性能优化[J]. 热力发电, 2023, 52(5): 92-99.
- WEI Le, SU Shaoxin, FANG Fang, et al. Optimization of automatic generation control response performance of flywheel-thermal power system based on load forecasting [J]. Thermal Power Generation, 2023, 52(5): 92-99.
- [19] 弓林娟, 王文毓, 高耀岩, 等. 面向大规模新能源消纳的火电机组平行控制[J]. 动力工程学报, 2023, 43(2): 136-142.
- GONG Linjuan, WANG Wenyu, GAO Yaokui, et al. Parallel control of thermal power unit for large-scale renewable energy accommodation[J]. Journal of Chinese Society of Power Engineering, 2023, 43(2): 136-142.
- [20] YAO J, CHEN J, XU J, et al. Nonlinear modeling and control of ultra-supercritical power plants with multi-level energy storage [J]. Applied Thermal Engineering, 2025: 128727.
- [21] HONG F, JI W M, LIANG L, et al. A new assessment mechanism of primary frequency regulation capability for a supercritical thermal power plant in deep peaking[J]. Energy Science & Engineering, 2023, 11(2): 547-564.