

多尺度注意力加权的CNN-BiGRU光伏功率预测方法

李子阳, 朱以墨, 迟晨, 张依林, 刘甲辉, 张丽艳*

大连交通大学, 轨道智能工程学院, 电子与通信工程系, 辽宁 大连

收稿日期: 2026年5月25日; 录用日期: 2026年6月17日; 发布日期: 2026年6月29日

摘要

光伏功率预测是保障光伏电站高效运维与电网优化调度的核心技术, 其精度直接影响新能源消纳水平与并网安全性。针对传统CNN-BiGRU模型在光伏功率预测中存在的单尺度注意力表征能力不足、特征融合维度受限及网络结构布局不合理等问题, 文章提出一种基于多尺度注意力加权的CNN-BiGRU光伏功率预测模型。该模型以“局部特征提取-多尺度特征融合-时序依赖建模-精准预测输出”为主干结构, 首先通过卷积神经网络提取输入时序数据的局部空间特征; 其次设计并行多尺度注意力模块, 采用3、5、7三种卷积核分别捕捉短期波动、中期趋势与长期周期特性, 并在192维高维特征空间中进行自适应加权, 显著提升模型对复杂气象条件下功率变化的拟合能力; 随后将优化后的特征输入BiGRU网络以建模双向时序依赖关系, 最后通过全连接层输出单步预测结果。实验结果表明, 所提模型在保证轻量化的同时有效降低了预测误差, 相较于传统CNN-BiGRU及现有改进模型, 在多项评价指标上均取得更优表现, 验证了多尺度注意力机制与结构优化策略的有效性。

关键词

光伏功率预测, 卷积神经网络, 双向门控循环单元, 多尺度注意力, 特征融合

Photovoltaic Power Prediction Method Based on Multi-Scale Attention Weighted CNN-BiGRU

Ziyang Li, Yimo Zhu, Chen Chi, Yilin Zhang, Jiahui Liu, Liyan Zhang*

Department of Electronic Communications, School of Railway Intelligent Engineering,
Dalian Jiaotong University, Dalian Liaoning

Received: May 25, 2026; accepted: June 17, 2026; published: June 29, 2026

*通讯作者。

文章引用: 李子阳, 朱以墨, 迟晨, 张依林, 刘甲辉, 张丽艳. 多尺度注意力加权的 CNN-BiGRU 光伏功率预测方法[J]. 智能电网, 2026, 16(3): 94-100. DOI: 10.12677/sg.2026.163011

Abstract

Photovoltaic power prediction is a core technology for ensuring efficient operation and maintenance of photovoltaic power stations and optimized dispatch of power grids, and its accuracy directly affects the level of renewable energy accommodation and grid-connected security. To address the limitations of traditional CNN-BiGRU models in PV power prediction, such as insufficient single-scale attention representation capability, restricted feature fusion dimensions, and unreasonable network structural layout, this paper proposes a CNN-BiGRU photovoltaic power prediction model based on multi-scale attention weighting. The proposed model follows a main structure of “local feature extraction, multi-scale feature fusion, temporal dependency modeling, and accurate prediction output”. Firstly, a Convolutional Neural Network (CNN) is utilized to extract local spatial features from the input time-series data. Secondly, a parallel multi-scale attention module is designed, employing convolution kernels of sizes 3, 5, and 7 to capture short-term fluctuations, medium-term trends, and long-term periodic characteristics, respectively. Adaptive weighting is performed in a 192-dimensional high-dimensional feature space, significantly enhancing the model’s fitting capability for power variations under complex meteorological conditions. Subsequently, the optimized features are fed into a Bidirectional Gated Recurrent Unit (BiGRU) network to model bidirectional temporal dependencies. Finally, the single-step prediction results are output through a fully connected layer. Experimental results demonstrate that the proposed model effectively reduces prediction errors while maintaining a lightweight architecture. Compared with traditional CNN-BiGRU and existing improved models, the proposed method achieves superior performance across multiple evaluation metrics, verifying the effectiveness of the multi-scale attention mechanism and structural optimization strategies.

Keywords

Photovoltaic Power Prediction, Convolutional Neural Network (CNN), Bidirectional Gated Recurrent Unit (BiGRU), Multi-Scale Attention, Feature Fusion

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

光伏功率预测作为保障光伏电站高效运维、电网优化调度及新能源消纳的关键技术，其预测精度直接关系到光伏能源规模化并网的安全性与经济性[1]。近年来，深度学习模型凭借强大的非线性拟合与时序建模能力，已成为光伏功率预测领域的主流研究方法。其中，卷积神经网络与双向门控循环单元(CNN-BiGRU)的组合模型，因兼具局部空间特征提取与时序双向依赖建模的优势，受到了广泛关注与应用[2]。

尽管现有研究已取得一定进展，但传统 CNN-BiGRU 及其改进模型在面对复杂多变的光伏出力场景时仍显不足，主要体现在以下两个方面：1) 注意力机制多局限于单尺度设计，难以同时刻画光伏功率序列中短期剧烈波动、中期变化趋势与长期周期性规律并存的多尺度特性[3]；特征融合维度较为单一，信息容量受限，导致模型在复杂气象条件下对功率突变的拟合能力偏弱[4]；2) 网络结构布局不尽合理，注意力模块位置设置缺乏针对性，未能充分实现关键特征筛选，易造成时序信息在深层传递过程中的损耗[5]。上述问题共同制约了模型在复杂场景下的预测精度与泛化性能。为此，本文提出一种基于多尺度注意力加权的 CNN-BiGRU 光伏功率预测模型。

2. 多尺度注意力加权的 CNN-BiGRU 模型

本研究基于多尺度注意力加权的 CNN-BiGRU 模型如图 1 所示，主要以“CNN 局部特征提取 - 多尺度特征融合 - 时序依赖建模 - 预测输出”为核心逻辑，整体分为 CNN 特征提取模块、多尺度注意力加权模块、BiGRU 时序建模模块及全连接输出模块四个核心模块，各模块间通过特征维度适配实现端到端的光伏功率预测。该模型通过引入多尺度注意力机制以增强对不同时间粒度特征的感知能力，利用卷积核提取多频特征并进行空间加权，优化多层次特征融合策略以提升信息利用率，并重新设计网络结构，注意力模块前置至 CNN 与 BiGRU 之间。减少时序信息传递损耗，既保留了完整空间特征，又减轻了时序建模负担。

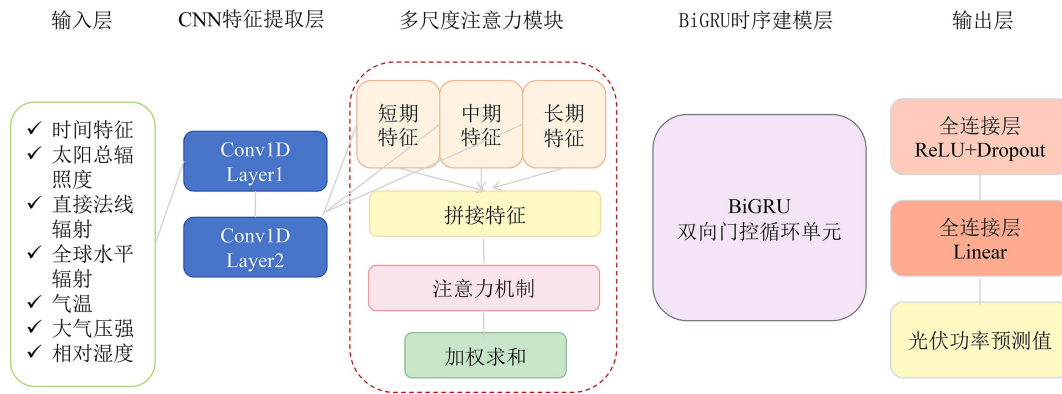


Figure 1. Multi-scale attention weighted CNN-BiGRU model
图 1. 多尺度注意力加权的 CNN-BiGRU 模型

2.1. 输入层的数据预处理

本文采集的光伏数据集时间范围为 2019 年 1 月 1 日至 2020 年 12 月 31 日，数据频率为 15 分钟，共 70,176 个样本，涵盖辐照度、气象及功率特征。预处理采用 K 近邻算法补全缺失值，并剔除异常样本，该方法在近期光伏数据修复研究中被证明能有效保持数据分布特性，适用于多变量时序填补[6]。数据显示，异常值集中于辐照度与功率变量(太阳总辐照度异常 9295 个，直接法线辐射与散射辐照度各 4202 个，功率异常 9382 个)，气象要素(温湿度、气压)无异常，这种辐照度与功率变量易出现异常、气象稳态变量较干净的现象与现有光伏实测数据质量分析结论一致[7]。据此剔除异常行后，保留有效样本 58,633 个，剔除了 11,543 个异常样本，提升了数据纯净度。批量剔除多变量耦合异常的样本行能避免错误标签干扰深度学习模型收敛，是提升光伏预测数据集质量的关键步骤[8]。

2.2. CNN 局部特征提取

光伏功率数据不仅包含时间序列的动态特性，其与气象特征(如辐照度、温度)之间也存在复杂的非线性关系，利用卷积神经网络(CNN)有效挖掘这种高维非线性映射已成为近年来的主流做法[9]。为了高效地提取这些局部特征，本文首先引入了卷积神经网络(CNN)模块。该模块由两个卷积层堆叠而成，旨在从原始输入序列中自动学习深层、抽象的特征表示[10]。给定一个输入序列，其中 $X \in R^{T \times F}$ ， T 为输入时间步长， F 为特征维度(包括气象特征和时间编码特征)。CNN 局部特征提取过程如下。

首先，将输入序列沿特征维度进行一维卷积，如式(1)所示。第一层卷积核尺寸为 3，输出通道数为 32，

$$H_1 = \text{ReLU}\left(\text{BatchNorm}\left(\text{Conv1D}_{3 \rightarrow 32}\left(X^T\right)\right)\right) \quad (1)$$

其中, $X^T \in R^{T \times F}$ 表示将原始输入转置以匹配 Conv1D 的输入格式。通过第二层卷积进一步提炼特征, 如式(2)所示, 其输出通道数扩展至 64。

$$H_2 = \text{ReLU}(\text{BatchNorm}(\text{Conv1D}_{32 \rightarrow 64}(H_1))) \quad (2)$$

得到了一个包含丰富局部信息的特征图 $H_2 \in R^{64 \times T}$ 。

2.3. 多尺度注意力加权

光伏功率序列中同时包含了短期波动、中期趋势和长期周期性等多种尺度的模式。传统的单一尺度卷积或注意力机制难以全面捕捉这些不同时间尺度的特征[4] [11]。为此, 本文提出了一个多尺度注意力模块, 旨在并行提取并自适应地融合多尺度时序特征。

该模块首先通过三个并行的、不同感受野的一维卷积层来提取多尺度特征, 分别如式(3)、式(4)和式(5)所示。

$$f_3 = \text{ReLU}(\text{BatchNorm}(\text{Conv1D}_{K=3, P=1}(H_2))) \quad (3)$$

$$f_5 = \text{ReLU}(\text{BatchNorm}(\text{Conv1D}_{K=5, P=2}(H_2))) \quad (4)$$

$$f_7 = \text{ReLU}(\text{BatchNorm}(\text{Conv1D}_{K=7, P=3}(H_2))) \quad (5)$$

其中, f_3 、 f_5 、 $f_7 \in R^{H \times T}$ 分别表示从 CNN 特征图 H_2 中提取的短、中、长期特征, H 为隐藏层维度(本文设为 64), k 和 p 分别代表卷积核尺寸和填充大小。

随后, 将这三个尺度的特征在通道维度上进行拼接, 形成一个多尺度特征, 如式(6)所示。

$$M \in R^{3H \times T}: M = \text{Concat}(f_3, f_5, f_7) \quad (6)$$

为了实现对不同时序位置及多尺度特征重要性的自适应评估, 本文构建了多尺度特征自适应加权模块。该机制基于拼接融合后的高维特征表示, 通过全连接层映射计算各时间步的原始注意力得分, 并引入 Softmax 函数将其归一化为概率分布形式, 如式(7)和式(8)所示, 从而实现关键时序信息的精准聚焦。

$$e_t = v^T \tanh(W_a m_t + b_a) \quad (7)$$

$$\alpha_t = \frac{\exp(e_t)}{\sum_{j=1}^T \exp(e_j)} \quad (8)$$

其中, $m_t \in R^{3H}$ 是多尺度特征矩阵 M 在时间步 t 的向量, $W_a \in R^{H \times 3H}$ 和 $b_a \in R^H$ 是可学习的权重矩阵和偏置, $v \in R^H$ 是注意力向量。最终, 通过加权求和, 得到一个包含全局多尺度信息的上下文向量 C , 如式(9)所示。

$$C = \sum_{t=1}^T \alpha_t m_t \quad (9)$$

该模块通过并行多尺度卷积和自适应注意力加权, 能够有效聚焦于对预测最关键的时空模式, 显著增强了模型的特征表达能力。

2.4. BiGRU 时序特征提取

光伏功率预测不仅依赖于过去的信息, 未来的信息也对其当前状态有重要影响。为了充分利用序列的双向上下文信息, 本文将多尺度注意力模块输出的上下文向量输入到一个双向门控循环单元(BiGRU)中[12]。BiGRU 由一个前向 GRU 和一个后向 GRU 组成, 分别从正反两个方向处理序列信息。其核心是

门控机制，通过更新门和重置门来控制信息的流动和记忆的更新，从而有效缓解长序列中的梯度消失问题[13]。前向和后向的 GRU 的计算分别如式(10)、式(11)所示。

$$\vec{h}_t = \text{GRU}_{\text{forward}}(c_t, \vec{h}_{t-1}) \quad (10)$$

$$\overleftarrow{h}_t = \text{GRU}_{\text{backward}}(c_t, \overleftarrow{h}_{t+1}) \quad (11)$$

其中， c_t 是在时间步 t 的输入特征。最终的隐藏状态由前向和后向的隐藏状态拼接而成 BiGRU，如式(12)所示。

$$h_t = [\vec{h}_t; \overleftarrow{h}_t] \quad (12)$$

通过这种双向结构，BiGRU 模块能够同时捕捉到时间步 t 之前和之后的依赖关系，为预测提供更全面的时序信息。

2.5. 全连接层的预测输出

最后，将 BiGRU 最后一个时间步的输出 $h_T \in R^{2H}$ 输入到一个全连接层，以生成最终的预测值。该模块通过一个两层的前馈神经网络，将高维的隐藏状态映射到目标输出空间，即下一时刻的光伏功率预测值计算如式(13)所示。

$$\hat{y} = \text{Linear}(\text{ReLU}(\text{Linear}(h_T))) \quad (13)$$

3. 实验及结果分析

3.1. 实验配置

本实验硬件配置为 Intel Core Ultra7 155H 处理器(16 核 22 线程，最高睿频 3.8 GHz)、16 GB DDR5 内存(5600 MHz)和 512 GB NVMe SSD，软件环境采用 Windows 11 家庭版操作系统。本文模型的参数设置如下：输入时间步长为 24，即使用过去 24 小时的数据预测未来 1 小时的功率值；批次大小为 64；CNN 模块隐藏层维度设置为 64；BiGRU 模块隐藏层维度设置为 64，层数为 2，双向结构的输出维度为 128；多尺度注意力模块中三种卷积核尺寸分别为 3、5、7，输出维度均为 64。为防止过拟合，在全连接层和 CNN 层后均设置了 Dropout 率为 0.2 的随机失活层。优化器采用 Adam，初始学习率为 0.001，权重衰减为 $1e-5$ 。

3.2. 与基础组合模型对比实验

为进一步凸显本文模型在 CNN-BiGRU 骨干网络基础上的综合改进效果，将本文模型与未引入注意力机制的基础 CNN-BiGRU 组合模型进行直接对比，如表 1 所示。结果表明，本文改进方案显著提升了预测精度。相较于基础 CNN-BiGRU，本文模型 MAE (降 26.6%)与 RMSE (降 28.8%)大幅降低， R^2 由 0.8553 提升至 0.9266。这揭示了基础模型虽兼顾了局部与时序特征，但因缺乏特征筛选机制，易受长序列中无关信息干扰，导致预测波动大且对拐点响应滞后。本文引入的多尺度注意力模块通过并行提取与自适应加权，有效滤除噪声并强化关键特征表达，成功解决了上述特征利用效率低的问题。

Table 1. Comparison results with benchmark combined models

表 1. 与基础组合模型对比实验结果

	MAE	RMSE	R^2
CNN-BiGRU	3.823	9.160	0.8553
CNN-BiGRU + 多尺度注意力	2.805	6.525	0.9266

3.3. 与传统注意力机制对比实验

表2对比了在CNN-BiGRU骨干网络中引入传统单尺度与本文多尺度注意力的性能差异。结果表明，多尺度设计优势显著：本文模型 MAE (2.805 MW)和 RMSE (6.525 MW)较传统方案分别降低 10.86%和 11.02%， R^2 提升至 0.9266，是唯一突破 0.92 的模型。这验证了传统单尺度机制难以兼顾光伏序列的短期突变与长期周期。本文多尺度模块通过并行构建不同感受野的特征分支，实现了多时间跨度依赖关系的自适应融合，显著增强了模型在复杂气象下的拟合能力与泛化性能。

Table 2. Comparison results with conventional single-scale attention

表 2. 与传统注意力对比实验结果

	MAE	RMSE	R^2
CNN-BiGRU + 传统注意力	3.146	7.332	0.9073
CNN-BiGRU + 多尺度注意力	2.805	6.525	0.9266

3.4. 与其他方法对比实验

如表3所示，本文提出的 CNN-BiGRU + 多尺度注意力模型在各项指标上均显著优于基准模型。具体而言，相较于单一的 LSTM 与 GRU 模型，本文模型的 MAE 分别降低了 13.8%和 20.3%，RMSE 分别降低了 11.5%和 25.2%， R^2 系数则提升至 0.9266，验证了多尺度注意力机制的有效性。与未加入注意力的 CNN-BiGRU 组合模型相比，本文模型的预测精度也有大幅提升(MAE 降低 26.6%，RMSE 降低 28.8%)。此外，相较于加入传统注意力的变体，本文模型通过多尺度特征加权机制，进一步将 R^2 提升了 2.1%，证明了其在捕捉光伏功率序列强非线性与多尺度耦合特性方面的优越性。

Table 3. Comparison results with other comparative methods

表 3. 与其他方法对比实验结果

	MAE	RMSE	R^2
LSTM	3.253	7.376	0.9062
GRU	3.518	8.720	0.8689
CNN-BiGRU	3.823	9.160	0.8553
CNN-BiGRU + 传统注意力	3.146	7.332	0.9073
CNN-BiGRU + 多尺度注意力	2.805	6.525	0.9266

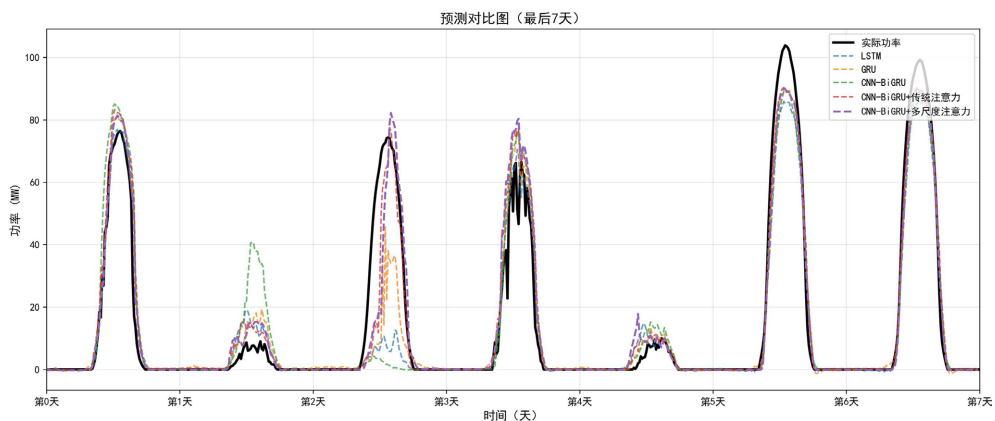


Figure 2. Photovoltaic power forecasting results

图 2. 光伏功率预测结果

测试集连续 7 天(168 h)预测结果如图 2 所示。由图可知, 基准 CNN-BiGRU 因缺乏特征筛选, 在午间峰值段波动剧烈且跟踪滞后; 单一 LSTM 与 GRU 模型虽输出平滑, 但存在欠拟合及对拐点响应迟钝的问题。引入注意力机制显著抑制了非必要波动。其中, 传统单尺度注意力在陡变时刻仍有偏差, 而本文模型拟合度最优, 早间爬坡与午间峰值跟踪精度最高。多尺度注意力并行捕捉了突变与周期特征, 显著提升了模型的鲁棒性与适用性。

4. 结论

本文给出的融合 CNN 特征提取、BiGRU 时序建模与多尺度注意力加权的光伏功率预测方法, 从整体性能、时序预测曲线及误差分布多维度量化验证了所提模型的优越性, 充分证明了多尺度注意力机制在提升光伏功率预测精度与鲁棒性方面的有效性。

参考文献

- [1] 吐松江·卡日, 雷柯松, 马小晶, 等. 基于 LSTM-Attention 和 CNN-BiGRU 误差修正的光伏功率预测[J]. 太阳能学报, 2024, 45(12): 85-93.
- [2] 张雯涵, 艾欣, 王昊洋, 等. 基于相似日划分与常春藤优化的 CNN-BiGRU-Attention 光伏功率预测研究[J]. 电网与清洁能源, 2025, 41(10): 119-127.
- [3] Zou, H., Yang, C., Ma, H., Zhu, S., Sun, J., Yang, J., *et al.* (2024) Short-Term Power Prediction of Distributed PV Based on Multi-Scale Feature Fusion with TPE-CBiGRU-SCA. *IET Generation, Transmission & Distribution*, **18**, 3200-3220. <https://doi.org/10.1049/gtd2.13266>
- [4] 陈殿昊, 臧海洋, 刘璟璇, 等. 基于多模态多尺度特征的超短期光伏功率预测[J]. 太阳能学报, 2025, 46(8): 472-480.
- [5] Amirteimoury, F., Memarzadeh, G., Noori, H., Keynia, F., Heydari, A. and Fekih, A. (2025) A Novel Multi-Scale Feature Attention Network-Based Prediction Model for the Accurate Forecasting of Photovoltaic Power Generation. *IEEE Access*, **13**, 165173-165188. <https://doi.org/10.1109/access.2025.3611168>
- [6] Lee, D.S., Lai, C.W. and Fu, S.K. (2024) A Short- and Medium-Term Forecasting Model for Roof PV Systems with Data Pre-Processing. *Heliyon*, **10**, e27752. <https://doi.org/10.1016/j.heliyon.2024.e27752>
- [7] Yan, K., Liu, J., Zhang, J., Yang, F., Gao, Y. and Du, Y. (2025) Robust Photovoltaic Forecasting under Severe Data Missingness via Multi-Domain Collaboration and Covariate Interaction. *Applied Energy*, **401**, Article ID: 126771. <https://doi.org/10.1016/j.apenergy.2025.126771>
- [8] Cai, J., Cai, Y., Yan, Y., Chen, L. and Zhang, X. (2026) Synergistic Cyclic Optimization Strategy for the Data Screening and Forecasting of Solar Power, Wind Power, and Electricity Load. *Renewable Energy*, **256**, Article ID: 124524. <https://doi.org/10.1016/j.renene.2025.124524>
- [9] Ren, X., Zhang, F., Yan, J. and Liu, Y. (2024) A Novel Convolutional Neural Net Architecture Based on Incorporating Meteorological Variable Inputs into Ultra-Short-Term Photovoltaic Power Forecasting. *Sustainability*, **16**, Article No. 2786. <https://doi.org/10.3390/su16072786>
- [10] 朱文志, 郭力, 刘一欣, 等. 基于改进 CNN-Autoformer 网络的光伏功率短期概率预测方法[J]. 太阳能学报, 2026, 47(3): 678-689.
- [11] Zhang, Y., Hu, C., Yang, Y., Jiao, F., Zhang, X. and Zhou, Y. (2026) Short Term Photovoltaic Power Prediction Using Multi-Scale Time and Frequency Features. *Solar Energy*, **310**, Article ID: 114478. <https://doi.org/10.1016/j.solener.2026.114478>
- [12] 孔令文, 陈思安, 耿若曦, 等. 基于自注意力机制-双向 GRU 的光伏功率间接预测模型[J]. 控制工程, 2026, 33(1): 73-79.
- [13] Dai, Y., Yu, W. and Leng, M. (2024) A Hybrid Ensemble Optimized Bigru Method for Short-Term Photovoltaic Generation Forecasting. *Energy*, **299**, Article ID: 131458. <https://doi.org/10.1016/j.energy.2024.131458>
- [14] 刘国海, 孙文卿, 吴振飞, 等. 基于 Attention-GRU 的短期光伏发电功率预测[J]. 太阳能学报, 2022, 43(2): 226-232.
- [15] 马越, 郭哲卓, 李尚明, 等. 黄土高原集中式光伏电站功率预测模型适用性评价[J]. 延安大学学报(自然科学版), 2026, 45(1): 89-97.