

# 基于新型神经网络的风电机负荷预测研究

钱 凯<sup>1\*</sup> 曾昌学<sup>1</sup> 李文韬<sup>1</sup>  
QIAN Kai ZENG Changxue LI Wentao

## 摘 要

风电机组负荷预测对于电网的稳定安全运行至关重要,因此文章提出了一种新型风电负荷预测方法,在引入低质量测量值作为外源信息的基础上,通过6个预测模型进行处理,从而实现提前1h的增强预报。评估的模型分为两组:第一组是作为参考的持久性模型和ARIMA模型,第二组是基于神经网络的其余4个模型。模型比较通过两种程序实现。一方面,评估4个质量指标(皮尔逊相关系数、一致指数、平均绝对误差和平均平方误差),另一方面,进行方差分析检验和多重比较程序。结果表明,对于持久性模型和ARIMA模型,隐层中有9个神经元的反向传播网络获得的改进对数(平均绝对误差、平均平方误差)分别为23.92%~47.48%、23.19%~45.54%,能够改善风电机组负荷估计和预测。

## 关键词

风电负荷预测;时间序列预测;人工神经网络;现场测量;外源信息

doi: 10.3969/j.issn.1672-9528.2025.11.012

## 0 引言

将风电场并入电网已成为电网稳定性的一个重要问题。电网稳定性的重要问题。由于风的动态特性是混乱的,因此将风力发电场并入电网已成为电网稳定性的一个重要问题。因此,以风力发电为基础的供电并不稳定,需要插入预测技

术<sup>[1]</sup>。短期预测问题的存在风电场的性能未得到优化,这会导致经济损失,并迫使电力系统监管机构制定详细的超规格规划。在风能占总发电量比例较高的地区,优化尤为关键<sup>[2]</sup>。

由于风的动态特性具有随机性,风电场接入电网已成为影响电网稳定性的重要问题。因此,基于风能的电力供应并不稳定,这就需要引入预测技术。短期预测通常基于时间序列分析,从较为简单的模型(自回归模型、线性多元回归、持久性模型)到复杂的结构(计算流体动力学、神经网络、

1. 华电(云南)新能源发电有限公司 云南昆明 650228

## 参考文献:

- [1] 夏金锋. 基于FPGA的实时目标跟踪设计与实现[D]. 南京: 南京信息工程大学, 2022.
- [2] 李刚, 邱尚斌, 林凌, 等. 基于背景差法和帧间差法的运动目标检测方法[J]. 仪器仪表学报, 2006(8): 961-964.
- [3] 王磊. 基于FPGA的视频移动目标检测系统的研究与实现[D]. 济南: 济南大学, 2010.
- [4] 袁国武, 陈志强, 龚健, 等. 一种结合光流法与三帧差分法的运动目标检测算法[J]. 小型微型计算机系统, 2013, 34(3): 668-671.
- [5] 赵建. 基于三帧差法的运动目标检测方法研究[D]. 西安: 西安电子科技大学, 2013.
- [6] 杨丹, 戴芳. 运动目标检测的ViBe算法改进[J]. 中国图象图形学报, 2018, 23(12): 1813-1828.
- [7] 张文喆. 基于FPGA的运动目标检测与跟踪系统设计实现[D]. 吉林: 吉林化工学院, 2024.
- [8] COMANICIU D, MEER P. Mean shift: a robust approach

toward feature space analysis[J]. IEEE transactions on pattern analysis and machine intelligence, 2002, 24(5): 603-619.

- [9] BASAR T. A new approach to linear filtering and prediction problems[C]//Twenty-Five Seminal Papers (1932-1981). Piscataway: IEEE, 2001: 167-179.

## 【作者简介】

郝子坤(2001—), 男, 山东泰安人, 硕士研究生, 研究方向: 智能信息处理与运动目标检测, email: 1292097677@qq.com.

卫建华(1968—), 男, 陕西西安人, 硕士, 副教授, 研究方向: 信息与信息处理、网络通信, email: 596445158@qq.com.

罗鑫迪(1999—), 男, 陕西咸阳人, 硕士研究生, 研究方向: 智能信息处理与图像识别, email: 1721880376@qq.com.

(收稿日期: 2025-05-22 修回日期: 2025-11-13)

模糊逻辑)。针对风电功率预测的问题,许多学者进行了研究。霍子宇等人<sup>[3]</sup>提出了一种计及风电出力概率分布尾部运行风险的两阶段鲁棒机组组合模型及其优化方法,将风电出力预测偏差尾部概率特性纳入机组组合模型中,优化系统风险裕度,提升系统安全性和经济性。谷亚停等人<sup>[4]</sup>提出了一种基于维度转换的深度学习风电功率预测模型—WFDT模型,该模型能够更好地处理复杂多变的风电数据,表现出更优的特征表达能力和预测精度,为风电场的管理和运营提供更准确可靠的决策支持。杨炎昆等人<sup>[5]</sup>提出了一种基于 Stacking 算法的短期风电功率预测集成学习方法,以降低风电出力预测的误差。郎伟明等人<sup>[6]</sup>提出一种基于自适应图卷积网络 (AGCN) - 长短期记忆 (LSTM) 网络的海上风电场功率超短期概率预测模型,形成的动态图拓扑可以更好地揭示海上风电场变化的空间关系,改进后的分位数回归模型有效避免了分位数预测存在的曲线交叉问题。上述研究表明,针对风电功率预测的复杂性与不确定性,学界正通过多维技术路径系统性地优化预测机制,在提升预测精度、稳定性与可靠性的同时,有效平衡电网运行风险与经济性,为高比例新能源接入下的电力系统安全调控提供了多样化解决方案。

本研究利用目标区域周围标准气象站的外源数据,来改进基于人工神经网络的短期预测。本文的创新之处在于,这些气象站的数据目前尚未被世界气象组织纳入考量。从这个角度而言,本文为将这些数据纳入提供了科学建议和方法。

## 1 目标区域和实验数据

本研究的实施过程如下:传统测量站正逐渐被先进的高质量风电负荷测量取代,以提高数据可靠性,这些数据构成了人工神经网络的一个输入阵列。本研究的具体流程如图 1 所示。

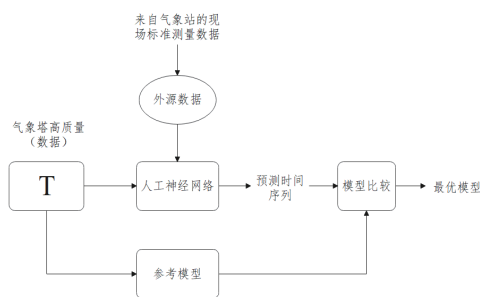


图 1 研究路线

目标塔高 10 m,位于国内某 A 地,每 10 min 采集一次数据,数据采集时间跨度为 2020 年 9 月至 2021 年 1 月。为了提高预测准确性,在半径 40 km 范围内选择了 10 个距离最近的站点,这些站点最初用于测量农业变量,每小时进行一次测量。由于大多数站点位于露天环境,风电负荷记录的可靠性不高,受障碍物影响较大。表 1 总结了各站点相对于目标站点的地理坐标、风电负荷、距离和相关系数。

表 1 本研究中使用的台站信息

| 位置  | 纬度      | 经度       | 海拔 /m | 风电负荷 /MW | 相关系数 |
|-----|---------|----------|-------|----------|------|
| A 地 | 41.0 °N | 113.1 °E | 1 300 | 150      | 1.00 |
| B 地 | 43.2 °N | 93.5 °E  | 800   | 120      | 0.72 |
| C 地 | 40.8 °N | 114.9 °E | 1 500 | 180      | 0.68 |
| D 地 | 40.1 °N | 96.9 °E  | 1 400 | 130      | 0.65 |
| E 地 | 33.4 °N | 120.1 °E | 5     | 250      | 0.58 |
| F 地 | 37.9 °N | 106.2 °E | 1 100 | 140      | 0.51 |
| G 地 | 37.5 °N | 121.4 °E | 50    | 200      | 0.49 |
| H 地 | 45.6 °N | 122.8 °E | 200   | 150      | 0.63 |
| I 地 | 21.9 °N | 111.9 °E | 10    | 300      | 0.32 |
| M 地 | 25.5 °N | 103.8 °E | 1 900 | 200      | 0.55 |
| N 地 | 30.0 °N | 122.1 °E | 20    | 180      | 0.47 |

## 2 算法介绍

本研究探讨了 6 种预测方法。其中 2 种(持久性模型和 ARIMA 模型)作为参考模型,其余基于人工神经网络的模型是本研究的主要关注点。持久性模型是预测时长在 3~6 h 内最常用的参考预测方法<sup>[7]</sup>,假设在时刻  $t$  的预测值与上一次测量值相似,既无需估计参数,也不需要外源变量。

本研究中使用的人工神经网络类型包括:自适应线性神经元、多层神经网络和径向基函数神经网络<sup>[8]</sup>。ADALINE 网络没有隐藏层,比前馈网络更简单,如图 2 所示。其权重和偏差通过 Widrow-Holf 规则进行训练,该规则可最小化均方误差。

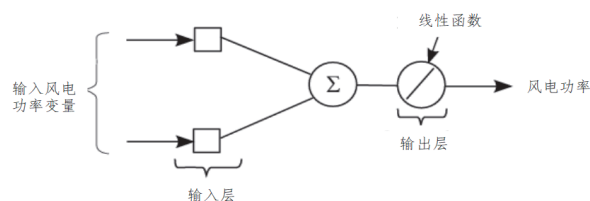


图 2 简单 ADALINE 网络的拓扑结构

多层神经网络,如图 3 所示。基于误差反向传播算法,是应用最广泛的方法<sup>[9]</sup>。该网络是一个动态系统,会随着学习规则发生变化,通过顺序寻找权重来编码知识。网络具有泛化能力,需要对其进行评估。本研究选择 Levenberg-Marquardt 算法,以最小化每一步输出之间差异的均方误差。

径向基函数神经网络的形式与两层多层网络类似,如图 4 所示<sup>[10]</sup>。其根本区别在于隐藏神经元,基于输入与突触向量(称为质心)之间的距离进行运算。径向基函数神经元具有局部响应特性,只有当输入向量与神经元的质心位于输入空间的相邻区域时,才会以明显的强度响应。

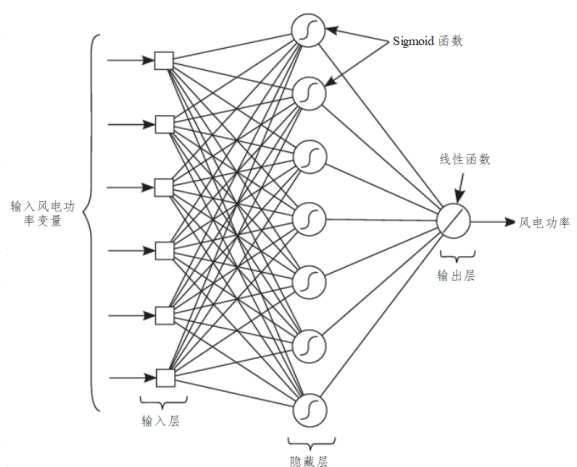


图3 单隐藏层神经网络示例

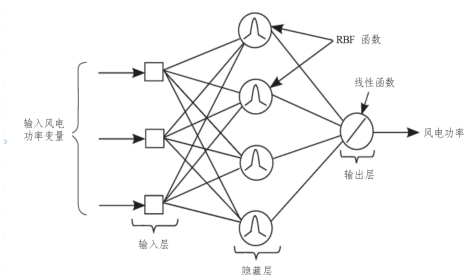


图4 径向基函数（RBF）神经网络的拓扑结构

### 3 建模与仿真

第一步是选择评估模型的质量指标：皮尔逊相关系数（ $\rho$ ）、一致性指数（IOA）、平均绝对误差（MAE）和均方误差（MSE）<sup>[11]</sup>。在选择质量指标后，使用目标站点（A地）获取的风电负荷时间序列对参考模型（持久性模型和ARIMA模型）进行评估。

$$\rho = \frac{\sigma_{y\hat{y}}}{\sigma_y \sigma_{\hat{y}}} \quad (1)$$

$$IOA = 1 - \frac{\sum_{t=1}^N (\hat{y}_t - y_t)^2}{\sum_{t=1}^N (|\hat{y}_t - y_t| + |y_t + \hat{y}_t|)^2} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N (|\hat{y}_t - y_t|) \quad (3)$$

$$MSE = \frac{1}{N} \sum_{t=1}^N (|\hat{y}_t - y_t|^2) \quad (4)$$

持久性模型的实施没有任何问题。然而，ARIMA模型需要初步确定最适合目标时间序列的阶数 $p$ 、 $d$ 和 $q$ 。为此，需要评估并绘制自相关系数（ACF）和偏自相关系数（PACF），结果如图5所示，自相关系数随着时间滞后的增加而衰减，但在许多时间滞后处仍具有正相关性，这表明需要进行一阶差分。图6展示了一阶差分时间序列的模型选择结果。通过对该图的观察，确定了3个潜在模型：ARIMA(2,1,0)，因为偏自相关系数在二阶滞后之后截尾；ARIMA(0,1,2)，因为自

相关系数在二阶滞后处衰减；ARIMA(2,1,2)，因为自相关系数和偏自相关系数的衰减均从二阶滞后开始。

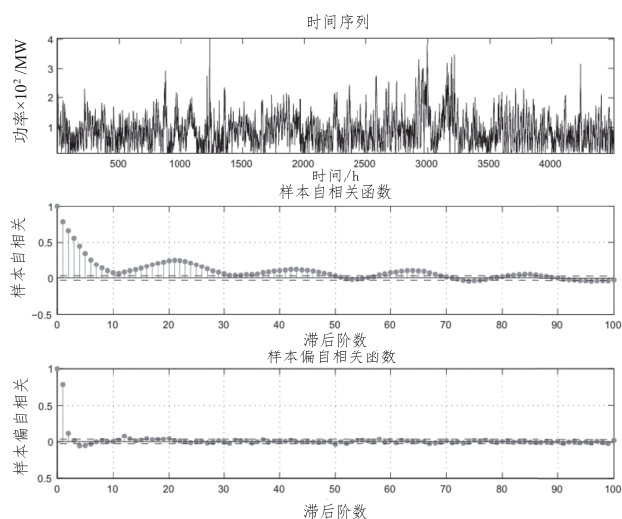


图5 国内A地的风电负荷时间序列图、样本自相关函数图和样本偏自相关函数图

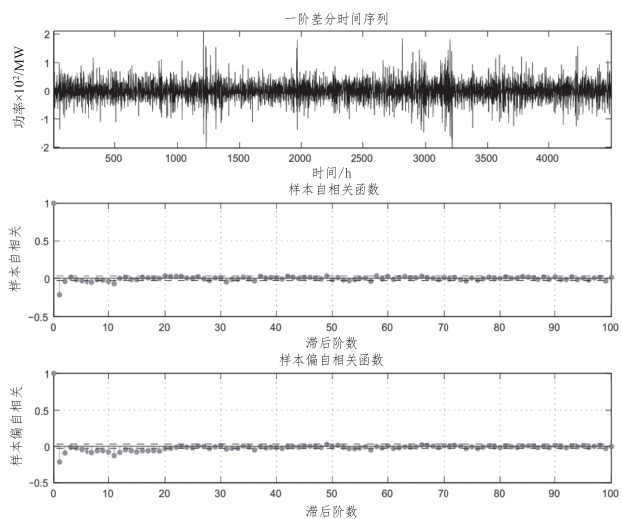


图6 国内A地经差分处理后的原始风电负荷时间序列图、样本自相关函数图和样本偏自相关函数图

ARIMA模型过程的最后一步是诊断检验，在这一步中验证残差的假设，只有通过检验，模型才能用于预测<sup>[12]</sup>。本文测试了以下模型：线性网络（LIN）、具有一个和两个隐藏层的反向传播网络（BP1和BP2），以及径向基函数网络（RBF）。考虑以下前提条件：

（1）第2节中的参考站点数据被用作外源变量，以改进预测。

（2）对数据进行归一化处理，使其在区间[-1,1]内，以加快计算速度。

（3）对测试网络进行训练，直到验证误差开始增加。此时停止训练，并在测试集中评估网络的性能。每个模型进行100次实验，取最佳结果。

表 2 收集了神经网络的参数, 包括网络架构和激活函数。选择 BP1 和 BP2 模型隐藏层神经元数量范围的规则如下: 第一个隐藏层的神经元数量是输入层和输出层神经元数量的平均值, 第二个隐藏层的神经元数量是第一个隐藏层神经元数量的一半。

表 2 所用网络模型的参数

|              | 线性       | 反向传播 1      | 反向传播 2      | 径向基函数   |
|--------------|----------|-------------|-------------|---------|
| 隐藏层          | —        | 1           | 2           | 1       |
| 第一层隐藏层的神经元数量 | —        | [4-10]      | [4-10]      | [1-150] |
| 第二层隐藏层的神经元数量 | —        | —           | [2-5]       | —       |
| 传递函数         | —        | S 型函数       | S 型函数       | 高斯函数    |
| 传递函数输出层      | 线性函数     | 线性函数        | 线性函数        | 线性函数    |
| 训练算法         | 威德罗 - 霍夫 | 列文伯格 - 马夸尔特 | 列文伯格 - 马夸尔特 | k 均值    |
| 扩展常数         | —        | —           | —           | [1-20]  |

4 结果与分析

本研究测试的模型配置超过 3 000 种, 涵盖不同的模型类型 (LIN、BP 或 RBF)、各层神经元数量以及模拟中使用的外源变量: 其中 93% 的模型性能优于持久性模型, 80% 的模型性能优于 ARIMA(2,1,0) 模型。为便于比较, 表 3 展示了生成最小均方误差的 10 个最佳模型在测试集上的质量指标结果, 以及参考模型 (持久性模型和 ARIMA(2,1,0) 模型) 的相关值。

如表 3 所示, 基于低质量站点数据的模型结果显著优于参考模型: 持久性模型和 ARIMA(2,1,0) 模型。所有测试模型中表现最佳的是具有一个隐藏层且包含 5 个神经元的反向传播网络 (BP1)。该模型的输入有 6 个, 其中一个是目标站点的风电负荷, 其余 5 个是距离最近的 5 个站点的外源风电负荷值。相对于持久性模型, 该模型的平均绝对误差降低了 23.92%, 均方误差降低了 47.48%。相对于 ARIMA(2,1,0) 模型, 改进幅度分别为 23.19% 和 45.54%。

表 3 参考模型与 6 个最佳模型在测试集上的参数及结果

| 模型    | 外部负荷 | 外部方向 | nh1 | nh2 | R       | IOA     | MAE     | MSE     |
|-------|------|------|-----|-----|---------|---------|---------|---------|
| 持久性   | —    | —    | —   | —   | 0.649 2 | 0.764 2 | 0.233 4 | 0.100 5 |
| ARIMA | —    | —    | —   | —   | 0.645 2 | 0.761 7 | 0.232 3 | 0.096 9 |
| BP1   | 8    | 0    | 9   | —   | 0.762 0 | 0.813 0 | 0.178 7 | 0.052 8 |
| BP2   | 5    | 0    | 6   | 5   | 0.760 0 | 0.814 2 | 0.176 8 | 0.053 0 |
| BP2   | 5    | 0    | 7   | 4   | 0.851 7 | 0.810 4 | 0.184 4 | 0.053 4 |
| BP1   | 7    | 0    | 9   | —   | 0.852 4 | 0.809 1 | 0.181 3 | 0.053 3 |
| BP2   | 9    | 0    | 4   | 4   | 0.853 6 | 0.807 3 | 0.186 5 | 0.053 9 |

为完成本研究, 进行了基于多重比较的统计分析, 以选择最佳模型。该分析包括两个测试: 参数检验 (方差分析, ANOVA) 和 Bonferroni 校正<sup>[13]</sup>。方差分析用于评估两个以上样本之间的差异, Bonferroni 校正用于获得最佳平均模型。表 4 展示了各模型、相应的平均误差以及检验结果。该表按均方误差从小到大排序。

表 4 检验结果

| 模型         | 误差均值     | 无显著差异的模型     |
|------------|----------|--------------|
| 1          | 0.057 25 | 1,8,3,7,10,6 |
| 3          | 0.057 58 | 1,8,3,7,10,6 |
| 2          | 0.057 64 | 1,8,3,7,10,6 |
| 5          | 0.057 67 | 1,8,3,7,10,6 |
| 4          | 0.057 73 | 1,8,3,7,10,6 |
| ARIMA(210) | 0.096 94 | 1,8,3,7,10,6 |
| 持续性模型      | 0.100 53 | 1,8,3,7,10,6 |

图 7 展示了最佳模型的性能, 同时也展示了两个参考模型 (持久性模型和 ARIMA 模型)。

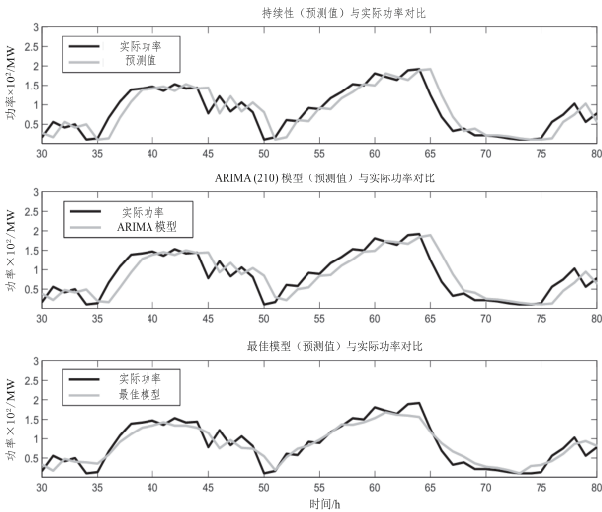


图 7 原始时间序列以及使用所得到的最佳模型进行预测的数据

数据来自测试集, 为更清晰地展示结果, 仅显示了 50 h 的数据。正如预期的那样, 基于人工神经网络的最佳模型生成的时间序列明显比参考模型更准确。

最后, 对最佳架构 (BP1-5 个隐藏神经元) 进行分析。评估基于外源变量的数量。首先, 仅将目标站点记录的时间序列作为唯一输入 (不使用外源变量), 然后逐个引入外源时间

序列以评估其作用。如图8所示,通过纳入外源测量数据,平均绝对误差和均方误差指标得到了改善。使用距离最近的6个站点的风电负荷记录时,效果最佳。

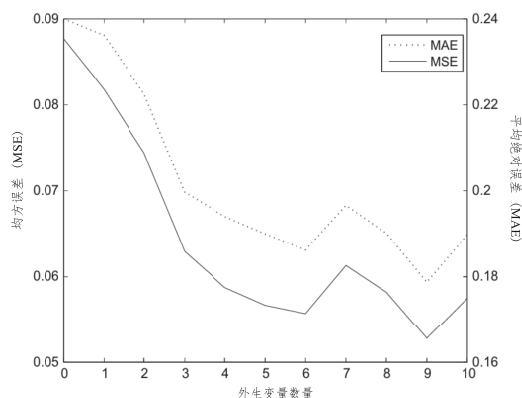


图8 使用最佳模型得到的原始时间序列和预测时间序列

## 5 结论

本研究的主要贡献在于,利用周围农业测量站获取的低质量风电负荷数据,可显著提高目标区域的风电负荷预测准确性。此外,通过测试不同的统计结构,找到了一种优化利用这些外源站点信息的模型。具体而言,最佳模型是具有一个隐藏层和5个神经元的反向传播网络。该模型使用5个外源风电负荷值,根据四个质量指标(皮尔逊相关系数、一致性指数、平均绝对误差和均方误差,显著优于两个参考模型。对于持久性模型,改进幅度分别为23.92%和47.48%;相对于ARIMA(2,1,0)模型,改进幅度分别为23.19%和45.54%。

从研究结果可以看出,在预测模型中应用低质量数据的优势显著。因此得出结论,那些因不符合世界气象组织要求而通常未被纳入预测模型的低质量站点数据,有助于提高先进目标预测准确性。尽管本研究并未声称其结果具有普遍适用性,但提供了有力的实践证据,表明这些数据具有改善估计和预测的潜力。

## 参考文献:

- [1] 符杨,任子旭,魏书荣,等.基于改进LSTM-TCN模型的海上风电超短期功率预测[J].中国电机工程学报,2022,42(12):4292-4303.
- [2] 张智浩,王聪,张宏立,等.时间生成对抗网络样本扩增下的风电功率预测[J/OL].电力系统及其自动化学报,1-12[2025-10-03].<https://doi.org/10.19635/j.cnki.csu-epsa.001588>.
- [3] 霍子宇,鲁宗相,乔颖,等.计及风电出力预测偏差概率分布尾部运行风险的鲁棒机组组合优化[J].电网技术,2025,

49(5):2014-2023.

- [4] 谷亚停,闫斯哲,田茂再.基于维度转换的深度学习风电功率预测研究[J].武汉理工大学学报(信息与管理工程版),2025,47(1):101-108.
- [5] 杨炎昆,李玉玲,杨仕友.短期风电出力预测的融合模型比较[J].电工技术,2024(24):75-78.
- [6] 郎伟明,麻向津,周博文,等.基于LSTM和非参数核密度估计的风电功率概率区间预测[J].智慧电力,2020,48(2):31-37.
- [7] 刘福才,窦金梅,王树恩.基于MFOA和LW的混沌时间序列鲁棒模糊预测[J].智能系统学报,2014,9(4):425-431.
- [8] 刘福才,朴春俊,裴润.模糊模型辨识中模糊聚类方法应用分析[J].系统工程与电子技术,2002(5):35-37.
- [9] 张家安,黄晨旭,李志军.考虑局部条件特征的风电功率短期预测[J].太阳能学报,2024,45(12):220-227.
- [10] 骆钊,吴谕侯,朱家祥,等.基于多尺度时间序列块自编码Transformer神经网络模型的风电超短期功率预测[J].电网技术,2023,47(9):3527-3537.
- [11] 王渝红,史云翔,周旭,等.基于时间模式注意力机制的BiLSTM多风电机组超短期功率预测[J].高电压技术,2022,48(5):1884-1892.
- [12] 邢培宇,魏云冰,胡骅,等.基于mRMR和SVMD-TPA-BiLSTM的短期风电功率预测[J].上海工程技术大学学报,2024,38(4):414-421.
- [13] 彭毅.基于DBO-PNN的短期风电功率预测模型[J].中国新技术新产品,2024(24):27-29.

## 【作者简介】

钱凯(1990—),通信作者(email:wyr@ipzhongnan.com),男,云南红河人,本科,工程师,研究方向:通信自动化网络安全。

曾昌学(2000—),男,四川乐山人,专科,助理工程师,研究方向:自动化控制。

李文韬(2001—),男,云南大理人,本科,助理工程师,研究方向:自动化控制。

(收稿日期:2025-05-09 修回日期:2025-11-11)