

基于分解技术的 IZOA-Transformer-BiGRU 短期风电功率预测*

蒲晓云 杨靖 杨兴 宁媛

(贵州大学 电气工程学院 贵阳 550025)

摘要: 准确的风电功率预测对于保障电网平稳运行和提升风资源利用效率具有重要意义。针对风电功率数据的非平稳性和间歇性等特征,本文提出了一种结合数据分解技术的 IZOA-Transformer-BiGRU 组合预测模型,以提升短期风电功率预测的精度和可靠性。首先,采用能量差值法确定变分模态分解(VMD)的子模态数,将具有较强随机波动性的原始风电功率分解为一系列相对平稳的子序列,从而更加充分地提取时序特征。其次,构建 Transformer-BiGRU 模型,引入多头注意力机制并行处理多个特征之间的交互关系,并利用 BiGRU 捕捉时序序列间的前后依赖性,从而提升预测性能。为了进一步优化模型性能,采用融合 Singer 混沌映射、透镜折射反向学习和单纯形法策略的改进斑马优化算法(IZOA),对 Transformer-BiGRU 模型的隐藏层神经元数、初始学习率、正则化系数和多头注意力头数四个关键超参数进行优化。最后,通过 IZOA-Transformer-BiGRU 对分解后的各子序列进行预测,经过叠加重构得到最终的预测结果。实验结果表明,与单一 BiGRU 模型相比,所提模型的决定系数提升了 5.10%,平均绝对误差、均方根误差以及平均绝对百分比误差分别降低了 56.17%、54.58%、54.55%,具有较高的预测精度。

关键词: 风电功率预测;变分模态分解;Transformer;双向门控循环单元;能量差值法;斑马优化算法

中图分类号: TN91;TM614 **文献标识码:** A **国家标准学科分类代码:** 470.4047

IZOA-Transformer-BiGRU short-term wind power prediction based on decomposition technique

Pu Xiaoyun Yang Jing Yang Xing Ning Yuan

(School of Electrical Engineering, Guizhou University, Guiyang 550025, China)

Abstract: Accurate wind power prediction is crucial for ensuring the stable operation of power grids and improving the efficiency of wind resource utilization. To address the non-stationary and intermittent characteristics of wind power data, this paper proposes a combined IZOA-Transformer-BiGRU prediction model based on data decomposition techniques to enhance the accuracy and reliability of short-term wind power forecasting. First, the energy difference method is employed to determine the number of sub-modalities for variational mode decomposition, which decomposes the original wind power with strong random fluctuations into a series of relatively stable sub-sequences, enabling better more effective extraction of temporal features. Next, the Transformer-BiGRU model is constructed, incorporating a multi-head attention mechanism to process interactions between multiple features in parallel, while the BiGRU component captures temporal dependencies within the sequence, thus enhancing prediction performance. To further improve the model's forecasting accuracy, an improved zebra optimization algorithm, integrating singer chaotic mapping, lens refraction-based learning, and the simplex method, is developed to optimize four key hyperparameters of the Transformer-BiGRU model: the number of hidden layer neurons, initial learning rate, regularization coefficient, and the number of attention heads. Finally, the IZOA-Transformer-BiGRU model predicts the subsequences derived from VMD, and the final prediction is reconstructed through aggregation. Experimental results show that, compared to the standalone BiGRU model, the proposed model improves the coefficient of determination by 5.10% and reduces the mean absolute error, root mean square error, and mean absolute percentage error by 56.17%, 54.58% and 54.55%, respectively, demonstrating its high prediction accuracy.

Keywords: wind power prediction; variational modal decomposition; Transformer; bidirectional gated recurrent unit; energy difference method; zebra optimization algorithm

0 引言

随着全球对化石能源消费的依赖加剧,环境问题日益

凸显,开发可再生能源已成为全球共识。风能作为一种清洁、储量丰富的可再生能源,受到广泛关注。截至2023年底,全球风电装机容量达到923 GW^[1]。然而,风能固有的

收稿日期:2024-10-17

* 基金项目:国家自然科学基金(62261005)、贵州省科技支撑项目(黔科合[2022]一般017,黔科合 ZK[2022]135)、贵州电科院2023年计及显著季节负荷与光伏波动性的单相调容变压器关键技术示范项目(K23-0109-014)资助

波动性、间歇性等特点,使得大规模风电并入电网时带来严峻挑战。因此,准确的短期风电功率预测在保障电网平稳运行和降低调度成本方面具有重要意义^[2]。

近年来,国内外学者在短期风电功率预测领域开展了大量研究,主要包括物理法、统计法和深度学习法^[3]。物理法依赖于数值天气预报系统^[4],需要大量气象、地形等数据支持,但模型建立复杂且计算成本高。统计法通过历史风电功率和气象规律建立非线性映射关系,具有一定灵活性,但在面对非平稳复杂数据时精度有限。近年来,支持向量回归^[5]、门控循环网络(gated recurrent unit, GRU)^[6]等深度学习模型因其出色的数据表征和挖掘能力被广泛应用于风电功率预测。文献[7]采用了双向门控循环网络(bidirectional gated recurrent unit, BiGRU)模型,通过双向信息处理充分捕捉数据特征,提高了预测精度。然而, BiGRU 在应对多维复杂数据时,其特征提取能力有限,难以平衡数据的全局性与局部性特征。为提升模型性能,研究者尝试将深度学习模型与其他方法相结合,发挥模型的互补优势。例如,文献[8]提出了 GRU-CNN 组合模型,通过卷积神经网络(convolutional neural networks, CNN)提取时序和空间特征,同时引入时序注意力机制优化特征提取能力,但在捕捉高频波动方面仍显不足。文献[9]将时间卷积神经网络(temporal convolutional networks, TCN)与 Transformer 相结合,利用局部特征提取与自注意力机制增强对关键特征的捕捉,但在应对大规模非平稳数据时,模型对复杂波动和突变的适应能力不足。

面对复杂波动数据,研究者通过引入数据分解技术降低了输入数据的复杂度和非平稳性,从而提升预测效果。例如,文献[10]提出结合经验模态分解(empirical mode decomposition, EMD)和 GRU 模型的方法,通过分解非平稳信号提升预测精度,但 EMD 易受模态混叠和端点效应的影响。相比之下,文献[11]利用变分模态分解(variational mode decomposition, VMD)有效解决了模态混叠问题,并进一步挖掘了信号的潜在信息。然而, VMD 在分解前需要合理设置子模态数,直接影响分解效果和预测性能。此外,深度学习预测模型的参数选择也通常依赖人工经验或反复试验,增加了优化成本。为此,采用智能优化算法进行超参数寻优已成为解决参数设定困难的有效方法之一。文献[12]采用灰狼优化算法对 LSTM 模型进行超参数优化,提升了预测性能,但易陷入局部最优。文献[13]通过结合混沌映射和高斯变异策略改进了麻雀优化算法,有效增强了算法全局搜索能力,提升了模型的预测效果。

针对单一模型特征提取能力不足、风电功率数据复杂波动以及模型参数优化繁琐的问题,本文提出了一种基于分解技术和改进斑马优化算法(improved zebra optimization algorithm, IZOA)的 Transformer-BiGRU 短期风电功率预测模型。首先,采用能量差值法计算 VMD

的子模态数,将原始风电功率分解为一系列频率各异且相对平稳的子序列,从而更全面地提取时序特征信息。其次,构建 Transformer-BiGRU 组合模型,结合 Transformer 的多头注意力机制挖掘全局特征,利用 BiGRU 捕捉时序数据中的长期依赖关系,增强模型的预测性能。为了进一步提升预测效果,采用融合 Singer 混沌映射、透镜折射反向学习和单纯形法的改进斑马优化算法,对 Transformer-BiGRU 模型的 4 个关键超参数进行优化。最后,将 VMD 分解得到的子序列分别输入 IZOA-Transformer-BiGRU 模型进行预测,并通过叠加重构得到最终的预测结果。实验表明,本文所提模型显著提高了短期风电功率预测的精度,验证了其有效性与应用价值。

1 变分模态分解

1.1 变分模态分解

变分模态分解是一种完全非递归的信号分解技术,将非线性、具有波动性的时序序列进行分解,得到若干个不同频率的子序列。

1)将原始输入信号 $f(t)$ 自适应分解为 R 个模态分量,构造变分约束模型:

$$\min_{(u_r) \dots (\omega_r)} \left\{ \sum_{r=1}^R \left\| \delta(t) \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_r \right] e^{-j\omega_r t} \right\|_2^2 \right\} \quad (1)$$

$$s.t. \sum_{r=1}^R u_r(t) = f(t) \quad (2)$$

式中: $\delta(t)$ 为单位脉冲信号, u_r 为各层模态分量, ω_r 为各层模态分量对应中心频率, R 为分解层数。

2)将约束问题转化为非约束问题进行求解,构建扩展的拉格朗日函数:

$$L(u_r, \omega_r, \lambda) = \alpha \sum_{r=1}^R \left\| \delta(t) \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_r \right] e^{-j\omega_r t} \right\|_2^2 + \left\| f(t) - \sum_{r=1}^R u_r(t) \right\|_2^2 + \langle \lambda(t), f(t) - \sum_{r=1}^R u_r(t) \rangle \quad (3)$$

式中: α 为二次惩罚因子, λ 为拉格朗日乘子。

3)通过交替方向乘子法对 u_r 、 ω_r 和进行更新迭代,完成非约束问题求解。

$$\lambda^{(n+1)}(\omega_r) = \lambda^{(n)}(\omega_r) + \tau(x(\omega_r) - \sum_{r=1}^R u_r^{(n+1)}(\omega_r)) \quad (4)$$

4)设置迭代停止条件,直到达到收敛精度或者最大迭代次数得到拉格朗日函数的最优解,收敛条件如式(5)所示。

$$\epsilon = \frac{\sum_{r=1}^R \|u_r^{n+1} - u_r^n\|_2^2}{\|u_r^n\|_2^2} \quad (5)$$

式中: ϵ 为收敛精度, n 为迭代次数。

1.2 分解模态数

经 VMD 分解后可得到有限个子模态序列,当每个子

序列满足正交性时,说明原始序列能量和经过VMD分解得到的子序列能量之和是一致的。由于 R 值过大时会产生虚拟分量,会导致各分量的能量之和过高,故采用能量差值法^[14]计算 R 值,通过计算不同分解模态数的能量差,选择其最小值为最佳分解模态数,计算公式如下:

$$E = \sqrt{\frac{\sum_{t=1}^{\tau} f^2(t)}{n}} \quad (6)$$

$$\eta = \frac{|E_r - E_{r-1}|}{E_{r-1}} \quad (7)$$

式中: E 为各序列能量之和, τ 为序列长度, η 为能量差值。

2 IZOA-Transformer-BiGRU 模型

2.1 气象特征选取

风电功率预测结果受多种气象因素影响。本文选取贵州某风电场2021年2月的气象和功率数据作为研究对象。其中,包含风机侧出风速、高度80 m测风塔风速、风向、温度、湿度、气压共6气象特征。然而,若将影响较小的冗余气象特征选作输入,不仅会复杂模型结构,还会影响训练效率和准确度。因此,利用最大信息系数法计算每个气象特征与实际风电功率之间的相关系数,选取相关系数较大的气象因素作为风电功率预测模型的输入。最终筛选气象特征为风机侧出风速、高度80 m测风塔风速、风向。

2.2 数据预处理

由于风电相关数据采用SCADA系统传感器自动采集,会存在一些异常和缺失值。因此,在风电功率预测前,需要对数据进行预处理。

1)异常值清洗:通过对基于密度的带噪声空间聚类(density based spatial clustering of applications with noise, DBSCAN)算法中的邻域半径和邻域密度阈值进行联合调参,将密度堆积的点分类为形状不同的数据簇,剔除噪声点。

2)缺失值填补:由于时序序列之间相邻数据具有较强的相关性,对原始缺失值采用线性插值法进行填补。

3)数据归一化:为减少量纲对预测结果的影响,对功率和气象数据进行max-min归一化处理。

2.3 改进斑马优化算法

斑马优化算法^[15]是源于模拟斑马觅食和防御两种行为模式所提出,具有模型参数较少、寻优效率高等特点,但面对复杂场景仍存在种群初始化分布不均、易陷入局部最优等不足。

1)Singer混沌映射:由于ZOA算法在初始阶段采用随机变量生成种群,导致初始个体具有较大的随机性和不确定性,一定程度上影响了算法寻优性能。Singer混沌映射具有数学形式简单、遍历性和规律性等特点,按照一定规律不重复地遍历搜索空间以增加种群多样性。利用这个特点可以改善初始种群在搜索空间上的分布质量,加强算法全

局搜索能力。

$$m_{i+1} = \mu(7.86m_i - 23.31m_i^2 + 28.75m_i^3 - 13.302875m_i^4) \quad (8)$$

式中: $m \in (0.9, 1.08)$, $\mu \in (0, 1)$ 。

2)透镜折射反向学习:在算法觅食阶段确认种群最优个体作为先锋斑马,每次迭代更新都会引导其他个体向最优解移动,导致迭代后期容易使种群陷入局部最优,故引入透镜折射反向学习提高种群多样性,增强算法跳出局部最优能力。透镜折射反向学习结合凸透镜成像现象和反向学习,利用光线进入不同传输介质时传播方向改变的原理。根据当前位置生成反向解,扩大搜索范围,并通过比较选择最优解进行迭代寻优。其中, x 、 $\tilde{x}$ 是坐标轴 $[a, b]$ 内的两个点,根据式(10)生成反向解。

$$p = \frac{\sin(\theta)}{\sin(\xi)} = \frac{((a+b)/2 - x)/\bar{L}}{(x - (a+b)/2)/L} \quad (9)$$

$$\tilde{x} = \frac{a+b}{2} + \frac{a+b}{2p} - \frac{x}{p} \quad (10)$$

式中: θ 、 ξ 为入射角和反射角, L 、 $\bar{L}$ 分别为入射光线与折射光线长度, p 为折射率。

3)单纯形法:为进一步提升算法后期收敛速率和精度,在防御阶段引入具有计算量小、收敛速度快等优点的单纯形法策略。通过建立一个 $s+1$ 个顶点的多面体,并计算其适应度值大小,得出最优点 X_g 、次优点 X_b 和最差点 X_w ,再根据其位置计算出中点位置 X_c 。利用反射、扩张、收缩等策略不断优化最差解,从而接近最优点。

(1)反射操作: $X_d = X_c + \phi(X_c - X_w)$, ϕ 为反射系数,本文取1。

(2)扩张操作:当 $f(X_w) > f(X_c)$ 时, γ 为反射系数, $X_c = X_c + \gamma(X_d - X_c)$,本文取2。

(3)内压缩操作:当 $f(X_g) > f(X_d)$ 时, $X_t = X_c + \beta(X_w - X_c)$, β 为压缩系数,本文取0.5。

(4)外压缩操作:当 $f(X_w) < f(X_t) < f(X_g)$ 时, $X_s = X_c - \beta(X_w - X_c)$ 。

采用6种基准测试函数^[16]与ZOA、GWO、MTBO算法进行对比,验证改进算法的有效性。测试函数收敛曲线如图1所示。在求解不同测试函数时,IZOA算法的收敛速度更快,曲线更加平滑。在 $F_1 \sim F_3$ 单峰测试函数中,IZOA算法收敛速度和精度有明显提升,并在 F_2 测试函数中提前达到最优值0。在 $F_4 \sim F_6$ 多峰测试函数中,IZOA算法能有效搜索整个解空间,解决了其他算法过早陷入局部最优,收敛精度缓慢甚至停滞的问题。因此,多策略协同改进算法IZOA具有良好的鲁棒性。

2.4 Transformer-BiGRU 模型

门控循环单元是循环神经网络的一种变体结构,它简化了LSTM网络的门控机制,从而减少训练参数和计算复杂度。但GRU网络通过门控机制控制时序信息流动时,只能获取到时间维度内单向传播的信息。BiGRU^[17]是由

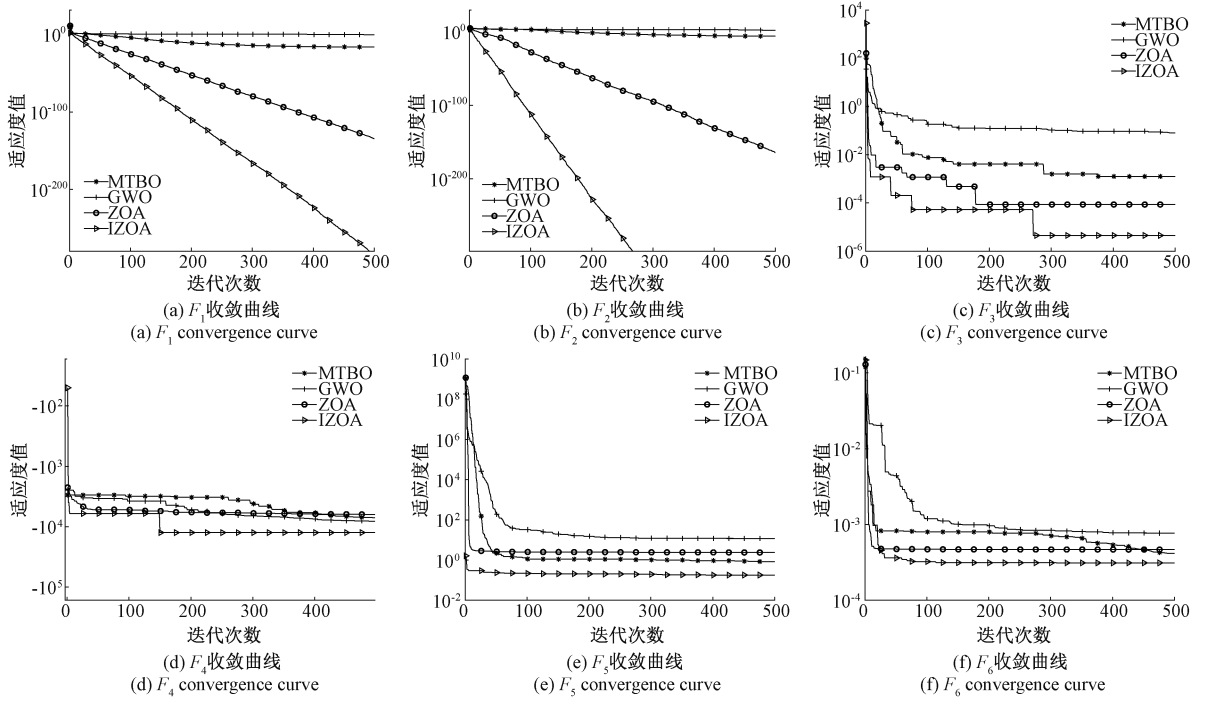


图 1 测试函数收敛曲线

Fig. 1 Convergence curves of the test function

两个独立的 GRU 网络组成,可以更全面地捕获数据中的时序关系和多变量之间的相互影响结构。在输入相同的时间序列后,分别从正向和反向处理输入序列中每个位置的上下文信息,然后将隐藏向量拼接作为最终输出,进一步加强时间序列特征提取,增强网络的记忆能力。

在每个时刻中网络隐藏状态由两个不同方向的 GRU 网络共同迭代,相应表达式如下:

$$\begin{cases} \vec{h}_t = \text{GRU}(x_t, \vec{h}_{t-1}) \\ \vec{h}_t = \text{GRU}(x_t, \vec{h}_{t-1}) \\ h_t = u_t \vec{h}_{t-1} + v_t \vec{h}_{t-1} + b_t \end{cases} \quad (11)$$

式中: $\vec{h}_t$ 、 $\vec{h}_t$ 为正向和反向 GRU 网络的 t 时刻输出状态, u 、 v 为正向、反向隐藏层对应的权重, b 为偏置矩阵。

Transformer^[18] 是利用注意力机制来提高模型训练速度和性能的模型。其核心机制为自注意力机制,在模型处理不同时间点的序列时,将输入序列的每个部分与其他部分进行关联和注意,使模型更好捕捉长距离依赖关系。

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (12)$$

式中: softmax 为归一化函数, $\sqrt{d_k}$ 为缩放因子。

多头注意力机制是由多个的自注意力组成,允许在多个子空间内进行并行计算。每个头专注于输入的不同特征信息或上下文关系以捕获关键信息,从而减少非必要因素的影响。首先,每个头部进行线性变换得到不同的查询(Query)、键(Key)和值(Value),并行计算这些向量的注意

力得分,以确定输出的权重。然后,将各个头的计算结果进行维度拼接,再经过线性变换产生最终的输出。这样,模型可以将来自不同头部的信息整合在一起,提高模型对输入数据的空间利用率。

$$\text{MultiHead}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Concat}(\text{head}_1, \dots, \text{head}_M) \mathbf{W}^{\text{output}} \quad (13)$$

$$\text{head}_M = \text{Attention}(\mathbf{Q}\mathbf{M}_M^Q, \mathbf{K}\mathbf{M}_M^K, \mathbf{V}\mathbf{M}_M^V) \quad (14)$$

式中: Concat 为维度拼接函数, $\mathbf{W}^{\text{output}}$ 为变换矩阵。

为提高风电功率预测效果,将多个气象特征和风电功率作为输入,搭建 Transformer-BiGRU 组合模型。Transformer-BiGRU 网络结构如图 2 所示。首先,采用 Transformer 网络的多头注意力机制计算不同输入特征之间的权重,自适应识别并聚焦于重要信息,充分提取多个气象特征和风电功率之间的关系;再输入到 BiGRU 网络,利用序列中的前后信息,动态捕捉非线性时序序列间的长期依赖关系,增强对数据的建模能力。通过结合 Transformer 的多头注意力机制和 BiGRU 网络的优势,构建组合模型能有效提取多特征数据的深层次信息,从而更加适合复杂的多特征输入的风电功率预测任务。

2.5 VMD-IZOA-Transformer-BiGRU 功率预测模型

鉴于风电功率序列具有一定波动性,本文采用分解技术和参数优化的 Transformer-BiGRU 组合预测模型以提高功率预测精度。组合模型预测流程如图 3 所示。具体步骤如下:

1) 对原始数据进行异常清洗、插值填补和归一化处

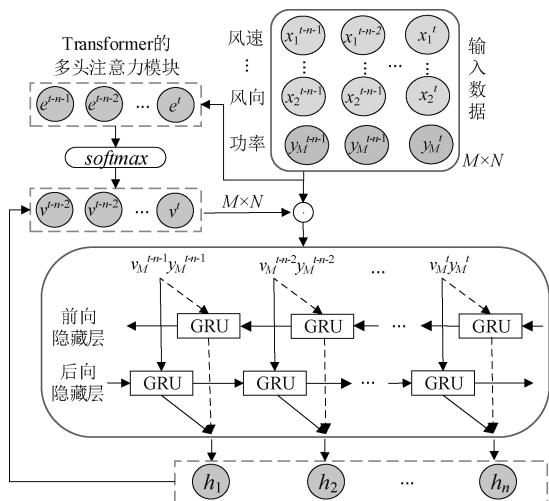


图2 Transformer-BiGRU网络结构

Fig. 2 Transformer-BiGRU network architecture

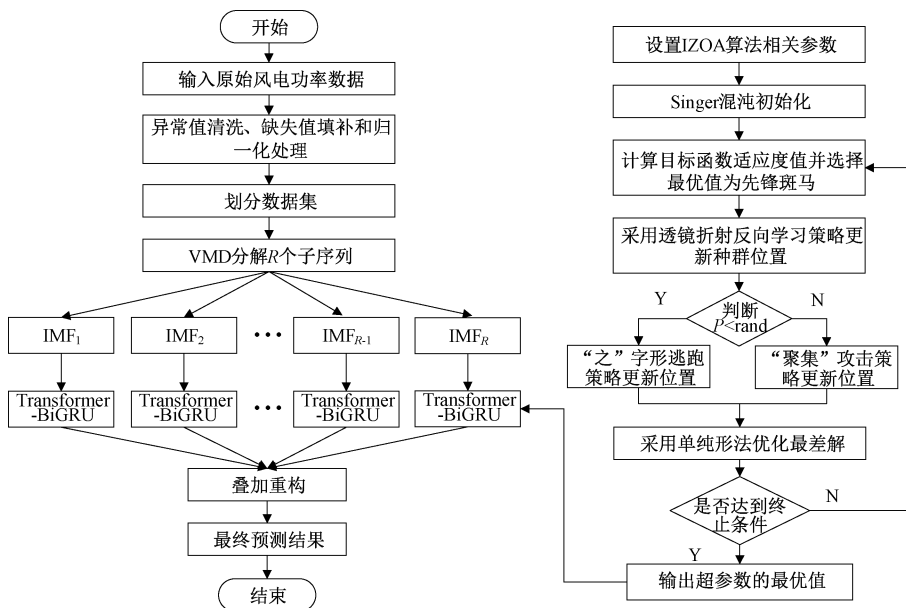


图3 基于VMD-IZOA-Transformer-BiGRU组合模型预测流程

Fig. 3 Prediction process based on VMD-IZOA-Transformer-BiGRU combined Models

$$\sigma_{RMSE} = \sqrt{\frac{1}{q} \sum_{i=1}^q (\hat{y}_i - y_i)^2} \quad (15)$$

$$\sigma_{MAE} = \frac{1}{q} \sum_{i=1}^q |y_i - \hat{y}_i| \quad (16)$$

$$\sigma_{MAPE} = \frac{100\%}{q} \sum_{i=1}^q |(y_i - \hat{y}_i) / y_i| \quad (17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^q (\hat{y}_i - y_i)^2}{\sum_{i=1}^q (\bar{y}_i - y_i)^2} \quad (18)$$

式中: q 为样本容量, y_i 为实际值, $\hat{y}_i$ 为预测值, $\bar{y}_i$ 为真实值的均值。

理,同时采用最大信息系数法筛选气象特征,并划分数据集。

2) 利用能量差值法确定 VMD 分解个数,采用 VMD 分解将原始风电功率序列分解为 R 个子序列,构造输入时序序列。

3) 分别将 VMD 分解得到的子序列输入到 Transformer-BiGRU 网络中,并利用 IZO 算法对四个超参数进行寻优,将寻优结果代入预测模型。

4) 将各个子序列预测结果进行叠加重构得到最终的功率预测值,并根据评价指标验证预测效果。

2.6 评价指标

为合理评价预测模型精度,选取 4 种误差评价指标^[19]进行分析,分别为均方根误差 (root mean square error, RMSE)、平均绝对误差 (mean absolute error, MAE)、平均绝对百分误差 (mean absolute percentage error, MAPE) 和决定系数 (R-square, R^2)。

3 算例仿真分析

3.1 变分模态分解结果分析

本文采用贵州某风电场实测数据,装机容量为 15 MW,该数据集包含风电功率和 6 个气象变量。数据收集于 2021 年 2 月 1 日~28 日,采样间隔为 15 min。将数据集进行划分,训练集和测试集的比例为 8:2。经相关性分析,选用风机侧出风速、高度 80 m 测风塔风速和风向作为气象特征输入。

为进一步挖掘风电功率的时序特征,利用 VMD 分解将原始数据分解成一系列具有不同频率和振幅的子模态

序列。在 VMD 分解前需要确定合适的分解子模态数 R 值。采用能量差值法确定 R 值,通过式(6)、式(7)计算出各模态数的能量。子模态能量差值如图 4 所示。当 $R=7$ 时能量差值最小,设置 VMD 分解层数为 7,故将原始功率序列分解为 7 个不同频率尺度且相对平稳子序列。

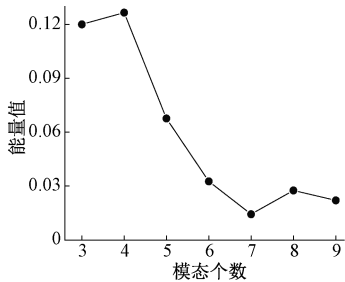


图 4 子模态能量差值
Fig. 4 Submode energy difference

采用 VMD 分解将序列分解为 7 个子模态分量,VMD 分解结果如图 5 所示。经分解后的 IMF 分量,相较于未经分解的功率曲线,其波动范围小、平稳性好,能更好体现 IMF 分量从高频到低频的时序特性,有利于深层次挖掘风电功率数据的变化规律。

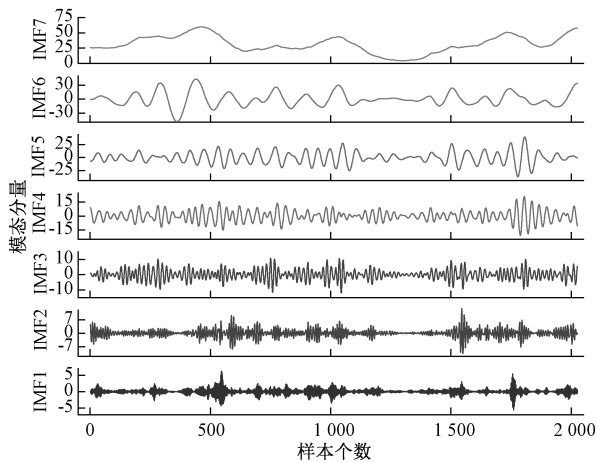


图 5 VMD 分解结果
Fig. 5 VMD decomposition results

3.2 IZOA-Transformer-BiGRU 模型超参数优化

Transformer-BiGRU 模型预测能力与隐藏层神经元数、初始学习率、正则化系数和 Transformer 中多头注意力头数 4 个超参数存在密切联系。隐藏层神经元数直接影响了模型的表达能力和运算效率;初始学习率决定了每次参数更新的步长,调整模型训练速度;正则化系数用于控制模型复杂度,防止对训练数据过度拟合,提高泛化能力;多头注意力能并行处理注意力分布,提高模型学习效率。经过调整 Transformer-BiGRU 模型超参数,可以有效提高高功率预测精度。

因此,利用 IZOA 算法对模型超参数进行寻优。采用 Adam 优化器,以均方根误差函数作为适应度函数,确定初

始种群数量为 15,迭代次数为 30,搜索空间维度为 4 维。设置 4 个超参数:隐藏层神经元数、初始学习率、正则化和多头注意力头数系数范围为:[10, 30]、[0.001, 0.01]、[0.000 1, 0.1]和[1, 8]。通过训练集不断寻优,直到达到最大迭代次数后,IZOA 算法寻优结束。当隐藏层神经元数为 12、初始学习率为 0.005 9、正则化系数为 0.001 53 和多头注意力头数为 4 时,模型预测效果最好。将返回超参数的最优组合输入到 Transformer-BiGRU 模型中进行功率预测。

3.3 预测结果对比分析

1)单一预测模型单步预测实验

为验证本文所提 VMD-IZOA-Transformer-BiGRU 组合模型对于风电功率预测性能,分别与传统单一模型(CNN、LSTM、BiGRU)及较为先进的单一预测模型(Transformer、Seq2Seq、WaveNet)进行对比实验。其中,Seq2Seq 模型采用双向 LSTM 作为 Encoder,带注意力机制的单向 LSTM 作为 Decoder;WaveNet 模型通过因果卷积和扩张卷积有效捕捉长时间依赖特征;Transformer 模型基于多层编码器和解码器组成,利用多头注意力机制生成预测结果。单一预测模型单步预测实验曲线对比如图 6 所示,其误差评价指标如表 1 所示。

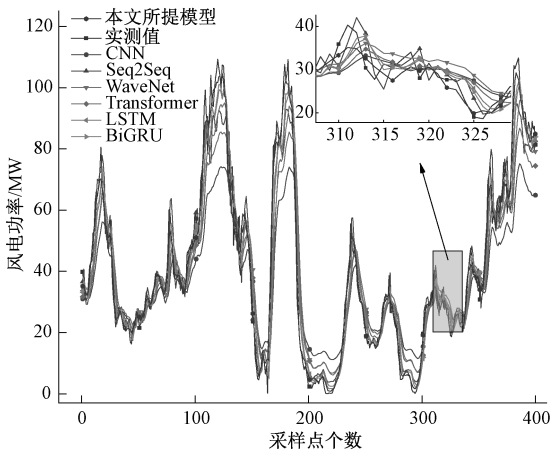


图 6 单一预测模型单步预测曲线对比
Fig. 6 Comparison of single-step prediction curves for a single prediction model

表 1 单一预测模型单步预测误差评价指标对比
Table 1 Comparison of indicators for evaluating single-step prediction errors of single prediction models

预测模型	RMSE	MAE	MAPE	R ²
CNN	12.042	10.235	20.449	0.757
Seq2Seq	11.246	9.487	19.746	0.783
WaveNet	10.967	8.949	17.098	0.813
Transformer	10.381	8.465	14.569	0.849
LSTM	8.985	6.758	11.916	0.916
BiGRU	7.137	5.112	10.330	0.941
本文所提模型	3.128	2.322	4.695	0.989

由图 6 可知,6 种单一模型对风电功率预测均表现出一定的效果。其中,BiGRU 的预测曲线与实测曲线偏差较小,表明其相较于 LSTM 通过更少的参数实现了更优的预测性能;Transformer 和 WaveNet 在复杂时序功率任务中展现了更强的曲线拟合能力,相比于 Seq2Seq 具有显著的优势。然而,单一模型在处理风电功率复杂波动时仍存在局限性,例如对突变点和拐点反应较为迟钝,容易出现相位偏移现象,导致预测结果滞后。相较之下,组合模型的预测曲线与实测曲线的贴合度更高,尤其在功率波动较大时,能够精确追踪变化趋势。这说明组合模型通过结合单一模型的优势,充分捕捉了功率波动极值等关键特征,从而实现更加准确的预测结果。

从表 1 分析可知,单一模型中采用双向 GRU 结构的 BiGRU 网络四种指标都获得了较好的效果。本文所提模型相较于 BiGRU 网络 RMSE、MAE、MAPE 分别降低了 56.17%、54.58%、54.55%, R^2 提升了 5.10%。这表明,本文模型在单一模型基础上通过组合数据分解技术和优化算法,预测精度有了显著提升,充分展现了其在风电功率复杂波动任务中的优势。

2)组合预测模型单步预测实验

将本文所提模型与 Transformer-BiGRU、VMD-BiGRU、VMD-Transformer-BiGRU、VMD-GWO-Transformer-BiGRU 和 EMD-IZOA-Transformer-BiGRU 五种组合预测模型进行对比。组合预测模型单步预测曲线对比如图 7 所示,其误差评价指标如表 2 所示。

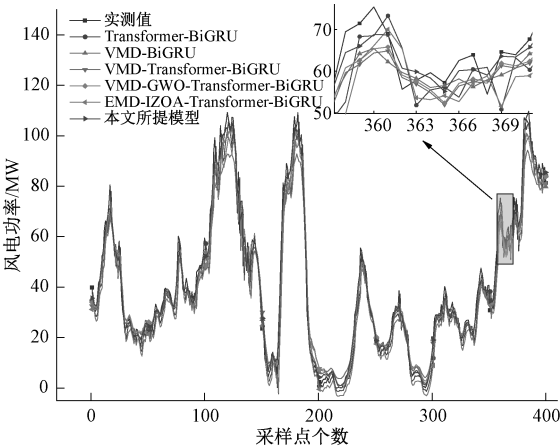


图 7 组合预测模型单步预测曲线对比

Fig. 7 Comparison of single-step prediction curves of combined prediction models

由图 7 可知,加入 VMD 分解能有效解决拟合曲线相位滞后的问题,并在融合注意力机制后预测曲线偏差减少。为获得更加适合预测模型的超参数组合,采用 IZOA 寻优后的预测曲线拟合效果更好。通过局部放大图可知,即使在实测曲线出现剧烈波动的情况下,所提组合模型在拐点处等相对预测难度较大的位置也能更好的进行追踪,

表 2 组合预测模型单步预测误差评价指标对比

Table 2 Comparison of single-step prediction error evaluation metrics for combined prediction models

预测模型	RMSE	MAE	MAPE	R^2
Transformer-BiGRU	6.425	4.644	8.865	0.957
VMD-BiGRU	5.263	3.928	7.897	0.968
VMD-Transformer-BiGRU	4.651	3.485	6.883	0.975
VMD-GWO-Transformer-BiGRU	4.347	3.274	6.606	0.978
EMD-IZOA-Transformer-BiGRU	6.083	4.355	8.763	0.961
本文所提模型	3.128	2.322	4.695	0.989

曲线拟合效果明显优于其他模型。

通过分析表 2 可知,与单一模型 BiGRU 相比,Transformer-BiGRU 网络 RMSE 降低了 9.98%,说明引入 Transformer 能增强模型捕捉复杂序列间的长期依赖关系;VMD-BiGRU 网络 RMSE 降低了 26.26%,说明分解算法对非平稳数据进行处理后能有效帮助模型提取时序特征,故组合模型结合各自优势可以一定程度提高模型预测效果。在组合模型中,采用 VMD 分解的本文所提模型相较于 EMD 分解的组合模型,RMSE、MAE、MAPE 分别降低了 48.58%、46.68%、46.42%, R^2 提升了 2.91%,表明 VMD 分解不仅可以解决预测曲线相位偏移的问题,并且预测精度比采用 EMD 分解组合模型更高。由于预测模型参数设定十分繁琐,需要人工反复尝试,因此采用改进智能算法对模型参数进行寻优。利用 IZOA 算法的本文所提模型相较于采用 GWO 算法进行参数优化的模型 RMSE、MAE、MAPE 分别降低了 28.04%、29.08%、28.93%, R^2 提升了 1.16%,说明 IZOA 算法能自适应搜索最优解,有效减少人为经验的干扰,提高模型预测精度。

3)组合预测模型多步预测实验

为进一步检验模型的预测性能,采用 CNN-LSTM、VMD-BiGRU、VMD-Transformer-BiGRU 和 VMD-IZOA-Transformer-BiGRU 四个模型进行多步预测实验,从而更全面地评估模型在处理长期依赖、误差传播和趋势捕捉方面的能力。组合预测模型多步预测误差评价指标对比如表 3 所示,3 种误差评价指标可视化如图 8 所示。

在多步预测实验中,随着预测步长的增加,预测精度和稳定性困难逐步增大。这需要预测模型提取更深层次的潜在特征,4 种模型都表现出不同程度的预测精度降低。在一步到三步预测中,CNN-LSTM 模型的 RMSE 分别为 7.137、8.918 和 10.382,预测误差不断增加。Transformer-BiGRU 模型的误差指标较 CNN-LSTM 更低,证明结合 Transformer 的关键特征捕捉能力和 BiGRU 物理处理序列时间依赖性的优势,能更加准确的捕捉序列的动态变化。与其他模型对比,在加入 VMD 分解后可以降低时间序列的复杂性,提

表 3 组合预测模型多步预测误差评价指标对比

Table 3 Comparison of multi-step prediction error evaluation metrics for combined prediction models

预测模型	RMSE			MAE			R ²		
	1 步	2 步	3 步	1 步	2 步	3 步	1 步	2 步	3 步
CNN-LSTM	6.894	8.432	9.573	4.621	5.473	6.412	0.952	0.937	0.918
Transformer-BiGRU	5.263	7.352	9.454	3.928	5.859	7.473	0.968	0.943	0.927
VMD-Transformer-BiGRU	4.651	6.453	8.239	3.485	4.634	5.796	0.975	0.954	0.945
本文所提模型	3.128	5.473	7.201	2.322	3.491	5.037	0.989	0.966	0.952

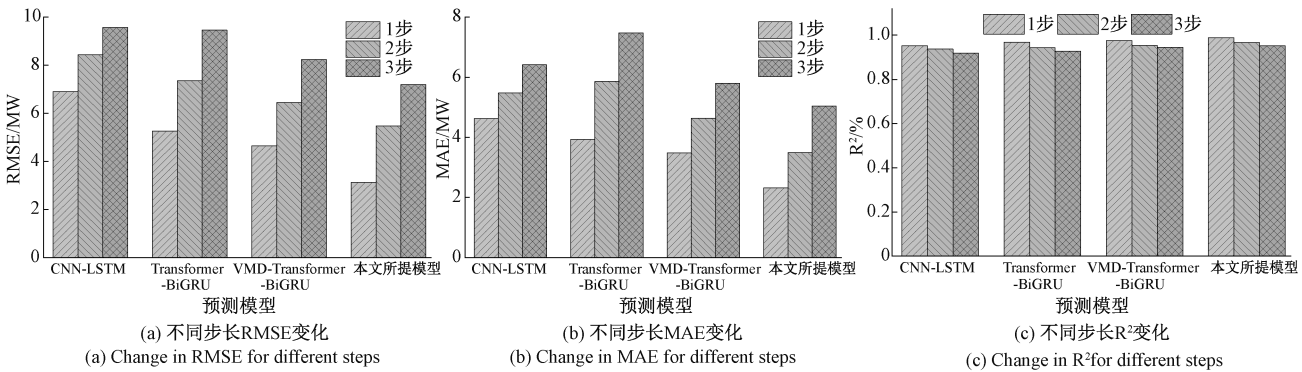


图 8 组合预测模型多步预测误差评价指标可视化

Fig. 8 Visualization of multi-step prediction error evaluation metrics for combinatorial prediction models

高模型对主要趋势的学习能力。利用 IZOA 算法自动寻找最优的参数组合,提高参数选择的客观性和准确性。因此,加入数据分解和参数寻优的组合模型在多步预测中产生可靠的预测结果。

实验表明,本文提出的基于数据分解的 IZOA-Transformer-BiGRU 短期风电功率组合预测模型能够降低风电功率预测过程中的波动性、提高时序序列之间特征提取能力,在一定时间尺度内具有较好的预测效果和稳定性。

4) VMD-IZOA-Transformer-BiGRU 预测模型泛化性实验

为验证本文所提模型的泛化性,采用新疆某风电场的数据进行仿真实验。该风电场的装机容量为 250 MW,选取了 2022 年 2 月的历史风电功率和气象数据,按照原始数据划分方式进行实验。在数据处理中,通过能量差值法确定 VMD 的子模态数量 $R=8$,将原始风电功率数据分解为 8 个相对平稳子序列。在对比实验中,分别采用 WaveNet、BiGRU、CNN-LSTM、Transformer-BiGRU、EMD-Transformer-BiGRU 和 VMD-Transformer-BiGRU 六中模型进行对比分析。泛化性实验预测曲线对比如图 9 所示,其误差评价指标对比如表 4 所示。

从图 9 的预测曲线可以看出,本文提出的 IZOA-Transformer-BiGRU 模型能够更好的追踪实际功率的变化趋势,预测值与实际值之间的偏差最小,曲线拟合程度显著提升。表 4 的误差评价指标对比进一步验证了该模

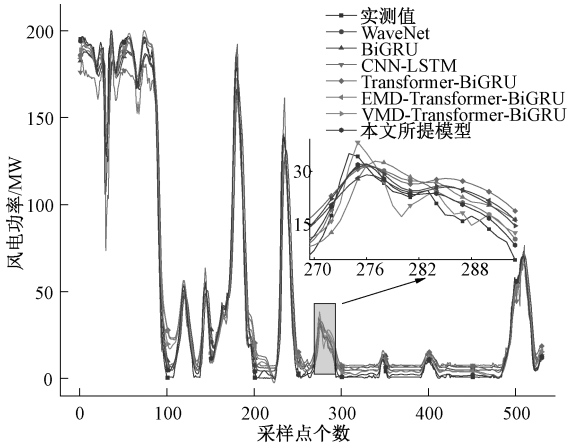


图 9 泛化性实验预测曲线对比

Fig. 9 Comparison of experimental prediction curves for generalizability

表 4 泛化性实验误差评价指标对比

Table 4 Comparison of experimental error evaluation metrics for generalizability

预测模型	RMSE	MAE	MAPE	R ²
WaveNet	15.313	9.839	9.013	0.948
BiGRU	13.792	9.321	8.698	0.960
CNN-LSTM	13.242	8.721	7.346	0.964
Transformer-BiGRU	10.508	7.945	4.963	0.972
EMD-Transformer-BiGRU	9.464	6.534	5.436	0.980
VMD-Transformer-BiGRU	7.474	5.196	4.015	0.988
本文所提模型	4.008	2.359	2.399	0.997

型的优越性。与BiGRU模型相比,本文所提模型RMSE、MAE、MAPE分别降低了70.94%、70.49%、72.42%, R^2 提升了3.83%。结果表明,本文提出的结合变分模态分解技术和改进IZOA算法的模型,能够准确地捕捉复杂的风电功率波动特征,在不同地理位置和数据分布下均表现良好,具有较强的泛化性和适用性。

4 结 论

本文考虑风电功率随机性和波动性的问题,提出了变分模态分解和改进斑马优化算法的Transformer-BiGRU组合模型,用于短期风电功率预测。首先,采用基于能量差法的VMD分解技术,充分提取功率序列的时序特征,降低数据波动,解决传统单一模型预测曲线相位滞后的问题。然后,通过利用Transformer的多头注意力机制并行考虑多个输入特征之间的交互关系,并结合BiGRU网络捕捉时序序列的前后依赖关系,结合两者优势,构建了Transformer-BiGRU组合模型,从而提升了模型预测性能。进一步,针对预测模型参数设置繁杂导致预测精度有限的问题,采用改进的IZOA算法对模型四个超参数进行寻优,降低模型预测误差。仿真对比实验结果表明,提出的VMD-IZOA-Transformer-BiGRU模型在预测精度上较传统单一模型有了显著提升,验证了组合模型方法的可行性。通过与其他模型进行单步、多步预测以及泛化性实验,进一步验证了该模型具有较高的准确性和泛化性。本文所提模型在风电功率预测任务中虽具有较高的准确性和泛化能力,但模型未考虑不同季节、天气等因素,后续工作可开展针对不同特定场景的研究。

参考文献

- [1] NJINDAN B I. Climate change, energy security risk, and clean energy investment[J]. Energy Economics, 2024, 129: 107225.
- [2] 黎静华, 骆怡辰, 杨舒惠, 等. 可再生能源电力不确定性预测方法综述[J]. 高电压技术, 2021, 47(4): 1144-1157.
LI J H, LUO Y CH, YANG SH H, et al. Review of uncertainty forecasting methods for renewable energy power[J]. High Voltage Engineering, 2021, 47(4): 1144-1157.
- [3] CHRIFI A L, DRID S, OURIAGLI M, et al. Overview of photovoltaic and wind electrical power hybrid systems[J]. Energies, 2023, 16(12): 4778.
- [4] CHEN Y X, BAI M L, ZHANG Y L, et al. Multivariable space-time correction for wind speed in numerical weather prediction(NWP) based on ConvLSTM and the prediction of probability interval[J]. Earth Science Informatics, 2023, 16(3): 1953-1974.
- [5] ZHENG Y, GE Y S, MUHSEN S, et al. New ridge regression, artificial neural networks and support vector machine for wind speed prediction[J]. Advances in Engineering Software, 2023, 179: 103426.
- [6] 孟安波, 陈顺, 王陈恩, 等. 基于混沌CSO优化时序注意力GRU模型的超短期风电功率预测[J]. 电网技术, 2021, 45(12): 4692-4700.
MENG AN B, CHEN SH, WANG CH EN, et al. Ultra-short-term wind power prediction based on chaotic CSO optimized temporal attention GRU model[J]. Power system Technology, 2021, 45(12): 4692-4700.
- [7] LIU L, LIU J C, YE Y, et al. Ultra-short-term wind power forecasting based on deep Bayesian model with uncertainty[J]. Renewable Energy, 2023, 205: 598-607.
- [8] 吴晓刚, 阎洁, 葛畅, 等. 基于改进GRU-CNN的风光水一体化超短期功率预测方法[J]. 中国电力, 2023, 56(9): 178-186, 205.
WU X G, YAN J, GE CH, et al. Ultra-short-term power forecasting method for wind-solar-hydro integration based on improved GRU-CNN[J]. Electric Power, 2023, 56(9): 178-186, 205.
- [9] 徐钊, 谢开贵, 王宇, 等. 基于TCN-Wpsformer混合模型的超短期风电功率预测[J]. 电力自动化设备, 2024, 44(8): 54-61.
XU T, XIE K G, WANG Y, et al. Ultra-large-capacity offshore wind power converter based on single-stage multilevel switching conversion[J]. Electric Power Automation Equipment, 2024, 44(8): 54-61.
- [10] 孟鑫禹, 王睿涵, 张喜平, 等. 基于经验模态分解与多分支神经网络的超短期风功率预测[J]. 计算机应用, 2021, 41(1): 237-242.
MENG X Y, WANG R H, ZHANG X P, et al. Ultra-short-term wind power prediction based on empirical mode decomposition and multi-branch neural network[J]. Journal of Computer Applications, 2021, 41(1): 237-242.
- [11] 史彭珍, 魏霞, 张春梅, 等. 基于VMD-BOA-LSSVM-AdaBoost的短期风电功率预测[J]. 太阳能学报, 2024, 45(1): 226-233.
SH P ZH, WEI X, ZHANG CH M, et al. Short-term wind power prediction based on VMD-BOA-LSSVM-AdaBoost[J]. Acta Energiæ Solaris Sinica, 2024, 45(1): 226-233.
- [12] 程杰, 陈鼎, 李春, 等. 基于GWO-CNN-BiLSTM的超短期风电预测[J]. 科学技术与工程, 2023, 23(35): 15091-15099.
CHENG J, CHEN D, LI CH, et al. Ultra-short-term

- wind power prediction based on GWO-CNN-BiLSTM[J]. Science Technology and Engineering, 2023, 23(35): 15091-15099.
- [13] WANG J N, ZHU H Q, ZHANG Y J, et al. A novel prediction model for wind power based on improved long short-term memory neural network[J]. Energy, 2023, 265: 126283.
- [14] 赵杰, 解则晓, 刘世萱. 基于能量熵 VMD 最优分解与 GRU 循环神经网络的潮汐预测精度提升方法研究[J]. 仪器仪表学报, 2023, 44(12): 79-87.
- ZHAO J, XIE Z X, LIU SH X. Tide prediction accuracy improvement method research based on VMD optimal decomposition of energy entropy and GRU recurrent neural network [J]. Chinese Journal of Scientific Instrument, 2023, 44(12): 79-87.
- [15] TROJOVSKA E, DEHGHANI M, TROJOVSKY P, et al. Zebra optimization algorithm: A new bio-inspired optimization algorithm for solving optimization algorithm[J]. IEEE Access, 2022, 10: 49445-49473.
- [16] 张开伟, 文中, 杨生鹏, 等. 基于改进 Cao 算法的 SSA 与误差修正的超短期风电功率预测[J]. 国外电子测量技术, 2024, 43(8): 37-46.
- ZHANG K W, WEN ZH, YANG SH P, et al. Ultra-short-term wind power forecasting through improved Cao algorithm for SSA and error correction [J]. Foreign Electronic Measurement Technology, 2024, 43(8): 37-46.
- [17] 白隆, 俞斌, 高峰, 等. 基于 ICEEMDAN 和 TCN-AM-BiGRU 的短期光伏功率预测[J]. 电子测量技术, 2024, 47(9): 61-69.
- BAI L, YU B, GAO F, et al. Short-term photovoltaic power prediction based on ICEEMDAN and TCN-AM-BiGRU [J]. Electronic Measurement Technology, 2024, 47(9): 61-69.
- [18] 朱相华, 智敏, 殷雁君. 基于 2D CNN 和 Transformer 的人体动作识别[J]. 电子测量技术, 2022, 45(15): 123-129.
- ZHU X H, ZHI M, YIN Y J. Human action recognition based on 2D CNN and Transformer [J]. Electronic Measurement Technology, 2022, 45(15): 123-129.
- [19] 章志晃, 徐启峰, 林穿. 基于 SSA-BiLSTM-AM 的短期风电功率预测[J]. 国外电子测量技术, 2023, 42(3): 46-51.
- ZHANG ZH H, XU Q F, LIN CH. Short-term wind power prediction based on SSA-BiLSTM-AM [J]. Foreign Electronic Measurement Technology, 2023, 42(3): 46-51.

作者简介

蒲晓云, 硕士研究生, 主要研究方向为功率预测和智能优化算法等。

E-mail: 3503295417@qq.com

杨靖, 博士, 教授, 主要研究方向为群体智能优化和物联网技术与应用等。

E-mail: jyang7@gzu.edu.cn

杨兴, 博士, 副教授, 主要研究方向为宽带信号合成技术和集成电路设计。

E-mail: xyang@gzu.edu.cn

宁媛(通信作者), 硕士, 教授, 主要研究方向为智能信息处理技术研究等。

E-mail: ee_yining@gzu.edu.cn