



DOI:10.12404/j.issn.1671-1815.2404251

引用格式:刘俊宏,富斯源,王亚君.基于优化 TCN 组合模型的短期光伏功率预测[J].科学技术与工程,2025,25(15):6378-6388.

Liu Junhong, Fu Siyuan, Wang Yajun. Short-term PV power prediction model based on optimized TCN combination model [J]. Science Technology and Engineering, 2025, 25(15): 6378-6388.

基于优化 TCN 组合模型的短期光伏功率预测

刘俊宏,富斯源,王亚君*

(辽宁工业大学电子与信息工程学院,锦州 121000)

摘要 为提高多输入特征下光伏发电功率模型的短期预测精度,提出了一种基于优化时域卷积网络超参数的光伏功率预测组合模型(LGGWO-TCN-MHSA)。该模型集改进灰狼优化算法(levy gold gray wolf optimization, LGGWO)、时域卷积网络(temporal convolutional network, TCN)和多头自注意力机制(multi-head self-attention, MHSA)于一体。首先,采用斯皮尔曼相关系数法提取对光伏功率影响较大的主要特征,并输入至 TCN 预测模型;然后,将提出的多策略改进灰狼优化算法 LGGWO 应用于 TCN 内部进行超参数优化,改善模型预测性能;最后,将预测值输入至多头自注意力模型中进一步提升预测精度。实验采用澳大利亚原始光伏数据进行验证,通过与卷积神经网络(convolutional neural networks, CNN)、长短期记忆神经网络(long short-term memory, LSTM)等六组模型进行对比,所提模型在测试数据集上的平均绝对误差(mean absolute error, MAE)和均方根误差(root mean square error, RMSE)分别降低了 2.03%~82.0% 和 10.5%~80.1%,结果表明:所提方法具有较高的预测精度和良好的稳定性。

关键词 光伏发电;光伏功率短期预测;改进灰狼优化算法;时域卷积网络;多头自注意力机制

中图分类号 TM615 TP615; **文献标志码** A

Short-term PV Power Prediction Model Based on Optimized TCN Combination Model

LIU Jun-hong, FU Si-yuan, WANG Ya-jun*

(School of Electronic and Information Engineering, Liaoning University of Technology, Jinzhou 121000, China)

[Abstract] To improve the short-term prediction accuracy of photovoltaic power generation models with multiple input features, a photovoltaic power prediction ensemble model LGGWO-TCN-MHSA based on optimizing TCN hyperparameters was proposed. The model integrated the levy gold grey wolf optimization (LGGWO), temporal convolutional network (TCN), and multi-head self-attention mechanism (MHSA). First, the Spearman correlation coefficient method extracted the main features that significantly affect photovoltaic power, which were then fed into the TCN prediction model. Then, the proposed multi-strategy LGGWO was applied to the TCN for hyperparameter optimization, which improved the model's prediction performance. Finally, the predicted values were input into the multi-head self-attention model to further boost prediction accuracy. The experiment was verified using original Australian photovoltaic data. By comparing with six groups of models including convolutional neural networks (CNN) and long short-term memory neural networks (LSTM), the mean absolute error (MAE) and root mean square error (RMSE) of the proposed model on the test data set were reduced by 2.03%~82.0% and 10.5%~80.1%, respectively. The results show that the proposed method has high prediction accuracy and good stability.

[Keywords] photovoltaic power; PV power short-term forecast; improved grey wolf optimization; temporal convolutional network; multi-head self-attention

2022 年,太阳能光伏发电量增加了创纪录的 270 TWh (增长率高达 26%),总量达到近 1 300 TWh。

截至 2023 年 12 月底,全国太阳能发电装机容量约 6.1 亿 kW,同比增长 55.2%。这一增长率符

合 2030 年前实现碳达峰,2060 年前实现碳中和的目标^[1]。随着光伏发电在电力系统中接入比例的增加,其固有的波动性和随机性给光伏发电的安全稳定运行带来了相应的技术挑战^[2]。准确地预测光伏发电功率,进而确定合理的运行方式和短期调

收稿日期:2024-06-07 修订日期:2024-10-29

基金项目:辽宁省教育厅重点攻关项目(JYTZD2023083)

第一作者:刘俊宏(1999—),男,汉族,吉林德惠人,硕士研究生。研究方向:基于深度学习的光伏功率预测。E-mail:1319374997@qq.com。

*通信作者:王亚君(1978—),女,满族,辽宁葫芦岛人,博士,教授。研究方向:电力电子技术及应用。E-mail:wj_lg@163.com。

度计划,对以新能源为主体的新型电力系统安全稳定运行具有重要的现实意义。

目前传统的光伏功率预测的研究方法主要分为物理方法和统计方法两类。物理方法需要建立复杂的物理模型,模型参数确定更为烦琐耗时,很难处理复杂情况,预测精度有限。统计方法通常采用传统方法处理光伏发电历史数据,如卡尔曼滤波、贝叶斯回归等,针对非线性数据的预测效果较差。随着计算机领域的不断发展,机器学习逐渐崭露头角,主要包括支持向量机、随机森林以及极限学习机等^[3-4],在此基础上也衍生出了以反向传播神经网络(back propagation, BP)、卷积神经网络(convolutional neural networks, CNN)、长短期记忆神经网络(long short-term memory, LSTM)为代表的神经网络深度学习方

文献[5]提出将 CNN 与 LSTM 模型相组合来提高预测模型精准度;文献[6]利用注意力机制与 LSTM 相结合来进行光伏功率预测,取得较高的模型精度与稳定性。文献[7]利用粒子群优化算法(particle swarm optimization, PSO)对 CNN 进行参数寻优,结果显示预测精度有所提高。文献[8]利用改进麻雀优化算法(improved sparrow search algorithm, ISSA)对最小二乘支持向量机(least squares support vector machine, LSSVM)参数进行寻优,获得较高的精度。文献[9]利用鲸鱼优化算法(whale optimization algorithm, WOA)对 LSTM 超参数变化寻优,效果较好。文献[10]提出一种多策略改进的灰狼优化算法,利用伪反向学习和布谷鸟搜索机制改善算法的全局寻优能力,结果显示改良后算法拥有更快收敛性,寻优结果更精确。

针对多维度非线性的光伏功率预测问题,提出一种基于 LGGWO-TCN-MTSA(levy gold gray wolf optimization, temporal convolutional network, and multi-head self-attention)的光伏功率预测模型。首先,为改善组合模型预测性能,提升模型训练速度,提出一种多策略改进灰狼优化算法 LGGWO,并对 TCN 内部的超参数进行寻优;然后,为进一步提升模型预测精度,降低预测误差,将预测值输入至多头自注意力模型中加权处理;最后,以真实数据进行预测仿真实验,并与当前相关的预测方法进行对比,验证本文所提模型的有效性。

1 组合模型原理

1.1 变量相关性分析

光伏发电受多种数据因素影响,其中最为重要的便是以辐照度为主的天气因素。为了分析不同

天气变量与发电功率之间的关系,引入 Spearman 相关性系数对涉及的原始数据中的 9 种特征量进行相关性分析,根据结果保留相关性高特征因素,相关性系数计算式为

$$S = 1 - \frac{6 \sum_{i=1}^n [R(x_i) - R(y_i)]^2}{n(n^2 - 1)} \quad (1)$$

式(1)中: S 为相关性系数; $R(x_i)$ 和 $R(y_i)$ 分别为特征 x_i 和 y_i 在各自数列中的排序。

相关性分析结果如表 1 所示,可以看出,光伏发电功率与全球水平辐射和贴片平均温度的相关性系数最大,其次为扩散水面辐射;风速和相对湿度处于中等相关性;风向和降雨量呈弱相关。综上,选取贴片平均温度、全球水平辐射、扩散水平辐射和历史发电功率作为输入特征量。

表 1 天气特征与 Spearman 相关性系数
Table 1 Weather feature and Spearman coefficient

天气特征	与光伏功率的相关性系数
风速	0.265
贴片平均温度	0.656
相对湿度	-0.334
风向	-0.084
降雨量	0.063
全球水平辐射	0.916
扩散水平辐射	0.542

1.2 时域卷积网络

时域卷积网络(TCN)由卷积神经网络(CNN)发展而来,借鉴了其在图像处理领域的思想,并应用与时间序列数据的处理之上。相比于循环神经网络(recurrent neural network, RNN)和 LSTM,TCN 利用膨胀因果卷积和残差链接两个特点,在降低计算复杂度的同时解决了梯度消失与梯度爆炸等问题,提高了时间序列数据训练的效率与模型性能。

膨胀因果卷积如图 1 所示,其中 d 为膨胀因子,其规定了数据点间的间距大小。与传统的卷积核相比,膨胀因果卷积通过引入膨胀系数,在相同的卷积层下,拥有更大的感受野,从而能够捕获更多的时间序列信息。因此 TCN 只需相对较少的层数,就能实现比 CNN 更广泛的感知范围,从而避免了信息的重复提取。

TCN 残差模块结构如图 2 所示,多个残差模块相连,通过跨层的方式传递信息,有效避免了传统神经网络中梯度消失和梯度爆炸问题,提高了模型的反馈与收敛速度。每个残差块包含两层卷积单元,数据首先由输入层进入残差块中,并通过一维膨胀因果卷积来提取时间序列数据中的序列信息。卷积计算后,数据经过权重归一化和激活函数处

理,最后应用 Special Dropout 正则化,减少模型计算量,防止模型过拟合,从而加快训练速度。

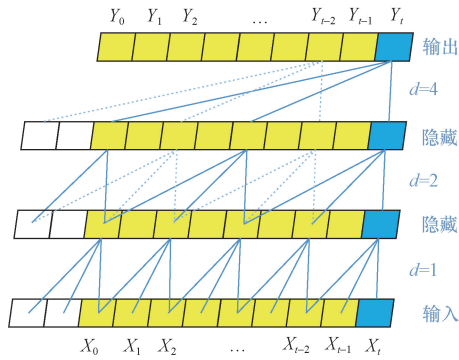


图1 膨胀因果卷积
Fig. 1 Dilated convolutions

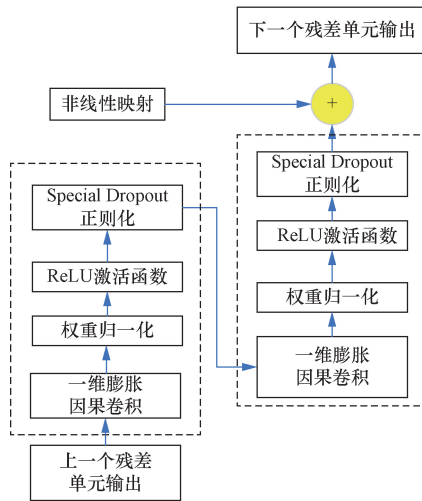


图2 残差模块结构
Fig. 2 Residual module

1.3 多头自注意力机制

自注意力机制是一种用于增强神经网络对输入数据中相关部分关注度的机制。其核心思想是通过加权求和的方式,使每个输出值都能关联于输入序列中的输入值,其计算公式为

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V} \quad (2)$$

式(2)中: $\frac{1}{\sqrt{d_k}}$ 为缩放因子,防止点积过大导致梯度消失。

首先对输入序列 x_1, x_2, \dots, x_T 进行线性变换,生成查询矩阵 \mathbf{Q} 、键矩阵 \mathbf{K} 和值矩阵 \mathbf{V} ;然后计算注意力得分矩阵 $a_{1,1}, a_{1,2}, \dots, a_{1,T}$ 并进行缩放,以防止得分过大;将缩放后的得分矩阵通过 softmax 函数转换为概率分布,以此表示各位置的注意力权重 $\hat{a}_{1,1}, \hat{a}_{1,2}, \dots, \hat{a}_{1,T}$,最后对权重进行加权求和,得到最终输出序列 b_1, b_2, \dots, b_T 。自注意力机制如图3所示。

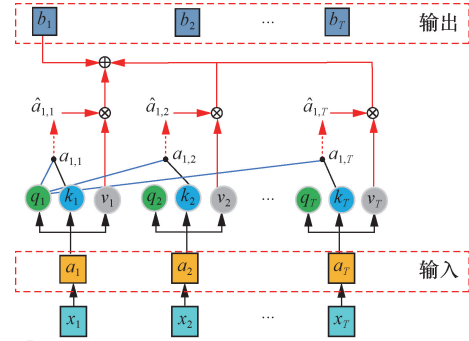


图3 单头自注意力机制原理
Fig. 3 Principle of self-attention mechanism

多头自注意力机制通过并行地应用多个独立的自注意力机制(即多个“头”),模型能够从不同的角度捕捉输入序列中的不同信息。这些不同的注意力头可以关注输入数据中的不同部分,从而使模型能够更好地理解和表示复杂的模式和关系。

1.4 多策略改进灰狼优化算法(LGGWO)

灰狼优化算法(gray wolf optimization, GWO)灵感来自于灰狼的捕猎行为和社会等级结构。该算法主要流程分为3个阶段:包围猎物、追踪猎物和攻击猎物。则灰狼捕食行为的数学模型可描述为

$$\mathbf{D} = |\mathbf{C}\mathbf{X}_p(t) - \mathbf{X}(t)| \quad (3)$$

$$\mathbf{X}(t+1) = \mathbf{X}_p(t) - \mathbf{A}\mathbf{D} \quad (4)$$

式中: t 为当前的迭代次数; \mathbf{X}_p 为猎物的位置向量; $\mathbf{X}(t)$ 为迭代 t 次后的灰狼位置向量; \mathbf{A} 、 \mathbf{C} 为向量系数,表达式为

$$\mathbf{A} = 2a\mathbf{r}_1 - a\mathbf{N} \quad (5)$$

$$\mathbf{C} = 2\mathbf{r}_2 \quad (6)$$

$$a = 2\left(1 - \frac{t}{T}\right) \quad (7)$$

式中: \mathbf{r}_1 和 \mathbf{r}_2 均为 $[0, 1]$ 的随机向量; \mathbf{N} 为元素全为 1 的向量; a 为收敛因子,随着迭代次数线性从 2 线性递减到 0; T 为种群最大迭代次数。

在狩猎时,由于 α, β, δ 狼对于猎物的潜在位置更具有感知性,所以灰狼会根据 α 狼、 β 狼和 δ 狼的位置来移动并更新位置,表达式为

$$\begin{cases} \mathbf{D}_\alpha = |\mathbf{C}_1\mathbf{X}_\alpha(t) - \mathbf{X}(t)| \\ \mathbf{D}_\beta = |\mathbf{C}_2\mathbf{X}_\beta(t) - \mathbf{X}(t)| \\ \mathbf{D}_\delta = |\mathbf{C}_3\mathbf{X}_\delta(t) - \mathbf{X}(t)| \end{cases} \quad (8)$$

$$\mathbf{X}(t+1) = (\mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3)/3 \quad (9)$$

尽管 GWO 在许多优化问题中表现出色,但依旧存在许多问题,如易陷入局部最优解,对性能参数十分敏感,以及收敛速度较慢。针对上述问题,对 GWO 进行改进,使用混沌映射初始化灰狼种群,提高搜索过程中的种群多样性。引入非线性收敛

因子、莱维飞行和黄金正弦策略,更好地平衡优化算法全局搜索和局部寻优能力。提出一种新的位置更新策略,配合步长欧式距离权重,有效解决算法寻优速度和寻优精度等问题。本文中在GWO上的改进主要如下。

(1)circle 映射。相比于随机分布的模式,作为混沌映射的典型代表,circle 映射可以使种群分布均匀,使寻找到的最优值概率增大,circle 映射数学表达式为

$$X_{t+1} = \text{mod}\left\{X_t + b - \left[\frac{a}{2\pi}\sin(2\pi X_t), 1\right]\right\} \quad (10)$$

式(10)中:mod()为取余函数; $a=0.3, b=0.3$ 。

(2)非线性收敛因子。传统灰狼算法中,A的取值决定了狼群全局搜索与局部寻优进程选择。由式(5)可知A的取值取决于收敛因子a的变化,因此,收敛因子a的取值影响着算法的全局搜索和局部寻优能力。现提出一种新的非线性收敛因子a,来平衡灰狼优化算法的全局搜索和局部搜索能力,式(7)更新为

$$a = \frac{w}{\sqrt{1 + \left(\frac{4t}{T}\right)^9}} \quad (11)$$

式(11)中:w为调节因子,本文中选取 $w=2$,使得收敛因子a从接近2的值非线性下降至接近0的值,非线性收敛因子和原收敛因子对比如图4所示。

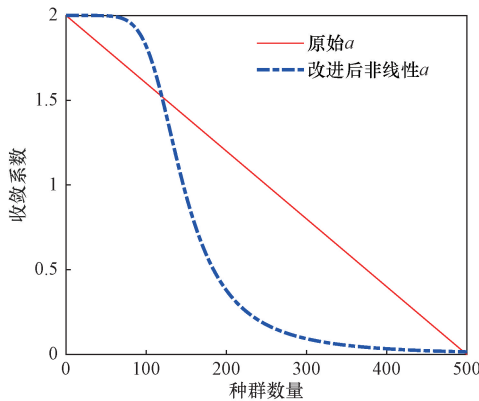


图4 非线性收敛因子对比

Fig.4 Comparison of nonlinear convergence factors

由图4可以看出原始收敛因子a在迭代过程中以固定速度降低,而本文中修改过后的非线性收敛因子,在算法迭代前期,减少速度相对较慢,使得改进后因子a相比原始值,能够保证较长时间的最大值,以提高全局搜索效率;在算法迭代后期,收敛因子减少速度加快,使得收敛因子a可以长时间维持较小值以提高算法的搜索精度。

(3)Levy 飞行与黄金正弦混合策略。为了提高算法搜索精度,受文献[11]启发,引入莱维飞行和

黄金正弦策略对算法寻优过程进行优化。 $|A| > 1$ 时采用莱维飞行优化算法全局搜索能力; $|A| < 1$ 时采用黄金正弦策略则提高算法后期局部寻优能力。

在数学上,莱维飞行模型是一种广义连续时间随机漫步过程,其特点是步长分布具有重尾特性。该特性使得搜索过程能够在短距离局部搜索和长距离全局探索之间取得平衡。将莱维飞行引入灰狼优化算法,可以有效降低灰狼优化算法陷入局部最优的风险,增强算法的全局搜索能力和优化效率,更新后的位置式为

$$X'_i(t) = X_i(t) - \theta \oplus \text{levy}(\lambda) \quad (12)$$

$$\theta \oplus \text{levy}(\lambda) \sim 0.01 [X_i(t) - X_a(t)] \frac{\mu}{|\nu|^{\frac{1}{\gamma}}} \quad (13)$$

式中: $X_i(t)$ 为第t代的第i个解; \square 表示点对点乘法; θ 表示控制步长的权重; $X_a(t)$ 为当前的最优解;levy(λ)表示莱维飞行搜索路径,其中 λ 满足 $1 < \lambda < 3$; γ 取值一般为1.5。模拟步长公式涉及的 μ 和 ν 服从正态分布,定义为

$$\mu \sim N(0, \sigma_\mu^2), \nu \sim N(0, \sigma_\nu^2) \quad (14)$$

$$\sigma_\mu = \left[\frac{\Gamma(1 + \gamma) \sin\left(\frac{\pi\gamma}{2}\right)}{\gamma \Gamma\left(\frac{1 + \gamma}{2}\right) \times 2^{\frac{\gamma-1}{2}}}\right]^{\frac{1}{\gamma}}, \sigma_\nu = 1 \quad (15)$$

黄金正弦策略是一种基于黄金分割和正弦函数的优化策略,该策略旨在通过引入黄金分割比例和正弦波形的变化,增强算法在搜索空间中的探索和利用能力,从而改善优化算法的局部搜索能力。黄金正弦的位置更新公式为

$$X'_i(t) = X_i(t) | \sin r_1 | + r_2 \sin r_1 | x_1 X_a(t) - x_2 X_i(t) | \quad (16)$$

式(16)中: r_1 为 $[0, 2\pi]$ 范围内的随机数; r_2 为 $[0, \pi]$ 范围内的随机数; $X_a(t)$ 为当前的最优解; x_1 和 x_2 为黄金分割系数,表达式为

$$\begin{cases} x_1 = a\tau + b(1 - \tau) \\ x_2 = a(1 - \tau) + b\tau \\ \tau = (\sqrt{5} - 1)/2 \end{cases} \quad (17)$$

式(17)中:a和b的默认值为 π 和 $-\pi$ 。

尽管上述两种策略可以带来位置的更新,但无法保证新解和原解的优劣,因此需要通过引入贪心机制来比较,从而保留适应度最好的解,贪心机制表达式为

$$X_i(t) = \begin{cases} X_i(t), & \text{fit}[X_i(t)] < \text{fit}[X'_i(t)] \\ X'_i(t), & \text{fit}[X'_i(t)] < \text{fit}[X_i(t)] \end{cases} \quad (18)$$

(4) 基于步长欧式距离的动态权重位置更新策略。由式(8)可知,原始GWO算法中,每个个体对 α, β, δ 狼具有相同的学习权重,这会妨碍新一代个体学习 α 狼的优越性从而使算法整体收敛速度变慢。文献[12]论证了步长欧式比例权重相比于适应度的优势,因此引入步长欧式距离的比例权重策略后,位置更新公式可表示为

$$\begin{cases} \sum \mathbf{X} = |\mathbf{X}_1| + |\mathbf{X}_2| + |\mathbf{X}_3| \\ W_1 = \frac{|\mathbf{X}_1|}{\sum |\mathbf{X}|}, W_2 = \frac{|\mathbf{X}_2|}{\sum |\mathbf{X}|}, W_3 = \frac{|\mathbf{X}_3|}{\sum |\mathbf{X}|}, \\ \sum |\mathbf{X}| \neq 0 \\ W_1 = W_2 = W_3 = \frac{1}{3}, \sum |\mathbf{X}| = 0 \end{cases} \quad (19)$$

$$\mathbf{X}(t+1) = \frac{W_1 \mathbf{X}_1 + W_2 \mathbf{X}_2 + W_3 \mathbf{X}_3}{3} \quad (20)$$

同时,受到文献[13]的启发,为了防止 α 狼、 β 狼和 δ 狼三者全部过早地陷入局部最优,导致算法停滞,因此引入动态权重因子 φ ,并再次加强 α 狼的优越性,因此最终的位置更新公式可表示为

$$\varphi = 1 - \frac{1}{2} \left(\frac{t}{T} \right)^2 \quad (21)$$

$$\mathbf{X}(t+1) = \frac{W_1 \mathbf{X}_1 + W_2 \mathbf{X}_2 + W_3 \mathbf{X}_3}{3} \varphi + \frac{\mathbf{X}_1}{2} \left(\frac{t}{T} \right)^2 \quad (22)$$

2 光伏功率短期预测模型

2.1 LGGWO 寻优模型验证

LGGWO 模型的参数寻优流程如下。

(1) 灰狼优化算法参数初始化,基于 circle 混沌映射初始化狼群。

(2) 计算种群适应度,选取适应度最优、次优和第三优个体,并记录位置。

(3) 迭代过程中根据 $|\mathbf{A}|$ 的值选择对最优狼进行混合策略调整位置。

(4) 根据位置更新公式对狼群个体进行位置更新。

(5) 循环直至迭代次数达到标准位置,输出最优解。

为了验证 LGGWO 的寻优性能,选用 CEC2005 中的 6 种基准测试函数对其进行仿真测试,函数覆盖单峰多峰等不同类型,函数设置如表 2 所示。LGGWO、GWO、WOA、DBO、NGO 主要参数设置保持一致:种群规模 $N=30$,最大迭代次数为 500。

从图 5 可以看出,LGGWO 在各类测试函数中的求解精度方面表现明显优于其他 3 种算法,收敛速度方面仅在 F6 函数中略慢;从表 3 可以看出,LGGWO 稳定性相较于其他算法效果更好,对于多峰函数能够较好地避免陷入局部最优,从而收敛于最优解。综合分析可得,相较于其余几种元启发式算法,LGGWO 拥有更为优秀的寻优性能。

表 2 标准测试函数

Table 2 Standard test function

函数名称	函数表达式	搜索空间	维数	理论最优值	峰值
Sphere	$f_1(x) = \sum_{i=1}^n x_i^2$	$[-100, 100]$	30	0	单峰
Schwefel's 2.22	$f_2(x) = \sum_{i=1}^n \left(\sum_{j=1}^i x_j \right)^2$	$[-100, 100]$	30	0	单峰
Quartic	$f_3(x) = \sum_{i=1}^n ix_i^4 + \text{dandom}[0, 1)$	$[-1.28, 1.28]$	30	0	单峰
Rastrigin	$f_4(x) = \sum_{i=1}^n ix_i^4 + \text{dandom}[0, 1)$	$[-5.12, 5.12]$	30	0	多峰
Ackley	$f_5(x) = -20 \exp \left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \right) - \exp \left[\frac{1}{n} \sum_{i=1}^n \cos(2\pi x_i) \right] + 20 + e$	$[32, 32]$	30	0	多峰
Penalized	$f_6(x) = 0.1 \{ \sin^2(3\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2 [1 + \sin^2(2\pi x_n)] \} + \sum_{i=1}^n u(x_i, 5, 100, 4)$	$[-50, 50]$	30	0	多峰

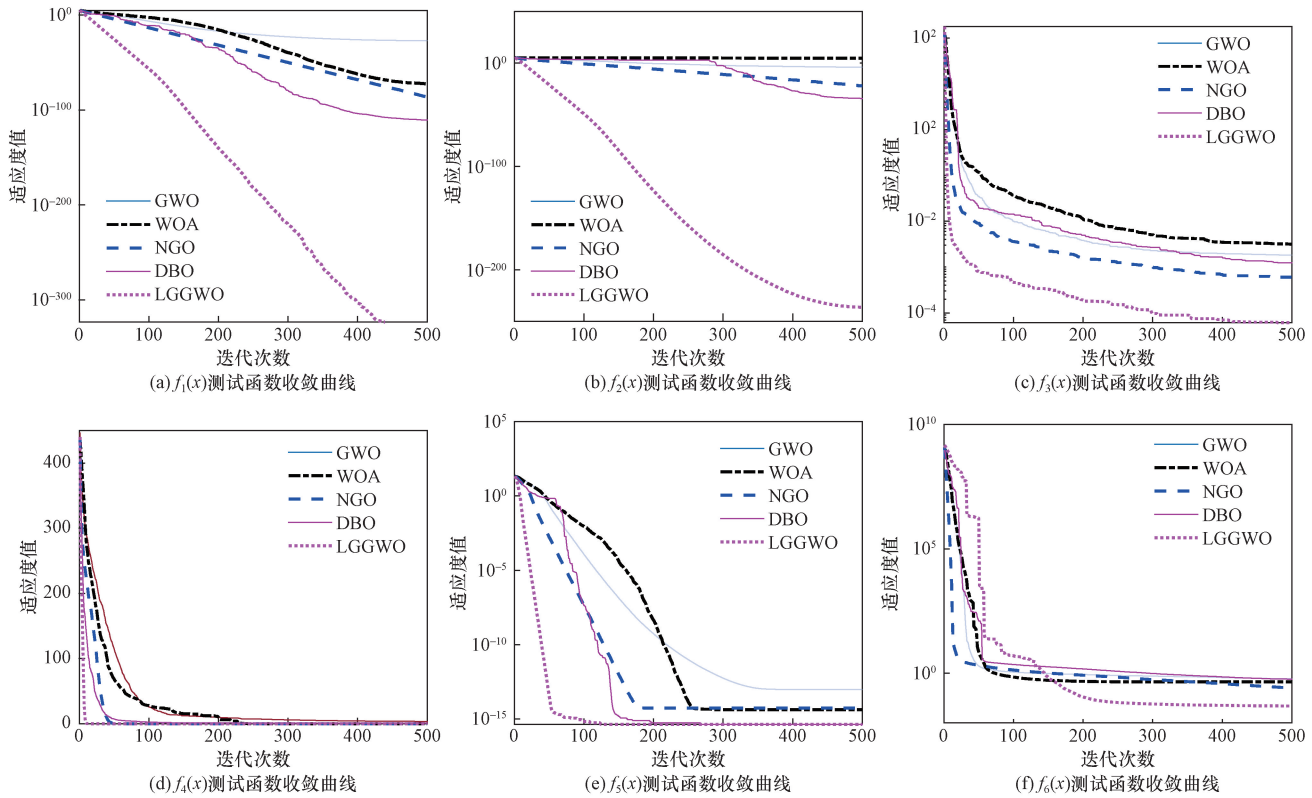


图5 5种算法对6个测试函数的收敛曲线

Fig. 5 Test function convergence curve comparison diagram

表3 优化算法测试结果

Table 3 Optimization algorithm test results

函数	统计指标	GWO	WOA	NGO	DBO	LGGWO
$f_1(x)$	最优值	1.1×10^{-26}	1.205×10^{-85}	1.394×10^{-89}	2.236×10^{-158}	0
	平均值	1.159×10^{-27}	8.320×10^{-73}	6.982×10^{-87}	1.951×10^{-116}	0
	标准差	2.398×10^{-27}	3.779×10^{-72}	2.335×10^{-86}	1.069×10^{-115}	0
$f_2(x)$	最优值	3.591×10^{-09}	898.727 6	1.208×10^{-29}	5.642×10^{-150}	8.449×10^{-301}
	平均值	5.996×10^{-6}	42 502.599	2.071×10^{-22}	1.942×10^{-38}	2.488×10^{-237}
	标准差	1.569×10^{-5}	14 467.977	6.000×10^{-22}	1.063×10^{-37}	0
$f_3(x)$	最优值	3.421×10^{-4}	6.150×10^{-5}	1.274×10^{-4}	4.095×10^{-5}	1.912×10^{-6}
	平均值	1.973×10^{-3}	1.869×10^{-3}	6.224×10^{-4}	1.322×10^{-3}	7.956×10^{-5}
	标准差	9.456×10^{-4}	2.134×10^{-3}	2.869×10^{-4}	1.034×10^{-3}	8.535×10^{-5}
$f_4(x)$	最优值	0	0	0	0	0
	平均值	2.937	0	0	2.421	0
	标准差	4.073	0	0	10.294	0
$f_5(x)$	最优值	7.505×10^{-14}	4.441×10^{-16}	3.997×10^{-15}	4.441×10^{-16}	4.441×10^{-16}
	平均值	1.03×10^{-13}	3.642×10^{-15}	6.01×10^{-15}	4.441×10^{-16}	4.441×10^{-16}
	标准差	$1.563 2 \times 10^{-14}$	2.158×10^{-15}	1.791×10^{-15}	0	0
$f_6(x)$	最优值	0.199 3	0.124 7	$3.821 6 \times 10^{-4}$	7.491×10^{-6}	5.099×10^{-6}
	平均值	0.618 2	0.489 8	0.302 1	0.619 7	0.131 9
	标准差	0.241 8	0.296 4	0.271 5	0.542 0	0.208 7

2.2 基于 LGGWO 算法优化 TCN 参数

控制 TCN 模型性能主要有以下几个参数。

(1) 卷积核大小。其决定了 TCN 模型在学习时间序列中的特征时的覆盖范围。通常较大的卷积核能够捕捉更长期的依赖关系,但可能会增加模型

的复杂度和计算成本。

(2) 残差链接数。残差连接通常用于加速 TCN 训练过程,缓解模型中梯度消失的问题,对模型的收敛速度和泛化能力有重要影响。

(3) 丢失率。丢失率是控制模型正则化的重要

参数,通常,丢失率的取值范围在0~1之间,常见的选择包括0.1、0.2、0.3等。设置低丢失率(如0.1)时,模型正则化效果较弱,有助于保持更多的特征信息,适合处理复杂的序列任务;设置高丢失率(如0.5)时,模型正则化效果强,可以有效防止过拟合,但可能导致模型学习困难,尤其是在数据量较少或模型较浅时。

因此为避免人为设置超参数导致的工作量增加、模型性能减弱等问题,使用LGGWO对TCN模型进行超参数寻优,流程图如图6所示,具体寻优流程见2.1节所示。

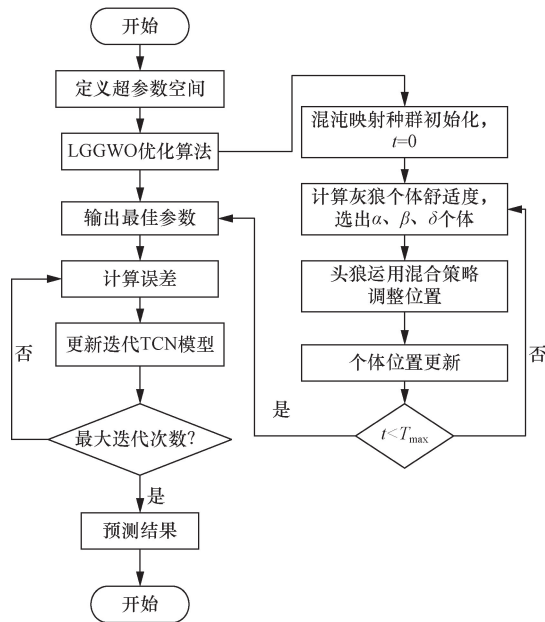


图6 LGGWO优化TCN流程图

Fig. 6 LGGWO optimizes TCN flow chart

2.3 LGGWO-TCN-MHSA 预测模型

本文中提出的基于LGGWO-TCN-MHSA的光伏功率短期预测流程如图7所示,具体步骤如下。

(1) 对数据进行特征提取与预处理。利用Spearman系数法筛选特征变量,同时对输入特征序列中异常值、缺失值进行数学处理。

(2) 将筛选处理后的特征数据归一化并划分数据集。其中80%为训练集,20%为测试集。

(3) 建立预测模型。将训练集序列输入至TCN-MHSA模块中,同时采用LGGWO算法对TCN内的超参数进行寻优以获得最佳的超参数序列。

(4) 将测试集数据输入最优超参数LGGWO-TCN-MHSA模型中,计算误差分析并得到输出测试集的预测结果。

(5) 利用RMSE、MAE、 R^2 作为误差评估标准,比较各模型的优劣,并进行评估分析。

在该组合模型中,时域卷积网络TCN负责模型

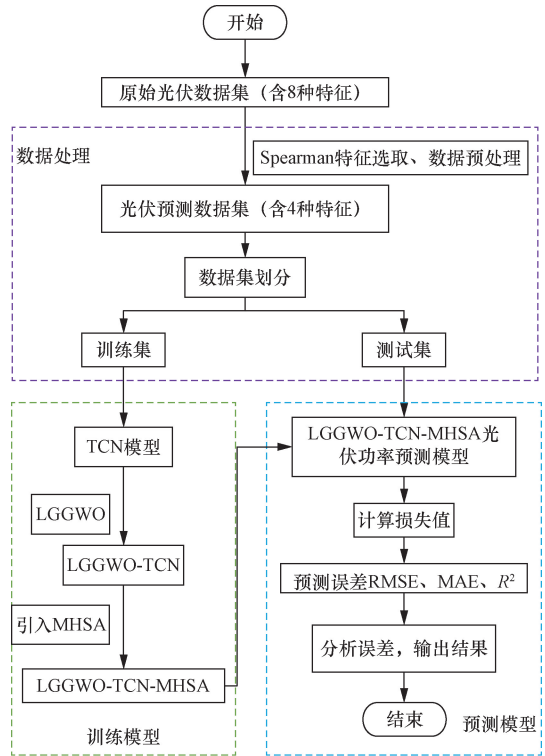


图7 LGGWO-TCN-MHSA 预测模型流程图

Fig. 7 LGGWO-TCN-MHSA predictive model flow chart

的主要时序预测部分;改进灰狼优化算法LGGWO负责对TCN模型内部超参数进行寻优,以提升TCN模型的预测性能;多头自注意力机制MHSA负责对TCN模型的输出结果进行加权处理,强调序列中关键的时间步信息,从而提升模型在时间序列预测任务上的表现力和准确性。

2.4 模型性能评估指标

为了评价预测模型的拟合优劣,本文中采用平均绝对误差MAE,均方根误差RMSE和决定系数 R^2 这3种不同指标作为模型性能的评估指标。具体计算公式为

$$\begin{cases} \text{MAE} = \frac{1}{m} \sum_{t=1}^m |y_t - y_t^*| \\ R^2 = 1 - \frac{\sum_{t=1}^m (y_t^* - y_t)^2}{\sum_{t=1}^m (y_t^* - \bar{y})^2} \\ \text{RMSE} = \sqrt{\frac{1}{m} \sum_{t=1}^m (y_t - y_t^*)^2} \end{cases} \quad (23)$$

式(23)中: m 为预测样本数量; y 为光伏发电功率实际值; y^* 为光伏发电功率预测值; \bar{y} 为光伏发电功率实际值的平均值。MAE和RMSE值越小表示预测效果越好; R^2 为决定系数,值越接近1表示预测值和真实值越接近,拟合效果越好。

3 算例分析

3.1 数据来源与数据预处理

算例数据选用澳大利亚 DKA 太阳能中心 (Desert Knowledge Australia Solar Centre) 某光伏电站(装机容量为 6.3 kW),采集时间为 2022 年 1、4、7、10 月这 4 个月的月中 15 d(即 11—25 日)的数据,采集时间为 00:00—24:00,采样周期为每 5 min 采集一次数据,每天共有 288 个时间节点,总计数据节点 17 280 个。

将筛选后的数据集进行数据提取,保留 06:00—18:55 的数据。针对数据中所涉及的异常值,需要进行剔除;数据中所涉及的缺失值采用缺失值前后数值的平均值进行填补,最后每日保留时间节点 156 个。数据集特征信息如表 4 所示。利用前文所提 Spearman 相关系数进行分析后,最终选择全球水平辐射、扩散水平辐射、贴片平均温度和光伏发电功率最为输入数据。

表 4 数据集特征信息

Table 4 Feature information of data set

数据集特征	单位
风速	m/s
贴片平均温度	℃
相对湿度	%
风向	(°)
降雨量	mm
全球水平辐射	W/m ²
扩散水平辐射	W/m ²
光伏发电功率	kW

3.2 超参数优化

TCN 模型内部超参数复杂且对性能影响较大,为避免人为设置超参数导致的工作量增加、模型性能减弱等问题,选用 LGGWO 对 TCN 内部进行超参数寻优。本实验选取卷积核数量、残差块个数以及 Dropout 率作为寻优的参数。设置 LGGWO 的基本参数,狼群数量为 10,最大迭代次数为 10 次,超参数优化结果如表 5 所示。

表 5 超参数优化结果

Table 5 Hyperparameter optimization results

参数	预设值	寻优值
卷积核数量	[16,64]	29
残差块个数	[1,5]	3
丢失率	[0.01,0.3]	0.12

3.3 对比模型参数设定

本次实验所选的对比模型,参数设置如表 6 所示,其中取预设值的平均值作为 TCN 单模型与 TCN-MHSA 组合模型的参数;LSTM 模型和 CNN 模型参数参考文献[14-15]中的数据进行设置。除了以下模型参数不同,运算过程中涉及的步长、学习率以及 Dropout 参数均保持相同,分别为 12、0.001 和 0.15。

表 6 对照模型参数设置

Table 6 Control group model parameter settings

模型名称	卷积核大小	卷积核数量/ LSTM 单元	激活函数
CNN	3×3	32	ReLU
LSTM	—	32	Tanh
TCN	3	32	ReLU

3.4 光伏功率预测

由于澳大利亚位于南半球,因此选择 1、4、7、10 月这 4 个月中每个月的月中 15 d 分别作为夏、秋、冬、春各季节的代表。以夏季数据集为例,前 80% 天数的数据作为训练集、后 20% 天数的数据作为测试集,其余 3 个季节数据集划分与夏季数据相同。

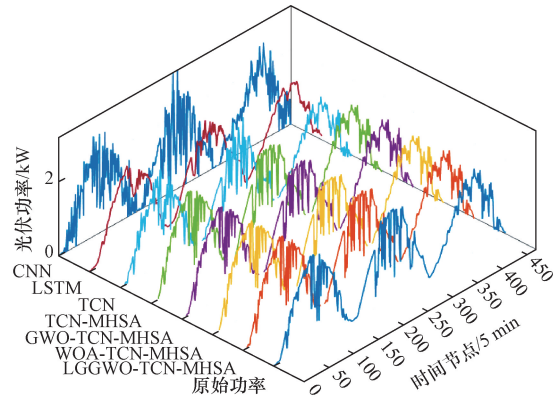
为了验证本文算法的有效性和优越性,将其和 CNN、LSTM、TCN、TCN-MHSA、GWO-TCN-MHSA、WOA-TCN-MHSA 等算法性能进行对比分析,结果如图 8 所示,预测误差统计结果如表 7 所示。

由表 7 中数据可知,在单一模型预测下,得益于膨胀卷积和残差结构,相比于 CNN 和 LSTM,TCN 在处理时序逻辑预测方面有明显的优势,其四个季节对应的预测误差均优于上述两模型。

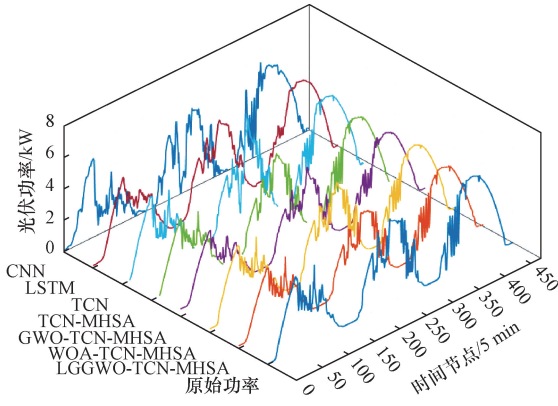
表 7 不同模型的四季预测误差

Table 7 Four seasons forecast error of different models

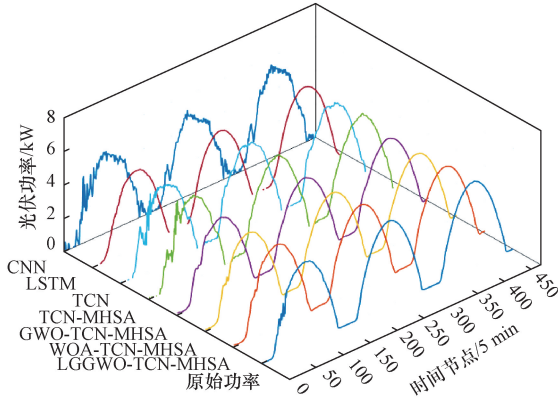
模型	夏季			秋季			冬季			春季			平均优化 时间/s
	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²	
CNN	0.683	0.981	0.712	0.592	0.886	0.722	0.381	0.521	0.896	0.454	0.665	0.883	—
LSTM	0.465	0.697	0.855	0.501	0.666	0.843	0.333	0.403	0.938	0.248	0.439	0.949	—
TCN	0.403	0.531	0.915	0.262	0.400	0.943	0.222	0.332	0.958	0.253	0.352	0.967	—
TCN-MHSA	0.170	0.213	0.987	0.331	0.386	0.947	0.122	0.212	0.983	0.137	0.165	0.991	—
GWO-TCN-MHSA	0.155	0.218	0.986	0.369	0.476	0.920	0.208	0.244	0.977	0.175	0.271	0.981	2 007.5
WOA-TCN-MHSA	0.198	0.244	0.982	0.280	0.334	0.960	0.138	0.162	0.990	0.169	0.194	0.990	1 715.5
LGGWO-TCN-MHSA	0.152	0.194	0.989	0.235	0.299	0.968	0.067	0.101	0.996	0.113	0.140	0.995	1 617.8



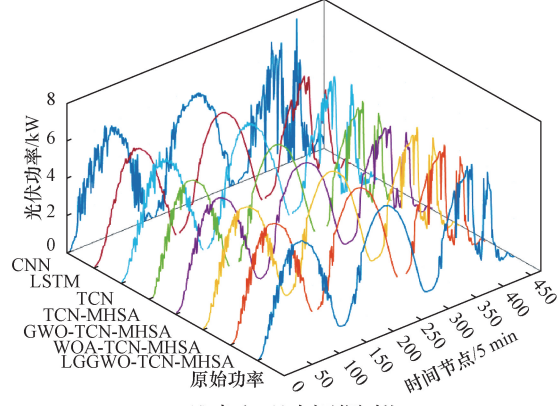
(a) 夏季三日功率预测对比



(b) 秋季三日功率预测对比



(c) 冬季三日功率预测对比



(d) 春季三日功率预测对比

图8 不同模型预测性能对比

Fig. 8 Comparison of prediction performance of different models

TCN-MHSA 组合模型,相比于单一 TCN 预测模型,在数据集波动性较大的夏秋两季,模型拟合相关系数 R^2 分别提高了 7.7% 和 5.3%,意味着预测值更贴近实际情况。MAE 降低了 57.7% 和 44.6%。RMAE 降低了 60% 和 52.9%,这表明组合模型相比于单一预测模型,拥有更好的抗扰动能力,多头注意力机制能够将 TCN 的输出加权处理,使得模型性能显著提高,精度提升,使得预测结果更贴近真实值。

针对本文中提出的 LGGWO-TCN-MHSA 模型中,改进灰狼优化算法 LGGWO 的优越性,本文中选择了原始灰狼算法 GWO 和鲸鱼优化算法 WOA 改进 TCN-MHSA 来作为复合模型的对照组。夏秋两季较为波动的数据使得 GWO 和 WOA 易陷入局部最优解,得到的所谓最优超参数代入 TCN-MHSA 中所输出的结果和对照组模型的输出结果相差不大,甚至秋季数据中 GWO-TCN-MHSA 模型的输出结果要略差于未寻优的 TCN-MHSA 模型。

由于 WOA 相比于 GWO 提出时间较晚,WOA-TCN-MHSA 模型全阶段优于 GWO-TCN-MHSA,二者均差于本文所提出模型。同时表格中也给出了三种优化算法的平均优化时间,LGGWO 的平均优化时长要短于两个对照组模型,说明本文所提优化算法改进了原始 GWO 算法易陷入局部最优解的弊端,能够对 TCN 模型内超参数进行寻优的同时,加快优化进程,使组合模型更好地捕捉光伏功率的变化规律,进而得到最接近实际值的预测曲线。

为了更好地验证本文方法的有效性,并评估各模型在极端天气下的光伏功率预测稳定性,选取 2016 年 1 月 24 日的光伏功率进行模型预测,结果如图 9 所示。

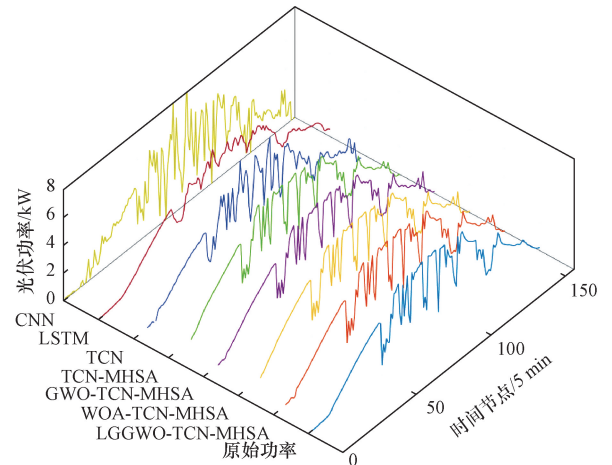


图9 极端天气模型性能对比

Fig. 9 Comparison of extreme weather model performance

从图 9 可以看出,在极端条件天气下(降雨及阴天),CNN、LSTM、TCN 模型预测误差较大。当原始数据波动较为严重即天气变化幅度较大时,单模型预测结果会随原始数据波动频繁,产生结果失真。结合注意力机制后,TCN 预测精度显著提升,数据波动减小,并且预测结果符合原始数据波动趋势。对 TCN 模型进行参数选优后,3 个结合优化算法的复合模型总体预测结果趋于一致,但在转折点处依旧会出现失真以及误差较大等情况。本文提出的组合模型相比于上述模型,整体预测误差均为最小,且模型拟合度最高。证明本文所提组合模型具有更好的稳定性以及更好的预测性能。

4 总结

针对原始数据复杂度较高条件下的光伏功率预测,本文提出了一种基于 LGGWO-TCN-MHSA 组合模型的短期光伏功率预测方法。通过对比不同模型之间的预测效果,证明所以方法的稳定性以及有效性,得到以下结论。

(1)单模型对比下,与 CNN 和 LSTM 相比,TCN 所得的预测误差较小,说明针对时序逻辑预测方面,膨胀卷积以及残差模块可以使 TCN 更好地抓住数据间的联系,提升预测精度。

(2)与单 TCN 模型相比,TCN-MHSA 所得预测误差均在一定程度上有所下降,说明注意力机制结合 TCN 能够更好地提升模型性能以及抗扰动力,从而提升预测精度。

(3)在 GWO 基础上融入混沌映射、非线性收敛因子和混合策略,有效平衡了算法的全局以及局部搜索能力,改善了易陷入局部最优的问题,加快了算法的寻优速度。将改进后的 LGGWO 与 TCN-MHSA 相结合,提高了组合模型的求解性能;相较于不含有优化算法的对比模型,本文所提方法的 MAE 降低了 10.6% ~ 82.0%, RMAE 降低了 11% ~ 80.1%;对比包含优化算法的复合模型,本文所提方法的 MAE 降低了 2.03% ~ 67.8%, RMAE 降低了 10.5% ~ 58.6%,优化速度相较原灰狼算法提升 19.4%,充分证实所提方法的有效性。

(4)通过对极端天气下的光伏功率进行预测和分析,本文所提方法相比于其他对照模型,预测结果与实际数据拟合度最高,预测精度最优。证明本文组合模型拥有较高稳定性,能够实现短期光伏发电功率的有效预测。

尽管本文所提模型在光伏功率全期预测上取得了一定成果,但仍存在一些不足之处,未来可以在以下几方面深入研究:一是扩大模型数据收集范

围,采用多种地区的光伏数据来源,以提高研究结果的可靠性和推广性;二是更新迭代原有优化算法,减少模型计算复杂度的同时,提升模型运行速度,实现应用效率提升。

参 考 文 献

- [1] 国家能源局. 2023 年全国电力工业统计数据[J]. 电力科技与环保, 2024, 40(1): 95.
National Energy Administration. 2023 national power industry statistics[J]. Electric Power Science and Technology and Environmental Protection, 2024, 40(1): 95.
- [2] 冯沛儒, 江桂芬, 徐加银, 等. 基于 CEEMD-BiLSTM-RFR 的短期光伏功率预测[J]. 科学技术与工程, 2024, 24(5): 1955-1962.
Feng Peiru, Jiang Guifen, Xu Jiayin, et al. Short-term PV power prediction based on CEEMD-BiLSTM-RFR[J]. Science Technology and Engineering, 2024, 24(5): 1955-1962.
- [3] 胡子延, 温蜜, 魏敏捷. 基于 ELM 和误差校正的风力发电区间预测方法[J]. 计算机仿真, 2024, 41(4): 69-74, 101.
Hu Ziyuan, Wen Mi, Wei Minjie. Wind power generation interval prediction method based on ELM and error correction[J]. Computer Simulation, 2024, 41(4): 69-74, 101.
- [4] 翟孝明. 基于支持向量机的区域电网光伏发电量预测研究[J]. 电工技术, 2024(3): 45-48.
Zhai Xiaoming. Research on regional power grid PV power generation prediction based on support vector machine[J]. Electrotechnics, 2024(3): 45-48.
- [5] Ibrahim M S, Gharghory S M, Kamal H A. A hybrid model of CNN and LSTM autoencoder-based short-term PV power generation forecasting[J]. Electrical Engineering, 2024, 106: 4239-4255.
- [6] Tai B, Yu L, Huang Y J, et al. Power prediction of photovoltaic power generation based on LSTM model with additive attention mechanism[C]//2023 7th International Conference on Smart Grid and Smart Cities (ICSGSC). New York: IEEE, 2023: 474-480.
- [7] 杨晶, 赵津蔓, 孟润泉, 等. 基于粒子群优化和卷积神经网络的电力系统运行状态辨识[J]. 电网技术, 2024, 48(1): 315-324.
Yang Jing, Zhao Jinman, Meng Runquan, et al. Identification of power system operating state based on particle swarm optimization and convolutional neural network[J]. Power System Technology, 2024, 48(1): 315-324.
- [8] 岳有军, 刘金林, 赵辉, 等. 考虑特征提取和优化 LSSVM 的短期光伏功率预测[J]. 陕西科技大学学报, 2023, 41(3): 200-207.
Yue Youjun, Liu Jinlin, Zhao Hui, et al. Consider feature extraction and optimize short-term PV power prediction of LSSVM[J]. Journal of Shaanxi University of Science and Technology, 2023, 41(3): 200-207.
- [9] 郝可青, 吕志刚, 邸若海, 等. 基于鲸鱼算法优化长短时记忆神经网络的锂电池剩余寿命预测[J]. 科学技术与工程, 2022, 22(29): 12900-12908.
Hao Keqing, Lü Zhigang, Di Ruohai, et al. Optimization of the remaining life prediction of lithium battery based on whale algorithm

- based on short-duration memory neural network[J]. Science Technology and Engineering, 2022, 22(29): 12900-12908.
- [10] 甘福宝, 王仲阳, 连寅行, 等. 基于改进灰狼优化算法的移动机器人路径规划方法[J]. 传感器与微系统, 2024, 43(8): 110-113.
- Gan Fubao, Wang Zhongyang, Lian Yinxing, et al. Path planning method for mobile robot based on improved gray wolf optimization algorithm[J]. Sensors and Microsystems, 2024, 43(8): 110-113.
- [11] 李阳, 李维刚, 赵云涛, 等. 基于莱维飞行和随机游动策略的灰狼算法[J]. 计算机科学, 2020, 47(8): 291-296.
- Li Yang, Li Weigang, Zhao Yuntao, et al. Grey wolf algorithm based on Levy flight and random walk strategies[J]. Computer Science, 2020, 47(8): 291-296.
- [12] 王秋萍, 王梦娜, 王晓峰. 改进收敛因子和比例权重的灰狼优化算法[J]. 计算机工程与应用, 2019, 55(21): 60-65, 98.
- Wang Qiuping, Wang Mengna, Wang Xiaofeng. Grey wolf optimization algorithm with improved convergence factor and proportional weight[J]. Computer Engineering and Applications, 2019, 55(21): 60-65, 98.
- [13] Yang W, Zhang Y, Zhu X, et al. Research on dynamic economic dispatch optimization problem based on improved grey wolf algorithm[J]. Energies, 2024, 17(6): 1491.
- [14] 邓皓云, 陈卓. 基于 EEMD-IWOA-TCN 的电网短期负荷预测[J]. 电力信息与通信技术, 2024, 22(1): 70-76.
- Deng Haoyun, Chen Zhuo. Short-term load prediction of power grid based on EEMD-IWOA-TCN[J]. Electric Power Information and Communication Technology, 2024, 22(1): 70-76.
- [15] 杨婷婷, 罗海玉, 李浩千, 等. 基于 IWOA 优化 TCN 模型的磨煤机故障预警研究[J/OL]. 华北电力大学学报(自然科学版), 1-9[2024-10-04]. <http://kns.cnki.net/kcms/detail/13.1212.TM.20240118.1430.002.html>.
- Yang Tingting, Luo Haiyu, Li Haoqian, et al. Research on coal mill fault warning based on IWOA optimized TCN model[J/OL]. Journal of North China Electric Power University (Natural Science Edition), 1-9[2024-10-04]. <http://kns.cnki.net/kcms/detail/13.1212.TM.20240118.1430.002.html>.