

基于IWOA-BiLSTM-MHSA神经网络的超短期风电功率预测

张州潼, 郭欢

(江汉大学人工智能学院, 湖北 武汉 430056)

摘要: 为提高超短期风电功率预测精度, 从超参数寻优与模型优化两个角度出发, 提出一种改进鲸鱼优化算法结合多头注意力机制的双向长短期记忆神经网络模型。首先, 在双向长短期记忆神经网络的基础上引入时序多头自注意力机制, 用于捕捉时序中更长距离的依赖关系, 增加模型的表征和泛化能力; 其次, 采用改进的鲸鱼优化算法对模型超参数进行寻优, 提出自适应参数和阈值策略, 增强算法的全局搜索能力和局部搜索速率; 最后, 基于新疆某地区风电场的实测数据展开算例分析。与多种传统机器学习和深度学习预测模型相比, 所提模型具有更高的预测精度, 特别是在数据集1中的MAPE相较LSTM、BP、RNN、GRU、BiLSTM模型分别降低了4.892%、7.722%、6.196%、3.864%、2.159%。

关键词: 双向长短期记忆网络; 多头自注意力机制; 风电功率预测; 改进鲸鱼优化算法

DOI: 10.11907/rjdk.241457

开放科学(资源服务)标识码(OSID):

中图分类号: TP183

文献标识码: A

文章编号: 1672-7800(2025)007-0046-08



Ultra-Short-Term Wind Power Prediction Based on IWOA-BiLSTM-MHSA Neural Network

ZHANG Zhoutong, GUO Huan

(Artificial Intelligence Academy, Jiangnan University, Wuhan 430056, China)

Abstract: To improve the accuracy of ultra short term wind power prediction, an improved whale optimization algorithm combined with multi head attention mechanism is proposed from the perspectives of hyperparameter optimization and model optimization, and a bidirectional long short-term memory neural network model is proposed. Firstly, based on the bidirectional long short-term memory neural network, a temporal multi head self attention mechanism is introduced to capture longer distance dependencies in the temporal sequence, increasing the model's representation and generalization ability; Secondly, an improved whale optimization algorithm is adopted to optimize the hyperparameters of the model, proposing adaptive parameter and threshold strategies to enhance the algorithm's global search capability and local search rate; Finally, based on the measured data of a wind farm in a certain region of Xinjiang, a case study analysis was conducted. It was found that compared with various traditional machine learning and deep learning prediction models, the proposed model has higher prediction accuracy, especially in dataset 1, where MAPE was reduced by 4.892%, 7.722%, 6.196%, 3.864%, and 2.159% compared to LSTM, BP, RNN, GRU, and BiLSTM models, respectively.

Key Words: BiLSTM; multi-head self-attention mechanism; wind power prediction; improved whale optimization algorithm

0 引言

在全球能源需求持续攀升的背景下,化石能源枯竭与

气候变暖的双重危机正对人类社会的可持续发展构成严峻挑战,新能源的发展迫在眉睫。风力发电作为主要的新能源发电途径之一,具有建设周期短、技术需求低、成本回收迅速以及环境污染少等优势,受到世界各国的青睐,已

收稿日期: 2024-05-14

扫描二维码阅读全文:

基金项目: 教育部人文社科基金项目(23YJCZH062)

作者简介: 张州潼(2000-),男,江汉大学人工智能学院硕士研究生,研究方向为人工智能;郭欢(1984-),女,博士,江汉大学人工智能学院副教授、硕士生导师,研究方向为预测理论与方法、复杂系统控制与优化决策。本文通讯作者: 郭欢。



成为发展最快的可再生能源之一^[1-2]。准确的风力发电功率预测不仅可以提高电力系统的稳定性和可靠性,而且可以帮助电力公司作出合理的运营决策,减少能源浪费,降低生产成本,提高经济效益^[3]。然而,与传统的化石燃料发电相比,风力发电的不稳定性使得准确预测其功率更具有挑战性^[4]。为确保电网正常运行并降低额外维护成本,研究高精度的风电场功率预测方法至关重要,尤其是针对超短期风电功率的预测^[5]。

目前,国内外学者提出许多预测风电功率的方法,主要分为物理方法、统计方法和机器学习方法^[6]。传统物理方法和统计方法通常依赖气象数据和统计模型进行分析,但在处理复杂气象系统与风力发电输出间的非线性关系时面临显著挑战^[7]。例如, Huang 等^[8]基于统计方法生成预测模型对短期风速进行预测,虽然取得了良好的预测效果,但是模型复杂度较高、计算效率较低。随着深度学习的不断发展,基于神经网络的风力发电预测方法成为研究热点^[9]。例如,朱乔木等^[10]使用长短期记忆神经网络(Long Short-Term Memory, LSTM)对风电功率进行超短期预测,预测效果优于人工神经网络(Artificial Neural Network, ANN)和支持向量机(Support Vector Machine, SVM),但风速的不确定性可能导致预测精度下降。为此,一些学者提出将降噪技术与深度学习模型相结合的预测方法^[11]。例如,曹有为等^[12]利用奇异谱分析(Singular Spectrum Analysis, SSA)对原始风电数据进行降噪处理,然后将处理后的数据用于 LSTM 模型训练,从而进行风电功率预测,但是由于 LSTM 结构的单向性无法考虑未来信息,导致预测精度较低。Graves 等^[13]提出的双向长短期记忆神经网络(Bidirectional Long Short-Term Memory, BiLSTM)可以捕获时间序列中的双向信息,预测精度高于 LSTM,但是捕捉长期关系的能力不足。LSTM 和 BiLSTM 优化了对局部时序依赖的处理,在处理较短的序列数据时展现出较好的时序信息挖掘能力,然而在应对风电功率这种长序列多维数据时可能会遗漏关键的序列结构信息,且难以捕捉到长期依赖关系^[14]。为此,姜旭初等^[15]结合 BiLSTM、卷积神经网络(Convolutional Neural Networks, CNN)与注意力机制进行短期风力发电预测,显著提升了预测性能,但模型最优超参数无法确定。刘爱国等^[16]采用遗传算法(Genetic Algorithm, GA)确定 SVM 的惩罚因子和核函数的最优超参数,但算法的迭代收敛速度仍有待提高。Yuan 等^[17]使用引力搜索算法(Gravitational Search Algorithm, GSA)优化最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)的核函数参数,为 LSSVM 模型选择最佳参数。Chen 等^[18]利用 LSTM 和反向传播神经网络(Back Propagation Neural Network, BPNN)构建短期风速预测模型,并利用麻雀优化算法(Sparrow Search Algorithm, SSA)对输入序列中不同复杂度的预测模型进行优化,但是模型的收敛速度和精度有待提高。

综上所述,目前风电预测还存在诸多挑战,单一深度学习模型不能综合考虑数据的局部和全局信息,预测精度差强人意^[19]。组合模型在风电功率预测精度和时长上具有显著优势,但性能在很大程度上受到超参数配置的影响,且最优超参数难以确定。当采用智能优化算法优化深度学习模型参数时,算法容易陷入局部最优,且优化速度较慢。针对上述问题,本文提出一种采用改进鲸鱼优化算法(Improved Whale Optimization Algorithm, IWOA)优化的基于 BiLSTM 与多头自注意力机制的模型 IWOA-BiLSTM-MHSA (IWOA-BiLSTM-Multi-Head Self-Attention) 模型。首先,在 BiLSTM 网络中添加后置的 MHSA 机制以增加模型对序列中重要信息的敏感度,允许模型更好地捕捉序列内的长距离依赖关系;其次,基于改进的鲸鱼优化算法(Improved Whale Optimization Algorithm, IWOA)对 BiLSTM-MHSA 模型进行超参寻优得到最优模型;最后,基于优化后的模型进行预测,与传统模型预测结果进行比较,验证所提模型的有效性。

1 模型构建

IWOA-BiLSTM-MHSA 模型流程如图 1 所示。首先,对收集到的数据进行预处理,并将其分为训练集和测试集。两层 BiLSTM 层能够分别捕捉输入序列的正向和反向模式,学习输入数据中的时序模式,具有长时记忆性。其次, MHSA 机制使模型能够集中处理序列中的关键部分,对于提高预测准确性尤其重要,尤其是在存在噪声或不相关信息的长序列中有助于提高模型对输入序列的理解与建模能力,挖掘深层次的时间相关性。同时在注意力机制中添加正弦位置编码,为模型提供序列位置信息,帮助其学习输入序列中不同位置的模式,增强了模型对序列结构的建模能力。最后,将 MHSA 层的输出传递给全连接层,该层负责输出风电功率预测结果,并使用测试集验证模型的有效性。

1.1 BiLSTM

传统的 LSTM 为单向结构,只能通过学习过去的信息来预测未来,但实际上时间序列的输出不止会受到过去的影响,而且会受到未来信息的影响^[20]。针对这一问题,对 LSTM 进行改良,将正向 LSTM 与反向 LSTM 组合成为 BiLSTM,其中正向 LSTM 使用过去信息预测未来信息,反向 LSTM 利用未来信息预测过去信息,共同影响 BiLSTM 的输出^[21]。BiLSTM 能分别捕捉输入序列的正向和反向模式,学习其中的时序模式,具有长时记忆性,适合处理时序数据^[22]。具体结构如图 2 所示,其中 $x_1 \sim x_t$ 为每一个时刻输入的数据, $h_1 \sim h_t$ 为相应的输出数据。BiLSTM 的输出过程为:

$$A_i = f_1(w_1 x_i + w_2 A_{i-1}) \quad (1)$$

$$B_i = f_2(w_3 x_i + w_4 B_{i+1}) \quad (2)$$

$$Y_i = f_3(w_5 A_i + w_6 B_i) \quad (3)$$

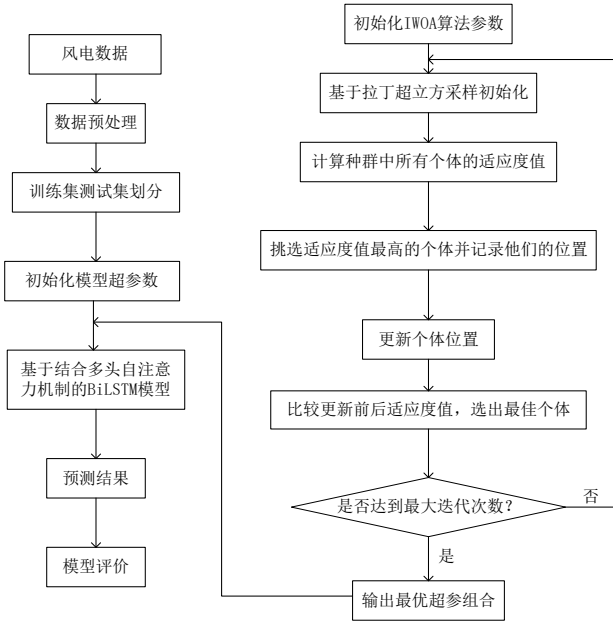


Fig. 1 Flow of IWOA-BiLSTM-MHSA model

图1 IWOA-BiLSTM-MHSA 模型流程

式中: $A_1, A_2, \dots, A_i, B_1, B_2, \dots, B_i$ 分别表示相应的前向和后向 LSTM 隐藏状态; f_1, f_2, f_3 为激活函数; $w_1, w_2, \dots, w_6$ 表示相应权重; $Y_1, \dots, Y_i$ 表示相应的输出。

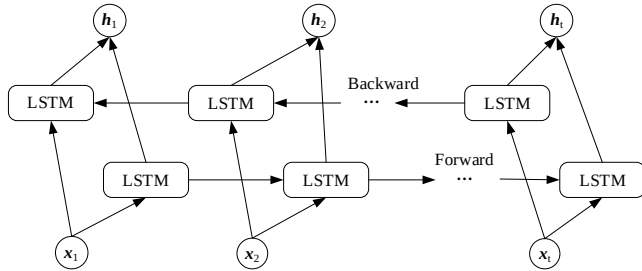


Fig. 2 Structure of BiLSTM

图2 BiLSTM 结构

1.2 MHSA

针对 BiLSTM 处理长序列能力不足的问题, 本文引入后置式 MHSA 机制, 使模型能够首先专注于局部特征提取, 然后捕捉时间序列中更长距离的依赖关系, 从而增强模型的表达能力和泛化能力^[23]。如图 3 所示, BiLSTM-MHSA 模型首先将 BiLSTM 每个时间步的输出作为一组时间序列数据输入 MHSA 机制层; 其次, 通过权重矩阵计算出每个时间步对应的 q, k, v 值; 再次, 通过计算每个时间步的输出 q 与其他时间步的输出 k 之间的点积, 并将结果与对应的 v 值相乘得到注意力矩阵, 该矩阵反映了不同时间步之间的重要关联信息; 最后, 将计算得到的注意力矩阵拼接起来, 形成一个大注意力矩阵, 用于对整个序列数据进行综合建模。BiLSTM-MHSA 模型将 BiLSTM 与 MHSA 有效结合起来, 能够更好地理解和建模序列数据的关联性, 具体结构如图 3 所示。

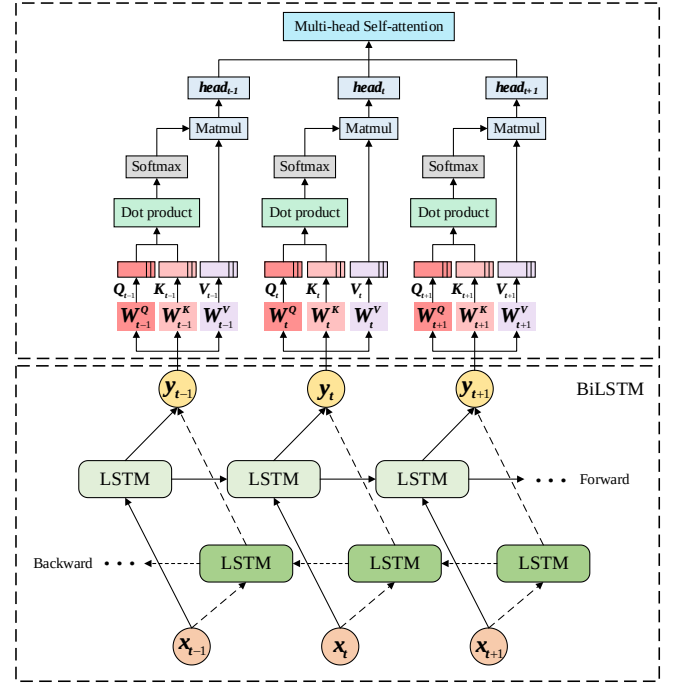


Fig. 3 Structure of BiLSTM-MHSA model

图3 BiLSTM-MHSA 模型结构

1.3 IWOA

1.3.1 WOA

WOA 是一种模拟鲸鱼群体捕食行为的优化方法, 主要包括围攻猎物、气泡网攻击和搜索猎物 3 个阶段。假设鲸鱼种群规模为 N , 搜索空间为 d 维, 第 i 只鲸鱼在 d 维空间中的位置表示为 $X_i = (X_i^1, X_i^2, \dots, X_i^d)$, $i = 1, 2, \dots, N$ ^[24]。每个鲸鱼的位置象征一个可能的解决方案, 通过持续更新这些位置, 在解的空间内寻找全局最优解。

在围攻猎物阶段, 初始最优解被设置为具有最高当前适应度的个体鲸鱼位置, 然后其他个体将尝试接近最优解并更新其位置。位置更新公式为:

$$D = |C \cdot X_{best}(t) - X(t)| \quad (4)$$

$$X(t+1) = X_{best}(t) - A \cdot D \quad (5)$$

式中: t 表示迭代次数; $X_{best}(t)$ 表示目前的全局最优位置; $X(t)$ 表示个体的当前位置; $A \cdot D$ 表示围攻猎物的步长。 A 和 C 的定义分别为:

$$A = 2a \cdot rand - a \quad (6)$$

$$C = 2 \cdot rand \quad (7)$$

式中: $rand$ 表示 $[0, 1]$ 内的随机数; a 为在迭代过程中随着迭代次数增加, 从 2 线性递减为 0 的收敛因子。表达式为:

$$a = (2 - 2t/T_{max}) \quad (8)$$

式中: t 和 T_{max} 分别为当前迭代次数和最大迭代次数。

气泡网攻击阶段包括收缩包围和螺旋更新位置两种行为, 二者同时进行, 需要引入随机概率 p 来作为判断的阈值。收缩包围机制通过减少线性收敛因子 a 来实现, 螺旋更新位置采用对数螺旋方程表示。

$$X(t+1) = \begin{cases} X_{best}(t) - A \cdot |C \cdot X_{best}(t) - X(t)|, p < 0.5 \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X_{best}(t), p \geq 0.5 \end{cases} \quad (9)$$

在搜索猎物阶段,WOA算法通过计算鲸鱼间的相对距离来更新其位置,从而实现对猎物的随机搜索($|A|>1$ 时)。公式为:

$$D = |C \cdot X_{rand} - X(t)| \quad (10)$$

$$X(t+1) = X_{rand} - A \cdot D \quad (11)$$

式中, X_{rand} 为当前种群中随机选择的鲸鱼位置。

1.3.2 改进措施

WOA虽然在解决优化问题上具有一定优势,但是在处理一些高维问题上存在收敛速度较慢、参数设置困难、局部搜索能力不足、多样性缺乏和参数依赖性高等不足。为此,本文提出IWOA,主要包含以下3个改进点。

(1)拉丁超立方采样种群初始化。种群多样性是避免算法陷入局部最优和加速全局搜索的关键因素。在原始WOA中,初始种群的生成通常是随机的,这可能导致种群多样性不足。为解决这一问题,本文基于拉丁超立方体抽样方法产生初始种群位置,保证全空间填充和抽样非重叠,从而使种群分布均匀,提高算法的多样性和搜索效率,加快收敛速度,并减少早熟收敛的风险。具体来说,定义参数空间,确定优化参数的范围,将其划分为 N 个相等的间隔(N 为生成的样本点数量)。在每个维度上,从每个间隔中随机选取一个值,既保证了每个间隔都被采样到,又使间隔内的采样点是随机的。然后将不同维度上的采样点组合起来,形成多维空间中的样本点,每个样本点是一个独立的个体或解。假设有一个参数 x 在区间 $[x_{min}, x_{max}]$ 内,若要生成 N 个样本点,对于第 i 个样本点,其值 x_i 通过式(12)进行计算。

$$x_i = x_{min} + \left(\frac{i - \text{random}(0,1)}{N} \right) \times (x_{max} - x_{min}) \quad (12)$$

重复该过程对每个参数进行操作,然后将所有参数的值组合起来,形成一个多维空间中的样本点。

(2)自适应参数。在原始WOA中,收敛因子 a 随着迭代次数线性递减至0,从而控制搜索步长,使算法在初始阶段具有较大的搜索范围(全局搜索),在迭代中期进入局部搜索阶段。为减少算法求解优化问题的耗时,同时增强其局部开发能力,提高解的精度,本文的收敛因子 a 使用一种非线性变化策略,该策略使参数向量 A 更早落入 $[-1,1]$ 的区间内,从而提升算法的求解精度与速度。

在螺旋更新过程中, b 是影响鲸鱼位置的一个重要参数。 b 取较大值时有利于搜索,取较小值时有利于算法收敛。本文将参数 b 映射到迭代次数 t 以实现自适应调整,在算法的初始阶段取较大 b 值以加快全局搜索速度;后期参数 b 随着迭代次数的增加而减小,从而加快局部搜索速度。 a 和 b 的表达式分别为:

$$a(t) = 2 \cdot (1 - \arctan(\sqrt{\frac{t}{\max_iteration}})) \quad (13)$$

$$b = 5 \cdot (1 - \frac{t}{\max_iteration}) \quad (14)$$

式中: t 为当前迭代次数; $\max_iteration$ 为最大迭代次数。

(3)自适应阈值。原始WOA中产生一个随机数 p ,取值在 $[0,1]$ 之间,采用这个随机数与50%进行比较,以判断鲸鱼是选择包围活动还是螺旋运动。然而,这种判断方式缺乏可靠性,很可能无法选择适合当前种群的运动方式。为此,本文引入自适应概率阈值 p_a 对WOA算法进行改进,将随机生成的 $[0,1]$ 范围内的随机数 p 与 p_a 进行比较,使鲸鱼有较大概率选择到适合当前群体的捕食策略。该改进措施改变了传统WOA中两种运动等概率同时进行的方式,而是变成了先进行包围活动,再进行螺旋运动,使鲸鱼种群能够更好更快地搜索到猎物,有利于提高算法的预测拟合度和收敛速度。同时,算法根据不同的状态在全局搜索与局部搜索之间进行平衡,具有更强的自适应性和搜索能力。自适应概率阈值 p_a 的表达式为:

$$p_a = 1 - \left[\frac{1}{3} \left(2 \times \frac{t^e}{\max_iteration^e} + \frac{t}{\max_iteration} \right) \right] \quad (15)$$

式中: t 为当前迭代次数; e 为控制参数; $\max_iteration$ 为最大迭代次数。因此,可将鲸鱼狩猎行为的公式改写为:

$$X(t+1) = \begin{cases} X_{best}(t) - A \cdot D, p < p_a \\ D' \cdot e^{bl} \cdot \cos(2\pi l) + X_{best}(t), p \geq p_a \end{cases} \quad (16)$$

自适应概率阈值 p_a 随着迭代次数的增大而减小,说明在迭代次数还很小时, p_a 值较大,此时 p 值会比 p_a 值小,那么鲸鱼会选择包围运动。相应地,在迭代次数很大时,鲸鱼会选择螺旋运动。

综上所述,IWOA算法流程如图4所示。

IWOA算法的伪代码为:

输入:鲸鱼种群规模 N ,迭代次数 t ,最大迭代次数 T_{max}

输出:最优位置 $X_{best}(t)$,最优值 $F(X_{best})$

1.初始化:使用拉丁超立方采样方法初始化 N 只鲸鱼,生成初始种群 x_i ;

2.迭代:

While $t < T_{max}$:

2.1 计算种群中每个个体的适应度值,记录当前最优解

2.2 更新全局最优位置 $X_{best}(t)$

2.3 依概率规则生成新种群:

For $i=1$ to N do:

2.3.1 根据公式(6)-(8),公式(13)-(14)计算更新收敛因子 a 、更新 A 和 C

2.3.2 if $p < p_a$

if $|A| < 1$:

按照公式(4)、(5)更新当前个体位置

else:

按照公式(10)、(11)更新当前个体位置

else:

按照公式(9)更新当前个体位置

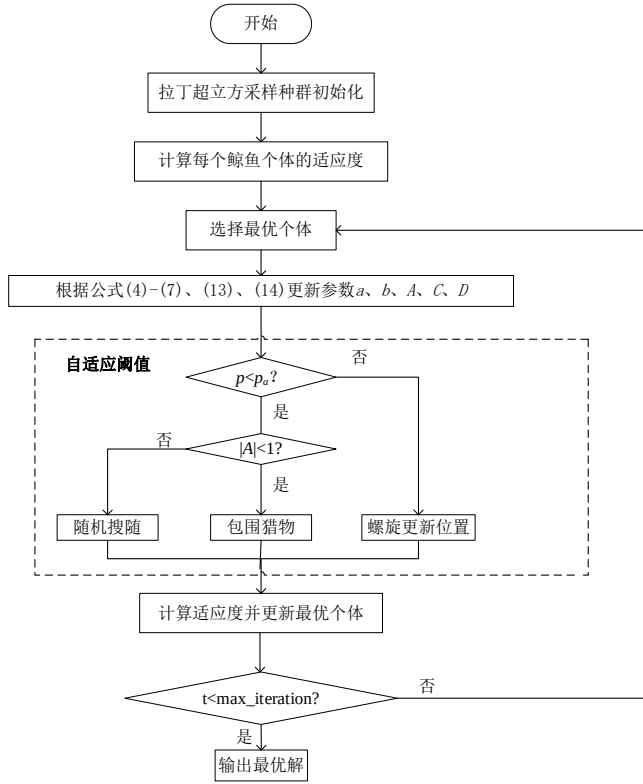


Fig. 4 IWOA algorithm flow

图4 IWOA算法流程

end for
 2.4 $t=t+1$
 end while
 3 返回:返回全局最优位置 $X_{best}(t)$ 和最优适应度值 $F(X_{best})$

2 实验方法与结果分析

2.1 实验环境

本文所有实验均在同一环境下进行,操作系统为 Windows11 (64 bit), 12Gen Inter(R)Core(TM)i7-12700H CPU、2.70 GHz 主频及 16 GB 内存, NVIDIA GeForce RTX3050 GPU, 计算环境 Python3.7, TensorFlow2.6.0。

2.2 IWOA 算法测试分析

采用基准测试函数对 IWOA 算法与传统 WOA、GA、PSO、SSA 算法进行比较,测试所提算法的全局搜索能力和寻优稳定性。基准测试函数详细信息如表 1 所示。其中, f_1 ~ f_4 为单峰函数, f_5 ~ f_6 为多峰函数。

为确保比较的公平性,每种算法的最大迭代次数设定为 500 次,种群规模设定为 50,进行 30 次独立运行。各算法在 6 个测试函数上的测试结果如表 2 所示。各算法收敛曲线比较如图 5 所示。

可以看出, IWOA 在寻优精度和收敛速度上均优于 WOA、GA、PSO、SSA 算法。IWOA 算法在求解函数 f_1 、 f_3 、 f_6 时寻优精度达到理论最优值 0;在求解函数 f_2 、 f_4 、 f_5 时虽未能达到理论最优值,但 IWOA 相较于 WOA、GA、PSO、SSA

Table 1 Benchmark function information

表1 基准测试函数信息

函数公式	维度	范围	最优值
$f_1(x) = \sum_{i=1}^n x_i^2$	30	$[-10, 10]$	0
$f_2(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $	30	$[-10, 10]$	0
$f_3(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$	30	$[-100, 100]$	0
$f_4(x) = \max_i \{ x_i \}$	30	$[-100, 100]$	0
$f_5(x) = \sum_{i=1}^n x_i^n - 10 \cos(2\pi x_i) + 10$	30	$[-5.12, 5.12]$	0
$f_6(x) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(\frac{x_i}{\sqrt{i}}) + 1$	30	$[-600, 600]$	0

Table 2 Test results

表2 测试结果

函数	参数	IWOA	WOA	GA	PSO	SSA
f_1	平均值	0	8.36×10^{-61}	32.07	23.33	1675.09
	标准差	0	4.69×10^{-62}	71.56	6.47	648.64
f_2	平均值	1.94×10^{-258}	3.71×10^{-45}	886.77	567.84	53.73
	标准差	0	5.31×10^{-41}	539.84	703.37	29.60
f_3	平均值	0	3.36×10^{-6}	17.34	971.41	24.71
	标准差	0	3.21×10^{-5}	12.98	912.40	79.88
f_4	平均值	4.38×10^{-5}	1.45	88.94	20.11	53.99
	标准差	4.98×10^{-5}	2.88	19.18	11.80	13.08
f_5	平均值	4.45×10^{-16}	2.80	19.80	19.83	10.19
	标准差	0	9.56	1.56	0.31	0.79
f_6	平均值	0	1.58	480.25	155.48	16.43
	标准差	0	2.69	34.25	62.47	5.43

算法有更高的寻优精度和更快的迭代速度。综合来看, IWOA 算法在求解优化问题时具有较为明显的优势。

2.3 评价指标

采用平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Squared Error, RMSE) 以及平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 作为预测算法的评价指标。MAE、MAPE、RMSE 值越小,表示模型预测效果越好。各指标的计算公式分别为:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (17)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (18)$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (19)$$

式中: y_i 表示第 i 时刻的真实值; $\hat{y}_i$ 表示第 i 时刻的预测值; N 表示样本总数量。

2.4 比较实验

为验证所提模型的优越性,选取新疆某地区 2019 年全年的风电场数据作为数据集 Dataset₁, 国华景峡北风电一场实测数据作为数据集 Dataset₂ 进行超短期风电功率预

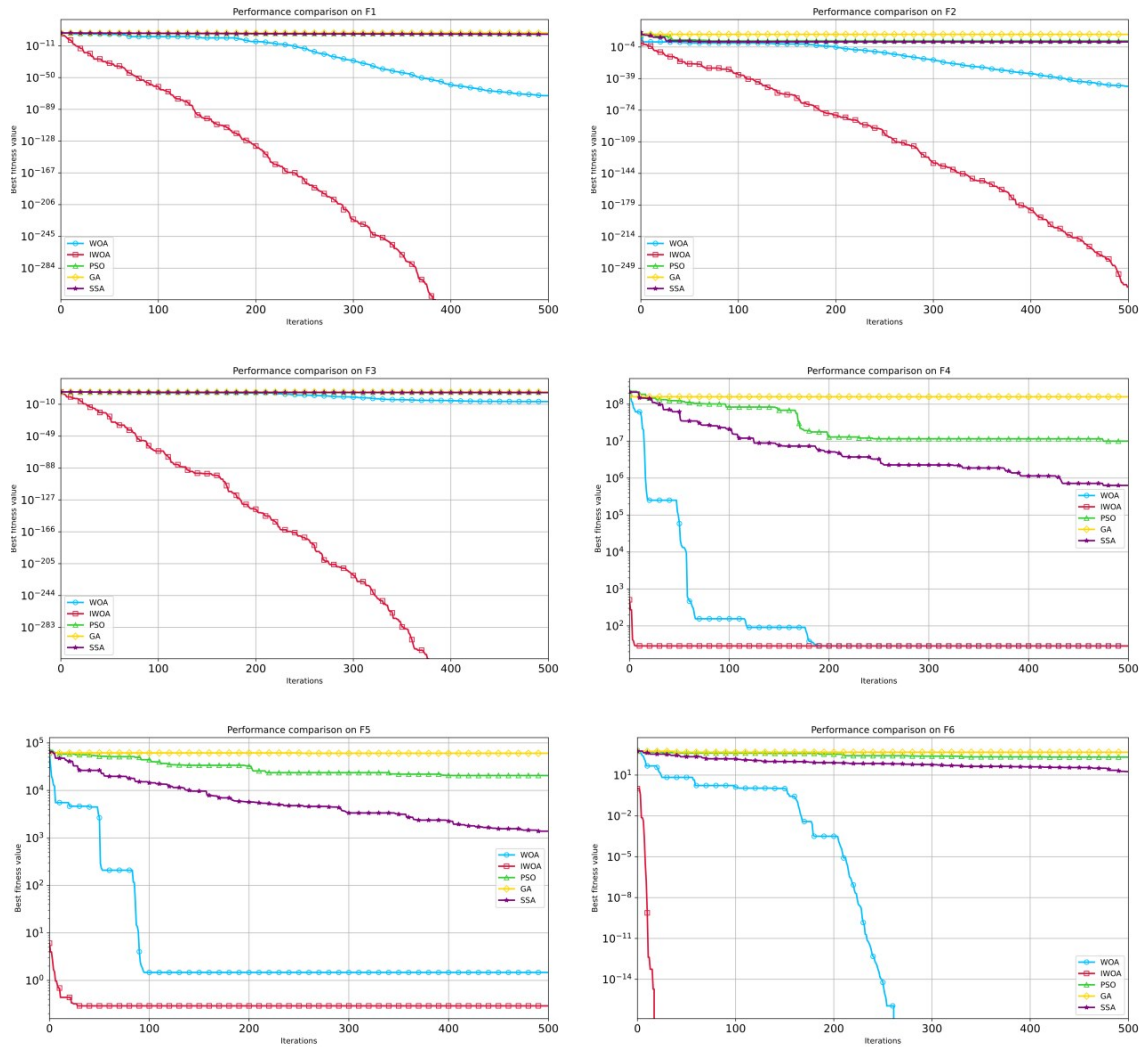


Fig. 5 Comparison of convergence curves of various algorithms

图5 各算法收敛曲线比较

测,选择 BiLSTM、BP(Back Propagation)、LSTM、RNN(Recurrent Neural Network)、GRU(Gated Recurrent Unit)作为对照模型。将两个数据集按照 8:2 的比例划分为训练集和测试集。为避免偶然性,各预测模型分别独立运行 30 次,并取平均值作为最终评价标准,实验结果均保留小数点后 3 位。各模型预测结果比较(某一天)如图 6 所示,预测评价指标比较如表 3 所示。

可以看出,在 Dataset_1 中,IWOA-BiLSTM-MHSA 模型的 RMSE 相较 LSTM、BP、RNN、GRU、BiLSTM 分别降低了 8.354、13.119、12.685、4.549、3.287,MAE 分别降低了 10.559、13.684、6.769、5.063、4.925,MAPE 分别降低了 4.892%、7.722%、6.196%、3.864%、2.159%;在 Dataset_2 中,IWOA-BiLSTM-MHSA 模型的 RMSE 相较 LSTM、BP、RNN、GRU、BiLSTM 分别降低了 6.965、20.336、8.939、3.884、2.347,MAE 相较 LSTM、BP、RNN、BiLSTM 分别降低了 3.658、14.013、1.633、0.989,但相较 GRU 增加了 0.468,MAPE 分别降低了 2.524%、17.899%、5.000%、1.545%、2.269%。综合来看,本文模型表现最佳。

2.5 消融实验

为验证所提模型各模块的有效性,进行以下消融实验。设置 3 种比较模型,分别为 BiLSTM、BiLSTM-MHSA、WOA-BiLSTM-MHSA,模型基础参数如表 4 所示。使用 IWOA 算法对 BiLSTM-MHSA 模型进行超参数优化,得出最优参数为第一层 BiLSTM 神经元个数为 217,第二层 BiLSTM 神经元个数为 256,MHSA 机制的参数为 num_heads=7, key_dim=10, value_dim=9。

各消融模型在两个数据集上的预测结果比较如图 7 所示,预测评价指标如表 5 所示。可以看出,IWOA 优化后的 BiLSTM-MHSA 模型在捕捉功率变化趋势方面表现更佳,预测曲线更为平滑。与 BiLSTM、BiLSTM-MHSA、WOA-BiLSTM-MHSA 模型相比,本文模型在 Dataset_1 中的 RMSE 分别降低了 3.287、2.030、1.483,MAE 分别降低了 4.925、1.603、1.436,MAPE 分别降低了 2.161%、1.306%、1.155%;在 Dataset_2 中的 RMSE 分别降低了 2.347、1.008、1.149,MAE 分别降低了 0.989、1.025、0.631,MAPE 分别降低了 2.269%、0.671%、1.008%。

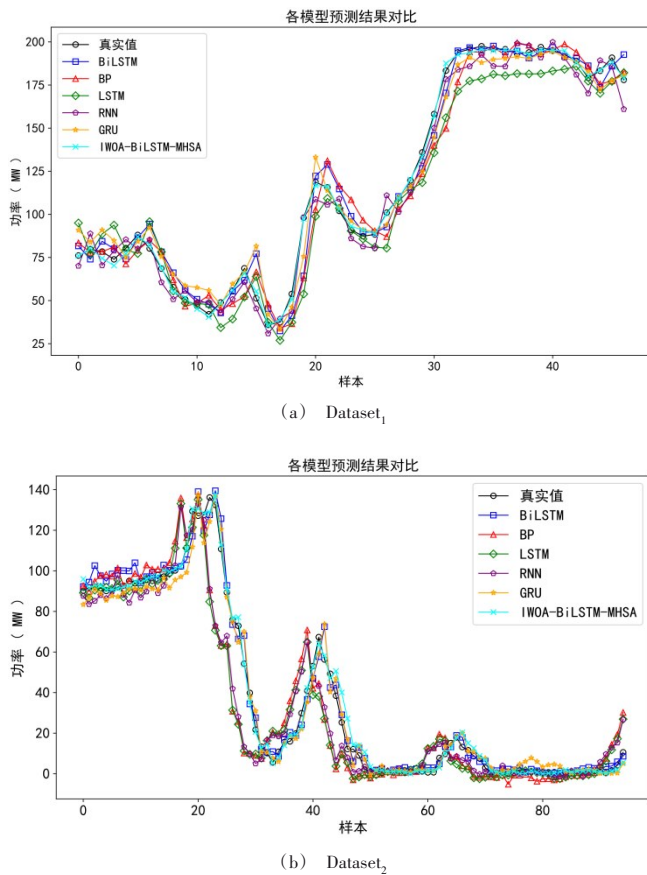


Fig. 6 Comparison of prediction results of various models

图6 各模型预测结果比较

Table 3 Comparison of predictive evaluation indicators among various models

表3 各模型预测评价指标比较

数据集	模型	RMSE/MW	MAE/MW	MAPE/%
Dataset ₁	LSTM	15.303	15.138	7.744
	BP	20.068	18.263	10.574
	RNN	19.634	11.348	9.048
	GRU	11.498	9.642	6.716
	BiLSTM	10.236	9.504	5.011
	IWOA-BiLSTM-MHSA	6.949	4.579	2.852
Dataset ₂	LSTM	15.975	10.358	3.017
	BP	29.346	20.713	18.392
	RNN	17.949	8.333	5.493
	GRU	12.894	6.232	2.038
	BiLSTM	11.357	7.689	2.762
	IWOA-BiLSTM-MHSA	9.010	6.700	0.493

3 结语

针对风力发电的波动性与不确定性导致难以捕捉风电数据中的时序关系这一问题,本文提出IWOA-BiLSTM-MHSA模型,实现了准确的超短期风力发电功率预测。其中,BiLSTM能有效捕捉时间序列数据中的前后依赖关系,提高模型对时间序列特征的学习能力;MHSA机制能使模型更好地捕捉时间序列中的复杂模式与依赖关系;IWOA

Table 4 Model basic parameters

表4 模型基础参数

参数	实验设置
BiLSTM1 神经元	128
BiLSTM2 神经元	64
多头自注意力机制(num_heads, key_dim, value_dim)	7, 8, 6
Dense 全连接层神经元	1
学习率	0.01
激活函数	relu, tanh, sigmoid
Epoch	300
Batch_size	512

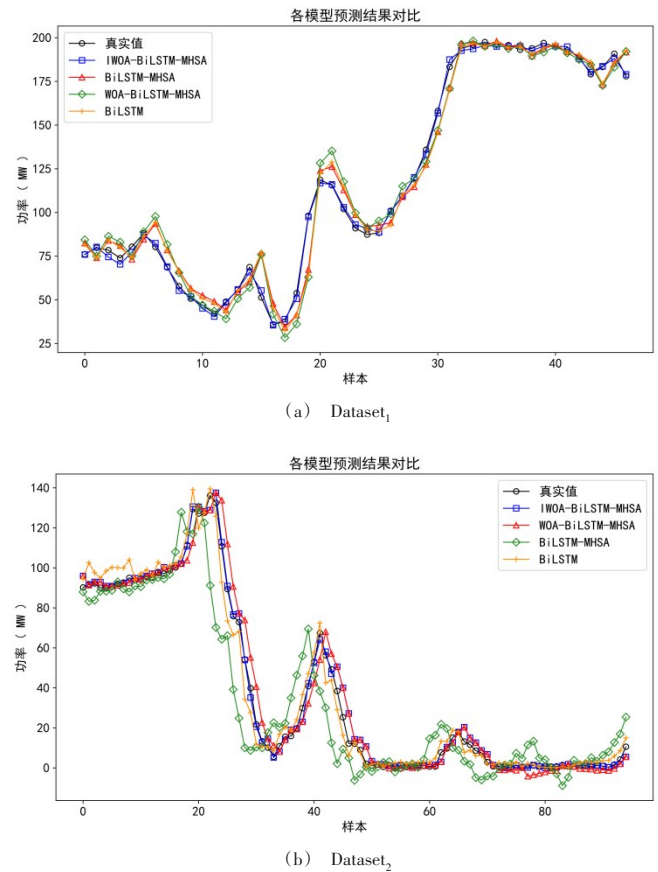


Fig. 7 Comparison of prediction results of various ablation models

图7 各消融模型预测结果比较

Table 5 Comparison of predictive evaluation indicators among various ablation models

表5 各消融模型预测评价指标比较

数据集	模型	RMSE/MW	MAE/MW	MAPE/%
Dataset ₁	BiLSTM	10.236	9.504	5.011
	BiLSTM-MHSA	8.979	6.182	4.158
	WOA-BiLSTM-MHSA	8.432	6.015	4.007
	IWOA-BiLSTM-MHSA	6.949	4.579	2.852
Dataset ₂	BiLSTM	11.357	7.689	2.762
	BiLSTM-MHSA	10.018	7.725	1.164
	WOA-BiLSTM-MHSA	10.159	7.331	1.501
	IWOA-BiLSTM-MHSA	9.010	6.700	0.493

引入拉丁超立方种群初始化、自适应参数以及自适应阈值策略,提高了算法的全局寻优能力及收敛速度。在真实数据集上的实验结果表明,该模型的预测精度相较多种传统

机器学习和深度学习预测模型有一定提高。后续拟将研究重心转向基于实时数据的调度模型,利用正在发生的天气数据和各机组的实时功率数据等,使模型更贴近实际应用场景,从而提高预测的实用性和调度的准确性。

参考文献:

- [1] ZHAO W, WEI Y M, SU Z. One day ahead wind speed forecasting: a resampling-based approach[J]. *Applied Energy*, 2016, 178: 886-901.
- [2] SHAHID F, ZAMEER A, MUNEEB M. A novel genetic LSTM model for wind power forecast[J]. *Energy*, 2021, 223: 120069.
- [3] HUANG Y H, SUN Y N, LI C, et al. Constructing method of short-term output scenarios for multi-regional wind power based on conditional generative adversarial network[J]. *Power System Technology*, 2023, 47(1): 63-77.
黄越辉,孙亚南,李驰,等. 基于条件生成对抗网络的多区域风电短期出力场景生成方法[J]. *电网技术*, 2023, 47(1): 63-77.
- [4] LI C Y, CHEN X, ZHANG P, et al. Multi-time-scale demand response dispatch considering wind power forecast error[J]. *Power System Technology*, 2018, 42(2): 487-494.
李春燕,陈骁,张鹏,等. 计及风电功率预测误差的需求响应多时间尺度优化调度[J]. *电网技术*, 2018, 42(2): 487-494.
- [5] SHI Z, LIANG H, DINAHAHI V. Direct interval forecast of uncertain wind power based on recurrent neural networks[J]. *IEEE Transactions on Sustainable Energy*, 2017, 9(3): 1177-1187.
- [6] ZHAO X, LIU J, YU D, et al. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data[J]. *Energy Conversion and Management*, 2018, 164: 560-569.
- [7] ZHANG D, CHEN B, ZHU H, et al. Short-term wind power prediction based on two-layer decomposition and BiTCN-BiLSTM-attention model[J]. *Applied Energy*, 2023, 285: 128762.
- [8] HUANG N, WU Y, CAI G, et al. Short-term wind speed forecast with low loss of information based on feature generation of OSVD[J]. *IEEE Access*, 2019, 7: 81027-81046.
- [9] MENG J J, JIANG X J, LI D C, et al. Wind speed prediction model along railway based on VMD-LSTM-WOA[J]. *Transducer and Microsystem Technologies*, 2023, 42(4): 152-156.
孟建军,江相君,李德仓,等. 基于VMD-LSTM-WOA的铁路沿线风速预测模型[J]. *传感器与微系统*, 2023, 42(4): 152-156.
- [10] ZHU Q M, LI H Y, WANG Z Q, et al. Short-term wind power forecasting based on LSTM[J]. *Power System Technology*, 2017, 41(12): 3797-3802.
朱乔木,李弘毅,王子琪,等. 基于长短期记忆网络的风电场发电功率超短期预测[J]. *电网技术*, 2017, 41(12): 3797-3802.
- [11] LIU H, MI X, LI Y. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network[J]. *Energy Conversion and Management*, 2018, 156: 498-514.
- [12] CAO Y W, YAN S H, LIU H T, et al. Short-term wind power forecasting method based on noise-reduction time-series deep learning network[J]. *Proceedings of the CSU-EPSA*, 2020, 32(1): 145-150.
曹有为,闫双红,刘海涛,等. 基于降噪时序深度学习网络的风电功率短期预测方法[J]. *电力系统及其自动化学报*, 2020, 32(1): 145-150.
- [13] GRAVES A, SCHMIDHUBER J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures[J]. *Neural Networks*, 2005, 18(5-6): 602-610.
- [14] YU M, NIU D, WANG K, et al. Short-term photovoltaic power point-interval forecasting based on double-layer decomposition and WOA-BiLSTM-Attention and considering weather classification[J]. *Energy*, 2023, 275: 127348.
- [15] JIANG X C, XU Y C, SONG C. A new method to predict short-term load of wind power[J]. *Journal of Beijing Normal University (Natural Science)*, 2022, 58(1): 39-46.
姜旭初,许宇澄,宋超. 短期风力发电负荷预测的新方法[J]. *北京师范大学学报(自然科学版)*, 2022, 58(1): 39-46.
- [16] LIU A G, XUE Y T, HU J L, et al. Ultra-short-term wind power forecasting based on SVM optimized by GA[J]. *Power System Protection and Control*, 2015, 43(2): 90-95.
刘爱国,薛云涛,胡江鹭,等. 基于GA优化SVM的风电功率的超短期预测[J]. *电力系统保护与控制*, 2015, 43(2): 90-95.
- [17] YUAN X, CHEN C, YUAN Y, et al. Short-term wind power prediction based on LSSVM-GSA model[J]. *Energy Conversion and Management*, 2015, 101: 393-401.
- [18] CHEN G, TANG B, ZENG X, et al. Short-term wind speed forecasting based on long short-term memory and improved BP neural network[J]. *International Journal of Electrical Power & Energy Systems*, 2022, 134: 107365.
- [19] LIU Y, GUAN L, HOU C, et al. Wind power short-term prediction based on LSTM and discrete wavelet transform[J]. *Applied Sciences*, 2019, 9(6): 1108.
- [20] WANG Y, ZOU R, LIU F, et al. A review of wind speed and wind power forecasting with deep neural networks[J]. *Applied Energy*, 2021, 304: 117766.
- [21] YANG Z M, PENG X S, LANG J X, et al. Short-term wind power prediction based on dynamic cluster division and BiLSTM deep learning method[J]. *High Voltage Engineering*, 2021, 47(4): 1195-1203.
杨子民,彭小圣,郎建勋,等. 基于集群动态划分与BiLSTM深度学习的风电集群短期功率预测[J]. *高电压技术*, 2021, 47(4): 1195-1203.
- [22] JASEENA K U, KOVOOR B C. Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks[J]. *Energy Conversion and Management*, 2021, 234: 113944.
- [23] ABBASIMEHR H, PAKI R. Improving time series forecasting using LSTM and attention models[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2022, 13: 673-691.
- [24] ZHANG B, WANG S, DENG L, et al. Ship motion attitude prediction model based on IWOA-TCN-Attention[J]. *Ocean Engineering*, 2023, 272: 113911.

(责任编辑:尹晨茹)