

基于 WOA-TCN 的太阳辐照度预测算法



牛怀广, 张鑫洋*, 尹龙

天津商业大学信息工程学院, 天津 300134

摘要: 太阳辐照度预测在多能协同供热系统的优化运行中扮演着关键角色。太阳辐照度作为影响供热系统效率的关键因素, 其预测的准确性直接关系到供热系统的运行效率和能源消耗。然而, 太阳辐照度具有时空间上的间歇性和非线性等不确定特性给现有的机器学习预测算法带来了挑战, 特别是传统超参数优化策略较为简单使其难以满足优化调控需求。针对该问题, 本文提出了一种基于鲸鱼算法优化时间卷积网络 (Whale Algorithm Optimization Temporal Convolutional Network, WOA-TCN) 的太阳辐照度预测模型。该模型充分利用了 TCN 时间卷积网络的并行处理能力与递归神经网络的时序建模功能, 通过模拟鲸鱼行为的群智能优化算法来优化 TCN 的超参数, 从而提高了模型的预测精度, 使得太阳辐照度的时空演变规律能够更准确地捕捉, 还能多能协同供热系统的优化运行提供有力的技术支持。并通过实验对比了 CNN、RNN 等算法的预测指标, 结果表明了基于 WOA-TCN 太阳辐照度预测模型的预测精度更高。

关键词: 太阳辐照度预测; 鲸鱼优化算法; 时间卷积网络; 深度学习

DOI: [10.57237/j.jest.2024.04.001](https://doi.org/10.57237/j.jest.2024.04.001)

Solar Irradiance Prediction Algorithm Based on WOA-TCN

Niu Huaiguang, Zhang Xinyang*, Yin Long

School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, China

Abstract: The prediction of solar irradiance plays a crucial role in the optimization operation of multi energy collaborative heating systems. As a key factor affecting the efficiency of the heating system, the accuracy of solar irradiance prediction is directly related to the operational efficiency and energy consumption of the heating system. However, solar irradiance has uncertain characteristics such as intermittency and nonlinearity in time and space, which poses challenges for existing machine learning prediction algorithms in hyperparameter optimization and makes it difficult to meet optimization and regulation requirements. To address this issue, we propose a solar irradiance prediction model based on Whale Algorithm Optimization Time Convolutional Network (WOA-TCN). This model fully utilizes the parallel processing capability of TCN time convolutional networks and the temporal modeling function of recurrent neural networks. The hyperparameters of the TCN are optimized by a swarm intelligence optimization algorithm that simulates the behavior of whales, which improves the prediction accuracy of the model, enables the temporal and spatial evolution patterns of solar irradiance to be captured more

基金项目: 大学生创新创业训练计划资助项目 (202310069029).

*通信作者: 张鑫洋, crozz26235120@163.com

收稿日期: 2024-06-21; 接受日期: 2024-07-13; 在线出版日期: 2024-07-18

<http://www.energysci.tech.org>

accurately, and also provides strong technical support for the optimal operation of the multi-energy cooperative heating system. The prediction indexes of CNN, RNN and other algorithms are also compared through experiments, and the results show that the prediction accuracy of the solar irradiance prediction model based on WOA-TCN is higher.

Keywords: Solar Irradiance Prediction; Whale Optimization Algorithm; Time Convolutional Network; Deep Learning

1 引言

随着全球气候变化和能源危机的加剧, 太阳辐射预测技术的发展对于实现可持续发展和减少化石能源的依赖具有举足轻重的作用[1]。太阳辐射预测算法作为一个重要的研究方向, 受到了越来越多学者的重视[2-4]。但由于太阳辐射受到天气、气候和地理位置等因素的影响, 具有很强的时空变异性[5]。此外, 模型和算法的局限性也影响着太阳辐射的预测精度。因此, 提高模型算法的预测精度与泛化能力对于太阳辐射预测系统起到至关重要的作用。

目前许多预测太阳辐射的方法都存在一些弊端。基于物理模型和统计模型的传统方法容易受外界因素和非线性特征值的影响[6]。而基于深度学习和机器学习的人工智能方法也会受其网络的局限性影响[7, 8]。

国内外的学者都对太阳辐照度的预测做了努力, 臧海洋[9]等人为提升短期太阳辐射预测的准确性, 提出一种基于 ICEEMDAN-LSTM 和残差注意力的短期太阳辐照度预测方法。该方法可有效提高短期太阳辐照度的预测精度。鹿晨东[10]等人提出了一种改良的深度混合模型, 即自编码器条件深度卷积生成对抗网络模型。研究表明该模型在提高太阳辐照度预测精度方面具有显著效果。官松泽[11]等人探讨了大型光伏发电并网对电力系统稳定性产生的影响, 并提出了一种新型的太阳辐照度超短期预测方案。结果显示, 相较于现有常用模型, 所提出的预测模型具有更高的预测精度。刘军[12]等人提出了一种新颖的太阳辐照度逐小时逐步预测框架, 该框架采用 CNN-LSTM-MLP 集成混合模型, 并结合误差校正和变分模式分解(VMD)。实验证明, 该模型在准确性和鲁棒性方面优于众多传统替代方法。Jalali [13]等人介绍了一种基于深度学习的多步超前方法, 实验结果表明, 该方法在预测性能方面具有优越性。顾琼[14]等人提出了一种历史适应差分进化与线性群体规模缩减(L-SHADE)和分解的方法, 以准确可靠地识别光伏模型在不同温度和辐照度下的未知参数。对比实验结果表明, 所提出的

L-SHADED 的均方根误差几乎比其他方法小 10 倍。

由于太阳辐照度时间序列具有非线性和非平稳性, 当前部分研究在预测精度和模型泛化能力方面尚不尽如人意。因此, 本文提出了一种基于鲸鱼优化算法和时间卷积网络的太阳辐射预测模型: WOA-TCN。该模型在预测太阳辐射时间序列问题中具备适应长期依赖关系、多尺度特征提取和正则化能力等优势能够更精准的预测太阳辐射。鲸鱼优化算法能更好地找出问题的最优解从而提高模型的鲁棒性和适应性。通过对该算法模型的设计与优化, 能够提高系统的运行效率和可靠性。

2 数据及目的

2.1 数据收集

国家太阳辐射数据库(the National Solar Radiation Database, NSRDB)为该研究提供了全球各个地区和各年份逐时的太阳辐射数据与气象数据。国家太阳辐射数据库是每小时气象数据值以及太阳辐射测量值: 全球水平辐照度 GHI (Global Horizontal Irradiance, GHI, W/m^2) 以及温度 ($^{\circ}\text{C}$), 相对湿度 (%) 等指标。以天津站点 (117.1951073E, 39.0856735N) 为例, 研究所使用的各种气象数据数据采集频率以每 60 分钟进行划分, 时间范围从 2020 年 1 月至 2020 年 12 月, 共计 8760 小时。

2.2 问题定义及目的

本文的研究目的是通过过去逐时的太阳辐射和气象数据, 对未来多个小时太阳辐射进行预测。通过输入过去 j 个时间步长的数据, 从而预测出未来每小时的太阳辐射。将输入特征表示为一个时间序列:

$$X = (x_1, x_2, K, x_t) \quad (1)$$

其中 x_t 是时间步 t 的输入特征向量。希望预测的目

标序列为:

$$Y = (y_{j+1}, y_{j+2}, \dots, y_{j+h}) \quad (2)$$

其中, h 表示预测的时间步数。

3 算法原理

3.1 数据预处理

数据预处理在数据分析和机器学习项目中扮演着至关重要的角色。一个高质量的数据预处理过程能够提高模型的性能,减少错误和偏见,从而使得分析结果更加可靠和准确。

为了提高模型的性能和泛化能力,我们需要进行特征选择来找到与拟合目标相关的特征,并降低输入数据的维度。特征选择的目的是避免模型过拟合,同时又保证模型的适用性[15]。

斯皮尔曼相关性分析 (Spearman's Rank Correlation) 是秩相关 (rank correlation) 的一种非参数度量 (nonparametric measure), 通常记为希腊字母 r , 用于评估两个变量之间的单调关系。

斯皮尔曼相关性分析是通过比较变量的等级而不是原始值来衡量相关性, 一个数的等级, 是指将其置于一系列数值中, 并按照数值大小进行升序排列后, 该数值所处的顺序位置。如果在排序过程中遇到数值相同的情况, 则该数值的等级将取这些相同数值所在位置的平均值。因此特别适用于非正态分布数据或数据中存在异常值的情况

$$r_{x,y} = 1 - \frac{6 \sum_{i=1}^n (x_i - y_i)^2}{n(n^2 - 1)} \quad (3)$$

其中, x_i 为输入第 i 个气象数据的等级, y_i 为输入第 i 个太阳辐射度的等级; $x_i - y_i$ 为两者之间的等级差; 数据 n 为数据点的数量。相关系数越大其相关性越强, 越小相关性越弱, 如公式(4)所示。

$$\begin{cases} 0.8 \leq |r_{x,y}| \leq 1 & \text{极强相关} \\ 0.6 \leq |r_{x,y}| < 0.8 & \text{强相关} \\ 0.4 \leq |r_{x,y}| < 0.6 & \text{中度相关} \\ 0.2 \leq |r_{x,y}| < 0.4 & \text{弱相关} \\ |r_{x,y}| < 0.2 & \text{无相关} \end{cases} \quad (4)$$

3.2 时间卷积网络

TCN 是一种用于序列建模的网络结构, 其主要包含膨胀卷积 (Dilated Convolution)、因果卷积 (Causal Convolution) 及残差连接 (Residual Connections) 三个核心部件。

膨胀卷积 (Dilated Convolution) 是一种特殊的卷积操作, 通过在卷积核的元素之间插入空洞 (即扩展步长), 从而增大感受野。与传统的卷积不同, 膨胀卷积可以在不增加计算复杂度的情况下扩展感受野, 从而能够捕捉更长时间范围内的依赖关系。对于一个一维序列输入 x , 膨胀卷积的输出 $y(i)$ 可以表示为:

$$y(i) = \sum_{k=0}^{k-1} x(i + r \cdot k) \cdot w_k \quad (5)$$

其中, k 是卷积核的大小, i 为网络层数, r 是膨胀因子, w_k 是卷积核的权重。

因果卷积 (Causal Convolution) 确保了卷积操作中, 输出时间步 t 的结果只依赖于当前及之前的输入, 避免了信息泄漏。这种设计对于时间序列预测等任务尤为重要, 因为它确保了预测模型不会使用未来的信息。

残差连接 (Residual Connections) 是一种在深层神经网络中常用的技术, 通过引入跨层的直接连接来缓解梯度消失问题, 促进更深层网络的训练。

残差连接在卷积层之间引入跳跃连接, 使得每一层的输出不仅仅是卷积操作的结果, 还加上了输入的直接路径。

TCN 结合了膨胀卷积、因果卷积及残差连接的优势, 能够有效地建模时间序列数据中的长距离依赖关系, 保持因果性, 并缓解深层网络的训练困难。这种设计使 TCN 在时间序列预测、信号处理及其他需要长时间依赖的任务中表现出色。

3.3 鲸鱼优化算法

鲸鱼优化算法 (Whale Optimization Algorithm, WOA) 是由 Mirjalili 和 Lewis 在 2016 年开发的一种新颖的群体智能优化技术[16]。

该算法通过模拟鲸鱼群体的搜索行为来解决优化问题, 其灵感源自对自然界中座头鲸群体狩猎行为。在 WOA 算法中, 每只座头鲸的位置代表一个可行解, 整个算法过程通过模拟鲸鱼的泡泡网捕食行为包含搜索觅食、收缩包围和螺旋更新位置三个阶段。

WOA 算法将鲸鱼捕食的目标视为潜在的最优解, 而鲸鱼的当前位置则被视为可能的解。在每次迭代过程中, 算法根据随机数 p 的值和系数向量 A 的大小来确定鲸鱼位置的调整方式, 这些调整方式包括三种不同的更新机制。随着迭代次数的增加, 鲸鱼群体逐渐向最优解收敛。WOA 算法的 3 种种群更新机制中, 搜索觅食机制的数学定义为:

$$X(t+1) = X_r(t) - A \cdot |C \cdot X_r(t) - X(t)| \quad (6)$$

式中 $X_r(t)$ 是从当前鲸鱼群体中随机选取的一个鲸鱼个体位置向量; $X(t)$ 是当前的鲸鱼个体位置向量, $|C \cdot X_r(t) - X(t)|$ 表示当前鲸鱼个体和随机选取鲸鱼个体之间的距离向量。系数向量 A 和 C 的定义为:

$$A = 2r \cdot a - a \quad (7)$$

$$C = 2r \quad (8)$$

在此式中控制参数向量 a 随着迭代次数的增加从 2 线性减小到 0; r 是一个分布于 $[0,1]$ 之间的随机向量。

收缩包围机制的数学模型定义如下:

$$X(t+1) = X_b(t) - A \cdot |C \cdot X_b(t) - X(t)| \quad (9)$$

式中 $X_b(t)$ 是当前鲸鱼群体中目标函数值最优的鲸鱼个体位置向量; $A \cdot |C \cdot X_b(t) - X(t)|$ 为当前鲸鱼个体的包围步长; A 越小时鲸鱼游走的步长越小。

螺旋更新位置阶段, 其数学模型可表示为:

$$X(t+1) = |X_b(t) - X(t)| \cdot e^{bl} \cdot \cos(2\pi l) + X_b(t) \quad (10)$$

式中 $|X_b(t) - X(t)|$ 表示当前鲸鱼个体与最佳位置鲸鱼之间的距离向量; 而常量系数 b 决定了鲸鱼在螺旋式前进时螺旋路径的形状, 当其值为 1 时, 路径就变成了标准的对数螺旋; l 是 $[-1,1]$ 之间的随机数, WOA 算法中鲸鱼个体位置向量 $X(t)$ 的每一个分量在种群更新时均独立进行。

综上所述, 通过将鲸鱼优化算法应用于 TCN 的超参数优化中, 可以利用其强大的全局搜索能力和自适应调节的特性, 更好地调整 TCN 模型的参数, 提高模型的预测性能和泛化能力, 从而使其更适应各种实际

任务和数据集。

4 基于 WOA-TCN 太阳辐照度预测模型

4.1 模型原理

TCN 需要优化的五个超参数分别为: 正则化参数、学习率、滤波器个数、滤波器大小和区块数。利用 WOA 算法对 TCN 超参数进行优化。每只鲸鱼的位置由一个 5 维向量表示, 其维数等于待优化的超参数个数。鲸鱼适应度值的计算方法如下: 以鲸鱼的位置参数作为模型的超参数, 模型建立后进行一次训练, 此时模型的识别率即为鲸鱼在当前位置的适应度。然后, 让每只鲸鱼根据自己的位置调整策略移动到新位置。上述过程称为一次搜索。设鲸鱼数量为 30 只, 搜索次数为 30 次。所有搜索结束后, 将选出适应度最高的鲸鱼, 其位置参数代表最优超参数。基于上述所选出的最优超参数构建 TCN 模型, 经过 50 次迭代最终得到训练好的模型。该算法流程图如图 1 所示。其优化和训练过程可分为以下五步:

- 1) 将数据集按照 8:2 的比例分为训练集与测试集。
- 2) 将种群初始化为 30 只鲸鱼。根据适应度计算公式评估每只鲸鱼的适应度值。
- 3) 鲸鱼根据围捕行为、泡泡网攻击和探索行为调整位置:
 - a) 围捕行为: 鲸鱼围绕猎物移动, 模拟捕食过程。
 - b) 泡泡网攻击: 鲸鱼通过创建泡泡网调整位置, 以更接近猎物。
 - c) 探索行为: 随机探索新位置, 避免陷入局部最优。
- 4) 用最优超参数构建基于 TCN 的预测模型。
- 5) 训练后利用测试集进行模型评估。

通过上述步骤, WOA 能够有效优化时序卷积网络的超参数。通过 WOA 优化的超参数, TCN 模型在太阳辐照度预测中可以表现出更高的准确性和稳定性。

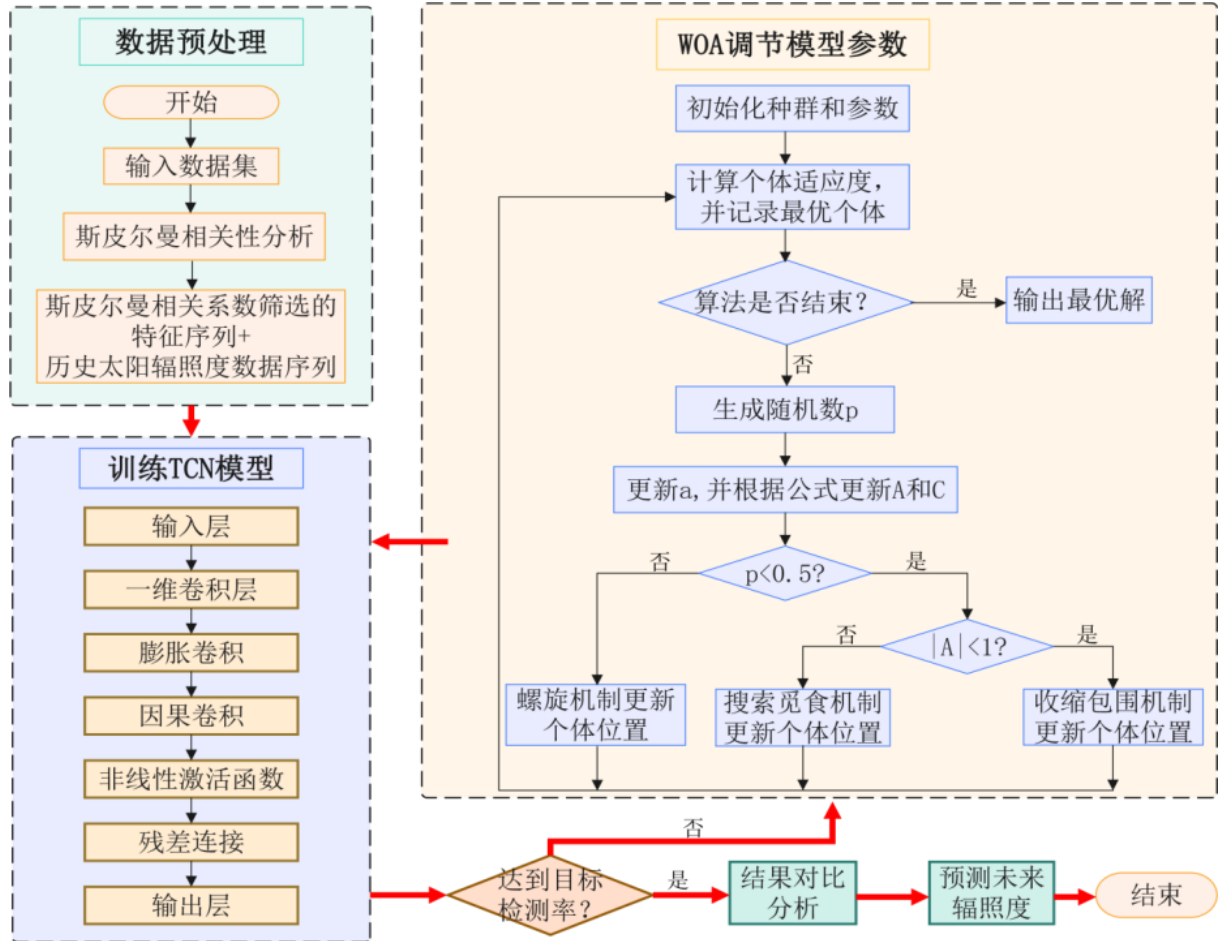


图 1 WOA-TCN 算法流程图

4.2 评价指标

为了全面而准确地评估各算法预测模型的性能表现, 本文采纳了一系列经典的评估指标。这些指标涵盖了均方误差 (Mean Square Error, MSE)、均方根误差 (Root Mean Square Error, RMSE)、平均绝对误差 (Mean Absolute Error, MAE) 和 (决定系数)。这些评价指标的数学公式如下:

$$E_{MSE} = \frac{1}{n} \sum_{t=1}^n (\hat{y}_j - y_j)^2 \quad (11)$$

$$E_{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_j - y_j)^2} \quad (12)$$

$$E_{MAE} = \frac{1}{n} \times \sum_{t=1}^n |\hat{y}_j - y_j| \quad (13)$$

$$E_{R^2} = 1 - \frac{\sum_{t=1}^n (y_j - \hat{y}_j)^2}{\sum_{t=1}^n (y_j - \bar{y})^2} \quad (14)$$

式中: y_j 表示太阳辐照度实测值; $\hat{y}_j$ 表示太阳辐照度预测值; $\bar{y}$ 表示太阳辐照度的样本均值; n 表示太阳辐照度的样本数。

5 实验和讨论

5.1 特征选择结果

图 2 展示了太阳辐照度与其他气象参数之间的相关性热图。从图中可以看出, 太阳辐照度与温度、相对湿度以及太阳天顶角的相关系数绝对值较高, 显示出显著的相关性。由此, 这些参数被确定为太阳辐照度预测模型的输入特征。

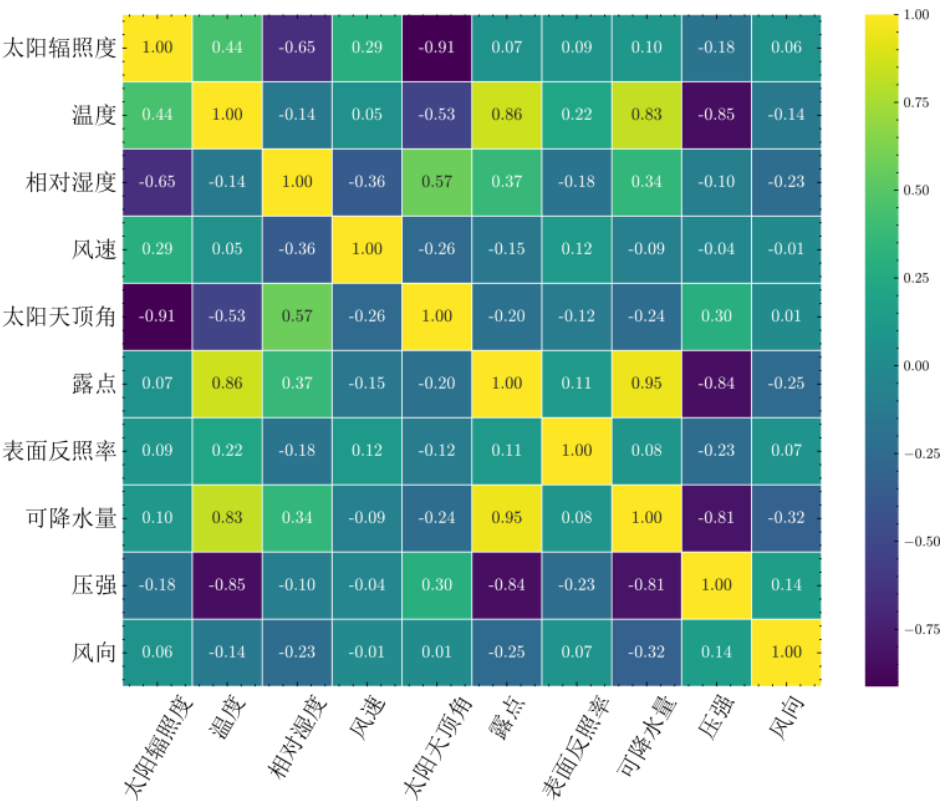


图 2 斯皮尔曼特征相关性热力图

5.2 WOA 超参数优化结果

为了提高预测结果的准确性，本文采用了鲸鱼优化算法对时序卷积网络模型超参数进行了寻优。WOA 更加高效地搜索模型参数空间，进一步提升 TCN 模型在预测任务中的性能。超参数寻优结果如表 1 所示。通过实验验证和性能评估，本文验证了 WOA 算法在 TCN 模型改进中的有效性和优越性。与传统的优化方法相比，融合了 WOA 算法的改进策略能够显著提升模型的预测性能。

表 1 WOA 对 TCN 超参数的寻优结果

参数名称	参数取值范围	参数值
正则化参数	[0.001, 0.1]	0.0914
学习率	[0.0001, 0.01]	0.0012
滤波器个数	[20, 100]	97
滤波器大小	[2, 10]	6
区块数	[2, 10]	2

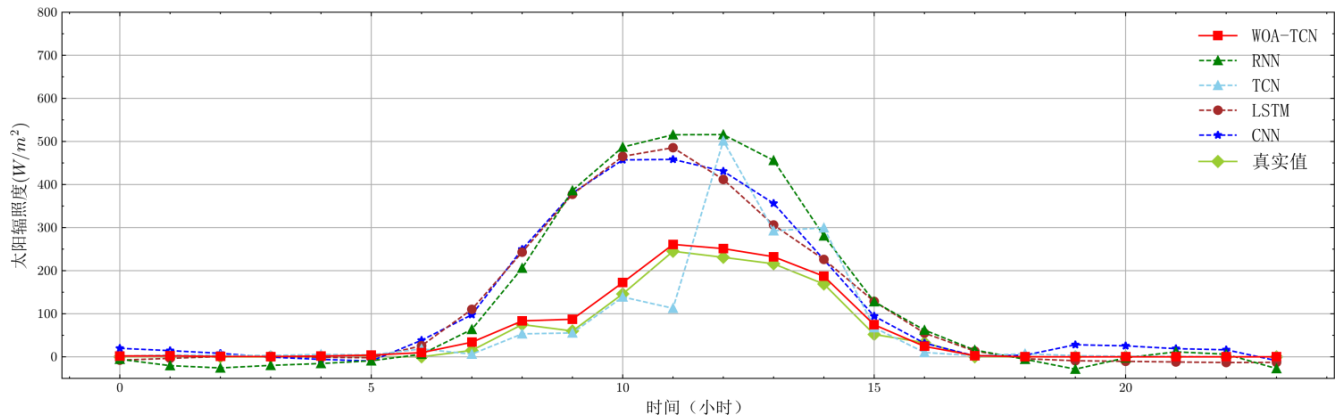
5.3 预测性能对比分析

为了证明 WOA-TCN 模型在太阳辐照度预测方面

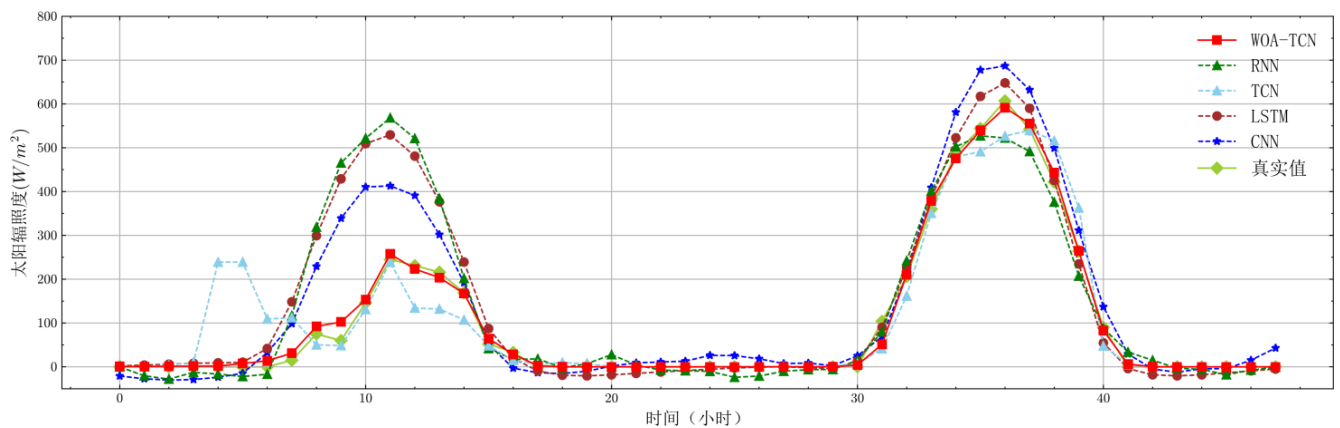
的优势，本研究将其与 RNN、TCN、LSTM 和 CNN 四种算法进行了深入的实验比较。图 3 展示了预测未来 24 小时、48 小时和 72 小时太阳辐照度的气象数据结果。图 3 显示出 WOA-TCN 模型在预测太阳辐照度方面具有最高的拟合精度。表 2 列出了 WOA-TCN 模型与其它算法在性能指标上的对比，包括均方误差（MSE）、均方根误差(RMSE)、平均绝对误差(MAE)和决定系数（R²）。WOA-TCN 在这些指标上均表现出最优的拟合性能，这表明该模型能够更有效地在解空间中寻找 TCN 超参数的最佳值，并高效地提取太阳辐照度的复杂时空特征。

表 2 WOA-TCN 预测模型与其它算法的性能统计指标对比

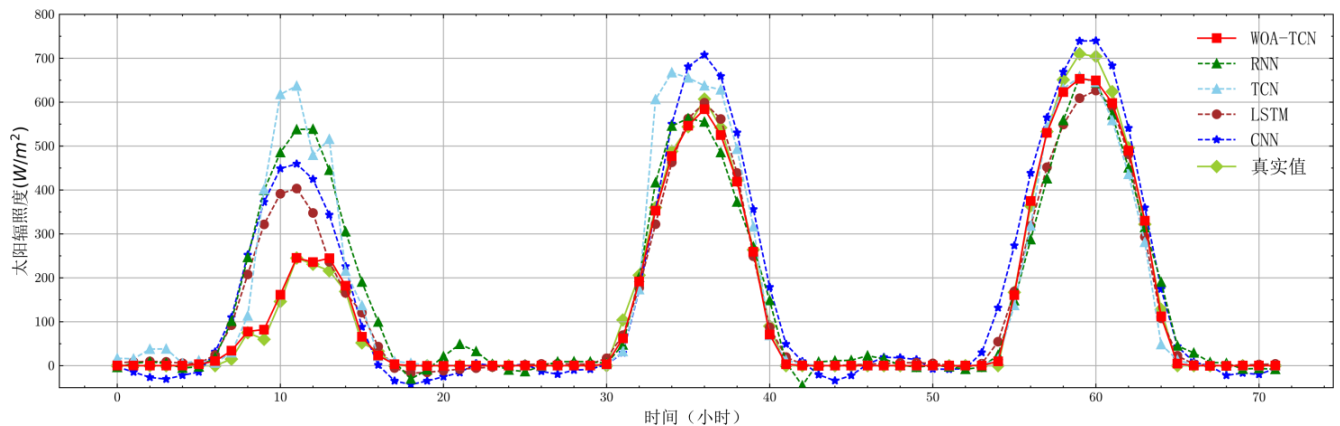
预测模型	MSE (W/m ²)	RMSE (W/m ²)	MAE (W/m ²)	R ²
WOA-TCN	1.97×10 ²	14.06	7.36	0.990
RNN	8.59×10 ³	92.70	50.56	0.807
TCN	1.12×10 ⁴	37.81	48.89	0.807
LSTM	3.41×10 ³	58.14	29.59	0.923
CNN	6.71×10 ³	81.89	51.47	0.850



a. 预测时间步长为 24 小时的太阳辐射度



b. 预测时间步长为 48 小时的太阳辐射度



c. 预测时间步长为 72 小时的太阳辐射度

图 3 不同预测时长下 WOA-TCN 模型与其它算法的预测结果曲线对比图

6 结论

太阳辐射度的准确预测对于提高清洁能源的使用效率和助力国家实现“碳中和”目标具有关键意义。然而，由于太阳辐射度的时空变化复杂，以及机器学习算法

在超参数调整时的复杂性，目前的预测精度往往难以满足实际需求。针对该问题，本文提出了一种结合鲸鱼优化算法和时间卷积网络的太阳辐射度预测模型（WOA-TCN）。该模型利用鲸鱼优化算法来调整时间卷积网络的超参数，以更精确地捕捉太阳辐射度的时空变化模式。通过对比实验，WOA-TCN 模型在预测

精度上超越了传统的RNN、TCN、LSTM和CNN模型,为太阳能的有效利用提供了重要的参考价值。

未来的研究可以进一步调整 WOA 算法的配置,优化模型结构和参数,以提高不同区域太阳辐照度预测的准确性和普适性。

参考文献

- [1] 陈怡然, 李进喜, 陈思洁, 等. 新能源环境下的太阳辐射预测系统 [J]. 计算机仿真, 2023, 40(05): 140-4.
- [2] 李津, 史加荣, 张琰妮, 等. 基于最大信息系数的短期太阳辐射协同估计 [J]. 太阳能学报, 2023, 44(09): 286-94.
- [3] 普智勇, 夏攀, 张璐, 等. 机器学习与统计方法在太阳能预报中的比较性分析 [J]. 太阳能学报, 2023, 44(07): 162-7.
- [4] 张晓英, 常正云, 罗童, 等. 基于聚类 EEMD -PCA -LSTM 与误差补偿的光热电站短期太阳直接法向辐射预测 [J]. 电气工程学报: 1-9.
- [5] 尹青, 张华, 何金海. 近 48 年华东地区地面太阳总辐射变化特征和影响因子分析 [J]. 大气与环境光学学报, 2011, 6(01): 37-46.
- [6] 达选芳, 李照荣, 王小勇, 等. 有云条件下太阳辐射短临预报订正技术研究 [J]. 干旱气象, 2021, 39(06): 1006-16.
- [7] 鲁玉军, 周世豪, 胡小勇. 基于 BP 神经网络和小波神经网络的太阳辐射强度预测 [J]. 软件工程, 2023, 26(01): 5-8+4.
- [8] 陈垒, 李杨露西. 基于实际天空图像和 CNN 卷积神经网络
- 的太阳辐射强度预测研究 [J]. 建筑科学, 2022, 38(08): 177-83.
- [9] 臧海祥, 张越, 程礼临, 等. 基于 ICEEMDAN-LSTM 和残差注意力的短期太阳辐照度预测 [J]. 太阳能学报, 2023, 44(12): 175-81.
- [10] 鹿晨东, 许英朝, 张帆, et al. 基于生成对抗网络的太阳辐照度预测模型 [J]. 厦门理工学院学报, 2023, 31(05): 17-24.
- [11] 官松泽, 唐钰本, 蔡争, et al. 基于 Kmeans++-Bi-LSTM 的太阳辐照度超短期预测 [J]. 太阳能学报, 2023, 44(12): 170-4.
- [12] LIU J, HUANG X Q, LI Q, et al. Hourly stepwise forecasting for solar irradiance using integrated hybrid models CNN-LSTM-MLP combined with error correction and VMD [J]. Energy Conv Manag, 2023, 280: 21.
- [13] JALALI S M J, AHMADIAN S, KAVOUSHI-FARD A, et al. Automated Deep CNN-LSTM Architecture Design for Solar Irradiance Forecasting [J]. IEEE Trans Syst Man Cybern -Syst, 2022, 52(1): 54-65.
- [14] GU Q, LI S J, GONG W Y, et al. L-SHADE with parameter decomposition for photovoltaic modules parameter identification under different temperature and irradiance [J]. Appl Soft Comput, 2023, 143: 15.
- [15] WANG P, XUE B, LIANG J, et al. Feature clustering-Assisted feature selection with differential evolution [J]. Pattern Recognit, 2023, 140: 13.
- [16] MIRJALILI S, LEWIS A. The Whale Optimization Algorithm [J]. Adv Eng Softw, 2016, 95(C): 51-67.