

基于多目标算法优化改进谱聚类和 BiLSTM-MHA 的光伏功率预测

唐晓乐¹, 卢 浩^{1,3}, 康滢婷³

(1. 新疆大学电气工程学院能源碳中和实验室, 乌鲁木齐 830047;

2. 西北能源碳中和教育部工程研究中心若羌能源产业研究院, 若羌 841800; 3. 新疆大学智能科学与技术学院, 乌鲁木齐 830047)

摘 要: 为提升典型天气下光伏功率预测的准确性与稳定性, 提出一种基于多目标算法(NSGA II)优化改进谱聚类和双向长短期记忆网络(BiLSTM)结合多头注意力机制(MHA)的光伏功率预测模型。首先, 对气象及历史光伏数据进行异常值检测与处理并分析主要影响特征。其次, 通过动态时间规整(DTW)对谱聚类中度矩阵的构建进行改进, 并利用 NSGA II 对相似性矩阵稀疏度与高斯核参数进行优化得到最优聚类模型, 将天气划分为晴天、阴天与雨天。最后, 建立典型天气下 NSGA II-BiLSTM-MHA 最优参数模型, 并与 4 种基准模型进行对比。结果表明, 在 3 种天气条件下, 所提模型的均方根误差(RMSE)与稳定性指标(SDEX)相比 SVR 分别降低 50.74%~62.95% 和 55.85%~60.08%, 决定系数(R^2)指标提升 8.99%~17.07%, 表明该模型在多种气象条件下均具备更高的预测准确性与稳定性。

关键词: 多目标优化; 光伏功率; 预测模型; 动态时间规整; 多头注意力机制

中图分类号: TM615; TP183

文献标志码: A

0 引 言

随着化石能源储量的不断减少以及大气污染问题的加重, 世界各国已意识到开发和利用新能源替代传统化石能源的紧迫性和重要性^[1]。然而光伏作为新能源中的主力能源, 其间歇性和波动性使其输出功率难以准确预测^[2]。在大规模并网过程中, 这些问题极可能导致电力供需失衡, 甚至威胁电力系统的安全与稳定^[3]。因此, 保证光伏功率预测的准确性与稳定性是保障新能源发展和电网运行安全的关键因素。现有光伏功率预测方法可分为物理方法与统计方法^[4]。物理预测法主要通过建立光伏发电系统和天气条件之间的数学模型, 通过模拟日间辐照、环境温度、相对湿度等气象因素对光伏发电进行预测^[5]。物理预测法对气象数据的要求较高且计算复杂。统计预测法则以气象及功率数据为输入, 通过统计分析和数据挖掘方法, 找出光伏功率与历史气象数据之间的规律性, 从而进行未来功率的预测。常见的基于统计方法的模型包括支持向量回归机^[6]、自回归整合移动平均^[7]、各类神经网络^[8]以及组合模型^[9]等。

目前, 组合模型的研究已成为学者们的关注热点, 尤其是基于深度学习的组合模型。文献[10]提出一种针对光伏集群的短期功率预测方法, 通过引入偏置补偿机制提升了长短期记忆网络(long short-term memory network, LSTM)的预测

性能; 文献[11]提出结合 LSTM-Attention 和卷积神经网络与双向门控循环单元结合的预测模型, 通过引入注意力机制(Attention)改善了 LSTM 在长序列中忽视关键信息的问题。后续组合模型则用于修正预测误差。文献[12]提出一种结合改进灰狼算法的 BiLSTM 短期光伏功率预测方法, 首先利用互信息方法对数据特征进行优化以减少数据噪声, 然后通过混合改进灰狼优化算法确定 BiLSTM 的最优超参数组合, 并引入注意力机制提升模型对关键时序信息的捕捉能力。然而, 尽管上述研究取得了显著成果, 但这些模型在应对复杂非线性长序列依赖关系时仍面临一定挑战。现有文大多专注于提高预测准确性, 往往忽视了模型的稳定性, 而稳定性对于长期有效的预测至关重要。

光伏功率预测受天气变化影响较大, 且不同天气条件下功率波动的特性各异, 准确识别并分离这些波动, 有助于减少模型在处理数据时的复杂度, 进而提升预测能力。文献[13]提出一种基于平衡迭代的层次聚类方法, 无监督地将历史气象数据聚类为 3 种典型天气类型; 文献[14]采用模糊 C 均值聚类(fuzzy C-means algorithm, FCM)算法, 通过模糊隶属度的方式对样本进行分组, 从而实现天气类型划分; 文献[15]提出基于图理论的相似天气划分方法, 通过构建天气特征向量与格拉姆矩阵, 实现了气象数据的图像化, 以此挖掘不同天气向量间的关系; 文献[16]采用改进的自组织数据迭代分析

收稿日期: 2025-02-18

基金项目: 新疆省重点研发计划(2022B01020-4); 国家自然科学基金(52266017); 新疆天山英才科技创新团队项目(2023TSYCTD0009); 新疆自治区研究生教育创新计划项目(XJ2023G052)

通信作者: 卢 浩(1989—), 男, 博士、教授, 主要从事光伏功率预测、多能互补综合能源系统、新能源发电等方面的研究。luhao@xju.edu.cn

算法对历史光伏发电功率数据进行了相似日划分,该算法通过自动分类历史数据,将相似的气象和发电特征划分为不同类别,为后续模型提供了适应性强的训练数据。然而,上述研究中的聚类方法大多依赖于欧氏距离来计算序列间的相关性,忽略了日照时长(有效功率点数)随季节变化而产生的差异。欧氏距离计算通常是基于逐个时间点的直接比较,当时间序列无法完全对齐时,欧式距离将比较不同时间点的数值。因此,基于欧氏距离的聚类方法往往不能反映序列之间的真实关系。此外,许多研究仅通过固定参数下的单一指标来衡量聚类效果,这使得聚类模型的性能无法得到充分发挥,从而影响了模型的整体表现。

为解决上述问题,本文提出一种基于多目标算法优化改进谱聚类(improved spectral clustering, ISC)和双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)结合多头注意力机制的光伏功率预测方法。首先,利用动态时间规整处理时间序列非线性对齐问题,构建度矩阵及拉普拉斯矩阵,其中相似矩阵稀疏度及拉普拉斯矩阵正则化参数通过 NSGA II 进行优化。其次,根据聚类结果将历史功率数据分别划分为晴天、阴天及雨天,分别建立不同天气下 NSGA II-BiLSTM-MHA 模型,其中 BiLSTM 充分考虑序列数据的前向和后向相关性,多头注意力机制(multi-head attention, MHA)可抑制表征学习过程中的非重要信息,最大程度减少不相关特征信息的影响,NSGA II 对模型参数进行优化,进一步提升模型的准确性和稳定性。最后,为验证所提模型的有效性,利用山东某光伏电站及气象站实际数据进行仿真实验。

1 理论基础

1.1 动态时间规整

动态时间规整(dynamic time warping, DTW)^[17]是一种相似距离计算方法,它通过非线性的方式对齐两个序列,从中找到最佳匹配路径,从而找出序列间总距离最小的相似时间序列。DTW 算法广泛应用于手写识别、噪声辨别、金融分析等领域,尤其适用于时间序列数据的比对与分析。给定两个时间序列 $X = \{x_1, x_2, \dots, x_n\}$ 与 $Y = \{y_1, y_2, \dots, y_m\}$, DTW 的目标通过一个最优路径将两个序列对齐,计算两者相似度。

1) 初始化相似矩阵:计算每对数据点 (x_i, y_j) 之间的欧氏距离 $d_{xy}(ij)$,并构建矩阵 D_{xy} ,定义为:

$$D_{xy} = \begin{bmatrix} d_{xy}(1,1) & d_{xy}(1,2) & \cdots & d_{xy}(1,j) \\ d_{xy}(2,1) & d_{xy}(2,2) & \cdots & d_{xy}(2,j) \\ \vdots & \vdots & \ddots & \vdots \\ d_{xy}(i,1) & d_{xy}(i,2) & \cdots & d_{xy}(i,j) \end{bmatrix} \quad (1)$$

2) 计算累计相似矩阵: C 为距离矩阵,其中每个元素 $C(i,j)$ 表示从 $(1,1)$ 到 (i,j) 的最小累计距离。累计距离的计算公式为:

$$C(i,j) = d_{xy}(i,j) + \min(C(i-1,j), \dots, C(i,j-1), C(i-1,j-1)) \quad (2)$$

式中: $C(i-1,j)$ 、 $C(i,j-1)$ 、 $C(i-1,j-1)$ ——上方、左侧与左上方相邻点的累计距离。

3) 最佳匹配路径:找出从 $(1,1)$ 到 (n,m) 的最佳匹配路径,匹配路径由累计矩阵中最小值的序列构成。

4) 计算相似度:根据最佳匹配路径上的距离之和计算两个序列的相似度,定义为:

$$W(X,Y) = \min \left(\sum_{r=1}^R d_{xy}(t_1^k, t_2^k) \right) \quad (3)$$

式中: R ——路径步数; t_1^k, t_2^k ——路径上每一步对应的时间点; $d_{xy}(t_1^k, t_2^k)$ ——每个时间点对的距离。

1.2 基于 DTW 的改进谱聚类

谱聚类(spectral clustering, SPC)^[18]是一种基于图论的聚类算法,通过构建相似性矩阵,利用图的拉普拉斯矩阵来揭示数据的内在结构。相比于传统的 K -均值聚类方法,SPC 在处理复杂和非线性结构的聚类任务时表现出更强的能力。它能够根据数据点的相似性自动发现合理的分组方式,无需事先指定聚类的形状和分布。ISC 通过引入 DTW 方法,取代传统 SPC 中使用欧氏距离构建度矩阵的方式,能更有效地处理时间序列之间的时间错位及非线性变化等问题。ISC 的基本步骤为:

1) 构建相似性矩阵:根据 DTW 计算出相似度矩阵 W ,其中 $W(i,j)$ 为第 i 个样本与第 j 个样本之间的相似度;

2) 构建度矩阵:度矩阵 S 为对角矩阵,其中每个对角线元素 $S(i,i)$ 表示第 i 个样本与其他样本的总相似度:

$$S(i,i) = \sum_j W(i,j) \quad (4)$$

3) 构建拉普拉斯矩阵: $L = S - W$,其中 L 为半正定对称矩阵,其特征值均为实数;

4) 计算特征向量:计算拉普拉斯矩阵 L 的前 k 个特征向量构成向量空间,并对该空间做标准化处理;

5) K -均值聚类:采用 K -均值方法对特征向量空间中的 k 个特征向量进行聚类。

1.3 NSGA II 算法

NSGA II 作为经典的多目标优化算法之一,被广泛应用于各领域的研究工作中^[19]。与其前身 NSGA 相比,NSGA II 在计算效率、解集质量以及多样性保持方面均表现出显著优势,这些特点使得 NSGA II 成为解决工程设计及能源系统优化等复杂问题的理想工具。NSGA II 算法的执行过程可分为以下 7 个关键步骤:

1) 初始化种群:根据决策空间上下界随机生成初始种群,所有个体均为搜索空间中的潜在解;

2) 交叉变异:利用交叉和变异操作生成新的子代个体,

促进解空间的探索;

3)快速非支配排序:根据种群个体间的支配关系将种群快速划分为多个前沿,所有个体依据其帕累托前沿(Pareto frontier, PF)等级进行排序;

4)拥挤度计算:利用欧式距离,计算每个个体与其他所有个体的距离之和,得到其在目标空间中的拥挤密度;

5)选择操作:通过二元锦标赛的选择策略,基于非支配等级和拥挤度值选出较好的个体进入下一代;

6)合并与更新:将父代和子代合并为一个候选种群,重新进行非支配排序和拥挤度计算,确保最优秀的个体进入下一代;

7)迭代终止:当算法达到最大迭代次数时,停止迭代并输出 PF,否则转到第 2)步。

1.4 双向长短期记忆网络

BiLSTM^[20]在 LSTM 基础上进行了改进,通过正向与反向两个方向连接隐藏层并进行堆叠,充分考虑了序列数据的前向和后向相关性。其中,前向网络与后向网络的结构相同,均包含 1 个存储单元和 3 个门(即遗忘门、输入门和输出门), F 、 I 和 O 分别为 3 个门的输出结果,计算公式定义如下:

$$I_t = \Pi(Y_i V_t + Q_i Z_{t-1} + b_i) \quad (5)$$

$$F_t = \Pi(Y_f V_t + Q_f Z_{t-1} + b_f) \quad (6)$$

$$O_t = \Pi(Y_o V_t + Q_o Z_{t-1} + b_o) \quad (7)$$

$$\tilde{E}_t = \tanh(Y_c V_t + Q_c Z_{t-1} + b_c) \quad (8)$$

$$E_t = F_t \otimes E_{t-1} + I_t \otimes \tilde{E}_t \quad (9)$$

$$Z_t = O_t \otimes \tanh(E_t) \quad (10)$$

式中: Π ——sigmoid 激活函数; Y ——输入 V_t 的权重矩阵; V_t 、 Z_t —— t 时刻的输入与循环输出状态; Q —— Z_{t-1} 的权重矩阵;下标 i, f, o, c ——输入门、遗忘门、输出门、记忆单元门; b ——偏置; $\tilde{E}_t$ ——候选记忆单元状态; E_t ——记忆单元; $\tanh$ ——双曲正切函数。

1.5 多头注意力机制

在时间序列预测中,模型很难建立起长序列的依赖关系,即当前时刻的信息中很难包含远距离的过去及未来信息。多头注意力机制可同时关注多个子空间,且可不受距离限制地关注到与当前所需最为相关的隐含知识,从而抑制表征学习过程中的非重要信息,最大程度减少不相关特征信息的影响。多头注意力模型通过给每个特征提取数据分配不同的权重来构建注意力矩阵。注意力模块的输入为 3 个大小与输入相同的张量 A 、 S 与 D 。通过点积计算即可得到重构后的注意力矩阵,计算方法为:

$$T(A, S, D) = \text{softmax}\left(\frac{AS^T}{\sqrt{d_s}}\right)D \quad (11)$$

式中:softmax——归一化指数函数; d_s ——调整因子, d_s 用于防止 AS 的内积过大,其大小为 S 向量维度的平方根。

上述为某一子空间中计算注意力矩阵的过程。多头注意力矩阵是将 S_p 个子空间的计算结果拼接起来,每个子空间的信息获取由多个注意力头同时并行计算。 S_p 个子空间为张量 A 、 S 与 D 沿相同方向分割后的 S_p 个片段。多头注意力矩阵定义为:

$$H(A, S, D) = \text{Concat}(h_1, h_2, \dots, h_{s_p})W_m \quad (12)$$

式中: h_i ——每个子空间中的注意力矩阵;Concat——连接函数; W_m ——线性变换矩阵。

2 预测模型总框架

2.1 基于 NSGA II 优化改进谱聚类和 BiLSTM-MHA 光伏功率预测模型

DTW 与 SPC 的结合能够充分挖掘功率曲线之间的相似性和差异性,有效区分不同类型的天气模式。基于 NSGA II 优化的 ISC 方法能更好地平衡聚类结果的内部紧密度与类间分离度,从而获得最优天气聚类结果。BiLSTM 与 MHA 的结合可突破距离限制,关注与当前所需最为相关的隐含知识,从而有效建立长期时序依赖关系。以均方根误差(root mean square error, RMSE)及稳定性指标(stability index, SDEX)作为两目标函数,引入 NSGA II 对模型参数进行优化,从而在提升模型预测准确性的同时增强其预测稳定性。综上,本文构建基于 NSGA II 优化 ISC 典型天气划分下的 NSGA II-BiLSTM-MHA 光伏功率预测模型,预测过程如图 1 所示。

1)以轮廓系数(silhouette coefficient, SC)与戴维森堡丁指数(Davies-Bouldin index, DBI)为目标函数,引入 NSGA II 多目标优化并进行初始化。以相似性矩阵的稀疏度及高斯核参数 σ 为决策变量,在提升 SC 的同时尽量减少 DBI,以实现最佳的聚类效果。计算公式为:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (13)$$

$$B = \frac{1}{k} \sum_{i=1}^k \max\left(\frac{S_i + S_j}{d(C_i, C_j)}\right) \quad (14)$$

式中: $a(i)$ ——样本点 i 到其簇内其他点的平均距离; $b(i)$ ——样本点 i 到最近其他簇的所有点的平均距离; S_i, S_j ——簇 i 和簇 j 的平均距离; $d(C_i, C_j)$ ——簇中心之间的距离。

2)通过 NSGA II 算法迭代输出 PF,并从中选取 SC 与 DBI 最大差值点为最优解;

3)输入光伏功率数据,并根据最优解建立 $W(i, j)$ 及拉普拉斯矩阵 $L = S - W$;

4)计算特征向量并对前 3 个特征向量进行 K -均值聚类,输出聚类结果,根据聚类簇中数据特征将数据划分为晴天、阴天和雨天;

5) 对数据进行异常值检测与处理, 并对处理后的完整数据进行相关性分析, 选出模型主要输入特征;

6) 构建 NSGA II-BiLSTM-MHA 模型, 利用 NSGA II 优化算法对模型的参数进行寻优, 其中均方根误差(RMSE)及稳定性指标(SDEX)为目标函数;

7) 利用 NSGA II 算法迭代输出 PF, 并选取拐点为最优解

建立最优参数预测模型;

8) 分别在 3 种天气下对 NSGA II-BiLSTM-MHA 模型进行训练, 输出预测结果;

9) 采用 RMSE、平均绝对误差(mean absolute error, MAE)、决定系数(coefficient of determination, R^2)以及 SDEX 评估模型的预测结果, 并对结果进行分析。

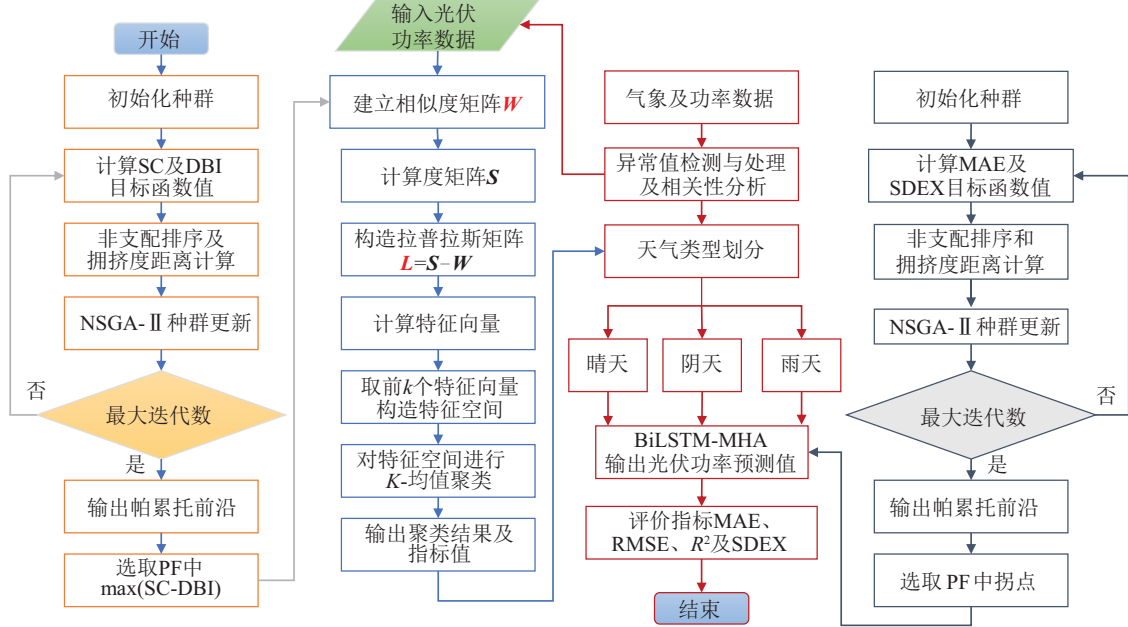


图1 预测模型流程

Fig. 1 Flow chart of prediction model

2.2 评价指标

为评估模型的预测性能, 分别引入 RMSE、MAE、 R^2 以及 SDEX 指标对模型的预测性能进行评估^[21]。其中 RMSE、MAE 用于衡量误差大小, 其量级与样本数据量级相关, 定义如下:

$$E_{\text{RMSE}} = \sqrt{\frac{1}{s} \sum_{i=1}^s (Y_{\text{real},i} - Y_{\text{pred},i})^2} \quad (15)$$

$$E_{\text{MAE}} = \frac{1}{s} \sum_{i=1}^s |Y_{\text{real},i} - Y_{\text{pred},i}| \quad (16)$$

式中: s ——预测样本数量; Y_{real} ——样本实际值; Y_{pred} ——预测值。

R^2 用于评估模型拟合程度, 其取值范围为 $[0, 1]$ 。SDEX 通过标准差函数 STD 评估预测稳定性, 定义为:

$$R^2 = 1 - \frac{\sum_{i=1}^s (Y_{\text{real},i} - Y_{\text{pred},i})^2}{\sum_{i=1}^s Y_{\text{real},i} - \bar{Y}_{\text{pred},i}} \quad (17)$$

$$U = \text{STD}(|Y_{\text{real}} - Y_{\text{pred}}|) \quad (18)$$

3 实例分析

3.1 数据集

以位于经度 $117^\circ 36' 0.39''$, 纬度 $35^\circ 58' 9.79''$ 的某光伏电

站数据以及当地气象数据作为实验数据集, 如图 2 所示。数据集中包含直射辐照度、总辐照度、散射辐照度 3 种辐射数据、气温、湿度、风速、风向、风力等级、阵风风速、降水量、云量以及历史功率。所有数据的时间分辨率均为 15 min。此外, 考虑到夜晚光伏发电功率为 0, 所以对历史光伏功率数据做夜间去零化处理。由于不同季节的日照时长具有一定差异, 因此日间光伏功率序列长度不等。

3.2 异常值检测及处理

光伏输出功率与辐照度通常呈显著线性关系, 然而实际光伏电站在运行过程中, 由于数据录入错误、仪器误差等因素, 导致原始数据出现异常值与缺失值。为提升数据可靠性, 需对原始数据进行缺失值的补充和异常值的替换。稳健回归方法中的随机样本一致性回归算法 (random sample consensus, RANSAC)^[22] 具有强大的异常值抵御能力, 它通过随机选择数据的子集来构建回归模型并对结果进行评估。RANSAC 算法会将不符合回归模型拟合的点自动识别为异常值, 不将其纳入模型训练中。进行 RANSAC 回归后, 通过计算每个数据点的回归残差来识别异常值。残差反映了回归模型的预测偏差, 若残差的绝对值超出 3 倍标准差 (即 3σ) 的范围, 则视为异常值。图 3 为异常值的筛选结果。

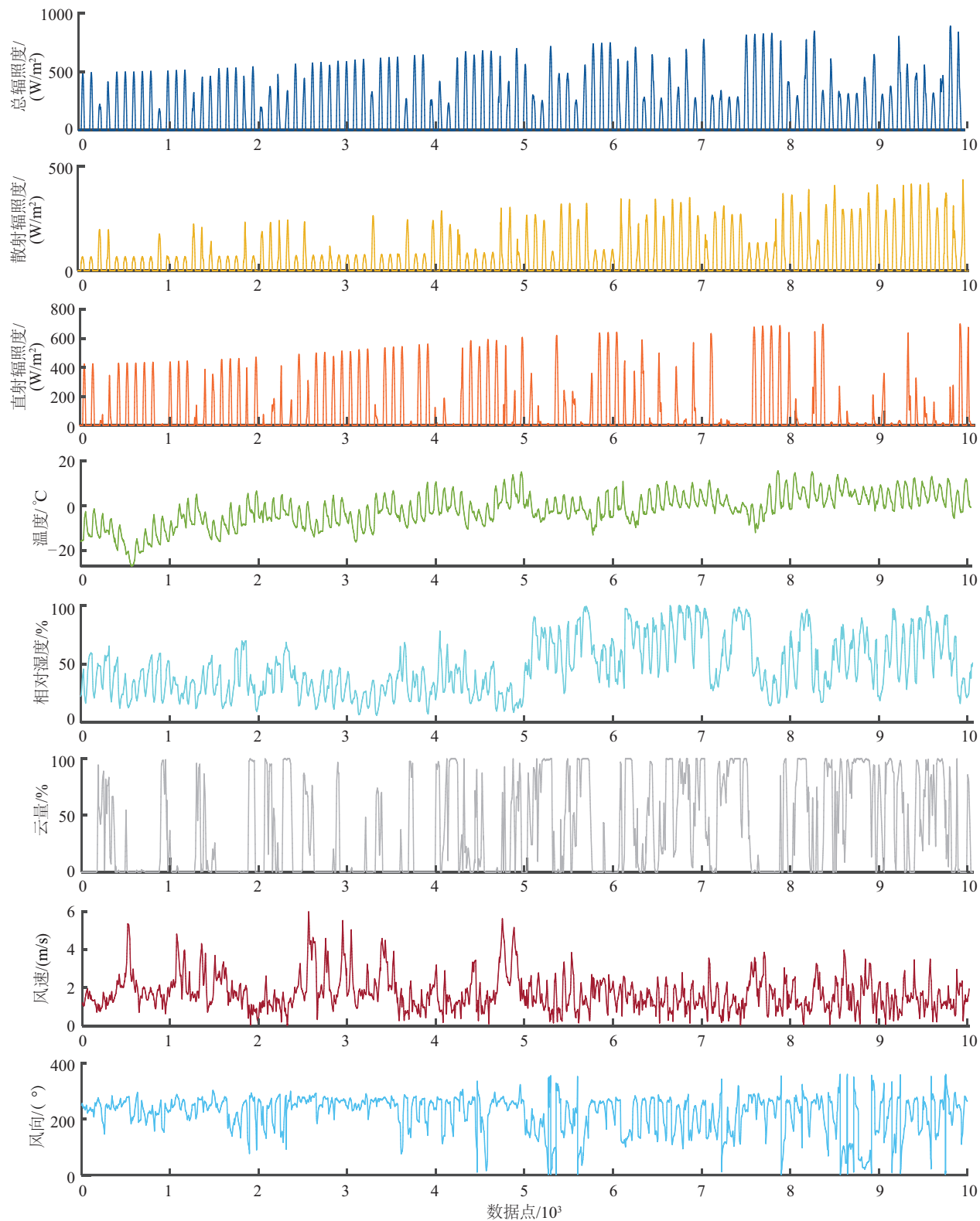


图2 数值天气时序数据

Fig. 2 Numerical weather time-series data

3.3 相关性分析

影响光伏输出的因素众多,除设备运行状况外,气象因素(如辐照度、温度和湿度等)通常是主要的影响因素。斯皮

尔曼等级相关系数(Spearman rank correlation, SRC)^[23]能够处理非线性关系以及顺序数据,可有效捕捉变量之间的单调关系。该方法通过对各数据点进行排序,计算排名之间的差

异从而得出相关系数。

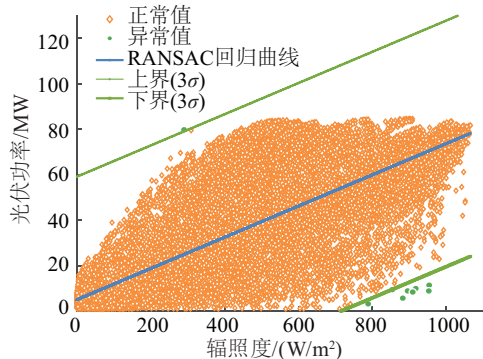


图3 RANSAC异常值检测结果

Fig. 3 RANSAC outlier detection results

图4展示了气象特征与历史功率之间的相关性大小,每个柱形图代表一种气象特征,幅值大小为其与光伏功率之间的相关系数 $|\rho_{xy}|$,反映了相关性程度。通过分析,选取 $|\rho_{xy}| > 0.3$ 的气象特征作为最终的输入特征,包括直射辐照度、总辐照度、散射辐照度、湿度以及云量。

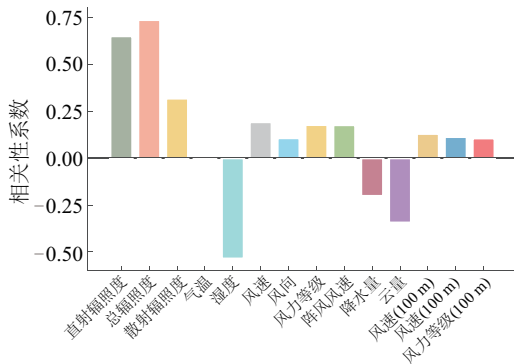


图4 气象特征与光伏功率之间的相关程度

Fig. 4 Degree of correlation between meteorological features and photovoltaic power

3.4 相似日划分结果

采用基于DTW及SPC的方法对典型天气进行划分,并通过NSGA II对模型进行优化,以提升聚类效果。考虑光伏预测中典型天气一般为晴天、阴天及雨天,因此聚类簇数设置为3。NSGA II中种群规模设置为100,最大迭代次数为50,交叉率及变异率分别为0.8和0.2。图5为获得的帕累托前沿,其中SC是衡量类内样本与类间样本距离的指标,范围为 $[-1, 1]$,越接近1表明聚类效果越好。DBI用于衡量类内距离和类间距离的比率,DBI越小代表类内距离越小的同时类间距离越大。因此,选取所获得前沿中SC与DBI的相对差值最大点为最优解建立最优聚类模型,此时相似性矩阵稀疏度与高斯核参数sigma分别为 $s = 5, \sigma = 0.2765$ 。

图6为划分后3种典型天气的功率平均值曲线,图7为

各天气下的功率曲线,其中晴天曲线较为平滑,阴天曲线表现出较强的波动性,雨天曲线除表现出较强的波动性特征之外,其功率幅值出较低。

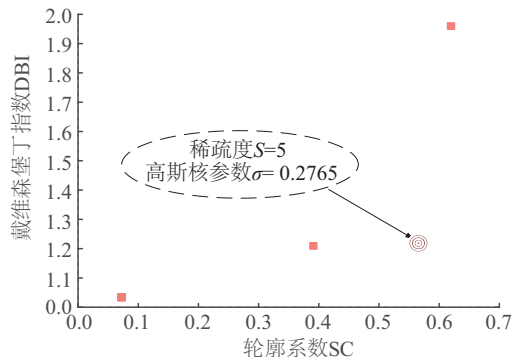


图5 NSGA II优化典型天气划分模型的帕累托前沿

Fig. 5 Pareto frontier of NSGA II optimized typical weather division model

为验证所提方法的有效性,选取K-均值、谱聚类SPC、基于DTW改进的ISC为基准模型进行对比。所有模型聚类簇数均设置为3,SPC采用欧氏距离平方计算相似度矩阵,采用SC及DBI指标对模型性能进行评估,结果如表1所示。所提模型具有最大的SC和较小的DBI指数,由图8可知,所提模型SC与DBI的相对差值最大,聚类性能最佳。

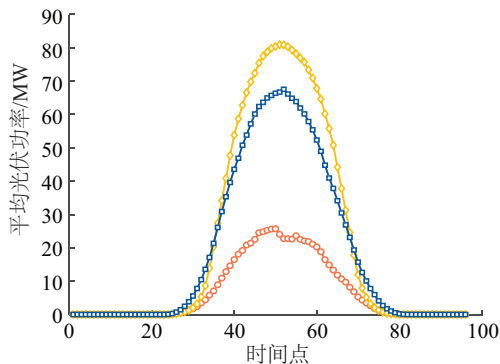
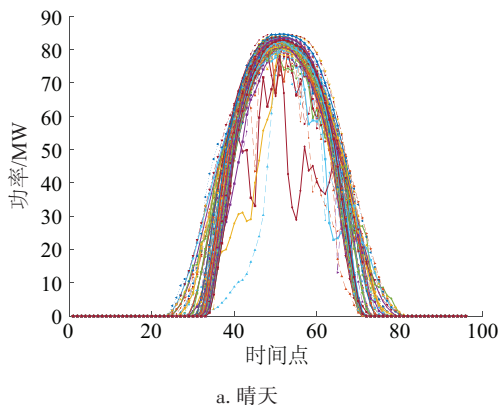


图6 3种典型天气下的日功率平均值曲线

Fig. 6 Average daily power curves under three typical weather conditions



a. 晴天

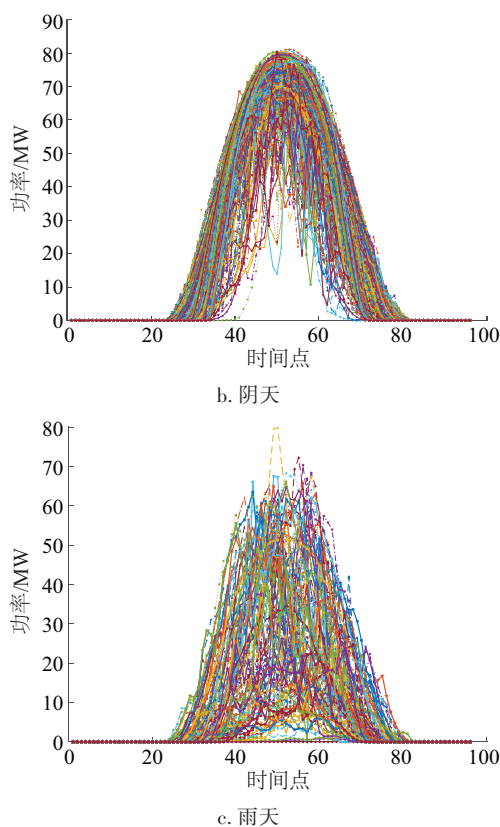


图7 3种典型天气下日光伏功率曲线

Fig. 7 Daily PV power curves under three typical weather conditions

表1 聚类模型性能比较

Table 1 Performance comparison of clustering models

模型	SC	DBI
K-均值	-0.3389	1.1978
SPC	-0.3036	1.1119
ISC	0.1513	1.7043
NSGA II-ISC	0.5643	1.2167

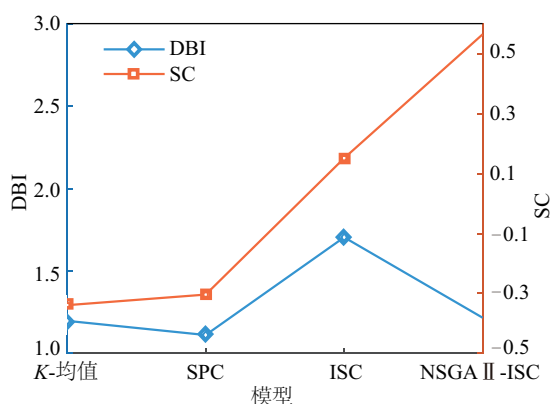


图8 SC与DBI相对差值比较

Fig. 8 Comparison of relative differences between SC and DBI

3.5 模型测试

3.5.1 参数设置

实验采用配置为: Win11 64bit 操作系统, Intel Core i7-11800H 处理器, NVIDIA GeForce GTX 1660 显卡, 编程软件为 Python 3.8 软件。每个模型运行 10 次后取平均结果。NSGA II 参数设置同 3.4 节。BiLSTM 中待优化参数分别为隐含层神经元数量、学习率、迭代次数以及批尺寸。参数含义以及优化范围如表 2 所示。

表2 BiLSTM主要的超参数

Table 2 Main hyperparameters of BiLSTM

超参数	超参数含义	参数范围
Nodes_hiddenlayer1	隐含层 1 的神经元数量	[5, 100]
Nodes_hiddenlayer2	隐含层 2 的神经元数量	[5, 100]
Learning_rate	学习率	[0.01, 0.3]
Epochs number	迭代次数	[50, 2000]
Batch_size	L1 正则化项的权重	[2, 100]

3.5.2 预测结果对比分析

采用山东某电站实际数据对天气类型划分下, NSGA II - BiLSTM-MHA 模型的有效性进行验证, 并通过支持向量回归 (support vector regression, SVR)、LSTM、BiLSTM、NSGA II - BiLSTM-Attention 共 4 种基准模型与所提模型进行对比, 结果如图 9 所示。从 3 种典型天气下的预测结果可看出, SVR 预测值与真实值的偏差最大且预测稳定性最差。BiLSTM 的预测效果优于 LSTM, 说明与传统单向 LSTM 相比, BiLSTM 提取特征能力更强。NSGA II - BiLSTM-Attention 的预测效果比 BiLSTM 更好, 与真实值之间的偏差较小, 稳定性更高, 证明采用 NSGA II 算法优化可有效提升模型的准确性与稳定性。NSGA II - BiLSTM-MHA 在 3 种天气情况下的表现效果最优, 与 BiLSTM 相比, NSGA II - BiLSTM-MHA 通过引入 NSGA II 多目标优化算法和多头注意力机制, 有效优化了参数选择, 并通过多头注意力关注数据的不同子空间, 进一步增强了特征捕捉能力。

采用 MAE、RMSE、 R^2 指标对模型性能进行评价, 结果如表 3 所示。在晴天类型下, NSGA II - BiLSTM-MHA 较 SVR 的 MAE、RMSE 及 SDEX 分别降低 56.30%、62.95% 和 59.69%, R^2 提升 10.00%, 较 NSGA II - BiLSTM-Attention 分别降低 13.31%、39.66%、34.57%, R^2 提升 3.13%。在阴天类型下, 所提模型相比 SVR 的 MAE、RMSE 及 SDEX 分别降低 53.33%、50.74%、55.85%, R^2 提升 17.07%, 较 NSGA II - BiLSTM-Attention 分别降低 22.73%、26.18%、32.96%, R^2 提升 4.35%。在雨天类型下, NSGA II - BiLSTM-MHA 较 SVR 的 3 种指标分别降低 56.14%、57.11%、60.08%, R^2 提升 8.99%。NSGA II - BiLSTM-MHA 较 NSGA II - BiLSTM-Attention 的 3 种指标分别降低 38.14%、32.68%、35.80%, R^2 提升 4.3%。

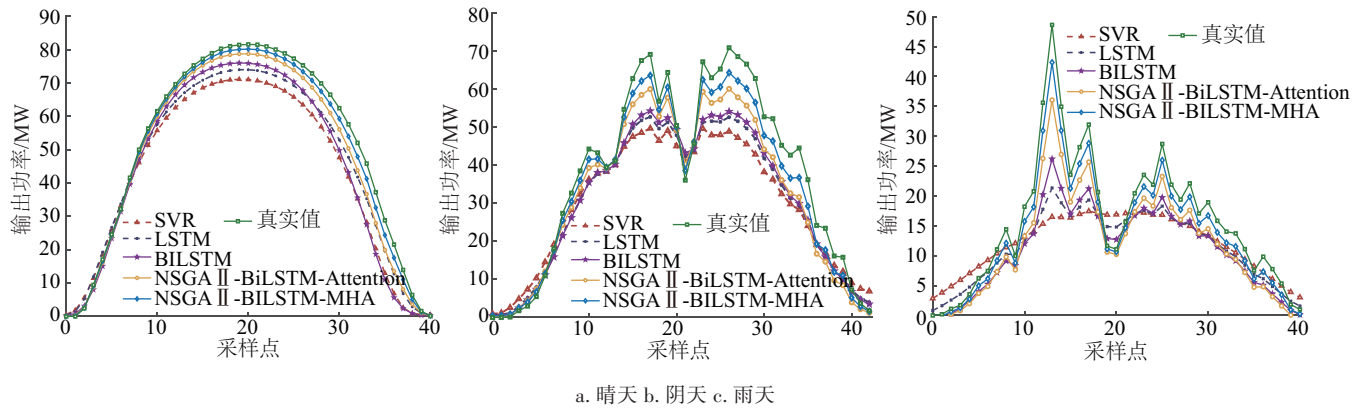


图9 光伏功率预测曲线

Fig. 9 Photovoltaic power prediction curves

表3 评价指标计算

Table 3 Evaluation index calculation

模型	MAE			RMSE			R^2			STD		
	晴天	阴天	雨天	晴天	阴天	雨天	晴天	阴天	雨天	晴天	阴天	雨天
SVR	6.11	7.50	4.40	9.69	11.45	8.79	0.90	0.82	0.89	8.31	13.41	10.42
LSTM	4.97	6.65	3.74	8.22	10.27	7.85	0.93	0.85	0.91	7.94	11.53	9.87
BiLSTM	4.18	6.58	3.51	6.32	9.60	6.91	0.96	0.87	0.92	7.31	10.62	7.65
NSGAII-BiLSTM-Attention	3.08	4.53	3.12	5.95	7.64	5.60	0.96	0.92	0.93	5.12	8.83	6.48
NSGAII-BiLSTM-MHA	2.67	3.50	1.93	3.59	5.64	3.77	0.99	0.96	0.97	3.35	5.92	4.16

4 结 论

[参考文献]

为促进新能源的高效利用,提高光伏发电系统功率预测的准确性与稳定性,提出一种基于多目标算法优化 ISC 和 BiLSTM-MHA 的光伏功率预测模型,用于提升短期光伏功率预测模型的性能。结合光伏电站实际数据对模型进行验证,得出以下主要结论:

1)提出的基于 NSGA II 优化基于动态时间规整改进谱聚类的方法通过引入 NSGA II 优化聚类参数,充分挖掘功率曲线之间的相似性和差异性,有效区分了不同类型的天气模式,为后续的建模、预测及模型性能验证提供更加细致和精准的数据基础。

2)采用 BiLSTM 模型,引入多头注意力机制,同时关注多个子空间,增强了模型的特征捕捉能力。通过 NSGA II 优化模型参数,有效提升了模型预测准确性与稳定性。

3)在 3 种典型天气下,所提模型相较于传统预测模型均表现出更优的预测准确性和稳定性,尤其在波动较大的雨天天气下提升更为显著。以雨天为例,与 SVR 模型相比,所提模型 MAE 降低 56.13%,RMSE 降低 57.11%,SDEX 降低 60.08%, R^2 提升 8.99%。

- [1] ZHANG M Y, HAN Y, WANG C Y, et al. Ultra-short-term photovoltaic power prediction based on similar day clustering and temporal convolutional network with bidirectional long short-term memory model: a case study using DKASC data [J]. Applied energy, 2024, 375: 124085.
- [2] LI C S, TANG G, XUE X M, et al. The short-term interval prediction of wind power using the deep learning model with gradient descend optimization[J]. Renewable energy, 2020, 155: 197-211.
- [3] 贺亦琛, 师长立, 郭小强, 等. 基于参数优化多核支持向量机的光伏功率预测算法[J]. 太阳能学报, 2024, 45(9): 394-404.
HE Y C, SHI C L, GUO X Q, et al. Photovoltaic power prediction algorithm based on parameter optimization of multi-kernel SVM[J]. Acta energiae solaris sinica, 2024, 45(9): 394-404.
- [4] 赖昌伟, 黎静华, 陈博, 等. 光伏发电出力预测技术研究综述[J]. 电工技术学报, 2019, 34(6): 1201-1217.

- LAI C W, LI J H, CHEN B, et al. Review of photovoltaic power output prediction technology [J]. Transactions of China Electrotechnical Society, 2019, 34 (6) : 1201-1217.
- [5] HENG J N, WANG J Z, XIAO L Y, et al. Research and application of a combined model based on frequent pattern growth algorithm and multi-objective optimization for solar radiation forecasting[J]. Applied energy, 2017, 208: 845-866.
- [6] ESEYE A T, ZHANG J H, ZHENG D H. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and meteorological information[J]. Renewable energy, 2018, 118: 357-367.
- [7] SHADAB A, AHMAD S, SAID S. Spatial forecasting of solar radiation using ARIMA model [J]. Remote Sensing Applications: Society and Environment, 2020, 20: 100427.
- [8] ALRASHIDI M, RAHMAN S. Short-term photovoltaic power production forecasting based on novel hybrid data-driven models[J]. Journal of big data, 2023, 10(1): 26.
- [9] MARKOVICS D, MAYER M J. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction [J]. Renewable and sustainable energy reviews, 2022, 161: 112364.
- [10] 卢俊杰, 蔡涛, 郎建勋, 等. 基于集群划分的光伏电站集群发电功率短期预测方法[J]. 高电压技术, 2022, 48 (5): 1943-1951.
- LU J J, CAI T, LANG J X, et al. Short-term power output forecasting of clustered photovoltaic solar plants based on cluster partition [J]. High voltage engineering, 2022, 48 (5): 1943-1951.
- [11] 吐松江·卡日, 雷柯松, 马小晶, 等. 基于 LSTM-Attention 和 CNN-BiGRU 误差修正的光伏功率预测[J]. 太阳能学报, 2024, 45(12): 85-93.
- TUSONGJIANG K, LEI K S, MA X J, et al. Photovoltaic power prediction based on LSTM-Attention and CNN-BiGRU error correction [J]. Acta energiae solaris sinica, 2024, 45(12): 85-93.
- [12] 赵如意, 王晓辉, 郑碧煌, 等. 基于特征优化和混合改进灰狼算法优化 BiLSTM 网络的短期光伏功率预测[J]. 电网技术, 2025, 49(1): 209-222.
- ZHAO R Y, WANG X H, ZHENG B H, et al. Short-term photovoltaic power prediction based on feature optimization and hybrid improved grey wolf algorithm-optimized BiLSTM network [J]. Power system technology, 2025, 49 (1): 209-222.
- [13] 董俊, 刘瑞, 束洪春, 等. 基于 BIRCH 聚类的 L-Transformer 分布式光伏短期发电功率预测[J]. 高电压技术, 2024, 50(9): 3883-3893.
- DONG J, LIU R, SHU H C, et al. Short-term distributed photovoltaic power generation prediction based on BIRCH clustering and L-Transformer [J]. High voltage engineering, 2024, 50(9): 3883-3893.
- [14] 杨国华, 张鸿皓, 郑豪丰, 等. 基于相似日聚类 and IHGWO-WNN-AdaBoost 模型的短期光伏功率预测[J]. 高电压技术, 2021, 47(4): 1185-1194.
- YANG G H, ZHANG H H, ZHENG H F, et al. Short-term photovoltaic power forecasting based on similar weather clustering and IHGWO-WNN-AdaBoost model [J]. High voltage engineering, 2021, 47(4): 1185-1194.
- [15] 吴春华, 董阿龙, 李智华, 等. 基于图相似日和 PSO-XGBoost 的光伏功率预测[J]. 高电压技术, 2022, 48 (8): 3250-3259.
- WU C H, DONG A L, LI Z H, et al. Photovoltaic power prediction based on graph similarity day and PSO-XGBoost [J]. High voltage engineering, 2022, 48(8): 3250-3259.
- [16] 方朝雄, 郑洁云, 张章煌, 等. 基于相似日与 VMD-DBO-KELM 的分布式光伏发电功率预测方法[J]. 高电压技术, 2025, 51(7): 3477-3487.
- FANG C X, ZHENG J Y, ZHANG Z H, et al. Distributed photovoltaic power prediction based on similar day and VMD-DBO-KELM [J]. High voltage engineering, 2025, 51 (7): 3477-3487.
- [17] 郭雯博. 基于相似日理论和 Tent-SSA-BP 模型的光伏发电功率预测研究[D]. 淮南: 安徽理工大学, 2024.
- GUO W B. Research on photovoltaic power prediction based on similar day theory and Tent-SSA-BP model [D]. Huainan: Anhui University of Science & Technology, 2024.
- [18] 李东颖. 基于 Bi-LSTM 负荷预测的光储充电站优化配置与控制研究[D]. 天津: 河北工业大学, 2022.
- LI D Y. Research on optimal configuration and control of optical storage charging station based on Bi-LSTM load forecasting [D]. Tianjin: Hebei University of Technology, 2022.
- [19] YUSUF A, BAYHAN N, TIRYAKI H, et al. Multi-objective optimization of concentrated Photovoltaic-Thermoelectric hybrid system via non-dominated sorting genetic algorithm (NSGA II) [J]. Energy conversion and management, 2021, 236: 114065.
- [20] 时培明, 郭轩宇, 杜清灿, 等. 基于 TCN-BiLSTM-Attention-ESN 的光伏功率预测[J]. 太阳能学报, 2024, 45(9): 304-316.
- SHI P M, GUO X Y, DU Q C, et al. Photovoltaic power prediction based on TCN-BiLSTM-Attention-ESN [J]. Acta energiae solaris sinica, 2024, 45(9): 304-316.
- [21] YU M, NIU D X, WANG K 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.
- [22] MARTÍNEZ-OTZETA J M, RODRÍGUEZ-MORENO I, MENDIALDUA I, et al. RANSAC for robotic applications: a survey[J]. *Sensors*, 2023, 23(1): 327.
- [23] DE WINTER J C F, GOSLING S D, POTTER J. Comparing the pearson and spearman correlation coefficients across distributions and sample sizes: a tutorial using simulations and empirical data [J]. *Psychological methods*, 2016, 21(3): 273-290.

PHOTOVOLTAIC POWER PREDICTION BASED ON IMPROVED SPECTRAL CLUSTERING AND BiLSTM-MHA OF MULTI-OBJECTIVE ALGORITHM OPTIMIZED

Tang Xiaole¹, Lu Hao¹⁻³, Kang Yanting³

(1. *Laboratory of Energy Carbon Neutrality, School of Electrical Engineering, Xinjiang University, Urumqi 830047, China;*

2. *Ruoqiang Energy Industry Research Institute, Engineering Research Center of Northwest Energy Carbon Neutrality, Ministry of Education, Ruoqiang 841800, China;*

3. *School of Intelligence Science and Technology, Xinjiang University, Urumqi 830047, China.*)

Abstract: To enhance the accuracy and stability of photovoltaic power forecasting under typical weather conditions, this paper proposes a forecasting model that integrates an improved spectral clustering method optimized by the NSGA II multi-objective algorithm with a BiLSTM network enhanced by a multi-head attention mechanism (MHA). Firstly, outlier detection and preprocessing are performed on meteorological and historical PV data, and key influencing features are identified. Then, the construction of the degree matrix in spectral clustering is improved using dynamic time warping (DTW), and NSGA II is employed to optimize the sparsity of the similarity matrix and the Gaussian kernel parameter, yielding an optimal clustering model that categorizes weather into sunny, cloudy, and rainy types. Finally, optimal NSGA II -BiLSTM-MHA models are established for each weather type and compared with four baseline models. Results show that, under three weather conditions, the proposed model achieves 50.74%-62.95% lower RMSE and 55.85%-60.09% lower SDEX than that of SVR, while improving the R^2 by 8.99%-17.07%.

Keywords: multiobjective optimization; photovoltaic power; prediction models; dynamic time warping; multi-head attention mechanism