

基于相似日聚类与 WOA-BiLSTM-Copula 算法的短期风光功率相关性概率区间预测方法

王凌梓, 沈海波, 邓力源, 刘显茁, 邓韦斯

(中国南方电网电力调度控制中心, 广州 510663)

摘要: 风电、光伏具有较强的随机性与波动性, 提高其预测精度对构建新型电力系统具有重要意义。位于同一地区的风光出力具有明显的时空相关性规律, 鉴于这些相关特征提出了一种基于相似日聚类与基于鲸鱼优化算法的双向长短期记忆神经网络 Copula (whale optimization algorithm-bidirectional long short-term memory neural network, WOA-BiLSTM-Copula) 算法的短期风光功率相关性概率区间预测模型。首先, 采用 K-means 聚类算法划分数值天气预报 (numerical weather prediction, NWP) 数据集, 依据 Kendall、Spearman 相关性系数提取具有相关性的风光联合出力典型相似日场景。其次, 针对相关性相似日场景, 采用非参数核密度估计法进行 Copula 建模, 确定风光出力最优 Copula 函数类型。之后, 训练鲸鱼算法优化的双向长短期记忆神经网络, 并对风电、光伏功率进行点预测。最后, 使用蒙特卡洛法对最优 Copula 函数采样, 基于风光点预测值生成相关性概率预测区间。仿真结果表明所提模型可以有效提取风光出力相关性特征, 与现有模型相比精度更高, 验证了模型的有效性。

关键词: 风光出力预测; 相似日聚类; 最优 Copula 函数建模; 鲸鱼优化算法; 双向长短期记忆神经网络; 概率区间预测

Short-Term Wind and Photovoltaic Power Correlation Probability Interval Prediction Method Based on Similar Day Clustering and WOA-BiLSTM-Copula Algorithm

WANG Lingzi, SHEN Haibo, DENG Liyuan, LIU Xianzhuo, DENG Weisi

(Power Dispatching and Control Center, CSG, Guangzhou 510663, China)

Abstract: Wind and photovoltaic power generation exhibit strong randomness and volatility. Improving the prediction accuracy is of great significance for constructing new power systems. Considering the certain correlation between wind and photovoltaic output in the same region, a short-term wind and photovoltaic power correlation probability interval prediction model based on similar day clustering and whale optimization algorithm-bidirectional long short-term memory neural network-Copula (WOA-BiLSTM-Copula) algorithm is proposed. Firstly, the K-means clustering algorithm is adopted to divide the numerical weather prediction (NWP) dataset, and wind and photovoltaic joint output typical similar day scenarios with correlation are extracted based on Kendall and Spearman correlation coefficients. Secondly, the non-parametric kernel density estimation method is used to establish Copula model for similar day scenes, and the optimal Copula function types of wind and photovoltaic output are determined. Then, the WOA is trained to optimize the BiLSTM and perform point predictions on wind power and photovoltaic power. Finally, the Monte Carlo method is used to sample the optimal Copula function, and the correlation probability prediction interval is generated based on wind and photovoltaic point prediction values. The simulation results show that the proposed model can effectively extract the correlation characteristics of wind and photovoltaic output, and the accuracy is higher than the existing models, which verifies the effectiveness of the model.

Key words: wind and photovoltaic power output prediction; similar day clustering; optimal Copula function modeling; whale optimization algorithm; bidirectional long short-term memory neural network; probability interval prediction

基金项目: 国家自然科学基金资助项目(41875118); 国家自然科学基金青年基金资助项目(41805047)。

Foundation item: Supported by the National Natural Science Foundation of China (41875118); the Youth Program of National Natural Science Foundation of China (41805047).

0 引言

随着环境污染、能源危机等问题的日益凸显,以风电、光电为代表的新能源发电技术凭借清洁、环保等优势发展迅速^[1]。2023年,可再生能源总装机容量年内达到1 450 GW,占全国发电总装机容量比重超过50%,全国风电光伏总装机容量突破1 000 GW,风电、光伏发电量占全社会用电量比重突破15%^[2]。但风能、光伏发电的不确定性和波动性给电力系统的安全运行和调度控制带来了挑战^[3],因此精准预测风电、光伏功率对保持电力系统稳定、提升新能源消纳能力有重要意义^[4]。

风光功率的预测方法按照预测形式主要可以分为点预测与概率预测^[5]。点预测即预测未来一段时间的风光功率期望值,忽略了更有价值的波动范围信息,鲁棒性较差。概率预测通过构造概率密度函数提供未来风光功率可能位于的区间,从而更好地描述其不确定性与未来变化趋势,为决策者提供更加可靠的信息,具有更加实际的意义。

文献[6]针对风电提出一种粒子群优化的核极限学习机区间预测模型,能为风电并网安全稳定运行提供决策支持。文献[7]等人提出一种基于鲁棒多标签生成对抗的风电场日前出力区间预测方法,并用东北某地区风电场数据验证了方法的优越性。文献[8]提出一种基于改进预处理方法和熵加权分位数回归网络(quantile regression-enhanced gated recurrent unit, QR-EGRU)相结合的点区间预测模型,具有较好的预测效果。文献[9]使用ISSA-BiGRU-Attention算法建立短期风功率点预测模型,并使用核密度估计算法生成概率预测区间。文献[10]提出了一种基于CEEMD-SA-DBN的光伏发电区间预测模型,消除了光伏电站地理位置的影响。文献[11]考虑光伏发电出力波动的季节分布特征,提出了一种基于季节分类的光伏发电功率非参数核密度估计区间预测方法。文献[12]针对光伏提出了基于长短时记忆神经网络(long short-term memory, LSTM)的短期区间预测模型,采用Bootstrap方法分析预测结果的误差分布,确定输出的区间范围。文献[13]提出了一种多相关参数情景生成的光伏发电日前区间预测方法,采用K-means聚类算法划分场景集,通过深度卷积生成对抗网络(deep convolution generative adversarial networks,

DCGAN)模型及分位数回归实现区间预测。

现有关于风电或光伏功率的区间预测研究大多基于点预测结果的预测误差分布,采用核密度估计、分位数回归等方法生成误差预测区间。此种方法假设风光发电是相互独立的,将风电、光伏作为独立分布变量,仅考虑各自单独的出力特征进行功率预测。然而,由于同一地区的风电场和光伏电站具有明显的空相关性规律,考虑风光相关性特征可以进一步提高模型预测精度^[14]。Copula理论是研究变量相关性的常用方法^[15]。文献[16]利用Gaussian-Copula函数提取影响风电和光伏发电的关键气象因子并使用LSTM神经网络实现中长期风光功率预测。文献[17-18]则使用Frank-Copula函数描述风光发电的相关性。目前研究中Copula函数建模时仅采用单一Copula模型对全部风光出力进行拟合,未考虑不同气象条件的影响,预测误差较大。

基于此,本文提出了一种基于相似日聚类与WOA-BiLSTM-Copula算法的短期风光功率相关性概率区间预测模型。首先,使用K-Means聚类算法结合相关性系数判断并构建了具有相关性的风光相似日场景集。然后,采用非参数核密度估计法对各相关性相似日场景进行最优Copula建模。最后,训练鲸鱼算法优化的双向长短时记忆神经网络进行风电、光伏功率点预测,并使用蒙特卡洛对最优Copula函数采样生成风光相关性概率预测区间。与现有研究进行了对比分析,验证了本文方法的有效性。

1 基本原理和数学建模

1.1 K-means聚类算法

通过对气象数据聚类预处理划分出具有相关性的风光联合出力相似日,对不同相似日样本数据分别建立模型进行预测,可以提高模型点预测精度。K-means聚类算法的基本思想是:在数据集中选取 k 个点作为聚类中心,按照方差最小标准对数据集中数据进行归类,然后通过迭代得到聚类结果,最终每类数据保持较大的相似度,具有类内紧凑、类间疏远的特点。具体处理步骤如下。

1) 设定输入数据集 $X = \{x_{1i}, x_{2i}, \dots, x_{mi}\}$, $i = 1, 2, \dots, n$, n 为样本数据对象, m 为每个样本数据的特征维度,输出簇为 $C = \{c_1, c_2, \dots, c_k\}$, k 为设定所需聚类数量;

- 2) 从 n 个样本数据中随机选取 k 个初始聚类中心;
- 3) 计算每个样本数据到聚类中心的欧式距离, 并将其划分到距离最近的聚类中心所在类;
- 4) 重新计算 k 个聚类中心;
- 5) 与之前的 k 个聚类中心比较, 若发生变化, 转 3), 否则转 6);
- 6) 若均值不再变化或达到预设迭代次数, 则输出聚类结果。

1.2 基于 WOA-BiLSTM 短期风光功率点预测模型

为了提高 BiLSTM 网络风光功率模型的预测精度, 引入 WOA 算法用于 BiLSTM 网络的超参数寻优, 从而提高风光功率点预测模型的精度。

1.2.1 双向长短期记忆网络

训练长短期记忆网络数据时只能用从前向序列中传来的历史信息, 导致遗忘较早学习的内容。而双向长短期记忆网络 (bidirectional long short-term memory, BiLSTM) 对 LSTM 进行优化改进, 由结构相同方向相反的两个 LSTM 组成, 结构图如图 1 所示, 可以从前后两个方向提取信息, 输出包含了输入数据过去和未来的信息, 从而提高预测精度^[19]。计算公式如下。

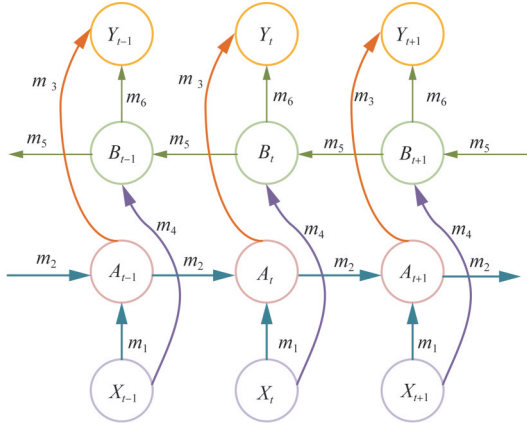


图 1 BiLSTM 神经网络结构

Fig. 1 Structure of BiLSTM memory neural network

$$\begin{cases} A_t = a(m_1 X_t + m_2 A_{t-1}), t = 1, 2, \dots, T_n \\ B_t = b(m_4 X_t + m_5 B_{t+1}), t = 1, 2, \dots, T_n \\ Y_t = c(m_3 A_t + m_6 B_t) \end{cases} \quad (1)$$

式中: X_t 为 t 时刻的输入数据; A_t 、 B_t 分别为 t 时刻向前和向后代代的 LSTM 隐藏状态; A_{t-1} 、 B_{t+1} 分别为在 $t-1$ 时刻的向前和在 $t+1$ 时刻的向后代代的 LSTM 隐藏状态; Y_t 为 t 时刻的输出数据; $m_1 \sim m_6$ 为

各层的权重; a 、 b 、 c 为各层的激活函数; T_n 为时间序列的总长度。

1.2.2 鲸鱼优化算法

鲸鱼优化算法 (whale optimization algorithm, WOA) 是一种新型启发式优化算法, 通过模仿鲸鱼捕猎而提出, 主要包含包围猎物、攻击猎物和寻找猎物。鲸鱼优化算法原理简单、收敛性好、稳定性强, 可为 BiLSTM 的超参数估计提供良好的算法基础^[20-22]。

在 WOA 中, 每只鲸鱼的位置都代表一个可行解。具体算法步骤如下。

1) 包围猎物

鲸鱼在捕食时首先要识别猎物位置并包围, 但猎物位置最初并不确定, 因此将鲸鱼当前最优位置设为目标猎物, 通过不断更新鲸鱼位置来向目标猎物靠近, 数学模型如下所示。

$$\begin{cases} D = |CB^*(t) - B(t)| \\ B(t+1) = B^*(t) - AD \\ A = 2\alpha z_1 - \alpha \\ C = 2z_2 \\ \alpha = 2 - \frac{2t}{T_{\max}} \end{cases} \quad (2)$$

式中: D 为鲸鱼与猎物之间的距离; t 为当前迭代数量; $B^*(t)$ 为鲸鱼当前最优位置向量; $B(t)$ 、 $B(t+1)$ 分别为鲸鱼当前和更新的位置向量; A 、 C 为向量系数。 z_1 、 z_2 为 0~1 之间的随机向量; α 为从 2 到 0 递减的值; $T_{\max}$ 为最大迭代数量。

2) 攻击猎物

当 $|A| < 1$ 时, 鲸鱼攻击方式有两种, 一种是包围攻击, 另一种是以猎物为中心螺旋逼近猎物, 分别有 50% 的概率, 其数学模型为:

$$B(t+1) = \begin{cases} B^*(t) - AD, p < 0.5 \\ B^*(t) + D_p e^{bl} \cos(2\pi l), p \geq 0.5 \end{cases} \quad (3)$$

式中: D_p 为螺旋攻击时使用的距离即鲸鱼位置与最佳猎物之间的距离; b 为定义螺旋线形状的常数; p 为 0 到 1 之间的随机值; l 为 -1 到 1 之间的随机值。

3) 寻找猎物

当 $|A| \geq 1$ 时, 鲸鱼未捕捉到猎物信息, 随机向更远处寻找猎物, 增强了算法的全局搜索能力。其数学模型如下。

$$\begin{cases} D = |CB_t - B(t)| \\ B(t+1) = B_t - AD \end{cases} \quad (4)$$

式中 B_r 为鲸鱼位置的随机向量。

1.3 基于Copula函数的概率区间预测模型

1.3.1 核密度估计法

核密度估计^[23]是估计概率密度函数的方法,不需要事先对样本数据的分布进行假设,可以直接对相似日中风光发电功率历史数据进行非参数估计,设 $f_k(m)$ 、 $f_k(n)$ 分别为点 x 风电概率密度函数估计值和点 y 光伏概率密度函数估计值。核密度估计法通过计算 x 、 y 邻域内各点到 x 、 y 的远近程度确定对 $f_k(m)$ 、 $f_k(n)$ 的贡献程度,公式如下。

$$\begin{cases} f_k(m) = \frac{1}{rk} \sum_{x=1}^{X_R} \Phi\left(\frac{m - M_r}{k}\right) \\ f_k(n) = \frac{1}{rk} \sum_{x=1}^{X_R} \Phi\left(\frac{n - N_r}{k}\right) \end{cases} \quad (5)$$

式中: m 为风机瞬时发电量; n 为光伏瞬时发电量; M_r 为第 r 个相似日逐时风机瞬时发电量; N_r 为第 r 个相似日光伏瞬时发电量; r 为第 r 个相似日; X_R 为相似日样本总数量; k 为带宽; $\Phi(\cdot)$ 为高斯核函数。

1.3.2 最优Copula函数建模

Copula 函数本质上是一种连接函数,可以将各相似日样本数据中风电、光伏联合分布函数与它们各自的边缘分布函数相连接。

对二元随机变量 (U, V) , Copula 函数为:

$$F(u, v) = C(F_U(u), F_V(v)) \quad (6)$$

式中: u 、 v 分别为风电、光伏的出力在各自边缘分布下的累积概率变量; $C(\cdot)$ 为二元 Copula 函数; $F_U(u)$ 为随机变量 U 的边缘累积分布函数; $F_V(v)$ 为随机变量 V 的边缘累积分布函数; $F(u, v)$ 为 (U, V) 的联合累积分布函数。

Copula 函数的类型丰富,常用的有椭圆族 Copula 和阿基米德 Copula,其中椭圆族 Copula 包括正态 Copula、t-Copula 等^[24-26],对应的二元函数 $C(u, v, \rho)$ 、 $C(u, v, \rho, k)$ 形式如下。

$$\begin{cases} C(u, v, \rho) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right] ds dt \\ C(u, v, \rho, k) = \int_{-\infty}^{t_k^{-1}(u)} \int_{-\infty}^{t_k^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1 + \frac{s^2 - 2\rho st + t^2}{k(1-\rho^2)}\right]^{-\frac{k+2}{2}} ds dt \end{cases} \quad (7)$$

式中: ρ 为相关系数; k 为自由度参数; s 、 t 为积分变量; $\Phi^{-1}(u)$ 为标准正态分布的反函数; $t_k^{-1}(u)$ 为 t 分布的反函数。

阿基米德 Copula 包括 Gumbel-Copula、Clayton-Copula、Frank-Copula 等^[27-28],对应的二元函数 $C(u, v, \theta)$ 形式如下。

$$\begin{cases} C(u, v, \theta) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}, \\ \theta \in (0, +\infty) \\ C(u, v, \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}, \theta \in (0, +\infty) \\ C(u, v, \theta) = -\frac{1}{\theta} \ln\left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right], \theta \in R \end{cases} \quad (8)$$

式中: θ 为依赖强度参数; R 为实数集。

为了得到各个相关性相似日场景中最优 Copula 函数类型,使用相关性系数判别法与欧式距离判别法,具体步骤参见文献[29]。

1.3.3 蒙特卡洛采样与概率预测区间生成

本文采用蒙特卡洛法对最优 Copula 函数采样,生成大量的风电、光伏相关性实际值功率样本。结合上文风光点预测值依据条件概率和大数定律分别求解,光伏功率点预测值 v' 条件下风电功率实际值 u 的条件概率,与风电功率点预测值 u' 条件下光伏功率实际值 v 的条件概率^[30]。

以求解风电功率概率区间为例:在 $V \leq v'$ 的条件下, Copula 函数中变量 U 分布函数为:

$$\begin{cases} C(\eta|w') = C(\eta, w')/w' \\ w' = F_V(v') \end{cases} \quad (9)$$

式中: η 为风电功率 u 的分位点; w' 为 v' 所对应的概率值。

对随机变量 u 、 v 求二阶导数得到条件概率密度函数为:

$$\begin{cases} f_{U|V}(u|v') = \frac{c(\eta, w') f_U(u) f_V(v')}{f_V(v')} = c(\eta, w') f_U(u) \\ c(\eta, w') = \partial^2 C(\eta, w') / \partial \eta \partial w' \end{cases} \quad (10)$$

式中: $f_U(u)$ 、 $f_V(v)$ 分别为实际风电功率 U 与点预测光伏功率 V 的边缘概率密度函数; $c(\eta, w')$ 为 Copula 概率密度函数。

则置信水平为 $(1-\alpha)\%$ 的风电功率预测区间为:

$$\begin{cases} P_l(u_i \leq U_i) = F_{U|V}(u_i \leq v') = \frac{\alpha}{2} \\ P_h(u_i \leq U_i) = F_{U|V}(u_i \leq v') = 1 - \frac{\alpha}{2} \end{cases} \quad (11)$$

式中： P_l 、 P_h 分别为风电功率预测的区间下限和上限； u_i 为第*i*时刻风电功率 U_i 在边缘分布下的累积概率； U_i 为第*i*时刻的风电预测功率； F_{UV} 为条件累积分布函数； α 为显著性水平。

2 基于 WOA-BiLSTM-Copula 的区间预测方法

2.1 预测流程

本文使用两步法实现风电与光伏的相关性概率区间预测。首先通过 WOA-BiLSTM 模型实现风电功率与光伏功率的点预测。然后，以点预测结果作为建模对象，采用蒙特卡洛法对最优 Copula 函数采样，分别计算风光实际功率条件概率并求得预测区间，最后结合点预测结果实现区间预测。本文所提方法的相关性概率区间预测流程图如图 2 所示。

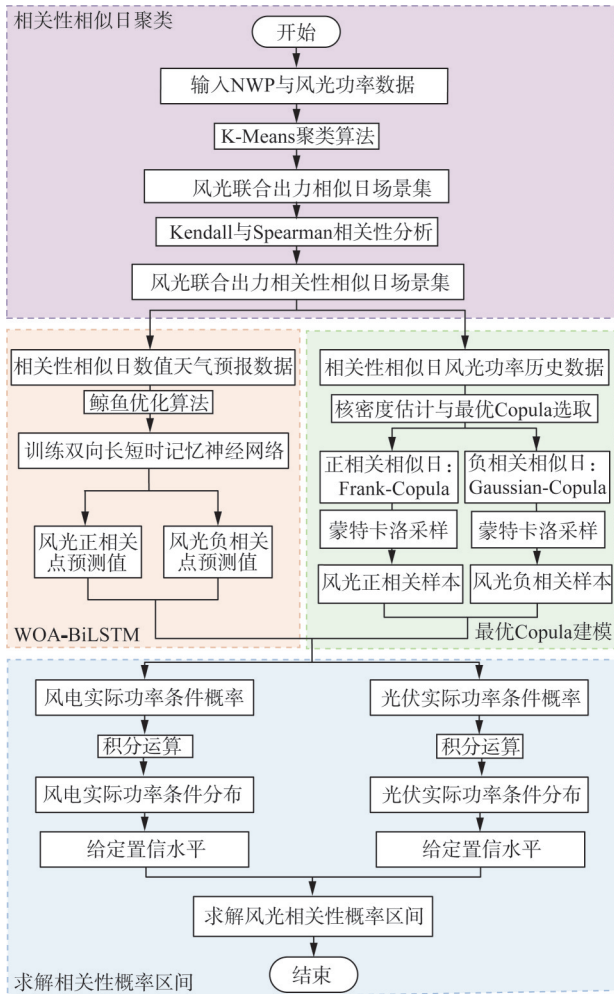


图2 WOA-BiLSTM-Copula 相关性概率区间预测流程图

Fig. 2 Flowchart of WOA-BiLSTM-Copula for correlation probability interval prediction

具体步骤如下。

1) 基于数值天气预报 (numerical weather prediction, NWP) 数据，使用 K-means 聚类算法对全年数据进行划分，构建具风光联合出力相似日场景集。

2) 根据相似日聚类场景结果分别计算各类型中风电、光伏历史功率数据的 Kendall 与 Spearman 相关系数，提取具有相关性的相似日场景。

3) 将 NWP 数据作为输入，使用 WOA 算法优化 BiLSTM 神经网络分别对各相关性相似日类型中风电、光伏功率进行点预测。

4) 使用非参数核密度估计法分别计算各相关性相似日聚类场景集中风电、光伏各自历史功率数据的概率密度函数。

5) 选取 5 种 Copula 函数构建各相关性相似日类型中风电、光伏的联合分布模型。

6) 参考文献[29]中最优 Copula 判断准则使用 Kendall、Spearman 秩相关系数、欧式平方距离做为指标为每一个相关性相似日类型选取最优 Copula 拟合函数。

7) 参考文献[31]使用蒙特卡洛法分别对各个最优 Copula 函数进行 1 000 次采样，根据风电点预测数据计算光伏实际功率条件概率、根据光伏点预测数据计算风电实际功率条件概率。

8) 根据给定置信水平对所得条件概率进行积分，求解得到风电与光伏的预测区间。

2.2 模型评价指标

为验证本文所提模型有效性，参考文献[4, 6]选用平均绝对百分比误差 δ_{MAPE} (mean absolute percentage error, MAPE)、平均绝对误差 δ_{MAE} (mean absolute error, MAE) 作为点预测评价指标，选用真实值落在预测区间内的概率 (prediction interval coverage probability, PICP)、预测区间归一化平均宽度 (prediction interval normalized averaged width, PINAW) 作为区间预测评价指标。

3 算例分析

3.1 数据来源

本文选取我国某风光同场发电站实测风电、光伏功率和数值天气预报数据进行仿真实验。其中数值天气预报数据包括 10 m 风速、30 m 风速、50 m 风速、气压、湿度、总辐射、直接辐射、散射辐

射。该电站风电、光伏装机均为 50 MW, 数据采样间隔为 15 min。为降低输入数据量纲影响, 在训练时进行归一化处理, 并随机选取 20% 数据作为测试集。

3.2 相关性相似日选取

使用 K-Means 算法对全年 NWP 数据聚类, 得到 4 种风光联合出力典型相似日, 各个相似日概率分别为 21 %、14 %、22 %、43 %。不同天气分型下风光出力如图 3 所示。

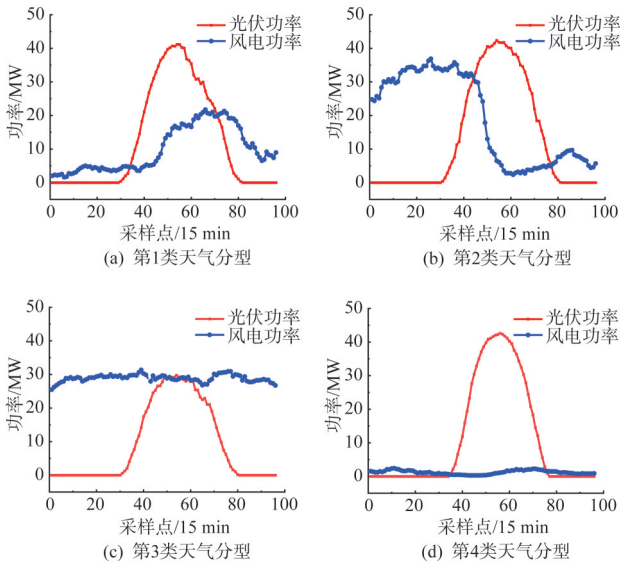


图3 风光联合出力相似日场景图

Fig. 3 Diagram of wind power and photovoltaic power generation in similar day scene

由图 3 可知, 所划分出的 4 种相似日类型表现出不同风光联合出力特征, 验证聚类有效性。在上述 4 种典型相似日类型中, 第一类代表风光叠加场景, 其风光 Kendall、Spearman 相关性系数分别为 0.218、0.307; 第二类代表风光互补场景, 其风光 Kendall、Spearman 相关性系数分别为 -0.187、-0.258; 第三类代表强风弱光场景, 其风光 Kendall、Spearman 相关性系数分别为 0.037、0.049; 第四类代表强光弱风场景, 其风光 Kendall、Spearman 相关性系数分别为 -0.021、-0.021。在前两种相似日类型中, 分别体现风光出力时空的正相关特征与负相关特征。在后两种相似日类型中相关性较弱, 风光出力较为独立。因此, 本文选取风光相关性较强的第一类与第二类相似日进行进一步研究。

3.3 最优 Copula 函数选取

使用所选取的相关性相似日场景中原始风光功率数据分别对 5 种 Copula 函数建模, 各个判别指标计算结果见表 1。

表 1 不同相似日类型最优 Copula 函数判别指标

Tab. 1 Optimal Copula function discriminant for different similar day types

类型	Copula 函数	Kendall	Spearman	欧式平方距离
风光正相关相似日	样本数据	0.217 4	0.306 7	
	Frank	0.216 4	0.320 4	0.093 8
	Clayton	0.327 4	0.470 6	0.381 3
	Gumbel	0.161 0	0.237 7	0.597 1
	Gaussian	0.183 7	0.272 7	0.291 4
	T	0.237 7	0.347 7	0.136 8
风光负相关相似日	样本数据	-0.186 8	-0.258 1	
	Frank	-0.197 5	-0.293 1	0.048 8
	Clayton	$7.254\ 3\times 10^{-7}$	$1.092\ 2\times 10^{-6}$	1.326 6
	Gumbel	$1.357\ 5\times 10^{-6}$	$2.051\ 0\times 10^{-6}$	1.326 6
	Gaussian	-0.173 9	-0.258 4	0.046 1
	T	-0.227 9	-0.332 9	0.087 8

在第一类相关性相似日场景中, Frank-Copula 的 Kendall、Spearman 相关性系数与原始样本最为接近, 且与样本数据的经验 Copula 函数的欧式距离最小, 根据最优 Copula 准则可以得第一类中最优 Copula 函数分别为 Frank-Copula。同理第二类中最优 Copula 函数为 Gaussian-Copula。

3.4 基于 WOA-BiLSTM 模型的点预测结果

区间预测与点预测结果密切相关, 为了验证本文所提出 WOA-BiLSTM 模型点预测性能的优越性, 还建立了基于 LSTM、BiLSTM 预测模型作为对照模型, 对 3 种模型的风电、光伏点预测结果进行了对比分析。其中, LSTM 和 BiLSTM 的神经元数量、学习率、迭代次数分别取 100、0.01、500, 本文所提 WOA-BiLSTM 算法的参数由 WOA 优化算法算出。在 WOA 算法中, 种群规模设置为 50, 收敛判据最大迭代次数为 150 次。各模型误差指标见表 2。

由表 2 可知, WOA-BiLSTM 模型在两种相关性相似日类型中都有较好的预测效果。在正相关相似日中相比于 LSTM 模型, 风电、光伏预测的 MAE 和 RAPE 指标分别降低了 33.39、28.72、36.36、34.69 个百分点。在负相关相似日中, 在

表 2 各模型点预测结果

Tab. 2 Point prediction results of each model

类型	项目	预测误差	LSTM	BiLSTM	WOA-BiLSTM
风光正相关相似日	风电	δ_{MAE}/MW	5.72	4.64	3.81
		$\delta_{MAPE}/\%$	6.72	5.54	4.79
	光伏	δ_{MAE}/MW	2.09	1.64	1.33
		$\delta_{MAPE}/\%$	5.13	4.03	3.35
风光负相关相似日	风电	δ_{MAE}/MW	7.34	6.73	5.25
		$\delta_{MAPE}/\%$	8.16	7.29	5.99
	光伏	δ_{MAE}/MW	2.26	1.92	1.45
		$\delta_{MAPE}/\%$	6.32	5.26	4.18

BiLSTM 模型基础上风电、光伏预测的 MAE 和 RAPE 指标分别又降低了 21.99、17.83、24.47、20.53 个百分点。可以看出本文模型在点预测上具有较高的精度。

3.5 区间预测结果分析

图 4—5 分别为本文 WOA-BiLSTM-Copula 模型在各相关性相似日场景下风电、光伏发电功率实际值预测值及概率区间预测结果图。

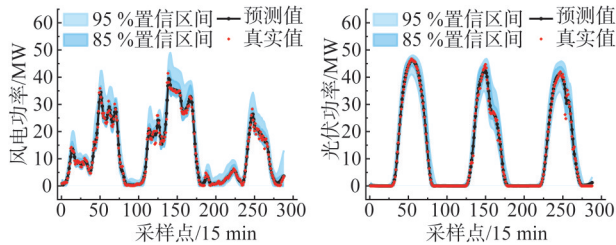


图 4 正相关相似日风光功率概率区间预测结果

Fig. 4 Probability interval prediction results of wind and photovoltaic powers on similar days with positive correlation

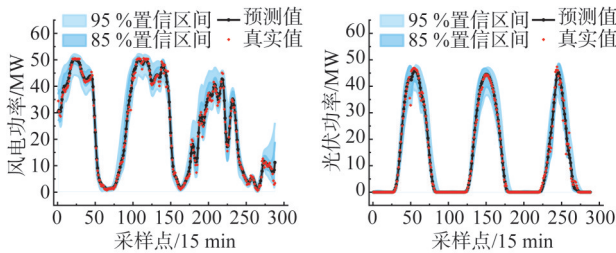


图 5 负相关相似日风光功率概率区间预测结果

Fig. 5 Probability interval prediction results of wind and photovoltaic powers on similar days with negative correlation

由图 4—5 可以看出，在风光具有明显相关性的两类相似日场景中，本文方法整体上能够紧密地跟随风光功率波动趋势，具有良好的预测效果。在

正相关相似日场景预测中，概率预测结果具有充足的正向偏差，用于保证风光功率正向波动时的覆盖率。在负相关相似日场景预测中，概率预测结果负向偏差占比增加，合理增加平均带宽，使得在风光剧烈波动时仍具有良好的预测效果。

为进一步验证本文所提方法在风光相关性场景下的预测优越性，将其与传统误差分布拟合方法进行对比。本文方法和核密度估计法在 85 %、95 % 置信水平下评价指标的预测结果，如表 3 所示。

表 3 各模型区间预测结果

Tab. 3 Interval prediction results of each model

类型	方法	项目	评价指标	85% 置信区间	95% 置信区间
风光正相关相似日	本文方法	风电	PICP 值/%	85.12	93.33
			PINAW 值/MW	16.61	20.13
		光伏	PICP 值/%	87.67	95.63
	核密度估计法	风电	PINAW 值/MW	7.59	14.94
			PICP 值/%	83.62	91.77
		光伏	PINAW 值/MW	18.25	22.43
风光负相关相似日	本文方法	风电	PICP 值/%	85.69	94.81
			PINAW 值/MW	9.27	15.46
		光伏	PICP 值/%	82.72	91.25
	核密度估计法	风电	PINAW 值/MW	19.51	24.26
			PICP 值/%	86.46	94.58
		光伏	PINAW 值/MW	7.82	10.78

由表 3 可知，在两种相关性相似日场景下，风电出力相较于光伏出力具有更大的波动性，因此，光伏预测区间具有更窄的平均带宽和更高的区间覆盖率。相比于核密度估计法，本文方法在不同场相似日场景及不同置信度下均可以很好地预测风电、光伏的波动出力区间，具有较高的可靠性。

4 结语

本文考虑风电与光伏出力的日相关性,提出基于WOA-BiLSTM-Copula模型的区间预测方法,得出以下主要结论。

1)使用K-Means算法对NWP数据进行聚类,提取风光联合出力相关性相似日数据集,提高了数据相关性特征与提高Copula建模的精准度。

2)训练WOA算法优化的BiLSTM神经网络进行风光功率预测,实现对风光出力时序关联性捕捉,使点预测结果更加精确、更具可靠性。

3)基于Copula函数的概率区间预测可以更好挖掘风电、光伏数据相关性特征,相较于现有方法具有更小的更窄的平均带宽和更高的区间覆盖率,验证了算法有效性。

参考文献

- [1] 钟嘉庆,李茂林,江静,等.基于Copula理论的风光出力预测误差分析方法的研究[J].电工电能新技术,2017,36(6): 39-46.
ZHONG Jiaqing, LI Maolin, JIANG Jing, et al. Method of wind/solar output forecast error analysis based on Copula theory[J]. Advanced Technology of Electrical Engineering and Energy, 2017, 36(6): 39-46.
- [2] 国家能源局.新华社:我国可再生能源发电总装机占比超过50%[EB/OL]. [2023-12-22]. <https://baijiahao.baidu.com/s?id=1785953559130581645&wfr=spider&for=pc>.
- [3] 盛四清,张立.考虑风光荷预测误差的电力系统经济优化调度[J].电力系统及其自动化学报,2017,29(9): 80-85.
SHENG Siqing, ZHANG Li. Economic dispatch of power system considering the prediction error of wind power, solar energy and load[J]. Proceedings of the CSU-EPSSA, 2017, 29(9): 80-85.
- [4] 商立群,李洪波,黄辰浩,等.基于多变量相空间重构和优化深度极限学习机的短期风电功率预测[J].南方电网技术,2023,17(2): 82-91.
SHANG Liqun, LI Hongbo, HUANG Chenhao, et al. Short-term wind power prediction based on multivariate phase space re-construction and optimized deep extreme learning machine[J]. Southern Power System Technology, 2023, 17(2): 82-91.
- [5] 吴问足,乔颖,鲁宗相,等.风电功率概率预测方法及展望[J].电力系统自动化,2017,41(18): 167-175.
WU Wenzu, QIAO Ying, LU Zongxiang, et al. Methods and Prospects for probabilistic forecasting of wind power[J]. Automation of Electric Power Systems, 2017, 41(18): 167-175.
- [6] 杨锡运,关文渊,刘玉奇,等.基于粒子群优化的核极限学习机模型的风电功率区间预测方法[J].中国电机工程学报,2015,35(S1): 146-153.
YANG Xiyun, GUAN Wenyuan, LIU Yuqi, et al. Prediction intervals forecasts of wind power based on PSO-KELM[J]. Proceedings of the CSEE, 2015, 35(S1): 146-153.
- [7] 潘霄,张明理,刘德宝,等.基于鲁棒多标签生成对抗的风电场日前出力区间预测[J].电力系统自动化,2022,46(10): 216-223.
PAN Xiao, ZHANG Mingli, LIU Debao, et al. Interval prediction of wind farm day-ahead output based on robust multi-label generative adversarial[J]. Automation of Electric Power Systems, 2022, 46(10): 216-223.
- [8] LIU T, QI S, QIAO X, et al. A hybrid short-term wind power point-interval prediction model based on combination of improved preprocessing methods and entropy weighted GRU quantile regression network[J]. Energy, 2024(288): 1-17.
- [9] CUI X, YU X, NIU D. The ultra-short-term wind power point-interval forecasting model based on improved variational mode decomposition and bidirectional gated recurrent unit improved by improved sparrow search algorithm and attention mechanism[J]. Energy, 2024(288): 1-19.
- [10] 杨茂,王凯旋.基于CEEMD-DBN模型的光伏出力日前区间预测[J].高电压技术,2021,47(4): 1156-1164.
YANG Mao, WANG Kaixuan. Day-ahead interval forecasting of PV power based on CEEMD-DBN model[J]. High Voltage Engineering, 2021, 47(4): 1156-1164.
- [11] SHI M, YIN R, YIN W, et al. Photovoltaic power interval forecasting method based on kernel density estimation algorithm[J]. IOP Conference Series: Earth and Environmental Science, 2020, 615(1): 5103-5114.
- [12] 宋绍剑,罗世坚,李国进,等.基于MPA-LSTM模型和Bootstrap方法的短期光伏功率区间预测[J].广西大学学报(自然科学版),2022,47(4): 986-997.
SONG Shaojian, LUO Shijian, LI Guojin, et al. A short-term PV power interval forecasting based on MPA-LSTM network model and Bootstrap method[J]. Journal of Guangxi University (Natural Science Edition), 2022, 47(4): 986-997.
- [13] WANG Z, WANG C, CHENG L, et al. An approach for day-ahead interval forecasting of photovoltaic power: a novel DCGAN and LSTM based quantile regression modeling method[J]. Energy Reports, 2022(8): 14020-14033.
- [14] MIN H, HE P, LI C, et al. The Temporal and spatial characteristics of wind-photovoltaic-hydro hybrid power output based on a cloud model and copula function[J]. Energies, 2024, 17(5): 1024-1035.
- [15] LI F, ZHANG J, LI G, et al. Cost reduction of a hybrid energy storage system considering correlation between wind and PV power[J]. Protection and Control of Modern Power Systems, 2016, 1(1): 1-9.
- [16] HAN S, QIAO Y, YAN J, et al. Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network[J]. Applied Energy, 2019(239): 181-191.
- [17] GAO J, YANG Y, GAO F, et al. Optimization of electric vehicles based on Frank-Copula-Glue CVaR combined wind and photovoltaic output scheduling research[J]. Energies, 2021, 14(19): 6080-6090.
- [18] GAO C, LIN J, ZENG F, et al. Wind-photovoltaic co-generation prediction and energy scheduling of low-carbon complex regional integrated energy system with hydrogen industry chain based on copula-MILP[J]. Applied Energy, 2022(328): 1-22.
- [19] 张晓艳,向勉,朱黎,等.基于MLP-BiLSTM-TCN组合的超

- 短期风电功率预测[J]. 湖北民族大学学报(自然科学版), 2023, 41(4): 513–519, 529.
- ZHANG Xiaoyan, XIANG Mian, ZHU Li, et al. Ultra short term wind power prediction based on MLP-BiLSTM-TCN combination[J]. Journal of Hubei Minzu University (Natural Science Edition), 2023, 41(4): 513–519, 529.
- [20] 谢丽蓉, 王斌, 包洪印, 等. 基于EEMD-WOA-LSSVM的超短期风电功率预测[J]. 太阳能学报, 2021, 42(7): 290–296.
- XIE Lirong, WANG Bin, BAO Hongyin, et al. Super-short-term wind power forecasting based on EEMD-WOA-LSSVM[J]. Acta Energaie Solaris Sinica, 2021, 42(7): 290–296.
- [21] 王东风, 张中印, 顾智勇, 等. 基于混合Copula函数及鲸鱼优化算法的风电场风速预测[J]. 电力科学与工程, 2022, 38(3): 33–40.
- WANG Dongfeng, ZHANG Zhongyin, GU Zhiyong, et al. Wind speed prediction of wind farm based on hybrid copula function and whale optimization algorithm[J]. Electric Power Science and Engineering, 2022, 38(3): 33–40.
- [22] 薛阳, 李金星, 杨江天, 等. 基于相似日分析和改进鲸鱼算法优化LSTM网络模型的光伏功率短期预测[J]. 南方电网技术, 2024, 18(11): 97–105.
- XUE Yang, LI Jinxing, YANG Jiangtian, et al. Short-term prediction of photo-voltaic power based on similar day analysis and improved whale algorithm to optimize LSTM network model[J]. Southern Power System Technology, 2024, 18(11): 97–105.
- [23] 宋宇, 李涵. 基于核密度估计和Copula函数的风、光出力场景生成[J]. 电气技术, 2022, 23(1): 56–63.
- SONG Yu, LI Han. Typical scene generation of wind and photovoltaic power output based on kernel density estimation and Copula function[J]. Electrical Engineering, 2022, 23(1): 56–63.
- [24] 邱宜彬, 李诗涵, 刘璐, 等. 基于场景D藤Copula模型的多风电场出力相关性建模[J]. 太阳能学报, 2019, 40(10): 2960–2966.
- QIU Yibin, LI Shihan, LIU Lu, et al. Correlation modeling of power output among multiple wind farms based on scenario-divine copula method[J]. Acta Energaie Solaris Sinica, 2019, 40(10): 2960–2966.
- [25] 杨天玥. 基于混合Copula函数的风电机组功率曲线多维度异常识别研究[D]. 沈阳: 沈阳工业大学, 2021.
- [26] 徐玉琴, 张林浩. 基于Copula函数的风速相关性建模及概率最优潮流分析[J]. 华北电力大学学报(自然科学版), 2013, 40(5): 54–59.
- XU Yuqin, ZHANG Linhao. Modeling of wind speed correlation and analysis of probabilistically optimal power flow based on Copula function[J]. Journal of North China Electric Power University, 2013, 40(5): 54–59.
- [27] 解陈力. 基于风电出力特性和Copula理论的风电时间序列生成[D]. 武汉: 华中科技大学, 2015.
- [28] 任相霖, 张粒子, 黄弦超. 计及预测误差时变相关特性的新型电力系统爬坡容量需求分析方法[J]. 南方电网技术, 2024, 18(1): 49–57.
- REN Xianglin, ZHANG Lizi, HUANG Xianchao. Analysis method for ramp capacity demand of new power system considering time-variant correlation characteristics of forecast errors[J]. Southern Power System Technology, 2024, 18(1): 49–57.
- [29] 林顺富, 刘持涛, 李东东, 等. 考虑电能交互的冷热电区域多微网系统双层多场景协同优化配置[J]. 中国电机工程学报, 2020, 40(5): 1409–1421.
- LIN Shunfu, LIU Chitao, LI Dongdong, et al. Bi-level multiple scenarios collaborative optimization configuration of CCHP regional multi-microgrid system considering power interaction among microgrids[J]. Proceedings of the CSEE, 2020, 40(5): 1409–1421.
- [30] 黄牧涛, 高素花, 汪旸, 等. 基于BRICH聚类 and Copula-Monte Carlo模拟的光伏发电功率概率预测[J]. 水电能源科学, 2023, 41(12): 220–224.
- HUANG Mutao, GAO Suhua, WANG Yanget al. Photovoltaic power probabilistic prediction based on BRICH clustering and Copula-Monte Carlo simulation[J]. 水电能源科学, 2023, 41(12): 220–224.
- [31] 杨茂, 董昊. 基于数值天气预报风速和蒙特卡洛法的短期风电功率区间预测[J]. 电力系统自动化, 2021, 45(5): 79–85.
- YANG Mao, DONG Hao. Short-term wind power interval prediction based on wind speed of numerical weather prediction and Monte Carlo method[J]. Automation of Electric Power Systems, 2021, 45(5): 79–85.

收稿日期: 2024-02-01;网络首发日期: 2024-11-20

作者简介:

王凌梓(1991), 女, 工程师, 硕士, 研究方向为电力气象应用, wanglz@csg.cn;

沈海波(1992), 男, 通信作者, 工程师, 博士, 研究方向为气候动力学、电力气象应用, shenhb1@csg.cn;

邓力源(1994), 女, 工程师, 博士, 研究方向为台风动力学、电力气象防灾减灾, dengly3@csg.cn。