文章编号:1674-2974(2024)02-0068-13 DOI:10.16339/j.cnki.hdxbzkb.2024227

基于新型相似日选取和VMD-NGO-BiGRU的短期光伏功率预测

王瑞,张璐婷,逯静†
(河南理工大学 计算机科学与技术学院,河南 焦作 454000)

摘 要:光伏功率预测在现代电力系统调度和运行中起着重要作用.针对光伏发电功率的多变性和复杂性,提出了一种基于新型相似日选取和北方苍鹰算法(Northern Goshawk Optimization, NGO)优化双向门控循环单元(Bidirectional Gated Recurrent Unit, BiGRU)的短期光伏功率预测方法.首先,利用斯皮尔曼相关系数选取主要气象因子,通过变分模态分解(Variational Mode Decomposition, VMD)将原始光伏功率和最大气象因子分解重构为一系列子信号.其次,通过构建新的评价指标筛选出相似日数据集,利用一组BiGRU建立以相似日子信号为网络输入的深度学习模型,并利用NGO对每个BiGRU网络的超参数进行有效优化.最后,对各子信号的预测结果进行综合,得到最终的光伏功率预测值.仿真结果表明,所提混合深度学习方法在预测精度和计算效率方面均优于其他方法.

关键词:光伏功率预测;变分模态分解;双向门控循环单元;北方苍鹰算法
中图分类号:TM615 **文献标志码**:A

Short Term Photovoltaic Power Prediction Based on New Similar Day Selection and VMD-NGO-BiGRU

WANG Rui, ZHANG Luting, LU Jing†
(School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China)

Abstract: Photovoltaic power prediction plays an important role in the scheduling and operation of modern power systems. Aiming at the variability and complexity of photovoltaic power generation, a short-term PV power prediction method based on new similar day selection and northern Goshawk optimization (NGO) to optimize bidirectional gated recurrent unit (BiGRU) is proposed. The main meteorological factors are selected with the Spearman correlation coefficient, and the original PV power and maximum meteorological factor are decomposed into a series of sub-signals by variational mode decomposition (VMD). Then, according to the construction of new evaluation indicators, the data set of similar days is screened out, a group of BiGRU is used to establish a deep learning model with similar day signals as network input, and NGO is used to effectively optimize the hyperparameters of each BiGRU network. Finally, the predicted value of PV power is obtained by synthesizing the predicted results of each sub-

* **收稿日期**:2023-07-29
基金项目:国家自然科学基金资助项目(62273133),National Natural Science Foundation of China(62273133);河南省科技攻关项目(222102210120),Scientific and Technological Breakthrough Foundation of Henan province(222102210120)
作者简介:王瑞(1977—),男,河北内丘人,河南理工大学 副教授,硕士
† 通信联系人,E-mail: lujing@hpu.edu.cn

signal. Simulation results show that the proposed hybrid deep learning method is superior to other methods in terms of prediction accuracy and computational efficiency.

Key words: photovoltaic power prediction; variational mode decomposition; bidirectional gated cycle unit; northern Goshawk algorithm

太阳能是可再生能源中最丰富、发展潜力最大的清洁能源,其开发与利用已经成为我国能源结构调整的重要方向[1].光伏发电具有间歇性、随机性和波动性,由此给电网的安全运行带来了一系列问题,受到了各界的广泛关注[2-3].随着新型电力系统建立,电网建设滞后于光伏电力的生产,“弃光”、“限电”问题加重,可能降低原始数据在预测模型中的可信度,从而进一步增加短期的预测难度[4].在“碳达峰”“碳中和”的战略背景下,我国电力结构清洁化趋势将更加凸显[5],光伏并网对电网的影响也越来越显著.开展光伏发电功率超短期预测,能有效降低光伏出力不确定性给电网带来的影响,提高光伏发电系统的可靠性[6].

预测光伏功率的方法多种多样,叶林等[7]基于气象因素与光伏功率波动特征的关联性,将天气过程划分为5种类型,并基于变分模态分解算法,将光伏功率分解为类晴空过程和波动过程,构建的预测模型提高了光伏功率预测准确性,验证了对天气进行划分的有效性.张金金等[8]通过模糊C均值聚类算法对数据进行聚类,获得较好的数据分类结果,但未考虑气象因素的影响.王福忠等[9]提出在第一阶段采用变分模态分解(VMD)分解原始功率并建立LSTM模型进行预测,第二阶段对误差序列进行误差补偿预测,与上一阶段预测结果叠加,得到最终预测结果.孟安波等[10]考虑到不同天气因素影响,通过加权模糊聚类对天气类型进行划分,筛选相似日数据集,采用纵横交叉算法优化改进的高斯过程回归模型的超参数,对光伏功率进行预测.杨国华等[11]采用模糊C均值聚类算法选取相似日,结合灰狼优化算法优化小波包神经网络和AdaBoost算法,建立预测模型.

基于上述分析,虽然预测精度和效率与早期的预测模型相比有了很大的突破,但在处理不同气候条件时,光伏功率预测模型的预测性能和鲁棒性仍然存在不足.为了从大量数据中选取与预测日相似的历史日,以及考虑到相关性越强的影响因素对于光伏功率预测结果影响越大,本文提出一种基于相似日选取,结合利用北方苍鹰优化算法优化BiGRU参数的混合预测模型.首先,通过斯皮尔曼相关系数法筛选出影响最大的气象因子,通过VMD分解将原始光伏功率和影响最大的气象因子分解重构为一系列子信号;其次,将气象因子分解后的分量重构,并对影响因素重新定义,根据构建新的评价指标筛选出相似日数据集;然后利用一组BiGRU,建立以相似日子信号为网络输入的深度学习模型,并利用NGO对每个BiGRU网络的超参数进行有效优化.最后通过算例仿真,证明了该模型的有效性.

1　算法原理

1.1　光伏发电功率影响因素

光伏功率主要受到辐照度、环境温度、湿度、气压等多种气象因素影响[12],不同的气象因素对光伏发电功率的影响程度也不同.对影响光伏发电的变量进行相关性分析,筛选出主要特征.

通常采用的皮尔逊相关系数需要满足正态分布,但是对于波动性比较强的光伏发电功率,大多数影响因素并不符合正态分布.图1是光伏功率的概率密度分布图,图中概率密度曲线为非正态分布,因此斯皮尔曼相关系数是相对较好的选择.斯皮尔曼相关系数的定义如下:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n\left(n^2 - 1\right)} \tag{1}$$

式中:d_i表示第i个数据对应的位次值之差,n表示总的观测样本数.

斯皮尔曼相关系数可以衡量两个变量之间的相关性,它将数据按序给出等级,若有数据数值相同时,则将计算数值等级的平均值作为等级数.同理,计算另一组数据的等级排序,得到等级大小,之后计算两组数据等级差,代入式(1),即可得出相关系数,可以方便应用于非正态性的数据.

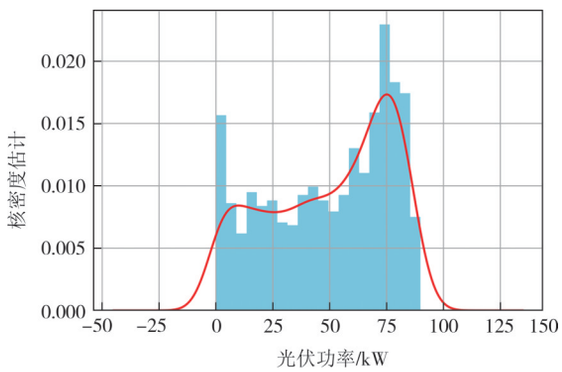


图 1 光伏功率的概率密度分布图

Fig.1 Probability density distribution diagram of photovoltaic power

本文选取 2022 年江苏某光伏电站实际数据进行相关性分析,影响因素有总辐射、散射辐射、环境温度、气压以及湿度.根据斯皮尔曼相关系数计算出各因素与光伏功率的相关系数.具体数据如表 1 所示.

表 1 各因素与功率的斯皮尔曼相关系数

Tab.1 Spearman correlation coefficient between various factors and power

影响因素	相关系数
总辐射	0.864
散射辐射	0.861
环境温度	0.111
气压	−0.012
湿度	−0.05

1.2 变分模态分解

VMD 为避免信号分解过程中出现模态混叠,在计算本征模态分量时舍弃了传统信号分解算法所使用的递归求解的思想,采用完全非递归的模态分解.与传统信号分解算法相比,VMD 拥有非递归求解和自主选择模态个数的优点,其核心思想是构建和求解变分问题[13].

1)构建变分约束模型:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_k \left\| \partial(t) \left[\left(\delta(t) + \dfrac{\mathrm{j}}{\pi t} \right) u_k(t) \right] \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \sum\limits_k u_k = f(t) \end{cases} \tag{2}$$

式中:$f(t)$为原始功率信号;u_k为各分量;ω_k为各分量对应中心频率;$\delta(t)$为脉冲信号.

2)通过引入拉格朗日算子λ和二次惩罚因子α,求解约束变分问题:

$$\begin{aligned} &L\left(\{u_k\},\{\omega_k\},\lambda\right) = \\ &\alpha \sum_k \left\| \left(\partial(t) \left[\left(\delta(t) + \frac{\mathrm{j}}{\pi t} \right) u_k(t) \right] \mathrm{e}^{-\mathrm{j}\omega_k t} \right\|_2^2 + \right. \\ &\left. \left\| f(t) - \sum_k u_k \right\|_2^2 + \left[\lambda(t), f(t) - \sum_k u_k \right] \right) \end{aligned} \tag{3}$$

对各分量u_k及ω_k通过交替方向乘子法交替更新,完成分解.

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum\limits_{i \neq k} \hat{u}_i(\omega) + \dfrac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha\left(\omega - \omega_k\right)^2} \tag{4}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left| \hat{u}_k^{n+1}(\omega)^2 \right| d\omega}{\int_0^\infty \left| \hat{u}_k^{n+1}(\omega)^2 \right| d\omega} \tag{5}$$

式中:n为迭代次数,$\hat{f}(\omega)$、$\hat{u}_i(\omega)$、$\hat{\lambda}(\omega)$、$\hat{u}_k^{n+1}$分别代表$f(t)$、$u(t)$、$\lambda(t)$、$u_k^{n+1}(t)$的傅里叶变换.

1.3 双向门控循环单元

BiGRU 模型采用了循环神经网络(RNN)的结构,其中每个 RNN 单元都是由一对门控循环单元(Gated Recurrent Unit,GRU)组成.与长短期记忆网络(Long Short-Term Memory,LSTM)相比,GRU 不仅减少了网络参数,而且保持了 LSTM 的良好性能[14].GRU 的结构如图 2 所示,计算公式如下:

$$r_t = \sigma\left(W_r x_t + U_r h_{t-1}\right) \tag{6}$$

$$z_t = \sigma\left(W_z x_t + U_z h_{t-1}\right) \tag{7}$$

$$\tilde{h}_t = \varphi\left(r_t \otimes U_h h_{t-1} + W_h x_t\right) \tag{8}$$

式中,r_t、z_t表示重置门和更新门;σ为 Sigmoid 激活函数;W和U表示权重矩阵;$\tilde{h}_t$表示中间记忆状态;$\varphi(\cdot)$表示非线性激活函数;$\otimes$表示哈达马积;h_{t-1}表示上一时刻的隐藏状态信息;x_t表示输入序列中的第t个元素.

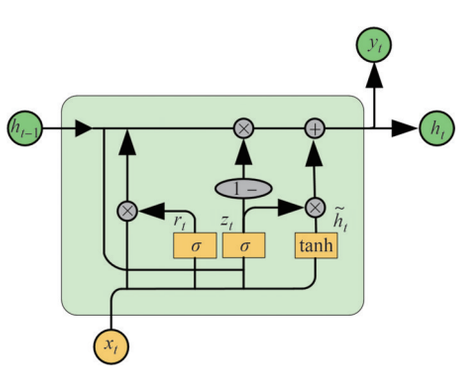


图 2 GRU 网络结构图

Fig.2 GRU network structure diagram

BiGRU由两个方向相反的GRU组成[15],在本模型中采用单层BiGRU.在BiGRU模型中,输入序列首先被分别输入到两个方向的RNN中.每个神经网络都包含一个隐藏层,用于存储序列中的信息.在前向RNN中,输入序列从左到右进行处理,每个时间步的输出结果都会被传递到下一个时间步.在后向RNN中,输入序列从右到左进行处理,每个时间步的输出结果也会被传递到下一个时间步.这样,每个元素在正向和反向的GRU中都被处理了一次,以提取更全面的特征信息.BiGRU结构如图3所示.

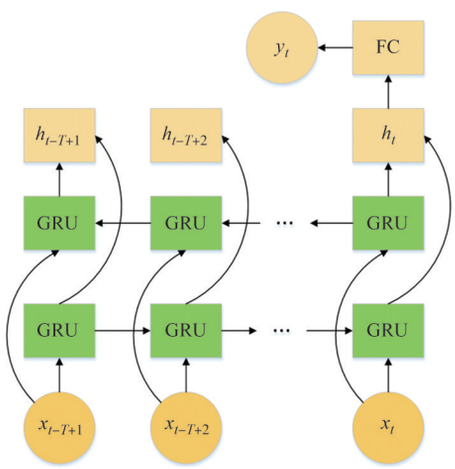


图3　BiGRU网络模型

Fig.3　BiGRU network model

BiGRU的计算公式如下:

$$\begin{cases}\overrightarrow{h}_t = \mathrm{GRU}(x_t, \overrightarrow{h}_{t-1}) \\ \overleftarrow{h}_t = \mathrm{GRU}(x_t, \overleftarrow{h}_{t+1}) \\ h_t = \alpha_t \overrightarrow{h}_t + \beta_t \overleftarrow{h}_t + b_t\end{cases} \tag{9}$$

式中:GRU(·)表示门控循环单元;x_t为输入;$\overrightarrow{h}_t$、$\overleftarrow{h}_t$分别为前向和后向隐藏层输出状态;α_t和β_t分别表示前向和后向隐藏层的输出权重;$\overrightarrow{h}_{t-1}$表示前向隐藏状态在前一个时刻的隐藏状态;$\overleftarrow{h}_{t+1}$表示后向隐藏状态在后一个时刻的隐藏状态;b_t为偏置值.

1.4　北方苍鹰算法

北方苍鹰优化算法于2022年提出,该算法模拟了北方苍鹰在捕猎过程中的行为,具体包括猎物识别与攻击、追逐及逃生等行为.北方苍鹰优化算法可包含两个阶段[16]:第一阶段,在识别出猎物后,它以高速向猎物移动攻击;在第二阶段,追逐猎取猎物.

1.4.1　初始化

初始化种群.以下矩阵代表的是北方苍鹰的种群数量矩阵:

$$\boldsymbol{X} = \begin{bmatrix}\boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_i \\ \vdots \\ \boldsymbol{X}_N\end{bmatrix}_{N\times m} = \begin{bmatrix}x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,m} \\ \vdots & & \vdots & & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,m} \\ \vdots & & \vdots & & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,m}\end{bmatrix}_{N\times m} \tag{10}$$

式中:$\boldsymbol{X}$为北方苍鹰的种群矩阵;$\boldsymbol{X}_i$为第i个北方苍鹰的位置;$x_{i,j}$为第i个北方苍鹰的第j维的位置;N为北方苍鹰的种群数量;m为求解问题的维度.

在北方苍鹰优化算法中,求解问题的目标函数可以用来计算北方苍鹰的目标函数值,表示如下:

$$F = \begin{bmatrix}F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N\end{bmatrix}_{N\times 1} = \begin{bmatrix}F(\boldsymbol{X}_1) \\ \vdots \\ F(\boldsymbol{X}_i) \\ \vdots \\ F(\boldsymbol{X}_N)\end{bmatrix}_{N\times 1} \tag{11}$$

式中:F为北方苍鹰种群的目标函数向量;F_i为第i个北方苍鹰的目标函数值.

1.4.2　识别猎物及攻击猎物

北方苍鹰在捕猎的第一阶段是随机选择一个猎物,并进行攻击.由于在搜索空间中对猎物的选择是随机的,因此这一阶段增强了NGO算法的勘探能力.该阶段对搜索空间进行全局搜索,以确定最优区域.在这一阶段,北方苍鹰进行猎物选择和攻击的行为,用式(12)~(14)描述:

$$P_i = \boldsymbol{X}_k (i = 1, 2, \cdots, N;\ k = 1, 2, \cdots, i-1, \cdots, N) \tag{12}$$

$$x_{i,j}^{\mathrm{new},P_1} = \begin{cases} x_{i,j} + r(P_{i,j} - Ix_{i,j}), F_{P_i} < F_i \\ x_{i,j} + r(x_{i,j} - P_{i,j}), F_{P_i} \geqslant F_i \end{cases} \tag{13}$$

$$\boldsymbol{X}_i = \begin{cases} \boldsymbol{X}_i^{\mathrm{new},P_1}, F_i^{\mathrm{new},P_1} < F_i \\ \boldsymbol{X}_i, F_i^{\mathrm{new},P_1} \geqslant F_i \end{cases} \tag{14}$$

式中:P_i为第i个北方苍鹰的猎物的位置;F_{P_i}为第i个北方苍鹰的猎物的位置的目标函数值;k是[1,N]范围内的随机整数;X_i^{new,P_1}为第i个北方苍鹰的新位置;$x_{i,j}^{\mathrm{new},P_1}$为第$i$个北方苍鹰的第$j$维的新位置;$F_i^{\mathrm{new},P_1}$为基于第一阶段更新后第$i$个北方苍鹰的目标函数值;$r$是[0,1]范围内的随机数;$I$为1或2的随机整数.

1.4.3　追逐及逃生

猎物受到北方苍鹰攻击会试图逃跑.因此,在追逐猎物的收尾过程中,北方苍鹰需要继续追逐猎物.北方苍鹰高速追击并抓捕猎物行为的模拟,提高了算法的局部搜索能力.假设这种狩猎活动接近于一个半径为R的攻击位置.在第二阶段中,用式(15)~

式(17)描述:

$$x_{i,j}^{\text{new},P_2} = x_{i,j} + R(2r-1)x_{i,j} \tag{15}$$

$$R = 0.02\left(1 - \frac{t}{T}\right) \tag{16}$$

$$\boldsymbol{X}_i = \begin{cases} \boldsymbol{X}_i^{\text{new},P_2}, F_i^{\text{new},P_2} < F_i \\ \boldsymbol{X}_i, F_i^{\text{new},P_2} \geqslant F_i \end{cases} \tag{17}$$

式中:$x_{i,j}^{\text{new},P_2}$为基于第二阶段更新后第i个北方苍鹰的第j维的位置;t为当前迭代次数;T为最大迭代次数;$\boldsymbol{X}_i^{\text{new},P_2}$为第$i$个北方苍鹰的新位置;$F_i^{\text{new},P_2}$为基于第二阶段更新后第$i$个北方苍鹰的目标函数值.

2 基于灰色关联度的相似日选取

2.1 灰色关联度

由于不同气象因素对光伏功率的影响程度不同,在构建新的综合评价指标时,应该对影响程度更大的气象因素赋予更大的权重.而灰色关联度就是一种有效确定权重的方法[17].

灰色关联度是基于各因素之间发展趋势的相似性和差异性,为衡量各因素之间的关联程度提供一种定量的度量.样本量多少、有无规律它都同样适用,而且计算量小,非常方便,并且量化结果与定性分析结果之间不会出现不一致的情况.定义关联系数为

$$\xi_i(k) = \frac{\min\limits_i \min\limits_k \left|x_0^{\,k} - x_i^{\,k}\right| + \rho \max\limits_i \max\limits_k \left|x_0^{\,k} - x_i^{\,k}\right|}{\left|x_0^{\,k} - x_i^{\,k}\right| + \rho \max\limits_i \max\limits_k \left|x_0^{\,k} - x_i^{\,k}\right|} \tag{18}$$

式中:ρ为分辨系数,取0.5;min min是两级最小差,max max是两级最大差;$x_0^{\,k}$代表第k个时刻光伏功率序列的值,$x_i^{\,k}$表示第k个时刻各个影响因素的值.

具体计算过程如下:

1)原始数据归一化处理.

$$x'_0(k) = \frac{x_0^{\,k}}{\overline{x_0}},\ x'_i(k) = \frac{x_i^{\,k}}{\overline{x_i}},\ k = 1,2,\cdots,n \tag{19}$$

$$i = 1,2,\cdots,m$$

式中:$\overline{x_0}$与$\overline{x_i}$分别为光伏功率序列和第i个影响因素的平均值,n为样本数量,m为影响因素个数.

2)求绝对差值矩阵.将高频分量、低频分量和复合天气因素与原始功率序列相应的元素做差后取绝对值,形成绝对差值矩阵.

$$\boldsymbol{\Delta}_i(k) = \left|x'_0(k) - x'_i(k)\right| \tag{20}$$

3)求矩阵关联系数.

$$\xi_i(k) = \frac{\boldsymbol{\Delta}_{\min} + \rho\boldsymbol{\Delta}_{\max}}{\boldsymbol{\Delta}_i(k) + \rho\boldsymbol{\Delta}_{\max}} \tag{21}$$

式中:$\boldsymbol{\Delta}_{\min}$与$\boldsymbol{\Delta}_{\max}$分别为绝对差值矩阵$\boldsymbol{\Delta}$的最小值与最大值.

4)计算关联度.根据关联系数确定关联度,即

$$r_i = \frac{1}{n}\sum_{k=1}^{n}\xi_i(k) \tag{22}$$

式中:r_i表示第i个影响因素与功率的关联度.

2.2 相似日的选取

首先,通过斯皮尔曼相关系数得到各天气因素与功率的相关系数,将相关系数最大的因素通过VMD分解成多个分量.其次,熵值表示序列的复杂程度,熵值越高代表序列越复杂,熵值越低代表序列越紧凑[18].本文使用排列熵对最大影响因素分解的分量进行划分,并将熵值大于0.6的划分为高频分量,熵值小于0.6的划分为低频分量,将剩余天气因素组合在一起构成复合天气因素.

考虑高频分量、低频分量和复合天气因素对原始功率的影响程度不同,计算各因素与功率的灰色关联度,对影响功率更大的因素赋予更大的权重.将各个影响因素的结果进行加权得到最终结果.

由式(22)求出m个影响因素与原始功率之间的关联度$r_1,r_2,\cdots,r_m$,确定权重为

$$w_i = \frac{r_i}{r_1 + r_2 + \cdots + r_m} \tag{23}$$

构建新的综合评价指数R_i:

$$R_i = w_1\cdot\mathrm{HF}_i + w_2\cdot\mathrm{LF}_i + w_3\cdot\mathrm{CIF}_i \tag{24}$$

式中w_1,w_2,w_3为权重,HF_i为高频分量,LF_i为低频分量,CIF_i为复合天气因素,R_i为综合评价指数值.

计算出综合评价指标R_i的值,并将R_i由大到小排序得到相似日;本文选取前20天的数据作为相似日数据集.

3 光伏功率预测模型

3.1 NGO优化BiGRU模型

为提升BiGRU模型的预测精度,本文利用北方苍鹰算法对BiGRU的超参数进行优化.考虑到学习率过低会使学习速度太慢,而学习率过高又很难收敛,因此采用LearningRateScheduler调度器,对学习率进行动态设置,学习率设置在[0.001,0.01].NGO算法同时优化了BiGRU的神经元个数和batch_size,神经元个数设置在[10,100],batch_size设置在[20,50].相关参数具体设置如表2所示,优化流程如图4

所示.对预测模型的基体优化过程如下:

1)初始化北方苍鹰的种群数目和位置,并计算北方苍鹰的目标函数值;

2)首先对搜索空间进行全局搜索,确定最优区域.随机选取猎物,通过式(12)计算猎物的位置,再由式(13)计算第一阶段北方苍鹰的位置;

3)判断北方苍鹰是否攻击猎物.若没有攻击则由式(13)继续更新北方苍鹰的位置;若对猎物发动攻击则由式(14)更新北方苍鹰的位置并跳出循环,进入第二阶段;

4)第二阶段北方苍鹰对猎物进行追逐和抓捕.由式(15)和式(16)计算第二阶段北方苍鹰的位置;

5)判断是否抓获猎物,若没有抓获猎物,则继续更新第二阶段北方苍鹰的位置;若抓获猎物,则输出北方苍鹰位置,即输出最优参数.

表2　参数设置

Tab.2　Parameter settings

参数	参数值
最优学习率	[0.001,0.01]
神经元个数	[10,100]
batch_size	[20,50]
时间步长	6
搜索维度	3
种群数量	30

3.2　基于相似日选取和VMD-NGO-BiGRU的短期光伏功率预测

基于相似日选取和VMD-NGO-BiGRU的短期光伏发电功率预测流程如图5所示,具体步骤如下:

1) 利用斯皮尔曼相关系数选出主要的气象因素;

2) VMD分解原始功率和影响最大的气象因素,并将影响最大的气象因素的分量通过排列熵划分为高频分量(HF)和低频分量(LF);

3) 将其它较不重要的气象因素表示为复合气象因子(CIF),构建新的综合评价指数,计算各因素与功率之间的灰色关联度r_i,得到权重w_i,使用式(24)计算R_i的值,并将R_i由大到小排序得到相似日;

4) 将分解后的原始功率序列的子分量分别输入经NGO优化的BiGRU网络模型中进行训练;

5) 将各分量的预测结果进行叠加重构,得到最终预测结果.

为验证本文所提方法的有效性,采用三种评价

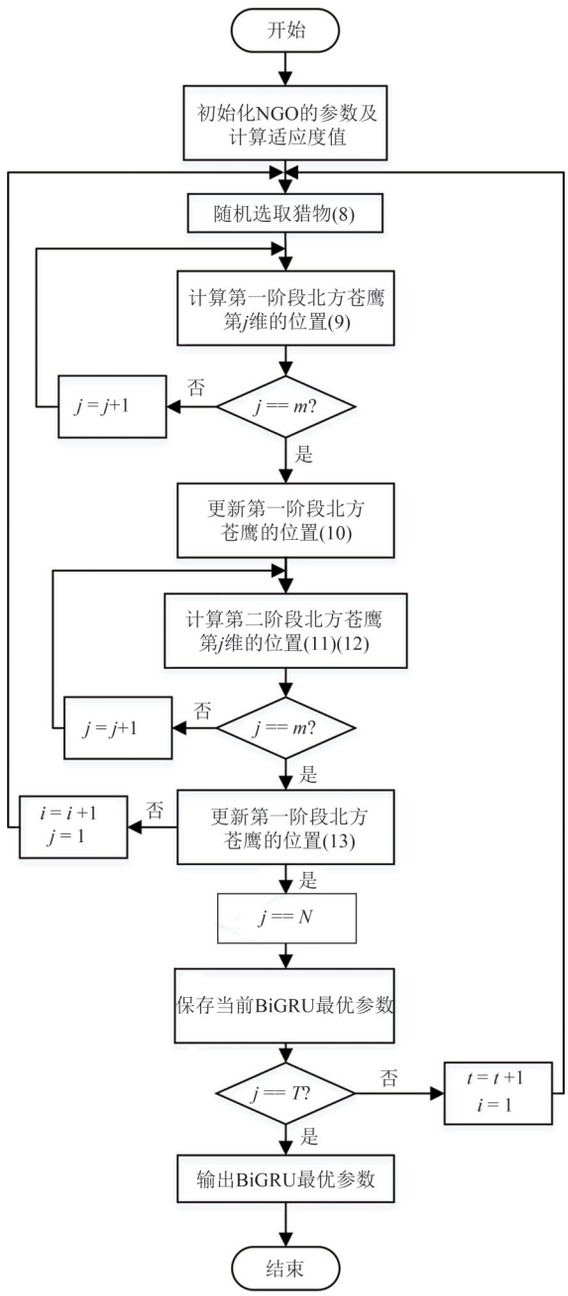


图4　NGO优化BiGRU流程图

Fig.4　NGO optimization BiGRU flowchart

指标对预测模型进行评价.分别是均方根误差(root mean square error,RMSE)、平均绝对误差(mean absolute error,MAE)和平均绝对百分比误差(mean absolute percentage Error, MAPE),具体公式如下:

$$E_{\mathrm{RMSE}}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}(i)-y(i)\right)^2} \tag{25}$$

$$E_{\mathrm{MAE}}=\frac{1}{N}\sum_{i=1}^{N}|\hat{y}(i)-y(i)| \tag{26}$$

$$E_{\mathrm{MAPE}}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}(i)-y(i)}{y(i)}\right|\times 100\% \tag{27}$$

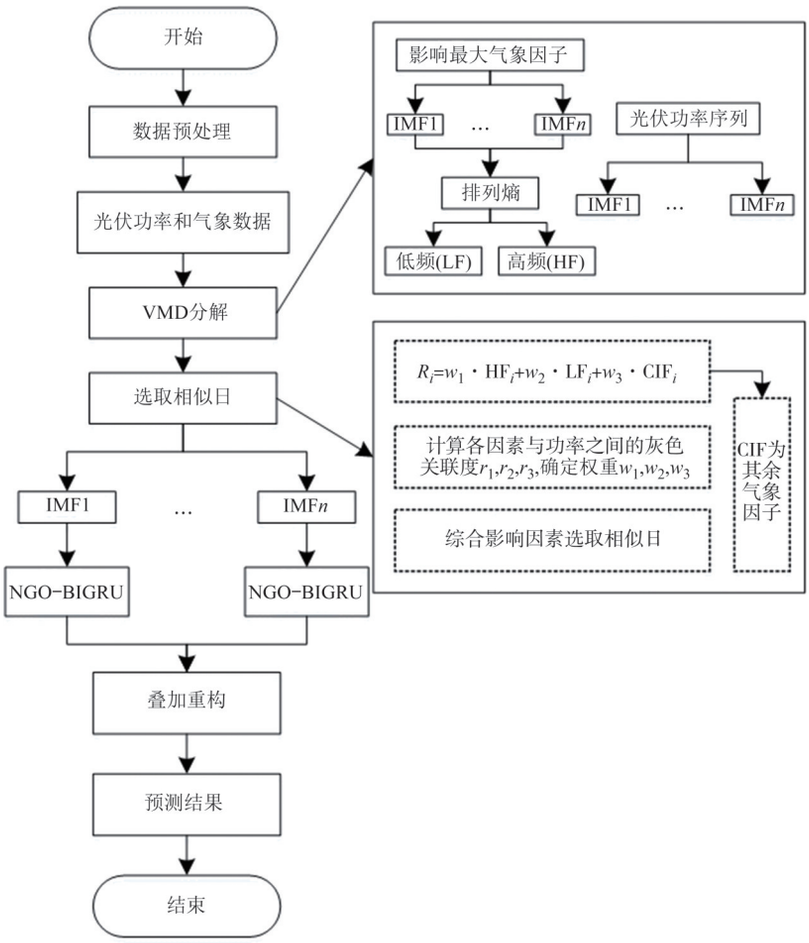


图5 预测模型流程图

Fig.5 Flow chart of this prediction model

式中:$\hat{y}(i)$表示光伏功率预测序列;$y(i)$表示原始光伏功率序列;N表示样本数量.

4 算例分析

本文选取2022年江苏某光伏电站一整年的数据进行分析.将天气划分为三种天气类型:晴天、阴天和雨天.每种天气类型都由综合评价指数选取前20天的相似日作为数据集,其中前18天作为训练数据,剩余2天作为测试数据.每日选取数据时间为7:00~18:45,每隔15 min记录一次,一天共有48条数据.为验证不同天气类型下相似日的效果,选取2022年8月26日(晴天)、2022年8月29日(阴天)和2022年9月20日(雨天)三天为待预测日进行实例分析.

4.1 VMD分解结果

VMD有两个重要的参数分解个数k和惩罚因子α,分解个数过大会造成过拟合,分解个数过小不能提取出有效信息,因此分解个数的选取要适中.以功率分解为例,如表3所示,观察中心频率可得,当k=6时IMF3和IMF4比较相近;当k=7时IMF4和IMF5比较相近,由于中心频率相近会产生模态混叠现象,不

表3 不同k值下分解相似日功率的中心频率

Tab.3 The center frequency of decomposed similar daily power under different k values

模态数	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
k=3	0.001 08	0.022 77	0.221 92				
k=4	0.000 17	0.022 52	0.130 38	0.341 76			
k=5	0.000 15	0.021 97	0.140 35	0.189 54	0.377 48		
k=6	0.000 14	0.021 98	0.138 39	0.152 59	0.420 36	0.282 17	
k=7	0.000 12	0.021 87	0.155 12	0.213 69	0.233 61	0.297 92	0.327 78

利于信息的提取,因此分解个数为5比较适当.同理可得总辐射分解个数应为6.VMD分解相似日功率和总辐射的结果分别如图6和图7所示.

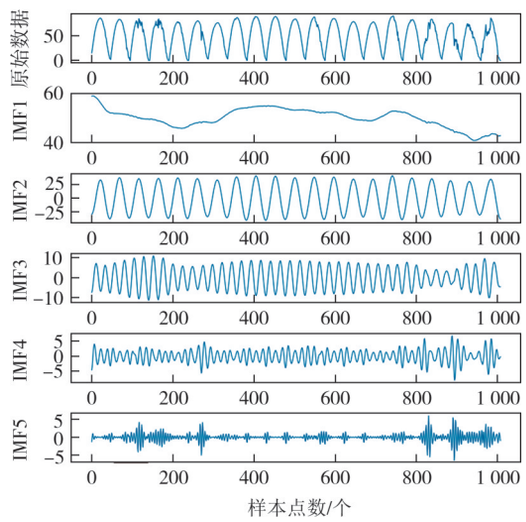


图6　VMD分解相似日功率

Fig.6　VMD decomposition similar daily power

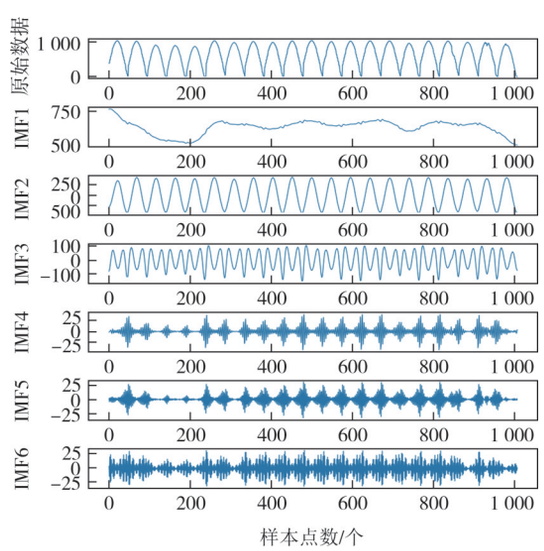


图7　VMD分解总辐射

Fig.7　VMD decomposition total radiation

4.2　相似日的选取结果

重构VMD分解总辐射的分量,将熵值小于0.6的划分为低频分量,大于0.6的划分为高频分量.由表4可知,将IMF2和IMF3重构划分为低频分量,剩余分量重构为高频分量.其余天气影响因素则通过归一化处理,加和求平均重构为复合天气因素.

本文旨在通过计算灰色关联度的值进行权重的分配,由式(22)计算出低频分量、高频分量和复合天气因素与功率的灰色关联度,如表5所示.由式(23)计算出各影响因素的权重如表6所示.

表4　VMD分解总辐射的分量熵值

Tab.4　Component entropy value of VMD decomposition total radiation

VMD分解分量	熵值
IMF1	0.716 34
IMF2	0.499 18
IMF3	0.592 23
IMF4	0.765 69
IMF5	0.963 65
IMF6	0.939 45

表5　各影响因素与功率的灰色关联度

Tab.5　Grey correlation degree between various influencing factors and power

功率影响因素	灰色关联度
低频分量	0.518 19
高频分量	0.627 68
复合因素	0.775 10

表6　各影响因素的权重

Tab.6　Weights of various influencing factors

功率影响因素	权重
低频分量	0.27
高频分量	0.33
复合因素	0.40

以2022年8月26日(晴天)为例,通过构建新的综合评价指数R_i,由式(24)计算出R_i的值并以降序排列的方式,筛选出前20天的相似日,前20天相似日的综合评价指数值如表7所示.

表7　相似日综合评价指数

Tab.7　Comprehensive evaluation index for similar days

R_i	评价指数	R_i	评价指数
R_1	0.952 6	R_{11}	0.925 8
R_2	0.948 6	R_{12}	0.924 9
R_3	0.948 2	R_{13}	0.923 9
R_4	0.945 5	R_{14}	0.921 2
R_5	0.945 1	R_{15}	0.921 1
R_6	0.937 6	R_{16}	0.919 5
R_7	0.931 6	R_{17}	0.915 9
R_8	0.929 6	R_{18}	0.914 7
R_9	0.929 5	R_{19}	0.914 1
R_{10}	0.925 9	R_{20}	0.913 4

为了验证选取的相似日与待预测日是否具有相似性,绘制如图8所示的20个相似日与待预测日的功率趋势图.图中虚线代表的是待预测日的功率曲线,由图可知所选取的20个相似日功率曲线的趋势

与待预测日基本保持一致,验证了所提方法选取相似日的有效性.

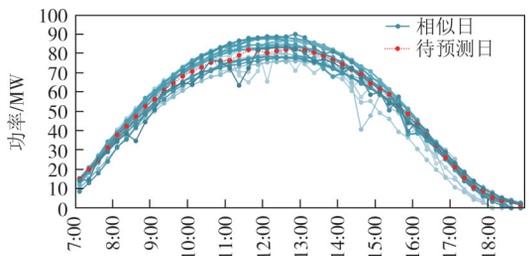


图 8 相似日与待预测日的功率趋势图

Fig.8 Power trend graph of similar day and day to be predicted

4.3 优化算法结果对比

为检验NGO优化算法的性能,将NGO优化算法与粒子群算法(Particle Swarm Algorithm,PSO)和鲸鱼优化算法(Whale Optimization Algorithm,WOA)进行对比.设置PSO个体最优系数c_1=1.0,全局最优系数c_2=1.5,惯性系数w=0.8.初始化设置三种优化算法的种群数量为30,迭代次数设置为50.从图9可以看出,PSO算法虽然收敛速度相对较快,但在三种优化算法中,适应度值最大,寻优能力最差;WOA算法是三种优化算法中收敛速度最慢的,适应度值相对于PSO较小,寻优能力次之;NGO算法在三种优化算法中收敛速度是最快的,适应度值也是最小的,寻优能力最强.

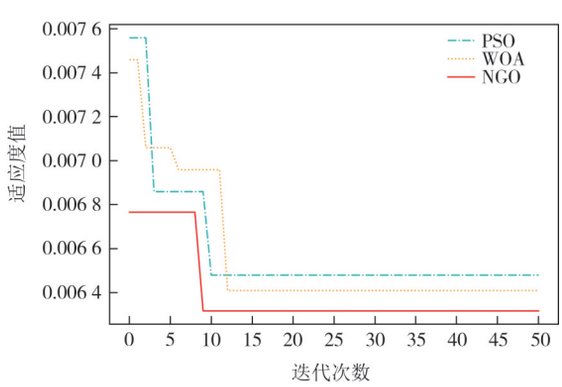


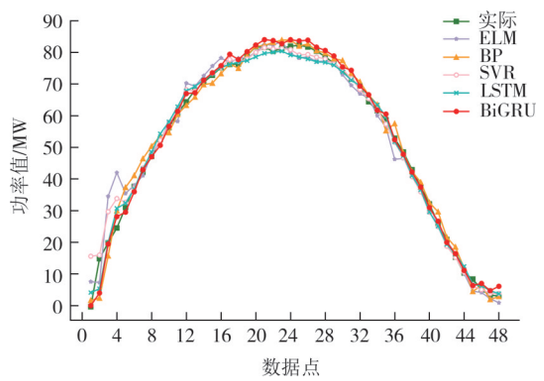
图 9 优化算法迭代图

Fig.9 Optimization algorithm iteration diagram

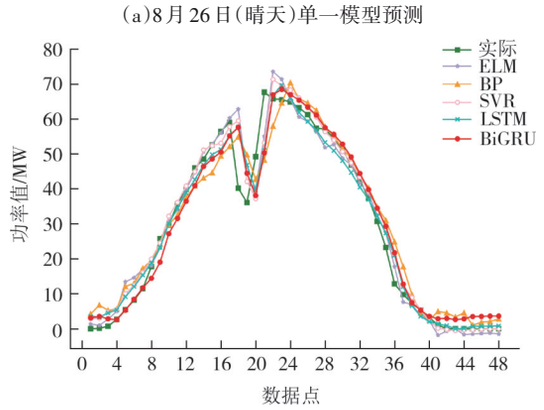
4.4 整体预测结果对比

本文从单一模型和组合模型两个方向进行实例验证,逐步验证在不同天气类型下所提方法的有效性.选用的数据集是通过综合评价指数选取的与待预测日相似的前20天的数据.

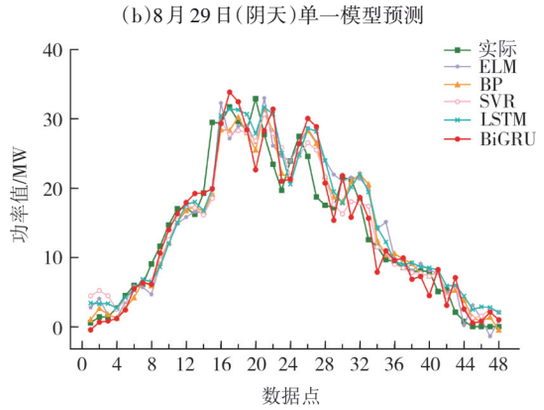
在单一模型预测中,选取极限学习机(Extreme Learning Machine, ELM)、BP神经网络(Back Propagation,BP)、支持向量回归(Support Vector Regression,SVR)、LSTM和BiGRU五种网络模型进行对比试验.由图10可知,在不同天气类型下,BiGRU模型预测值与真实值的拟合程度更好,三种误差均是最小的.由表8可知,在晴天情况下,BiGRU模型的RMSE、MAE和MAPE比LSTM分别提升了12.29%、19.93%和2.11%;在阴天情况下,BiGRU模型的RMSE、MAE和MAPE比LSTM分别提升了6.08%、13.78%和10.47%;在雨天情况下,BiGRU模型的RMSE、MAE和MAPE比LSTM分别提升了6.37%、29.02%和19.31%.说明筛选过相似日后,使用BiGRU模型进行功率预测效果会更好.



(a)8月26日(晴天)单一模型预测



(b)8月29日(阴天)单一模型预测



(c)9月20日(雨天)单一模型预测

图 10 单一模型不同天气类型下预测结果对比

Fig.10 Comparison of forecast results of single model under different weather types

表 8　单一模型评价指标

Tab.8　Single model evaluation index

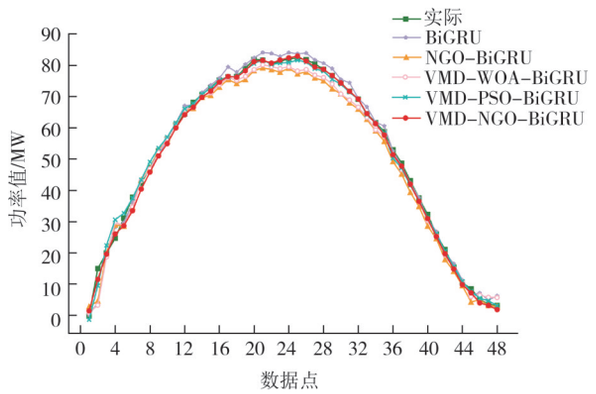
天气类型	单一模型	RMSE/MW	MAE/MW	MAPE/%
晴天	ELM	4.284 5	2.685 7	7.512 0
	BP	3.837 1	2.164 2	6.922 3
	SVR	3.305 4	1.768 5	6.580 3
	LSTM	2.429 8	1.745 2	5.995 0
	BiGRU	2.306 9	1.545 9	5.783 7
阴天	ELM	6.096 1	4.682 6	13.307 6
	BP	5.185 9	4.568 7	12.781 4
	SVR	5.951 4	4.159 2	11.422 7
	LSTM	5.840 1	4.054 2	11.253 1
	BiGRU	5.231 8	3.916 4	10.206 3
雨天	ELM	6.554 9	4.781 5	15.216 3
	BP	6.484 4	4.566 1	13.754 7
	SVR	6.470 1	4.350 6	12.680 3
	LSTM	6.208 6	3.951 0	12.198 5
	BiGRU	5.572 1	3.660 8	10.268 0

在组合模型预测中,选取NGO-BiGRU(方法1)与本文方法进行对比,证明对原始功率分解能够更大限度地提升模型的预测精度;为验证NGO算法的优越性,将本文方法与VMD-PSO-BiGRU(方法2)、VMD-WOA-BiGRU(方法3)进行对比.

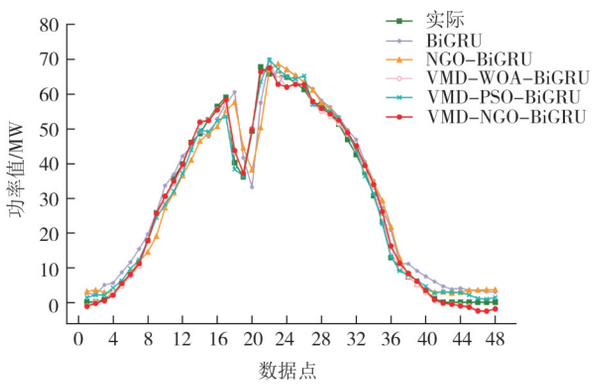
晴天时,由图11(a)可知,晴天时的光伏功率波动比较小,各方法预测曲线与真实功率曲线的拟合程度都比较好.从图12(a)可以看出,本文方法的三种误差最小,准确度最高.为了更加清晰的展示本文方法的精度,由表9可知,在晴天情况下,经过NGO优化的BiGRU比单一模型BiGRU的RMSE、MAE和MAPE分别提升了2.27%、6.31%和43.05%,说明NGO优化BiGRU模型是有一定效果的;本文方法比未经过VMD分解的方法1的RMSE、MAE和MAPE分别提升了11.56%、6.31%和38.91%,说明分解原始功率能够提取更多重要信息;本文方法与方法2相比RMSE、MAE和MAPE分别提升了5.5%、2.13%和11.03%,说明NGO算法优化BiGRU模型要优于PSO算法;本文方法与方法3相比RMSE、MAE和MAPE分别提升了3.11%、2.37%和7.40%.

阴天时,由图11(b)可知,阴天时的光伏功率波动略大,各方法的预测曲线与真实功率曲线的拟合程度相较于晴天略差.从图12(b)可以看出,虽然本文所提方法的三种误差仍是最小的,但是相对于晴天来说误差有所增大,这是由于光伏功率的发电受到天气因素的影响.由表9可知,在阴天情况下,经过NGO优化的BiGRU比单一模型BiGRU的RMSE、MAE和MAPE分别提升了4.12%、8.79%和14.92%;本文所提方法比未经过VMD分解的方法1的RMSE、MAE和MAPE分别提升了36.16%、21.46%和31.83%;本文方法与方法2相比RMSE、MAE和MAPE分别提升了8.99%、8.32%和11.05%;本文方法与方法3相比RMSE、MAE和MAPE分别提升了2.88%、2.19%和9.32%.

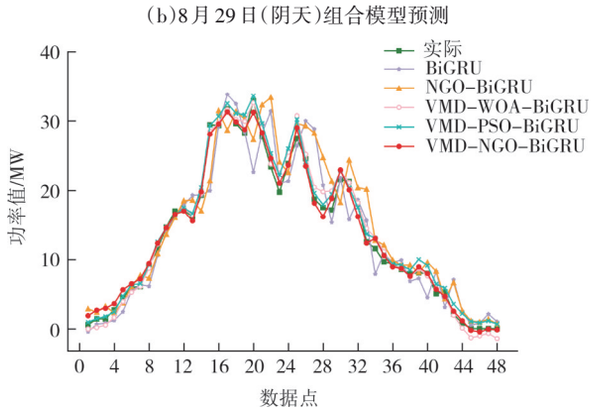
雨天时,由图11(c)可知,雨天时的光伏功率波



(a)8月26日(晴天)组合模型预测



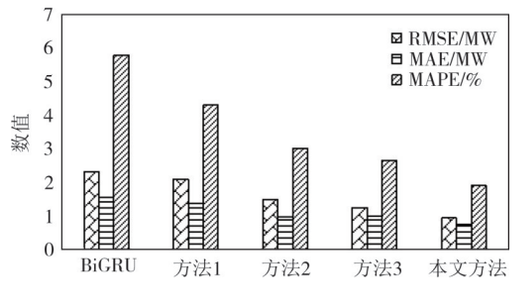
(b)8月29日(阴天)组合模型预测



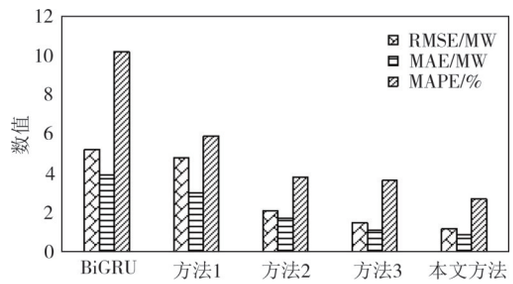
(c)9月20日(雨天)组合模型预测

图11　组合模型不同天气类型下预测结果对比

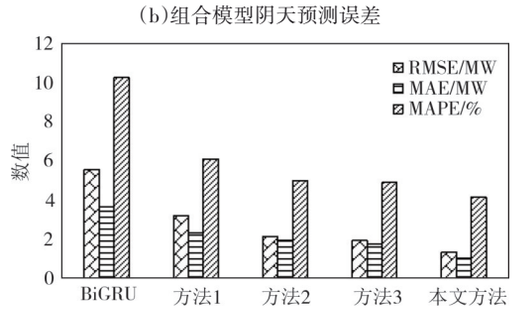
Fig.11　Comparison of prediction results of combined models under different weather types



(a)组合模型晴天预测误差



(b)组合模型阴天预测误差



(c)组合模型雨天预测误差

图 12 组合模型不同天气类型下评价指标

Fig.12 Evaluation indicators of combined models under different weather types

表 9 组合模型误差指标

Tab.9 Combined model error indicator

天气类型	组合模型	RMSE/MW	MAE/MW	MAPE/%
晴天	BiGRU	2.306 9	1.545 9	5.783 7
	方法 1	2.080 1	1.365 8	4.292 2
	方法 2	1.474 6	0.947 8	2.996 4
	方法 3	1.235 8	0.971 6	2.632 8
	本文方法	0.924 3	0.734 5	1.892 5
阴天	BiGRU	5.231 8	3.916 4	10.206 3
	方法 1	4.819 5	3.036 8	5.901 5
	方法 2	2.103 1	1.722 7	3.823 6
	方法 3	1.492 2	1.109 5	3.651 2
	本文方法	1.203 9	0.890 7	2.719 0
雨天	BiGRU	5.572 1	3.660 8	10.268 0
	方法 1	3.201 8	2.348 6	6.115 9
	方法 2	2.146 5	1.957 6	4.986 6
	方法 3	1.935 9	1.760 1	4.904 7
	本文方法	1.336 5	1.058 1	4.167 1

动较为剧烈,各方法预测曲线与真实功率曲线的拟合程度也相对较差.从图 12(c)可以看出,在误差评价指标中,本文所提方法的误差仍是最低的,但整体预测效果相较于晴天和阴天的预测效果略差一些.由表 9 可知,在雨天情况下,经过 NGO 优化的 BiGRU 比单一模型 BiGRU 的 RMSE、MAE 和 MAPE 分别提升了 23.70%、13.12% 和 41.52%;本文方法比方法 1 的 RMSE、MAE 和 MAPE 分别提升了 18.65%、12.91% 和 19.49%;本文方法与方法 2 相比 RMSE、MAE 和 MAPE 分别提升了 8.10%、8.99% 和 8.19%;本文方法与方法 3 相比 RMSE、MAE 和 MAPE 分别提升了 5.99%、7.02% 和 7.38%.

通过实例分析,验证了在各类天气类型下,组合模型的预测效果比单一模型的预测效果要好.通过 BiGRU 与方法 1 对比可知,经过 NGO 优化算法选取的参数进行功率预测,结果更为准确;通过本文方法与方法 1 对比可知,功率经过分解的模型预测精度要比未分解的精度高.将本文方法与方法 2 进行对比,可知 NGO 优化算法优化 BiGRU 的参数是优于 PSO 和 WOA 优化算法的。在晴天、阴天和雨天三种天气类型下,本文所提方法的 MAPE 分别为 1.8925%、2.7190% 和 4.1671%,均在 5% 以内,由此说明,经过 NGO 优化的 BiGRU 模型有更强的鲁棒性和预测性能.

为验证所提选取相似日方法对模型的预测精度是有所提升的,以 2022 年 8 月 26 日(晴天)的数据为例,将本文筛选相似日方法与未筛选相似日和用模糊 C 均值聚类(Fuzzy C-Means Algorithm,FCM)筛选相似日的方法进行对比.未筛选相似日的数据集为 2022 年江苏某光伏电站一整年中晴天天气类型的前 20 天的数据.FCM 选取相似日的数据集为筛选出的数据集中前 20 天的数据.三者均通过 VMD-NGO-BiGRU 模型进行预测,未筛选相似日(No_Similar_Day)、FCM 筛选相似日(FCM_ Similar_Day)和本文方法筛选相似日(Similar_Day)的预测结果如图 13 所示,评价指标如表 10 所示.由表 10 可知,FCM 筛选相似日比未筛选相似日的 RMSE 提升了 8.52%,MAE 提升了 10.8%,MAPE 提升了 2.19%;本文筛选相似日方法比 FCM 筛选相似日的 RMSE 提升了 14.96%,MAE 提升了 9.99%,MAPE 提升了 3.04%.由此可知,FCM 选取相似日相对于未筛选相似日的数据来说,预测结果精度有所提升;相较于 FCM 选取相似日,本文所提筛选相似日的方法对模型的预测精度有更大的提升.

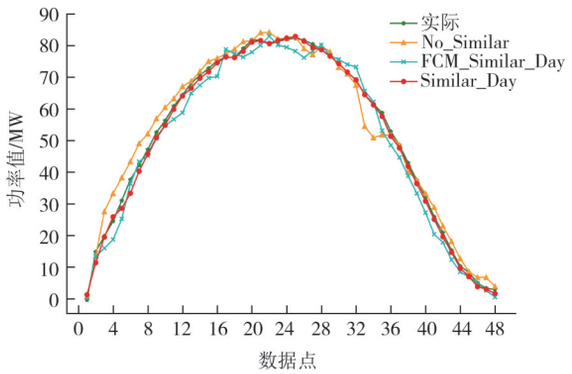


图 13　未筛选相似日预测结果

Fig.13　Unscreened similar day prediction results

表 10　未筛选相似日和筛选相似日误差指标

Tab.10　Unscreened similar days and screened similar day error indicators

	RMSE/MW	MAE/MW	MAPE/%
未筛选相似日	3.272 2	2.814 4	7.124 3
FCM筛选相似日	2.420 4	1.734 2	4.934 0
本文筛选相似日	0.924 3	0.734 5	1.892 5

5　结语

针对不同天气条件下,模型的预测性能和鲁棒性仍然存在不足的问题,本文提出了一种基于相似日选取和北方苍鹰算法优化双向门控循环单元的短期光伏功率预测模型,并得到如下结论:

1)本文构建的新的筛选相似日的评价指标筛选出的相似日相较于传统聚类方法更加适合待预测日,对模型预测精度有一定提升.

2)NGO算法与PSO优化算法相比具有更强的寻优性能,可以为BiGRU获取更优参数解,提高了模型的预测性能.

3)本文模型在不同天气影响下,对应的RMSE、MAE和MAPE均有较大幅度提升,预测误差值也比较小,充分凸显了该模型良好的预测精度.

仿真结果表明,经本文构建的筛选相似日的评价指标有效地筛选出了相似日数据集,通过模型进行训练,在不同天气类型下的误差评价指标均优于其他模型,具有良好的适应性.

参考文献

[1]　龚莺飞,鲁宗相,乔颖,等.光伏功率预测技术[J].电力系统自动化,2016,40(4):140-151.
GONG Y F, LU Z X, QIAO Y, et al. An overview of photovoltaic energy system output forecasting technology [J]. Automation of Electric Power Systems, 2016,40(4):140-151.(in Chinese)

[2]　董存,王铮,白捷予,等.光伏发电功率超短期预测方法综述[J/OL].高电压技术:1-13[2023-06-24].
DONG C, WANG Z, BAI J Y, et al. Review of ultra-short-term forecasting methods for photovoltaic power generation[J/OL].High Voltage Engineering:1-13[2023-06-24]. (in Chinese)

[3]　KIM G G, CHOI J H, PARK S Y, et al. Prediction model for PV performance with correlation analysis ofenvironmental variables [J]. IEEE Journal of Photovoltaics,2019,9(3):832-841.

[4]　MILAD H S A, FAROOQ U, EL HAWARY M E, et al. Neo-fuzzy integrated adaptive decayedbrain emotional learning network for online time series prediction[J]. IEEE Access, 2017, 5(1): 1037-1049.

[5]　刘卫亮,刘长良,林永君,等.计及雾霾影响因素的光伏发电超短期功率预测[J].中国电机工程学报,2018,38(14):4086-4095+4315.
LIU W L, LIU C L, LIN Y J, et al. Super short-term photovoltaic power forecasting considering influence factor of Smog [J]. Proceedings of the CSEE, 2018, 38(14): 4086-4095+4315. (in Chinese)

[6]　ADEL M, ALESSANDRO M P, EMANUELE O, et al. Advanced methods for photovoltaic output power forecasting: A review[J]. Applied Sciences,2020,10(2): 487.

[7]　叶林,裴铭,路朋,等.基于天气分型的短期光伏功率组合预测方法[J].电力系统自动化,2021,45(1):44-54.
YE L, PEI M, LU P, et al. Combination forecasting method of short-term photovoltaic power based on weather classification[J]. Automation of Electric Power Systems, 2021, 45(1): 44-54. (in Chinese)

[8]　张金金,张倩,马愿,等.模糊聚类-Elman神经网络短期光伏发电预测模型[J].电测与仪表,2020,57(12):46-51.
ZHANG J J, ZHANG Q, MA Y, et al. Fuzzy clustering-Elman neural network model for short-term photovoltaic power generation prediction [J]. Electrical Measurement & Instrumentation, 2020, 57(12):46-51. (in Chinese)

[9]　王福忠,王帅峰,张丽.基于VMD-LSTM与误差补偿的光伏发电超短期功率预测[J].太阳能学报,2022,43(8):96-103.
WANG F Z, WANG S F, ZHANG L. Ultra short term power prediction of photovoltaic power generation based on VMD-LSTM and error compensation [J]. Acta Energiae Solaris Sinica, 2022, 43(8):96-103. (in Chinese)

[10]　孟安波, 陈嘉铭, 黎湛联, 等. 基于相似日理论和CSO-WGPR的短期光伏发电功率预测[J]. 高电压技术, 2021, 47(4): 1176-1184.
MENG A B, CHEN J M, LI Z L, et al. Short-term photovoltaic power generation prediction based on similar day theory and CSO-WGPR[J].High Voltage Engineering, 2021, 47(4): 1176-1184. (in Chinese)

[11]　杨国华, 张鸿皓, 郑豪丰, 等. 基于相似日聚类和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. (in Chinese)

[12] 王清亮, 杨博, 应欣峰, 等. 非晴空条件下光伏发电短期功率预测方法[J]. 太阳能学报, 2022, 43(3): 188-196.
WANG Q L, YANG B, YING X F, et al. Short-term photovoltaic power forecasting method under non-clear sky condition[J]. Acta Energiae Solaris Sinica, 2022, 43(3): 188-196. (in Chinese)

[13] 商立群, 李洪波, 侯亚东, 等. 基于 VMD-ISSA-KELM 的短期光伏发电功率预测[J]. 电力系统保护与控制, 2022, 50(21): 138-148.
SHANG L Q, LI H B, HOU Y D, et al. Short-term photovoltaic power generation prediction based on VMD-ISSA-KELM [J]. Automation of Electric Power Systems, 2022, 50(21): 138-148. (in Chinese)

[14] 滕陈源, 丁逸超, 张有兵, 等. 基于 VMD-Informer-BiLSTM 模型的超短期光伏功率预测[J/OL]. 高电压技术: 1-12[2023-06-24].
TENG C Y, DING Y C, ZHANG Y B, et al. Ultra-short-term photovoltaic power prediction based on VMD-informer-BiLSTM model [J/OL]. High Voltage Engineering: 1-12 [2023-06-24]. (in Chinese)

[15] 董雪, 赵宏伟, 赵生校, 等. 基于 SOM 聚类和二次分解的 BiGRU 超短期光伏功率预测[J]. 太阳能学报, 2022, 43(11): 85-93.
DONG X, ZHAO H W, ZHAO S X, et al. Ultra-short-term forecasting method of photovoltaic power based on SOM clustering, secondary decomposition and bigru [J]. Acta Energiae Solaris Sinica, 2022, 43(11): 85-93. (in Chinese)

[16] 宋江涛, 崔双喜, 刘洪广. 基于二次分解 NGO-VMD 残差项与长短时记忆神经网络的超短期风功率预测[J]. 科学技术与工程, 2023, 23(6): 2428-2437.
SONG J T, CUI S X, LIU H G. Ultra-short-term wind power prediction based on the quadratic decomposition ngo-vmd residual term and long short-term memory (LSTM) [J]. Science Technology and Engineering, 2023, 23 (6): 2428-2437. (in Chinese)

[17] 叶家豪, 魏霞, 黄德启, 等. 基于灰色关联分析的 BSO-ELM-AdaBoost 风电功率短期预测[J]. 太阳能学报, 2022, 43(3): 426-432.
YE J H, WEI X, HUANG D Q, et al. Short-term forecast of wind power based on BSO-ELM-AdaBoost with grey correl ation analysis[J]. Acta Energiae Solaris Sinica, 2022, 43(3): 426-432. (in Chinese)

[18] 王琦, 季顺祥, 钱子伟, 等. 基于熵理论和改进 ELM 的光伏发电功率预测[J]. 太阳能学报, 2020, 41(10): 151-158.
WANG Q, JI S X, QIAN Z W, et al. Photovoltaic power prediction based on entropy theory and improved ELM [J]. Acta Energiae Solaris Sinica, 2020, 41(10): 151-158. (in Chinese)