

DOI:10.19651/j.cnki.emt.2313035

基于 SSA-BiLSTM 非线性组合方法的光伏功率预测*

袁建华^{1,2} 蒋文军^{1,2} 李洪强^{1,2} 徐杰^{1,2} 高延玲^{1,2}(1. 三峡大学电气与新能源学院 宜昌 443002; 2. 三峡大学电气与
新能源学院建设的省部级重点实验室(中心)宜昌 443002)

摘要: 采用多种模型进行线性组合来对光伏功率预测,能有效避免收敛性差、可靠性低等缺点。线性组合模型中,将单一模型之间简为线性关系能简化组合模型计算,但会使预测精度降低。针对此问题,提出一种基于麻雀搜索算法(SSA)优化双向长短期记忆网络(BiLSTM)非线性组合方法的预测模型。首先,利用基于核改进的模糊C均值聚类算法(KFCM)和变分模态分解(VMD)对原始数据样本进行预处理;然后,采用 Elman 和 SSA-BiLSTM 对经过预处理后的光伏功率进行建模预测;最后,通过麻雀搜索算法优化双向长短期记忆网络对两个单一模型进行非线性组合,建立短期光伏功率非线性组合模型。通过某个光伏电站实测数据建立对比算例,结果表明所提组合模型在不同天气下的 RMSE 和 MAE 平均值分别为 0.689 kW 和 0.540 kW,均优于其他对比模型,验证了所提组合模型的有效性和优越性。

关键词: 光伏功率预测;非线性组合方法;麻雀搜索算法;BiLSTM 网络

中图分类号: TP271 **文献标识码:** A **国家标准学科分类代码:** 470.4047

Photovoltaic power forecasting based on SSA-BiLSTM nonlinear combination method

Yuan Jianhua^{1,2} Jiang Wenjun^{1,2} Li Hongqiang^{1,2} Xu Jie^{1,2} Gao Yanling^{1,2}

(1. College of Electrical and New Energy, China Three Gorges University, Yichang 443002, China; 2. Provincial and Ministerial Key Laboratory (Center) of College of Electrical and New Energy of China Three Gorges University, Yichang 443002, China)

Abstract: The linear combination of various models can effectively avoid the disadvantages of poor convergence and low reliability for photovoltaic power forecasting. Simplifying the linear relationship between a single model in a linear combinatorial model can simplify the calculation of the combinatorial model, but reduce the prediction accuracy. Aiming at this problem, a prediction model based on Sparrow Search Algorithm (SSA) was proposed to optimize Bidirectional Long Short-Term Memory (BiLSTM) nonlinear combination method. Firstly, the Kernel-based Fuzzy C-means (KFCM) and Variational Modal Decomposition (VMD) are used to preprocess the original data samples. Then, using the Elman and SSA-BiLSTM forecast after photovoltaic (PV) power of pretreatment; Finally, the nonlinear combination of the two single models is optimized by the sparrow search algorithm to establish a nonlinear combination prediction model for short-term photovoltaic power. A comparative calculation example is established based on the measured data of a photovoltaic power plant, and the results showed that the average RMSE and MAE values of the proposed combined model under different weather conditions are 0.689 kW and 0.540 kW, respectively, which are superior to other comparative models, verifying the effectiveness and superiority of the proposed combined model.

Keywords: photovoltaic power forecasting; nonlinear combination method; sparrow search algorithm; BiLSTM network

0 引言

随着能源需求的快速增长和传统化石能源有限的储

量,光伏发电作为最具经济性和环保性的可再生能源之一,发展迅速并备受关注^[1]。然而,光伏发电的随机性和间歇性特点给光伏发电并网的稳定运行和管理造成了很大的挑

收稿日期:2023-03-08

* 基金项目:煤燃烧国家重点实验室开放基金(FSKLCCA1607)、梯级水电站运行与控制湖北省重点实验室基金(2015KJX07)、产学研协同培养研究生实践创新能力机制研究项目(SDYJ201604)资助

战^[2]。因此开发高精度的短期光伏功率预测模型,以便预测光伏发电的短期变化情况,这对提高光伏发电并网运行的安全和稳定至关重要。

目前,光伏发电功率的短期预测大多基于数据驱动模型,该类模型只需考虑选择合适的输入变量与输出变量,这对于短期光伏功率预测来说非常合适,常见的数据驱动模型包含支持向量机^[3]、随机森林^[4]和长短期记忆网络(long short-term memory, LSTM)^[5]等。近年来,Elman 神经网络因其独特的 4 层网络结构而具有较好的鲁棒性和强大的计算能力,因此在非线性时间序列预测领域得到了广泛应用。文献[6]基于 Elman 网络完成了不同天气下待测日的光伏功率短期预测并取得了较好的预测效果。文献[7]在利用模型预测光伏功率前,采用变分模态分解(variational mode decomposition, VMD)对光伏功率预处理。实验表明,加入 VMD 可有效挖掘非线性时间序列的特征信息,提高模型的学习效果。此外,由 LSTM 网络改进而来的双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)具有双向传递信息的能力,从而可以更全面地捕捉非线性时间序列中的信息,提高网络的学习效果,因而 BiLSTM 网络也被广泛应用于光伏功率预测领域。文献[8]利用 BiLSTM 为基础模型对光伏功率进行预测前,利用模糊 C 均值(fuzzy C-means, FCM)聚类算法对光伏发电功率历史数据进行天气分类,有效提高了模型的预测精度。文献[9]采用 BiLSTM 网络并结合小波分解取得了对风电、光伏功率不错的预测效果,但是 BiLSTM 的预测性能取决于其相关参数的选择。

上述文献采用的光伏功率预测方法都是基于单一模型,存在收敛性差、泛化能力弱等缺点。组合模型通过组合方法将单一模型各自的优势整合起来能有效避免单一预测模型自身存在的缺点,发挥各自的优势。文献[10]采用最优加权组合算法确定误差反向传播(back-propagation, BP)模型和支持向量机(support vector machines, SVM)模型的组合权重系数,以得到组合预测模型。文献[11]以误差倒数法来聚合加入卷积神经网络(convolutional neural network, CNN)的 LSTM 网络和极限梯度提升(extrme gradient boosting, XGBoost)两个单一模型,发挥 CNN-LSTM 模型在提取时间序列特征的优势和 XGBoost 模型较强的学习能力和鲁棒性,实验表明该组合模型预测精度好于单一模型。文献[12]以遗传算法赋予深度信念网络(deep belief network, DBN)和模糊模型(takagi-sugeno, T-S)在不同时刻下的组合权重系数,来整合 DBN 和 T-S 两个单一模型并获得最终预测结果。上述文献提出的组合模型中,采用的组合方法往往基于智能优化算法或非负约束理论,其本质是将单一模型的预测值与最终预测结果之间表征为线性关系以简化计算,而忽略其内部复杂的非线性关系导致预测精度降低。针对此问题,已开发出一些基

于非线性组合方法的组合模型,但这些组合方法大多基于机器学习(SVM)^[13]。

综上所述,本文提出一种基于麻雀搜索算法(sparrow search algorithm, SSA)优化 BiLSTM 网络非线性组合方法的短期光伏功率预测模型。首先考虑到不同天气下光伏功率输出的差异性和其自身非平稳性特点,采用基于核改进的模糊 C 均值聚类算法(kernel-based fuzzy C-means, KFCM)对原始数据样本按照不同天气类型分类,以及采用 VMD 算法对聚类后的不同天气类型下原始光伏功率数据进行分解;接着利用 Elman 和 SSA-BiLSTM 对分解后的光伏功率各子序列进行并行预测,利用 SSA 算法对 BiLSTM 相关参数寻优以提高网络学习能力;然后采用 SSA-BiLSTM 网络对两个单一模型进行非线性加权组合,得到组合预测模型;最后通过算例仿真证明了本文所提组合模型有效提高了预测精度。

1 建模理论

1.1 Elman 神经网络

Elman 神经网络是一种典型的局部循环神经网络,相比于普通神经网络,Elman 神经网络的隐藏层通过额外的承接层将自身的状态反馈回输入层^[14],从而在处理非线性时间序列数据时具有更强的学习能力和泛化能力,其数学表达式为:

$$y(k) = g(\omega^3 h(k) + b_2) \quad (1)$$

$$x_c(k) = \lambda x_c(k-1) + h(k-1) \quad (2)$$

$$h(k) = F(\omega^1 x_c(k) + \omega^2 u(k-1) + b_1) \quad (3)$$

式中: $y(k)$ 为 k 时刻输出层的输出向量; $g(*)$ 表示输出层的神经元激活函数; λ 为自连接的反馈增益; $x_c(k)$ 为 k 时刻隐藏层的输出向量; $F(*)$ 表示隐藏层神经元的激活函数; b_1 和 b_2 表示隐藏层和输出层的阈值; $u(k)$ 为输入向量。

1.2 麻雀所示算法优化双向长短期记忆网络

BiLSTM 是一种双向循环结构的神经网络,在传统的 LSTM 基础上增加了一个反向 LSTM 层,该结构允许 BiLSTM 网络能够同时考虑输入序列中的过去和未来信息^[15-16],从而具有更强的处理非线性关系的能力,典型的 BiLSTM 网络如图 1 所示。

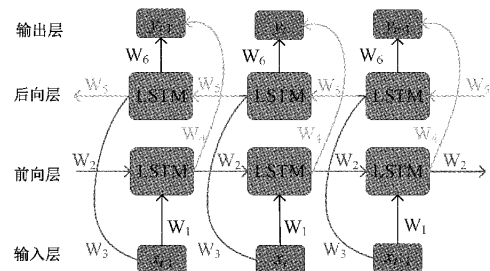


图 1 典型的 BiLSTM 网络

将 BiLSTM 应用于非线性时间序列预测领域时,选择合理的网络相关参数是关键。由于人为随机设置网络相关

参数可能会降低预测性能,因此一般采用智能优化算法来优化 BiLSTM 网络相关参数。本文选择 SSA 算法来优化网络相关参数,2020 年提出的 SSA 算法,它的灵感来自麻雀的觅食和反捕食行为^[17],并且将麻雀种群个体划分为发现者类型麻雀、跟随者类型麻雀和警戒者类型麻雀 3 类。与灰狼算法、蝙蝠算法和鲸鱼算法等近十年新提出的智能优化算法对比实验表明,SSA 算法在收敛速度、寻优能力和自适应能力等 3 个评价指标方面表现出最优的综合性能。由于 BiLSTM 网络中,BiLSTM 层神经元数量、学习率和训练轮次(Epoch)随机设置,可能导致网络训练过程中,出现过拟合、欠拟合以及训练时间过长或过短等情况发生,从而影响网络的预测效果,因此本文以 BiLSTM 层的神经元数量、学习率和训练轮次(Epoch)等作为 SSA 优化的目标参数,以网络的均方误差为麻雀搜索算法的适应度函数,在设定目标参数的取值范围后,麻雀种群个体的位置和目标参数被随机初始化,BiLSTM 网络更新过程中计算适应度函数值并对其排序,直到找到一组能使适应度函数值达到最小的目标参数组合。

2 基于 SSA-BiLSTM 非线性加权组合方法的光伏功率预测

2.1 模型构建

首先构建 Elman 和 SSA-BiLSTM1 两个单一预测模型。Elman 预测模型如图 2 所示,输入为总辐射、环境温度、相对湿度等气象数据以及历史光伏功率值,输出为预测功率值。SSA-BiLSTM1 模型与 Elman 模型的输入输出相同,并且是以 BiLSTM1 模型为基础,采用 SSA 算法优化搜索得到 BiLSTM1 模型中最佳的 3 个 BiLSTM 层的神经元数量、学习率、和 Epoch,图 3 为 BiLSTM1 模型结构。

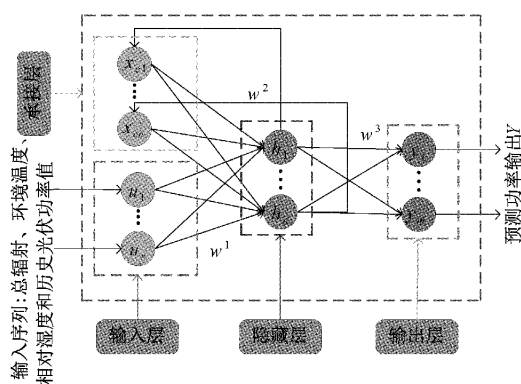


图 2 Elman 预测模型

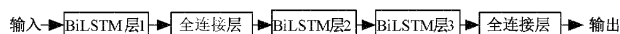


图 3 BiLSTM1 模型结构

绝大多数研究表明,组合模型通常比单一模型具有更好的预测性能,但组合模型的预测性能会受到组合方

法的影响^[18]。目前已有的组合模型大多采用智能优化算法或非负约束理论对单一预测模型进行线性加权组合,同时在确定各单一模型的权重系数时,采用的适应度函数为:

$$f = \frac{1}{n} \sum_{i=1}^n [\omega_1 \hat{y}_1(i) + \omega_2 \hat{y}_2(i) - y(i)]^2 \quad (4)$$

式中: n 为测试集样本总数; ω_1 和 ω_2 为单一模型组合的权重系数,且满足 $\omega_1 + \omega_2 = 1$; $\hat{y}_1(i)$ 和 $\hat{y}_2(i)$ 为 Elman 和 SSA-BiLSTM1 两个单一模型光伏功率的预测值; $y(i)$ 表示对应的光伏功率实测值。

然后根据智能算法或非负约束理论找到一组最小化适应度函数的权重系数 ω_1 和 ω_2 后,此时传统的线性组合模型可表示为:

$$F_s(i) = \omega_1 \hat{y}_1(i) + \omega_2 \hat{y}_2(i) \quad (5)$$

虽然线性组合方法可以在一定程度上聚合各单一预测模型的优势,提高预测性能,但是它仍然不能很好地提取和表征组合模型中,单一模型之间的非线性关系^[19]。目前已经开发出一些非线性组合方法,但这些方法大多基于机器学习。相较于机器学习方法,神经网络具有更强的提取和表征非线性关系的能力。因此本文采用一种基于 SSA-BiLSTM2 的组合方法,表征 Elman 和 SSA-BiLSTM1 两个单一预测模型的预测值与最终预测结果之间的非线性关系,如式(6)所示,SSA-BiLSTM2 是以 BiLSTM2 网络为基础,采用 SSA 算法优化搜索得到 BiLSTM2 网络中最佳的 2 个 BiLSTM 层的神经元数量、学习率和 Epoch,图 4 为 BiLSTM2 网络结构。

$$F_s(i) = H[\hat{y}_1(i), \hat{y}_2(i)] \quad (6)$$

式中: $F_s(i)$ 为光伏功率预测的最终结果; H 为 SSA-BiLSTM2 网络。



图 4 BiLSTM2 网络结构

2.2 模型求解

本文提出了一种基于 SSA-BiLSTM 非线性组合方法的预测模型,其求解过程如图 5 所示。首先本文选择总辐射、环境温度和相对湿度等影响光伏发电的强气象因素,以及光伏发电实测功率数据组成原始数据样本。通过采用聚类算法对原始数据样本进行聚类处理以降低天气因素对光伏功率输出的影响,以及采用分解策略降低光伏功率的非平稳性,挖掘光伏功率序列中潜在的信息,在此基础上求解出的预测模型可有效提高模型的适应性,因此先利用聚类算法和分解策略对原始数据样本进行预处理。

相较于传统的 K 均值硬聚类算法,基于 FCM 聚类算法在特征差异不明显的数据样本中也能实现较为柔性的模糊聚类划分^[20]。KFCM 在 FCM 的基础上通过引入合适的核函数以改进 FCM 对异常值和噪声比较敏感,收敛速度

慢等优点。为采用 KFCM 聚类算法对原始数据样本进行聚类处理,构造如式(7)所示的聚类特征:

$$X_i = [I_{mi}, T_{mi}, H_{mi}, I_{ni}, T_{ni}, I_{avi}, T_{avi}, I_{vai}, T_{vai}] \quad (7)$$

式中: I_{mi} 、 I_{ni} 、 I_{avi} 和 I_{vai} 为第 i 日总辐射的最大值、最小值、平均值和方差; T_{mi} 、 T_{ni} 、 T_{avi} 和 T_{vai} 为第 i 日环境温度的最大值、最小值、平均值和方差; H_{mi} 为第 i 日相对湿度的最大值。

接着采用文献[21]所述的 KFCM 聚类算法根据所构建的聚类特征,对原始数据样本进行聚类处理,并且根据总辐射这一气象数据呈现出的特点,将原始数据样本划分晴天、多云、雨天和突变天气 4 种类型数据样本。

然后鉴于 VMD 分解算法具有良好的分解效果和噪声鲁棒性^[22],因而采用 VMD 对不同天气类型下的原始光伏功率序列进行分解,如式(8)所示。

$$s. t. \sum_{k=1}^K u_k = P(t) \quad (8)$$

式中: u_k 为分解后的第 k 个子序列; $P(t)$ 为不同天气类型下的原始光伏功率序列。

对原始数据样本进行预处理后,即对 Elman 和 SSA-BiLSTM1 两个单一模型以及 SSA-BiLSTM2 网络的求解。对于 Elman 和 SSA-BiLSTM1 两个单一模型的求解,均以总辐射、环境温度以及相对湿度等气象数据以及分解后各序列的历史光伏发电功率值作为输入,分解后的光伏功率各子序列作为输出来并行训练这两个单一预测模型,在训练模型过程中最小化均方误差,以确定模型中各层的连接权重和偏置向量。SSA-BiLSTM2 网络的求解是以 Elman 和 SSA-BiLSTM1 模型的预测值作为输入,对应的光伏功率值作为输出来训练该网络,通过最小化均方误差,确定该网络中各层的连接权重和偏置向量。

综上,本文方法——基于 SSA-BiLSTM 非线性组合方法的预测模型得到求解后,通过 Elman 和 SSA-BiLSTM1 对分解后多个光伏功率子序列进行并行预测,然后将各子序列的预测值加和,得到在晴天、多云、雨天和突变天气 4 种类型天气下,两个单一模型的光伏发电功率预测值,最后基于 SSA-BiLSTM2 组合方法,对两个单一模型的预测值进行非线性加权组合,得到最终的光伏功率预测值。

2.3 模型评价指标

本文使用均方根误差 (RMSE) 和平均绝对误差 (MAE) 作为衡量不同模型预测性能的评价指标,具体数学表达式如下:

$$\gamma_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (D_i - \hat{D}_i)^2} \quad (9)$$

$$\gamma_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |D_i - \hat{D}_i| \quad (10)$$

式中: D_i 为光伏功率实际值; \hat{D}_i 为预测模型的光伏功率的计算值; n 为测试集的样本数。

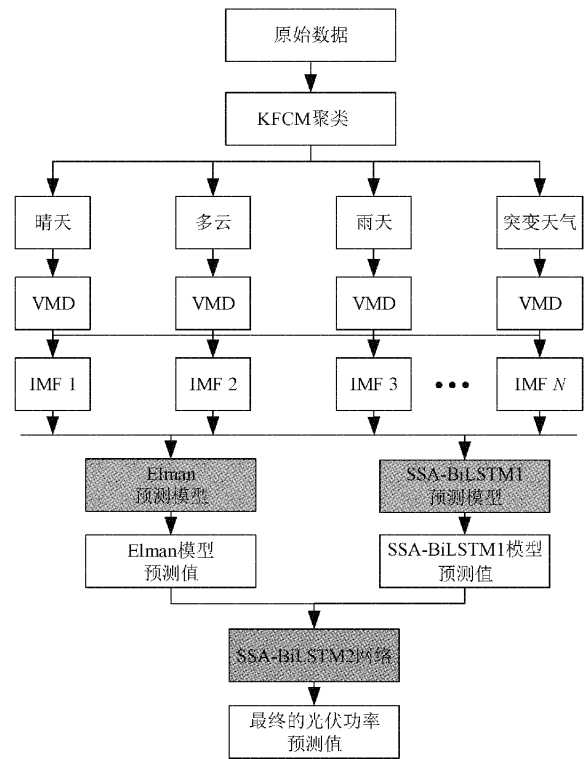


图 5 模型求解流程图

3 算例仿真分析

本文采用某个光伏电站 2020 年 6 月 1 日~9 月 30 日,共计 122 天的光伏电站实测光伏功率序列数据以及对应地区的气象数据,采样间隔均为 15 min,考虑到电站凌晨和晚上发电功率基本为零,因此保留每天 07:00~18:00 的数据,总计 45 个采样点。

3.1 数据预处理

基于所构建的每日气象特征向量,采用 KFCM 聚类算法对原始数据样本进行相似天气聚类,得到了 4 种不同的天气类型数据样本——晴天 48 天、多云 30 天、雨天 24 天和突变天气 20 天。将每种同天气类型数据样本的前 70% 数据设为训练样本,后 30% 数据设为测试样本。Elman、SSA-BiLSTM 在加入 KFCM 聚类算法前后的预测性能对比结果如表 1 所示。

表 1 聚类前和聚类后预测误差对比

预测模型	$\gamma_{\text{RMSE}}/\text{kW}$	$\gamma_{\text{MAE}}/\text{kW}$
Elman	7.63	5.81
SSA-BiLSTM	6.57	4.95
KFCM+Elman	6.15	4.68
KFCM+SSA-BiLSTM	5.32	4.19

由表格后 2 行的结果与前 2 行对比可知,加入 KFCM 聚类算法可有效加强数据之间的联系,提高模型的预测

性能。

基于KFCM聚类处理后,针对4种不同天气类型数据样本中的光伏功率序列,采用VMD进行分解,获得一系列子序列。为了充分挖掘光伏功率序列潜在的规律信息,以便模型具有更好的预测性能,模态分量 k 的设置显得尤为重要。本文采用中心频率法确定4种不同天气类型数据样本中的光伏功率序列最佳分解个数,晴天、多云、雨天和突变天气的最佳 k 值分别是6、6、5和4,其他相关参数设置如下:惩罚系数 $\alpha=2000$;初始中心频率 $\omega=1$;收敛误差 $\varepsilon=1\times 10^{-7}$ 。以雨天天气样本中的光伏功率序列为例,采用VMD分解方法,将其分解为一系列子序列,图6为雨天天气样本中的光伏功率序列以及分解后的各子序列。

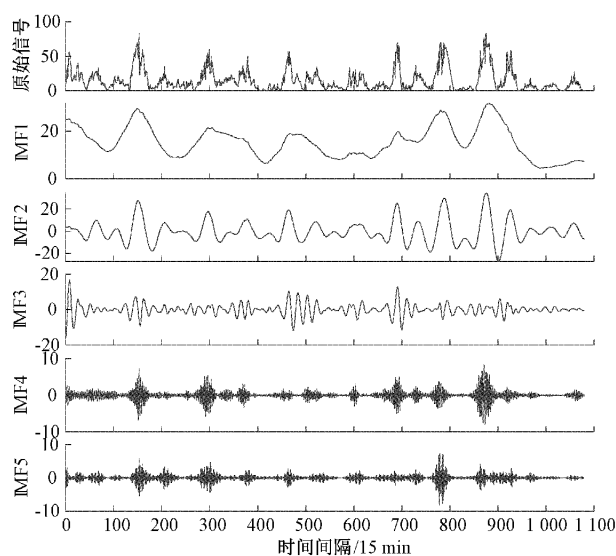


图6 雨天天气下光伏发电功率序列数据的VMD分解结果

从图6中可以看出,利用VMD分解方法对雨天天气下的光伏发电功率序列进行分解,得到了5个波动变化各异的子序列,其中IMF1是低频子序列,呈现出整体平稳的趋势,反映了雨天天气下光伏发电功率的总体波动情况,IMF2到IMF5为高频子序列,频率变化各有特点,可以反映出光伏发电功率在雨天天气下的局部变化特点。

综上所述,以雨天天气下的光伏发电功率序列为例,经过VMD分解后,得到一系列子序列,各子序列频率变化各有特点,没有出现明显的模态混叠,由此表明采用中心频率法确定雨天天气下光伏功率序列的分解个数 k 为5是较为合适的。

接着以雨天天气下的光伏功率数据为例,VMD+Elman、VMD+SSA-BiLSTM和本文方法在该数据下的损失函数曲线如图7所示。

由图7可知,本文方法相较于其他两种单一模型,收敛速率变快,即在同样的损失值下,本文方法所需的训练轮次

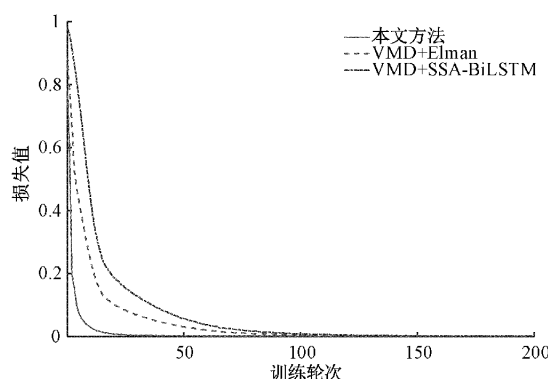


图7 雨天天气下各模型损失值对比

少于其他两种单一模型。

3.2 仿真结果分析

为了验证本文方法的预测性能,将本文方法与BiLSTM、Elman、SSA-BiLSTM、VMD+BiLSTM、VMD+Elman以及VMD+SSA-BiLSTM等6种单一模型进行对比,图8为本文方法与其他单一模型在晴天、多云、雨天和突变天气下的光伏功率预测效果对比,预测模型的相关参数如表2所示,为确保实验结果的可靠性,Elman和BiLSTM均采用试错法以确定最佳参数,麻雀搜索算法中,种群大小 $N=20$,迭代次数 $T_{max}=50$, $R_2=0.6$, $ST=0.4$,维度 $D=2$,警戒者和跟随者型麻雀分别占0.2和0.5,其余为发现者类型麻雀。

表2 预测模型相关参数设置

算法模型	网络结构	学习率	最大Epoch
Elman	182-71-10-1	0.000 1	500
BiLSTM	110-42-20-1	0.027 7	600
SSA-BiLSTM1	120-100-60-45-1	0.021 5	1 200
SSA-BiLSTM2	105-84-40-1	0.090 2	1 000

由图8(a)可知,处于晴天气象条件下,光伏发电功率曲线波动幅度较小,各预测模型得到的光伏功率预测曲线均基本贴近于光伏发电功率实测值曲线,这反映出各预测模型对晴天气象下的光伏功率具有较好的适应性,表现出较好的预测效果。

从图8(b)中可以看出,受多云气象条件下云层和太阳辐射较为复杂的变化情况影响,光伏发电功率曲线波动次数增多,此时各预测模型的拟合功率曲线与光伏功率实测值曲线偏差较大,对比的模型中,原始的Elman和BiLSTM模型追踪多云天气下的光伏发电功率变化效果较差,加入VMD分解后,得到的预测曲线与实际曲线的偏差要小于原始模型,与此同时SSA-BiLSTM和VMD+SSA-BiLSTM预测模型相较于BiLSTM和VMD+BiLSTM模型,得到的光伏功率预测曲线更为接近于实际功率曲线,这

表明经过麻雀搜索算法优化后的 BiLSTM 和 VMD+BiLSTM 模型具有更强的学习能力。

由图 8(c)可以知晓,受雨天气象条件下云量较多、辐射量较少的天气情况影响,光伏电站的整体光伏发电实测功率值小于晴天和多云天气类型,各预测模型得到的预测功率曲线与实测值曲线有较大偏离,但与其他单一模型相比,本文方法得到的功率预测曲线整体上更为接近于实测值曲线,由此表明,通过 SSA-BiLSTM 对 VMD+Elman 和 VMD+SSA-BiLSTM 两个单一模型进

行非线性加权组合,使得本文方法具有更强的预测性能。

由图 8(d)可知,气象条件未发生突变时,各模型追踪光伏功率变化效果较好,但气象条件发生突变后,光伏功率曲线频率波动变大,各预测模型得到的拟合功率曲线与光伏功率实测值曲线偏差较大,但相较于单一模型,本文方法得到的拟合功率曲线整体上更为接近于实测值曲线,这表明本文方法在突变天气时仍具有较强的抗干扰能力和泛化能力。

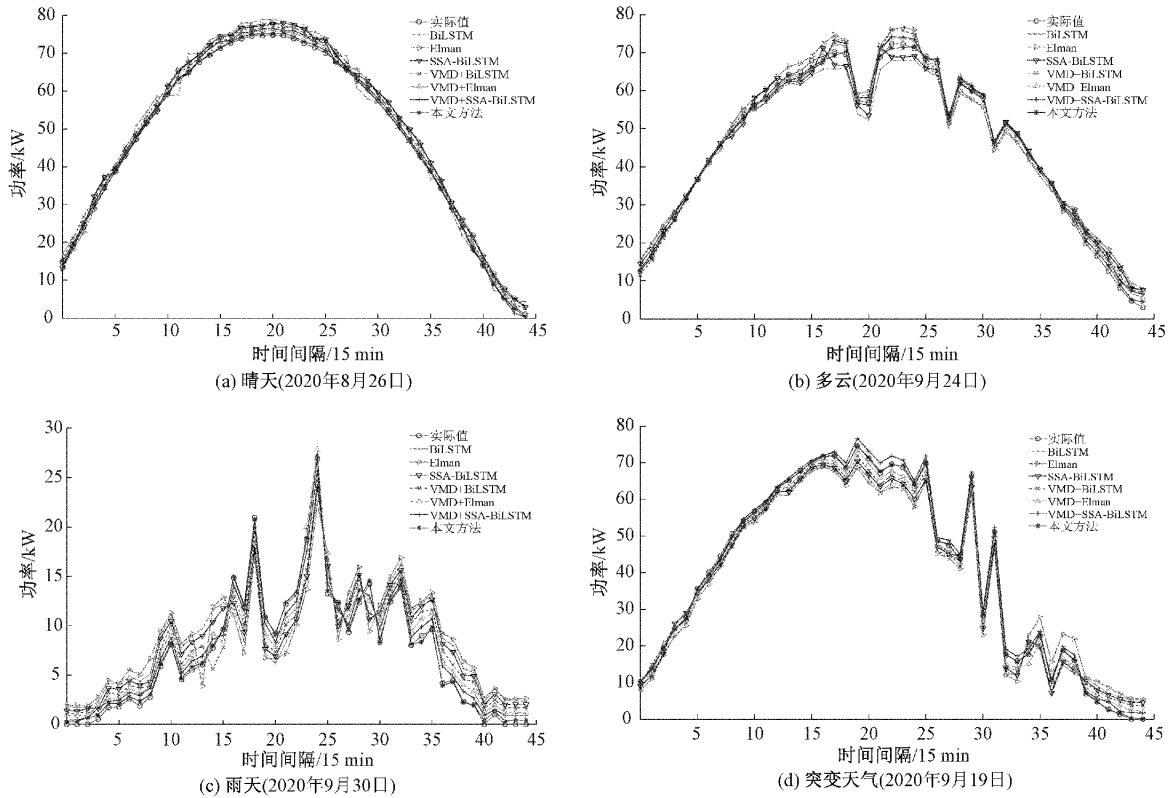


图 8 不同天气类型各模型预测结果

表 3 为本文方法与对比模型在 4 种天气类型测试样本数据中的预测误差对比。

表 3 不同模型的预测误差对比

预测模型	晴天		多云		雨天		突变天气	
	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW
BiLSTM	1.334	1.176	1.913	1.662	3.799	2.654	2.713	1.815
Elman	1.255	1.013	1.776	1.559	4.556	3.354	3.416	2.414
SSA-BiLSTM	0.798	0.679	1.346	1.064	2.845	1.977	2.312	1.485
VMD+BiLSTM	0.711	0.560	0.891	0.731	1.345	1.132	1.106	0.851
VMD+Elman	0.674	0.503	0.765	0.604	1.032	0.834	0.899	0.761
VMD+SSA-BiLSTM	0.702	0.542	0.803	0.629	0.986	0.791	0.868	0.736
本文方法	0.611	0.465	0.669	0.511	0.768	0.607	0.708	0.578

由表3给出的信息可知,相较于原始的Elman、BiLSTM模型,VMD+Elman和VMD+BiLSTM模型在4种天气类型测试集数据中取得的预测误差更小,这表明相较于直接将原始光伏发电功率数据输入到模型中进行预测,采用VMD分解策略将功率分解为一系列子序列后进行建模预测,可有效挖掘每个子序列的特征信息,使得模型映射输入与输出之间的非线性关系更加准确,同时有效避免原始光伏功率序列中,潜在的数据噪声的影响,提高预测精度,与此同时,将SSA-BiLSTM与BiLSTM以及VMD+SSA-BiLSTM与VMD+BiLSTM进行对应比较,SSA-BiLSTM、VMD+SSA-BiLSTM的预测误差小于BiLSTM和VMD+BiLSTM模型,由此说明,采用麻雀搜索算法优化双向长短期记忆网络的相关参数,可有效提高网络的预测性能,与此同时本文方法与对比的单一模型相比,在4种不同天气类型的测试样本数据中,得到的预测误差最小,相较于VMD+Elman,本文方法在晴天、多云、雨天和突变天气下的RMSE,分别降低了9.35%、

12.55%、25.58%和21.25%,MAE分别降低了7.55%、15.40%、27.22%和24.05%;相比于VMD+SSA-BiLSTM,本文方法在晴天、多云、雨天和突变天气下的RMSE,分别降低了12.96%、16.69%、22.11%和18.43%,MAE分别降低了14.21%、18.76%、23.26%和21.47%,也由此可以得出,采用SSA-BiLSTM组合方法对VMD+Elman和VMD+SSA-BiLSTM两个单一模型进行非线性加权组合,有效解决了单一模型对光伏功率预测时效果较差的问题,提高了预测性能。

为了验证本文采用的组合方法(SSA-BiLSTM)的优势,将本文采用的组合方法与最优加权组合方法^[10]、误差倒数法^[11]、遗传算法(genetic algorithm, GA)^[12]和SVM^[13]等4种不同组合方法进行对比,其中最优加权组合方法、误差倒数法和GA为线性组合方法,为保证实验的客观性,各组合方法聚合的单一模型都是VMD+Elman和VMD+SSA-BiLSTM,只是组合方法的不同,表4为基于不同组合方法的组合模型在4种天气类型测试样本数据中的实验结果。

表4 不同组合方法性能对比

组合方法	晴天		多云		雨天		突变天气	
	$\gamma_{\text{RMSE}}/\text{kW}$	$\gamma_{\text{MAE}}/\text{kW}$	$\gamma_{\text{RMSE}}/\text{kW}$	$\gamma_{\text{MAE}}/\text{kW}$	$\gamma_{\text{RMSE}}/\text{kW}$	$\gamma_{\text{MAE}}/\text{kW}$	$\gamma_{\text{RMSE}}/\text{kW}$	$\gamma_{\text{MAE}}/\text{kW}$
最优加权法 ^[10]	0.689	0.524	0.801	0.626	1.212	1.081	0.903	0.775
误差倒数法 ^[11]	0.679	0.510	0.796	0.620	1.172	0.943	0.890	0.753
GA ^[12]	0.670	0.501	0.780	0.616	1.012	0.825	0.866	0.731
SVM ^[13]	0.668	0.494	0.762	0.593	0.969	0.782	0.838	0.712
本文方法	0.611	0.465	0.669	0.511	0.768	0.607	0.708	0.578

由表4可知,采用不同组合方法的组合模型RMSE和MAE值都小于1.5,这表明采用组合方法都能很好地聚合VMD+Elman和VMD+SSA-BiLSTM两个单一预测模型各自的优势,实现在晴天、多云、雨天和突变天气下,对光伏功率变化情况良好的追踪,但结合表3可知,并不是所有组合方法在不同天气类型下都能有效聚合单一预测模型,取得相比单一模型更好的预测精度,与此同时,基于SVM和SSA-BiLSTM非线性组合方法的组合模型得到的RMSE和MAE值,明显小于基于最优加权组合方法、误差倒数法和GA的组合模型,由此表明,相较于线性组合模型中将单一模型之间描述为线性关系,基于非线性组合方法,将组合模型中单一模型之间表征为非线性关系能更好地聚合单一模型的各自优势,提高预测精度,并且相较于传统SVM非线性组合方法,本文采用的组合方法在晴天、多云、雨天和突变天气下的RMSE,分别降低了8.53%、12.20%、20.74%和15.51%,MAE分

别降低了5.87%、13.83%、22.38%和18.82%,也由此表明,本文采用的组合方法能更好地映射VMD+Elman和VMD+SSA-BiLSTM模型的预测值与最终预测结果之间复杂的非线性关系,使得本文方法具有更强的适应性。

为了进一步验证本文方法的泛化能力,以我国西部某个高纬度光伏电站(简称B地区)2020年1~5月的光伏功率及相关气象历史数据建立对比算例,该数据预处理方法与上文一致,以及预测模型相关参数的设置、对比的模型设置等。表5为不同模型在B地区不同天气类型测试集数据中的预测误差指标,表6为不同组合方法在B地区不同天气类型测试集数据中的预测误差指标。

由表5和6可知,相较于其他单一模型和基于其他不同组合方法的组合模型,本文方法在B地区4种不同天气类型测试样本数据中的RMSE和MAE值最小,这进一步表明本文方法具有较好的泛化能力。

表 5 不同模型在 B 地区不同天气类型测试集数据中的预测误差分析

预测模型	晴天		多云		雨天		突变天气	
	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW
BiLSTM	1.578	1.295	2.314	1.912	4.121	3.235	2.904	1.915
Elman	1.322	1.164	2.012	1.744	4.899	3.633	3.896	2.733
SSA-BiLSTM	0.894	0.795	1.973	1.698	3.227	2.543	2.634	1.618
VMD+BiLSTM	0.801	0.634	0.956	0.896	1.894	1.693	1.315	1.141
VMD+Elman	0.729	0.585	0.841	0.768	1.424	1.241	1.201	0.998
VMD+SSA-BiLSTM	0.743	0.622	0.882	0.799	1.202	1.195	0.983	0.897
本文方法	0.636	0.501	0.711	0.632	0.949	0.893	0.812	0.725

表 6 不同组合方法在 B 地区不同天气类型测试集数据中的性能分析

组合方法	晴天		多云		雨天		突变天气	
	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW	γ_{RMSE}/kW	γ_{MAE}/kW
最优加权法 ^[10]	0.738	0.613	0.877	0.790	1.578	1.452	1.143	0.973
误差倒数法 ^[11]	0.726	0.581	0.870	0.781	1.505	1.404	1.113	0.961
GA ^[12]	0.718	0.575	0.861	0.772	1.401	1.218	0.970	0.882
SVM ^[13]	0.701	0.558	0.832	0.754	1.211	1.198	0.964	0.859
本文方法	0.636	0.501	0.711	0.632	0.949	0.893	0.812	0.725

4 结 论

针对光伏发电功率组合预测模型中大多采用线性组合方法会导致组合模型预测精度降低的问题,本文提出一种基于 SSA-BiLSTM 非线性组合方法的预测模型,通过算例仿真可以获得以下结论:

基于 KFCM 聚类算法将光伏功率数据进行天气分类,可有效降低不同天气对光伏功率输出特性的影响,提高模型对光伏功率波动性的学习,以及在此基础上利用 VMD 对不同天气下的光伏功率分解,以提取光伏功率序列中的特征信息,为后续良好的预测奠定基础。

采用 SSA 算法对 BiLSTM 网络相关参数进行优化,以避免 BiLSTM 模型参数寻优难的问题,提高了模型的学习能力。

仿真结果表明,本文提出的基于 SSA-BiLSTM 非线性组合方法的预测模型,在不同天气下的预测误差优于单一模型,有效提升了预测精度,且验证了本文采用的组合方法的优越性。

参考文献

- [1] 朱正林,张冕.基于 AO 优化 VMD-CE-BiGRU 的光伏发电功率预测[J].国外电子测量技术,2022,41(10):56-61.
- [2] 丁明,虞海彪,刘练,等.基于多变量相空间重构和 RBF 神经网络的光伏功率预测方法[J].电子测量与仪器学报,2020,34(8):1-7.
- [3] 张雨金,杨凌帆,葛双治,等.基于 Kmeans-SVM 的短

期光伏发电功率预测[J].电力系统保护与控制,2018,46(21):118-124.

- [4] 刘志超,袁三男,唐万成.基于 BLSTM-随机森林的短期光伏发电输出功率预测[J].电源技术,2021,45(4):495-498.
- [5] 高超,孙谊娟,赵洪峰,等.基于 ICOA-LSTM 的短期负荷预测研究[J].电子测量技术,2022,45(13):88-95.
- [6] 李练兵,张佳,韩靖楠,等.基于 Elman 算法的光伏阵列的短期功率预测研究[J].太阳能学报,2017,38(6):1560-1566.
- [7] 张娜,任强,刘广忱,等.基于 VMD-GWO-ELMAN 的光伏功率短期预测方法[J].中国电力,2022,55(5):57-65.
- [8] 王开艳,杜浩东,贾嵘,等.基于相似日聚类和 QR-CNN-BiLSTM 模型的光伏功率短期区间概率预测[J].高电压技术,2022,48(11):4372-4388.
- [9] 谢小瑜,周俊煌,张勇军,等.基于 W-BiLSTM 的可再生能源超短期发电功率预测方法[J].电力系统自动化,2021,45(8):175-184.
- [10] 杨红军,王胜辉,李潇潇.基于最优加权组合模型的光伏出力预测[J].沈阳工程学院学报(自然科学版),2019,15(4):300-305.
- [11] 汤德清,朱武,侯林超.基于 CNN-LSTM-XGBoost 模型的超短期光伏功率预测[J].电源技术,2022,46(9):1048-1052.
- [12] 谭小钰,刘芳,马俊杰,等.基于 DBN 与 T-S 时变权重组合的光伏功率超短期预测模型[J].太阳能学报,

- 2021,42(10):42-48.
- [13] 王粟, 魏磊锋, 曾亮. 基于 GWO-SVM 与随机森林的组合光伏功率预测模型[J]. 昆明理工大学学报(自然科学版), 2021, 46(5): 82-88.
- [14] 贾志琴, 史元浩, 梁建宇, 等. 基于 Elman 神经网络的受热面积灰动态预测[J]. 电子测量与仪器学报, 2019, 33(9): 50-56.
- [15] 张加劲. 基于注意力机制和 CNN-BiLSTM 模型的航空发动机剩余寿命预测[J]. 电子测量与仪器学报, 2022, 36(8): 231-237.
- [16] 王义, 刘欣, 高德欣. 基于 BiLSTM 神经网络的锂电池 SOH 估计与 RUL 预测[J]. 电子测量技术, 2021, 44(20): 1-5.
- [17] 杨邓, 杨俊杰, 胡晨阳, 等. 基于改进 LSSVM 的短期电力负荷预测[J]. 电子测量技术, 2021, 44(18): 47-53.
- [18] CHEN J, ZENG G Q, ZHOU W, et al. Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization[J]. Energy conversion and management, 2018, 165: 681-695.
- [19] CHEN M R, ZENG G Q, LU K D, et al. A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM[J]. IEEE Internet of Things Journal, 2019, 6(4): 6997-7010.
- [20] 李飞, 王鸿玺, 谭阿峰, 等. 考虑用电行为分析的电力用户用电预测研究[J]. 电子测量技术, 2020, 43(1): 74-79.
- [21] 吴辰文, 李壮, 梁雨欣. 基于 Canopy 的 KFCM 聚类优化算法[J]. 西北大学学报(自然科学版), 2022, 52(3): 444-451.
- [22] 王冉, 后麒麟, 石如玉, 等. 基于变分模态分解与集成深度模型的锂电池剩余寿命预测方法[J]. 仪器仪表学报, 2021, 42(4): 111-120.

作者简介

袁建华, 博士, 副教授, 主要研究方向为光电子技术应用研究、无人机控制等。

E-mail: sdyjh@ctgu.edu.cn

蒋文军, 硕士研究生, 主要研究方向为新能源预测。

E-mail: 2500412214@qq.com