

面向光伏功率预测的残差深度学习模型

干逸飞¹ 吕品^{1*} 郑树泉²

¹(上海电机学院电子信息学院 上海 201306)

²(上海产业技术研究院 上海 201206)

摘要 为保证光伏功率预测模型在气象突变时具有较高的精度,提出用残差量化气象突变,并将其构造为一种新特征。应用最大信息系数(MIC)剔除无关的气象特征后,引入XGBoost模型得到残差序列。利用残差的自相关性,将上一时刻的残差作为当前时刻的新特征,构建面向光伏功率预测的残差深度学习模型。实验结果表明,在气象突变下,该模型能取得更高的精确度。

关键词 超短期光伏功率预测 XGBoost LSTM 预测修正

中图分类号 TP3 TM615

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.11.014

RESIDUAL DEEP LEARNING MODEL FOR PHOTOVOLTAIC POWER PREDICTION

Gan Yifei¹ Lü Pin^{1*} Zheng Shuquan²

¹(School of Electronic Information, Shanghai Dianji University, Shanghai 201306, China)

²(Shanghai Industrial Technology Institute, Shanghai 201206, China)

Abstract To ensure that the photovoltaic power prediction model has high accuracy when the meteorological abrupt changes, quantifying the weather abruptly by residuals is proposed and constructed as a new feature. After applying the maximum information coefficient (MIC) to eliminate the irrelevant meteorological features, the XGBoost model was introduced to obtain the residual series. Using the autocorrelation of the residuals, the residuals of the previous moment were used as the new features of the current moment to construct a deep learning model of residuals for photovoltaic power prediction. The experimental results show that the proposed model can achieve higher accuracy under sudden meteorological changes.

Keywords Ultra-short-term photovoltaic power prediction XGBoost LSTM Forecast correction

0 引言

近年来,作为实现“碳达峰”和“碳中和”目标的重要手段,光伏发电产业发展迅速,预计30年后,光伏发电量可占总耗电量的27%左右^[1-2]。然而,复杂多变的天气、云层的遮蔽和环境温度突变等偶然的气象突变因素,使得光伏输出功率具有很强波动性和随机性^[3-4],造成弃光电量居高不下,同时给光伏电站并网带来了安全隐患^[5]。因此,准确地预测光伏发电功率可以减少能源浪费,有益于电力系统的

稳定^[6-7]。

目前,主流的光伏发电功率预测技术分为物理建模方法和机器学习方法。物理建模方法需要相当详细的光伏电站参数,通过物理方程计算光伏发电功率,其缺陷在于预测精度较低、建模过程不方便,不同的组件需要单独建模,无法移植,难以大规模推广^[8]。机器学习方法的特点是不必考虑具体的物理细节,通过输入历史天气特征和历史发电功率,可以直接拟合出输入特征与输出光伏功率之间的隐含关系,易于应用,且预测精度较高^[9]。彭曙蓉等^[10]优化输入特征,提出了融合多种天气特征的XGBoost模型,利用相关系数法筛

选出与发电功率有较强相关性的特征,预测效果好于广泛使用的 SVM 模型。左远龙等^[11]提出一种 BP 神经网络的超参数优化方法,结果表明优化模型的预测结果与真实值的拟合程度很高。杨晶显等^[12]利用序列分解方法将光伏功率序列分解为多个分量,各自输入到基于注意力机制的 LSTM 中,通过增加模型数量和加入注意力层来提升模型的拟合精度,结果表明,提出的预测方法可以同时提高单点和多时间点的预测准确度。

上述方法从不同角度优化机器学习模型,在平稳的气象条件下得到了较高的预测精度,但是在面对区域性的天气突变时,均难以保证预测精度。为此,基于云图信息的预测方法受到广泛关注。文献[13]指出,短期内,光伏功率的异常波动主要由云层变化引起。陈志宝等^[14]从天空云图中获取与辐射变化相关的特征,采用 RBF 神经网络预测辐射值,然后通过辐射值计算光伏功率。这种间接预测法的精度明显好于无云图信息的传统模型,但不足之处在于观测设备成像范围和分辨率有限,而且在多云天气,会出现云增强现象^[15],很难得到准确和有效的云图信息。同时,云图信息也无法表示出其他降低预测模型精度的气象异常因素,例如数值天气预报中的不一致问题^[16]。

针对上述问题,本文利用残差来量化气象突变,并将其作为新特征来改善预测精度。直观上,残差主要源于数据集噪声、模型固有偏差和未知的气象突变,那么为使残差尽可能地代表未知的气象突变,就要最大程度地优化其余两个因素。首先应用最大信息系数(MIC)计算数据集中各个特征与光伏发电功率的相关性,剔除无关特征,降低数据集噪声,提高后续模型的预测精度。之后为了减少模型本身的固有偏差,应用泛化能力强、拟合精度高的 XGBoost 模型,将上一时刻 XGBoost 残差作为当前时刻的新特征输入到 LSTM 中,以此构建面向光伏功率预测的残差深度学习模型。实验结果表明:

1) 数据集经过 MIC 特征筛选后,各种机器学习模型的预测准确度都有所上升。

2) 实验得到的残差具有很高的 1 阶偏自相关性,这表明本文提出的用上一时刻的残差去近似下一时刻的残差是有意义的。

3) 相较于已有的方法,本文提出的残差深度学习模型,在气象突变的情况下能得到更高的预测准确度。

1 模型总述

1.1 数据处理

数据处理的流程如图 1 所示。

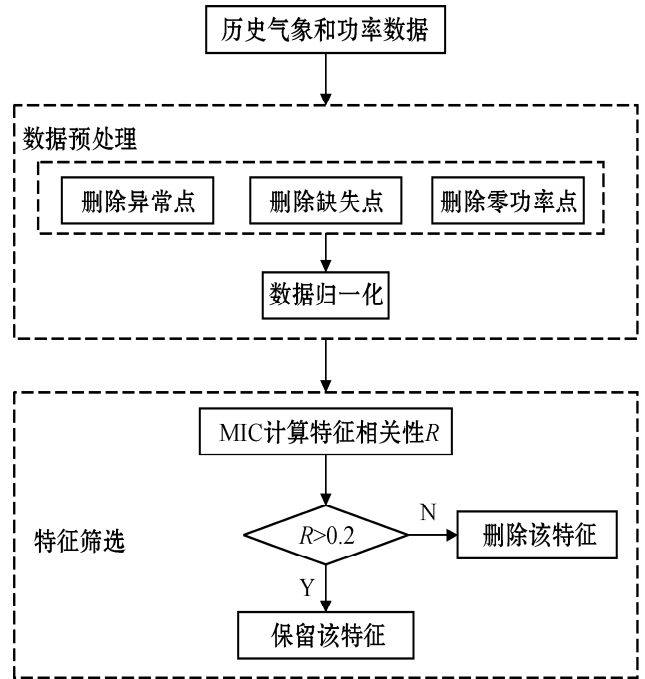


图 1 数据处理流程

1) 数据预处理。由于光伏发电的特点,原始的光伏数据集往往含有异常点、缺失点和大量的零功率点^[17]。这会对模型造成不良影响,需要提前加以去除。之后,需要对数据集中的气象特征进行归一化处理,以此加快模型的训练速度,防止 LSTM 模型无法收敛。归一化特征的计算式如下:

$$\mathbf{x}^* = \frac{\mathbf{x} - \min(\mathbf{x})}{\max(\mathbf{x}) - \min(\mathbf{x})} \quad (1)$$

式中: \mathbf{x} 代表特征向量, $\min(\mathbf{x})$ 和 $\max(\mathbf{x})$ 表示该特征的最小值和最大值,计算结果 \mathbf{x}^* 的范围会被缩放到 0 到 1 之间。

2) 特征筛选。无论是气象特征还是发电功率,都有很强的周期性和突变性,传统的线性相关系数很难衡量这样的数据中存在的非线性相关性。而 MIC 同时包含随机变量之间的线性和非线性相关性,可以很精准地剔除与发电功率无关的气象特征,降低数据集中的噪声,有利于提高模型准确度,从而提高残差中隐藏的气象突变的比例。本文设定的筛选阈值为 0.2。

1.2 模型构建

模型构建的流程如图 2 所示。

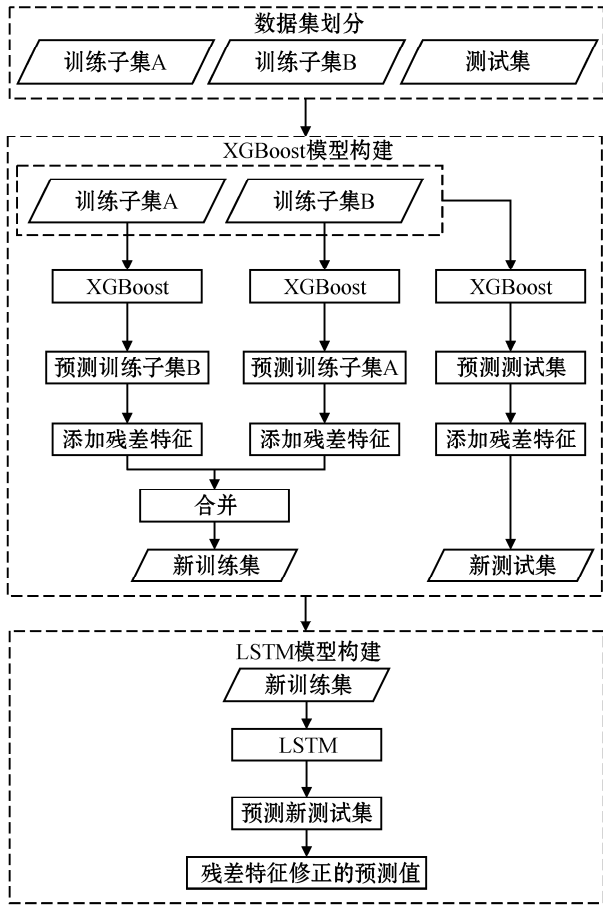


图 2 模型构建流程

1) 数据集划分。与传统的数据集划分不同的是,本文以一个发电日为最小分割单位,将数据集划分为训练集和测试集。这不是一种完全随机的划分方法。其原因是:(1) 实际应用中,光伏功率预测是基于时间顺序的连续预测;(2) 本文方法利用了残差的自相关性,数据需要满足时间上的连续性。此外,为了更全面地评价本文方法,还需要查阅历史天气数据,从中挑选出涵盖各种天气的发电日,形成全天候测试集。

为了获得残差特征,需要多个训练集。因为要使 LSTM 模型能够预测带有残差特征的测试集,训练集也必须要有残差特征。于是,将训练集划分为训练子集 A 和训练子集 B,并分别为它们构建各自的残差特征。

2) XGBoost 模型构建。XGBoost 模型的泛化能力和预测精度均比传统的机器学习模型好,已被广泛运用于各种机器学习竞赛中,并取得极好的效果。因此,利用高性能的 XGBoost 模型可以有效降低残差中隐藏的模型固有偏差,使残差特征更有效地量化气象突变。获得残差特征的步骤如下:(1) 用训练子集 A 训练 XGBoost 模型,然后用训练子集 B 作为测试集,取得训练子集 B 的残差特征。(2) 用训练子集 B 训练 XGBoost 模型,然后用训练子集 A 作为测试集,取得训练子集 A 的残差特征。(3) 用训练子集 A 和训练子

集 B 训练 XGBoost 模型,然后用全天候测试集作为测试集,取得全天候测试集的残差特征。

实验中,观察到残差序列的 1 阶偏自相关系数很高,故相邻时刻的残差值可以互相近似代替,但是为了防止未来信息泄露,只能用上一时刻的残差作为当前时刻的残差特征。所以在添加残差特征时,要将残差序列延后一个时刻,再与原始特征合并。添加残差特征后,形成新的训练子集 A、训练子集 B 和测试集。最后将新训练子集 A 和新训练子集 B 合并成为新训练集。

3) LSTM 模型构建。与 XGBoost 模型的单点预测不同,LSTM 模型擅长捕捉时间序列数据中的依赖关系,广泛应用于时间序列预测任务,效果显著^[18-19]。

LSTM 模型将在具有残差特征的新数据集上训练和预测,得到经过残差特征修正的预测值,以此提高预测准确度。

2 相关原理

2.1 MIC

互信息 (MI) 可以衡量随机变量 S 和随机变量 T 之间的相关程度^[20]。计算式如下:

$$I(S;T) = \sum_{s \in S} \sum_{t \in T} p(s,t) \log_2 \left(\frac{p(s,t)}{p(s)p(t)} \right) \quad (2)$$

式中: $p(s,t)$ 表示联合概率分布, $p(s)$ 和 $p(t)$ 分别是边缘概率分布。

但是,传统的 MI 没有上界,也无法归一化,对于不同的特征,产生的结果不能从数值的大小上进行比较。文献[21]提出的 MIC 方法解决了 MI 的这两个缺点,计算得出的 MIC 值位于 $[0,1]$ 之间,其值的大小,代表两个随机变量之间的相关性强弱。

MIC 用网格图来刻画随机变量所处的二维空间,同时在随机变量 S 和随机变量 T 的方向上划出一定的区域,划分区域的尺度可以有很多种,然后查看各区域中数据点的分布情况,并求出各区域的互信息,之后选择各区域中最大的互信息值,最后进行归一化。最大互信息系数的归一化公式如下:

$$mic(S;T) = \max_{a \times b < B} \frac{I(S;T)}{\log_2 \min(a,b)} \quad (3)$$

式中:随机变量 S 进行了 a 等分,随机变量 T 进行了 b 等分, B 是一个变量,通常设置为数据总数的 0.6 次方。

2.2 XGBoost

XGBoost 通过不断地生成新的分类回归树 (CART) 来预测上一棵树的残差,从而不断地改善模型预测效果。每个样本都会被加入到一个叶子节点中,一个叶

子节点上所有样本的均值称为叶子权重,最后将所有叶子权重求和,得到样本的预测值^[22]。样本 i 的预测结果表示为:

$$y_i^{(c)} = \sum_c f_c(x_i) \quad (4)$$

式中: C 为集成模型中树的数量, $f_c(x_i)$ 为第 c 棵树的叶子权重。

XGBoost 的迭代方程如下:

$$\begin{cases} y_i^{(0)} = 0 \\ y_i^{(c)} = y_i^{(c-1)} + f_c(x_i) \quad c > 0 \end{cases} \quad (5)$$

式中: $y_i^{(0)}$ 表示初始树, $f_c(x_i)$ 为第 c 棵树的叶子权重, $y_i^{(c)}$ 和 $y_i^{(c-1)}$ 表示迭代到第 c 棵树和第 $c-1$ 棵树时模型的输出值。

XGBoost 一方面采用二阶泰勒公式展开损失函数提高精确度,另一方面引入正则化项,表示对复杂度的约束,避免过拟合,表达式为:

$$\Omega(f_c) = \frac{1}{2}\lambda \|\omega\| + \gamma N \quad (6)$$

式中: λ 为 L2 正则化因子, ω 为叶子权重, N 代表叶子节点的个数, γ 为叶子节点的惩罚因子。

XGBoost 的目标函数表达式如下:

$$O^{(t)} = \sum_{i=1}^n l(y_i, y_i^{(t-1)} + f_k(x_i)) + \Omega(f_k) \quad (7)$$

将式(7)展开后得到:

$$O^{(t)} = \sum_{i=1}^n \left[g_i f_k(x_i) + \frac{1}{2} h_i f_k^2(x_i) \right] + \Omega(f_k) \quad (8)$$

式中: g_i 为损失函数的一阶导, h_i 为损失函数的二阶导, n 为样本数。

2.3 LSTM

LSTM 由 Hochreiter 等^[23] 提出,加入了细胞状态来携带可以穿越多个单元的信息,防止早期的信息在处理过程中消失。同时又引入 3 个门控结构,用来对细胞状态进行更新。LSTM 单元的内部结构如图 3 所示。

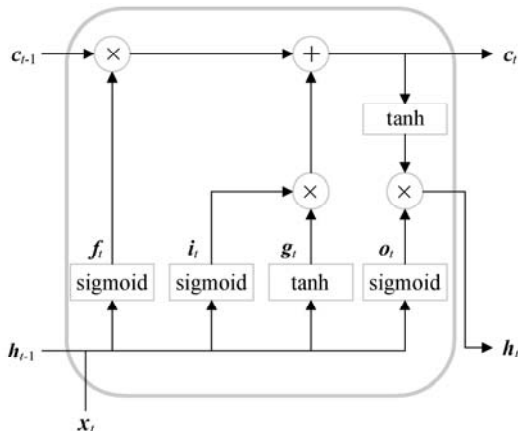


图 3 LSTM 单元结构

图 3 中, x_t 为当前的输入信息, h_{t-1} 和 h_t 分别表示上一时间步和当前时间步的隐藏状态,Sigmoid 和 tanh 为激活函数。 c_{t-1} 和 c_t 分别表示上一时间步和当前时间步的细胞状态, f_t 为遗忘门, i_t 为输入门, o_t 为输出门, g_t 为当前候选的细胞状态,

f_t 影响 c_{t-1} 中的信息变化,表达式为:

$$f_t = \text{sigmoid}(W_f[h_{t-1}, x_t] + b_f) \quad (9)$$

i_t 影响 g_t 中有多少信息会更新到 c_t 中,计算式如下:

$$g_t = \text{tanh}(W_g[h_{t-1}, x_t] + b_g) \quad (10)$$

$$i_t = \text{sigmoid}(W_i[h_{t-1}, x_t] + b_i) \quad (11)$$

当前的 c_t 更新式如下:

$$c_t = f_t \times c_{t-1} + i_t \times g_t \quad (12)$$

c_t 经过 f_t 和 i_t 的更新后,就可以通过 o_t 得到当前的 h_t 了,表达式分别如下:

$$o_t = \text{sigmoid}(W_o[h_{t-1}, x_t] + b_o) \quad (13)$$

$$h_t = o_t \times \text{tanh}(c_t) \quad (14)$$

式中: W 和 b 表示各个门控单元和隐藏状态的权重和偏置。

最后,当前的预测结果 y_t 可以根据目标任务的不同,通过变换 h_t 得到。

3 实验与结果分析

本文实验基于 Python 3.6.13、Keras 2.3.1 和 Scikit-learn 0.24.2。实验采用的数据来自华东地区某一风光混合发电站,记录了从 2020 年 6 月至 9 月的实时气象数据和光伏功率序列,每隔 15 分钟采样一次。其中环境气象特征有平均气压、平均温度、平均湿度、平均风速和风向等 5 种。太阳辐射特征有直射辐照度、散射辐照度和光照度等 3 种。在数据预处理阶段,剔除掉功率为 0 的数据点、缺失点和异常点,之后对 8 种气象特征进行归一化处理。

本文采用 3 种误差评估指标,计算式如下:

$$M_{AE} = \frac{1}{n} \sum_{i=1}^n |t_i - p_i| \quad (15)$$

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (t_i - p_i)^2} \quad (16)$$

$$M_{APE} = \frac{100}{n} \sum_{i=1}^n \left| \frac{t_i - p_i}{t_i} \right| \quad (17)$$

式中: p_i 为预测值, t_i 为实测值, n 为样本数。

3.1 气象突变

气象突变的类型有很多,而在本文采用的气象数据集中,没有标注气象变化的信息。因此本文是从光伏输出功率的波动角度,间接观察某一时刻的气象

突变。

图 4 和图 5 分别为数据集中 8 月 16 日和 9 月 2 日的光伏输出功率,通过查阅当地的历史天气数据得知当天的天气类型。

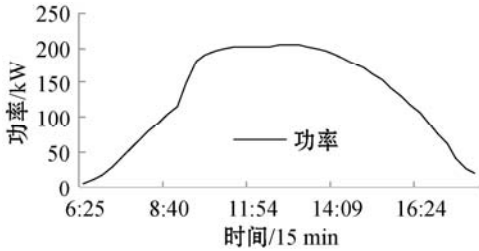


图 4 晴天的光伏输出功率

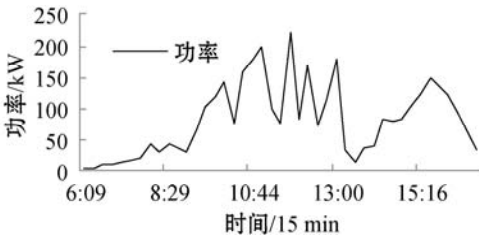


图 5 中雨转多云天的光伏输出功率

如图 4 所示,晴天时,光伏发电组件受到的干扰较少,功率值呈现平稳上升和下降,在中午达到峰值。

如图 5 所示,在中雨转多云的天气条件下,光伏输出功率呈现无序波动,在中午时甚至接近 0。

显然,晴天的气象干扰远远少于非晴天,因此光伏组件在晴天时的输出功率要比非晴天更加平稳。于是,气象突变可以用宽泛的天气类型来表示。在实验中,除晴天外,其余发电日均标记为气象突变日,用于后续的实验对比。

3.2 MIC 特征筛选

实验通过 Python 第三方库 minepy 中的 MINE 类来计算特征与发电功率间的 MIC,筛选阈值设置为 0.2,计算结果如表 1 所示。

表 1 特征与发电功率之间的 MIC

特征	MIC
气压	0.19
温度	0.22
湿度	0.36
风向	0.09
风速	0.14
光照强度	0.81
散射辐照度	0.62
总辐照度	0.89

从表 1 可知,有 3 个特征的 MIC 小于 0.2,分别是气压、风速和风向。为了验证 MIC 特征筛选法可以普

遍地提高各种模型的预测精度,本文采用 SVM、BP 和 XGBoost 进行对比实验。SVM 和 XGBoost 的参数设置为默认,BP 设置为 3 层,每层的神经元数分别为 256、256 和 1。特征筛选前后的预测效果对比如表 2 所示,MIC 前缀表示该模型在筛选后的数据集上训练并预测。

表 2 MIC 特征筛选前后模型精度的变化

模型	MAE/kW	RMSE/kW	MAPE/%
SVM	10.51	15.82	20.81
MIC-SVM	9.41	14.58	18.55
BP	8.47	12.89	15.79
MIC-BP	7.82	12.45	13.85
XGBoost	6.41	10.53	11.56
MIC-XGBoost	6.30	10.28	10.54

由表 2 可知,在剔除气压、风速和风向之后,各个模型的准确度均得到了较大的改善。同时,MIC-XGBoost 模型的三种误差指标均为最优。这说明:1) MIC 特征筛选方法通过降低数据集的维度和噪声,可以有效提升模型预测精度。2) XGBoost 模型作为一种集成模型,具有很强的学习、拟合能力,降低了单一模型固有的偏差。这表明本文采用的 MIC 特征筛选法和 XGBoost 模型可以有效地提高残差来源中未知气象突变的比例。

本文后续实验中采用的数据集全部是经过 MIC 特征筛选后的数据集。

3.3 残差分析

在序列数据中, x_i 和 x_{i+k} 之间的相关性受到中间序列值的影响,即同时存在间接相关性和直接相关性,而偏自相关函数(PACF)是剔除了间接相关性后的自相关性度量。

实验采用 Python 第三方库 statsmodels 中的 pacf 方法来计算 MIC-XGBoost 残差的偏自相关系数,计算结果如表 3 所示。

表 3 不同时滞的偏自相关系数

时滞	偏自相关系数
0	1.00
1	0.72
2	0.12
3	0.08
4	0.06
5	-0.03

由表 3 可知,残差的 1 阶偏自相关系数高达 0.72,相邻时刻的残差具有很强的相关性。从这个角度看,

用上一个时刻的残差来近似下一时刻的残差,作为量化下一时刻气象突变的新特征是有意义的。

3.4 残差特征

为了验证残差特征的有效性,实验中构建了 2 个 LSTM 模型,分别为原始 LSTM 模型和基于残差特征的 LSTM-Res 模型。以上 2 个模型的结构均为 2 个 LSTM 层和 1 个 Dense 层的线性叠加,每层神经元个数分别设置为 256、256 和 1,dropout 设置为 0.1。

考虑到光伏数据集较小且波动较大,实验采用 Full-Batch 训练,每批样本数 batch_size 设置为 10 000,迭代次数设置为 3 000,时间步长设置为 3。通过 Keras 回调函数保存最优模型。其余参数均为默认。以上 2 个模型在全天候测试集上预测效果的对比如表 4 所示。

表 4 模型的预测效果变化

模型	MAE/kW	RMSE/kW	MAPE/%
LSTM	5.82	8.90	10.68
LSTM_Res	5.53	8.14	9.84

由表 4 可知,LSTM-Res 相比于原始的 LSTM,三种模型评估指标分别减少 0.29 kW、0.76 kW 和 0.84 百分点。这表明残差特征可以提升模型预测效果。

为了直观地分析 LSTM-Res 模型在气象突变时的预测效果,将其在不同气象条件下的预测结果展示在图 6、图 7 和图 8 中,同时采用 XGBoost 和 LSTM 作为对比。

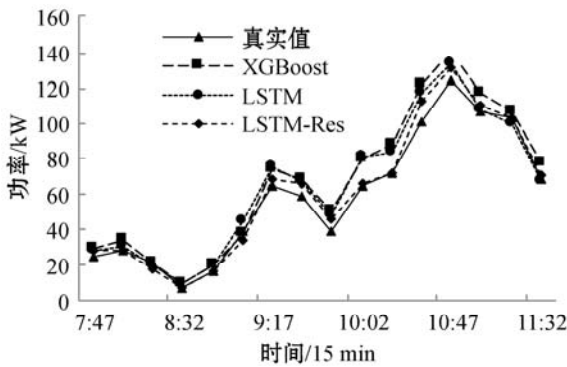


图 6 多云天的预测

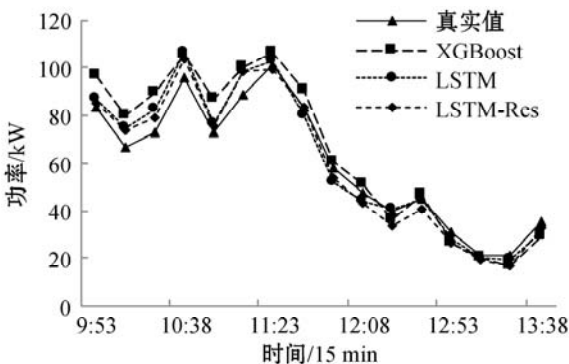


图 7 阴雨天的预测

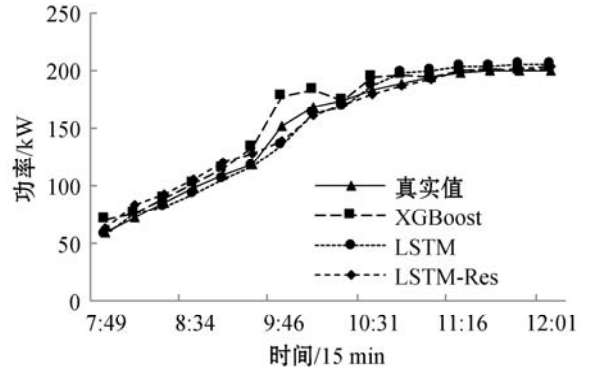


图 8 晴天的预测

如图 6 所示,在多云条件下,短期的气象突变影响较大。真实功率以曲折的方式上升,波动幅度较大。在 9:17 至 10:02 的光伏功率下降过程中,3 种模型均高估了光伏功率,但是 LSTM-Res 更贴近真实值。在 10:02 至 10:47 的光伏功率的上升过程中,XGBoost 和 LSTM 模型还是高估了真实值,而 LSTM-Res 模型能良好地拟合真实值。

如图 7 所示,在阴雨条件下,短期的气象突变影响非常大。真实功率在 9:53 至 12:08 这段时间内出现了无序波动,3 种模型均高估了真实值,其中 XGBoost 模型的预测效果最差,而 LSTM-Res 模型高估的幅度最小。此后,真实功率持续下降,各模型均能良好地拟合。

观察图 8 可知,在平稳的晴天,真实功率有序上升,各模型均能良好地拟合真实值,但是在 9:46 至 10:31,XGBoost 模型的预测功率明显偏离了真实功率。

各个模型在全天候测试集上的误差指标如表 5 所示。

表 5 3 种模型的预测精度

天气	模型	MAE/kW	RMSE/kW	MAPE/%
晴	XGBoost	5.62	8.43	9.68
	LSTM	5.37	7.62	8.74
	LSTM-Res	4.55	6.15	8.32
气象突变	XGBoost	6.54	10.86	10.67
	LSTM	5.97	9.30	12.01
	LSTM-Res	5.88	8.73	10.36
平均	XGBoost	6.30	10.29	10.41
	LSTM	5.82	8.90	10.68
	LSTM-Res	5.53	8.14	9.84

由表 5 可知,在不同的天气条件下,本文提出的

LSTM-Res 的预测精度都是最高的。平均来看,LSTM-Res 相比于 XGBoost 和 LSTM,MAE 分别减少 0.77 kW 和 0.29 kW, RMSE 分别减少 2.15 kW 和 0.76 kW, MAPE 分别减少 0.57 个百分点和 0.84 百分点。

4 结 语

本文提出利用残差来量化气象突变的方法。直觉上,短期内的气象变化是相似的、连续的。而通过计算得出,残差的 1 阶偏自相关系数很高。因此,相邻时刻的残差可以互相近似,但是为了避免未来信息泄露,只能用上一个时刻的残差来近似替代下一时刻的残差。由此衍生出将上一时刻的残差作为当前时刻的新特征的操作。

为了尽可能提高残差中隐藏的未知气象突变的比例,本文使用 MIC 筛选特征以及引入学习能力较强的 XGBoost 来生成残差。将残差作为 LSTM 模型的一种新特征,模型在气象突变的情况下,依然能良好地拟合光伏发电功率,同时预测误差较传统模型大幅下降,这表明本文提出的面向光伏功率预测的残差深度学习模型 LSTM-Res 具有更好的适用性。

参 考 文 献

[1] 吕鑫,祁雨霏,董馨阳,等. 2020 年光伏及风电产业前景预测与展望[J]. 北京理工大学学报(社会科学版),2020,22(2):20-25.

[2] McElroy M, Chen X. Wind and solar power in the united states: Status and prospects[J]. CSEE Journal of Power and Energy Systems,2017,3(1):1-6.

[3] Notton G, Voyant C. Forecasting of intermittent solar energy resource[M]//Advances in Renewable Energies and Power Technologies. Elsevier,2018.

[4] Wang J, Zhong H, Lai X, et al. Exploring key weather factors from analytical modeling toward improved solar power forecasting[J]. IEEE Transactions on Smart Grid,2019,10(2):1417-1427.

[5] 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.

[6] 陈懿,刘涤尘,吴军,等. 分布式光伏电源对配电网电压波动的影响研究[J]. 电测与仪表,2018,55(14):27-32.

[7] 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.

[8] Dolara A, Leva S, Manzolini G. Comparison of different physical models for PV power output prediction[J]. Solar Energy,2015,119:83-99.

[9] Voyant C, Notton G, Kalogirou S, et al. Machine learning methods for solar radiation forecasting: A review[J]. Renewable Energy,2017,105:569-582.

[10] 彭曙蓉,郑国栋,黄土峻,等. 基于 XGBoost 算法融合多特征短期光伏发电量预测[J]. 电测与仪表,2020,57(24):76-83.

[11] 左远龙,黄玉水,杨晓辉,等. 基于 PFA-MBAS-BP 神经网络模型的光伏发电短期预测[J]. 电力系统保护与控制,2020,48(15):84-91.

[12] 杨晶显,张帅,刘继春,等. 基于 VMD 和双重注意力机制 LSTM 的短期光伏功率预测[J]. 电力系统自动化,2021,45(3):174-182.

[13] Barbieri F, Rajakaruna S, Ghosh A. Very short-term photovoltaic power forecasting with cloud modeling: A review[J]. Renewable and Sustainable Energy Reviews,2017,75:242-263.

[14] 陈志宝,丁杰,周海,等. 地基云图结合径向基函数人工神经网络的光伏功率超短期预测模型[J]. 中国电机工程学报,2015,35(3):561-567.

[15] Gueymard A. Cloud and albedo enhancement impacts on solar irradiance using high-frequency measurements from thermopile and photodiode radiometers. Part 1: Impacts on global horizontal irradiance[J]. Solar Energy,2017,153:755-765.

[16] 郭换换,智协飞,段明铿,等. 数值天气预报中的不一致性问题综述[J]. 气象科学,2016,36(1):134-140.

[17] 邱明,鲁冠军,吴昊天,等. 基于数据清洗与组合学习的光伏发电功率预测方法研究[J]. 可再生能源,2020,38(12):1583-1589.

[18] 吕海灿,王伟峰,赵兵,等. 基于 Wide & Deep-LSTM 模型的短期台区负荷预测[J]. 电网技术,2020,44(2):428-436.

[19] 王依宁,解大,王西田,等. 基于 PCA-LSTM 模型的风电机网相互作用预测[J]. 中国电机工程学报,2019,39(14):4070-4081.

[20] Peng H, Long F, Ding C. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2005,27(8):1226-1238.

[21] Reshef D, Reshef Y, Finucane H, et al. Detecting novel associations in large data sets[J]. Science,334(6062):1518-1524.

[22] Chen T, Guestrin C. XGBoost: A scalable tree boosting system[C]//22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,2016:785-794.

[23] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation,1997,9(8):1735-1780.