

考虑季节特性与数据窗口的短期光伏功率预测组合模型

张静¹, 熊国江^{1,2}

(1. 贵州大学电气工程学院, 贵州 贵阳 550025;

2. 贵州大学勘察设计研究院有限责任公司, 贵州 贵阳 550025)

摘要: 光伏功率的间歇性和随机性因季节变化呈现出不同的特点, 考虑季节特性对提高光伏功率预测精度具有重要意义。因此, 文中提出一种考虑季节特性和数据窗口的短期光伏功率预测组合模型。首先, 通过皮尔逊相关系数法确定对光伏功率贡献度高的气象因素, 降低预测模型的输入特征维数。其次, 对比不同季节下不同模型的光伏功率预测精度, 选择光伏功率预测误差最小且相关性最低的2个模型构建组合模型, 即门控循环单元(gated recurrent unit, GRU)模型和极限梯度提升(extreme gradient boosting, XGboost)模型。然后, 分析历史气象数据中不同输入窗口对GRU-XGboost模型预测精度的影响, 确定最优数据窗口。最后, 在此基础上分别采用GRU和XGboost对光伏功率进行预测, 将2个预测结果加权组合得到最终预测结果。结果表明, 与其他模型相比, 所提模型具有更强的适应性和更高的预测精度。

关键词: 短期光伏功率预测; 季节特性; 数据窗口; 门控循环单元(GRU); 极限梯度提升(XGboost); 组合模型

中图分类号: TM615

文献标志码: A

文章编号: 2096-3203(2025)01-0183-10

0 引言

随着环境问题的日益突出和能源需求的不断增长, 太阳能因其高效、环保和分布广泛等特点得到大量应用。然而光伏功率的间歇性和随机性给电力系统的稳定性带来了严峻的挑战, 因此高精度的光伏功率预测对降低发电成本和制定合理调度计划具有重要意义^[1-9]。

根据光伏功率预测方法的不同, 可分为物理预测方法、数据驱动预测方法和组合预测方法。物理预测方法因模型不准确、基础信息有误等造成的系统性偏差, 导致预测精度通常不理想^[10]。随着人工智能的发展, 以机器学习算法为代表的驱动预测方法得到广泛关注, 机器学习算法可以充分挖掘数据的内在特征和隐藏的变化规律, 能更好地拟合复杂的非线性映射关系^[11-13]。文献[14-15]基于极限梯度提升(extreme gradient boosting, XGboost)建立了光伏功率预测模型, XGboost支持列采样和并行处理, 因此训练时间更短。文献[16]通过优化门控循环单元(gated recurrent unit, GRU), 解决了数据维度过高导致GRU预测性能下降的问题。文献[17]使用典型气象年方法生成典型气象年数据, 然后结合广义回归神经网络实现光伏功率预测。任何单一模型都会因其自身的局限性导致在某种场景下出现较大预测误差。GRU和长短期记忆(long

short term memory, LSTM)网络同为循环神经网络的变体, 不可并行的弊端导致二者在面对大规模样本时训练时间长。XGboost模型过度依赖训练样本的组成, 当出现训练集之外的数值时, 预测效果较差。组合预测方法可以综合每种模型的优点, 通过取长补短的方式充分发挥各模型的优势。文献[18]通过选择与测试日相似的历史数据作为模型的训练集, 基于卷积神经网络(convolutional neural network, CNN)和双向LSTM构建了光伏功率区间概率预测模型。文献[19]首先使用变分模态分解对原始光伏功率序列进行降噪处理得到若干子序列, 然后选择改进后的核极限学习机(kernel based extreme learning machine, KELM)对子序列分别预测并重构得到最终光伏功率预测值。

光伏功率通常与天气类型相关, 利用单一预测模型很难顾及多样化的天气类型, 因此需要构建一种精细化的光伏功率预测模型, 适应不同的预测场景。文献[20]将天气聚类为晴天、多云与阴雨天, 分析不同天气类型下气象因素与光伏功率的相关性, 选择相关性高的气象因素构建模型输入特征。文献[21]使用历史光伏数据确定历史天气类型, 从而提高历史天气数据的预报准确率, 利用LSTM实现光伏功率预测。文献[22]通过K-means++聚类算法将历史光伏数据按照天气类型进行分类, 采用CNN和LSTM构建光伏功率时间序列预测模型。除了天气类型, 历史气象数据输入的时间维度也会影响光伏功率预测精度。文献[23]将前一时刻历史气象数据进行预处理, 结合光伏功率数据构建光伏

功率预测模型的输入序列。

历史光伏数据包含历史气象数据和历史光伏功率数据,不同季节下气象因素差距很大,因此以全年光伏数据建模会出现模型泛化性不足的问题。同时,历史光伏数据蕴含信息的挖掘方式也会影响光伏功率预测模型的精度。文献[14-23]从天气分类和挖掘历史光伏数据内在规律2个方面提高光伏功率预测精度,但仍有一定的不足:(1)基于天气类型构建光伏功率预测模型能有效提高其泛化能力,但会因某些天气条件下数据量少导致建模困难;(2)将历史光伏数据作为光伏功率预测模型输入特征之一,考虑了历史光伏数据对未来光伏功率预测的指导作用,但并未考虑不同的数据窗口对光伏功率预测模型精度的影响。

针对以上存在的问题,文中提出一种考虑季节特性和数据窗口的GRU-XGboost短期光伏功率组合预测方法。首先,采用皮尔逊相关系数法对历史气象数据进行相关性分析,剔除相关性低的气象因素。其次,筛选出不同季节中泛化性最好、误差分布相关性最低的2个单模型构建组合模型。然后,研究不同的数据窗口与GRU-XGboost模型预测精度的关系。最后,2个单模型分别输入气象数据得到初步光伏功率预测结果,再根据单模型的权重系数计算最终光伏功率预测结果。实验结果表明,文中提出的光伏功率预测方法有更高的预测精度。

1 季节特性分析

光伏发电与气候、环境和地理位置等因素密切相关,其中气候是主要影响因素。不同季节的气候对光伏功率的影响不同,光伏功率通常随着气候的变化而变化^[24-25]。以宁夏某光伏场站数据作为模型输入样本,重点研究不同气象因素对光伏功率的影响和光伏功率的季节特性。

1.1 气象因素分析

气候是影响光伏发电的主要因素之一,气候变化可以通过温度、湿度、太阳辐照度和风速等气象因素表达。基于光伏发电的特性,文中采用皮尔逊相关系数法^[26]分析每个气象因素与光伏输出功率之间的相关程度。假设 X 和 Y 为2个样本:样本 X 包含 r 个观测值, $X=\{x_1, x_2, \dots, x_j, \dots, x_r\}$,样本 Y 包含 r 个观测值, $Y=\{y_1, y_2, \dots, y_j, \dots, y_r\}$,则皮尔逊相关系数法的表达式为:

$$S = \frac{\sum_{j=1}^r (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum_{j=1}^r (x_j - \bar{x})^2 (y_j - \bar{y})^2}} \quad (1)$$

式中: S 为皮尔逊相关系数; \bar{x} 为样本 X 的平均值; \bar{y} 为样本 Y 的平均值。

各气象因素与光伏功率的皮尔逊相关系数 S 如表1所示,| S |越大,则相关性越高。由表1可得,气温、云层不透明度、太阳散射辐射、太阳直接散射、太阳辐照度、相对湿度和天顶角与光伏功率相关性较高,因此选择这7个气象因素组成光伏功率预测模型的输入特征。

表1 气象因素与光伏功率的皮尔逊相关系数

Table 1 The Pearson correlation coefficient of meteorological factors and photovoltaic power

气象因素	S	气象因素	S
气温	0.458	大气可降水量	0.078
方位角	-0.013	相对湿度	-0.451
云层不透明度	-0.183	降雪深度	-0.118
露点温度	0.049	地面气压	-0.107
太阳散射辐射	0.741	天顶角	-0.837
太阳直接散射	0.874	高度10 m风向	0.040
太阳辐照度	0.999	高度10 m风速	0.082

1.2 光伏发电季节特性分析

光伏发电通过光电效应将光能转化为电能,太阳辐照度是光电效应的直接影响因素^[27]。太阳辐照度的大小受季节变化的影响,如冬季地表入射短波辐射较夏季低,因此晴空条件下冬季的光伏功率整体低于夏季。通过分析光伏功率的季节性变化和不同季节下光伏功率短时间内波动性,构建精细化光伏功率预测模型,图1为全年光伏功率与时间的三维曲线。其中,3月—5月为春季(1~92 d),6月—8月为夏季(93~184 d),9月—11月为秋季(185~275 d),12月—2月为冬季(276~365 d)。由图1可知,春季和夏季的光伏功率明显高于秋季和冬季,且秋季最低。光伏功率在春季末到夏季初最高,夏季随着温度的升高,光伏组件温度随之上升,导致光伏电池发热,进而影响光伏组件的发电效率。

为分析该地区不同季节下光伏功率的相对变化率,文中采用式(2)计算光伏功率的相对变化率,并绘制光伏功率相对变化率的频率分布图,如图2所示。由于日出和日落时刻光伏输出功率波动剧烈,同时其输出功率值较小甚至接近于0,不利于相对变化率的计算,故不考虑日出与日落1 h内光伏功率相对变化率。

$$\alpha = \frac{|P(t) - P(t-1)|}{P(t-1)} \quad (2)$$

式中: α 为光伏功率的相对变化率; $P(t)$ 为 t 时刻的光伏功率; $P(t-1)$ 为 $t-1$ 时刻的光伏功率。

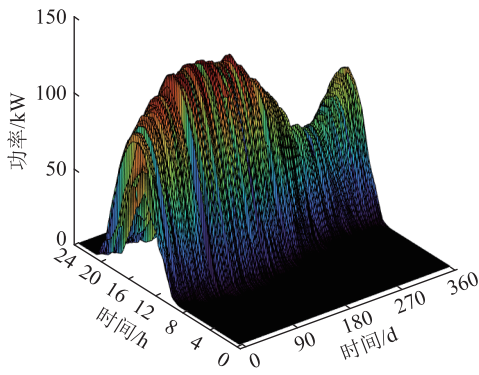


图1 光伏功率与时间三维曲线

Fig.1 Three-dimensional curves of photovoltaic power and time

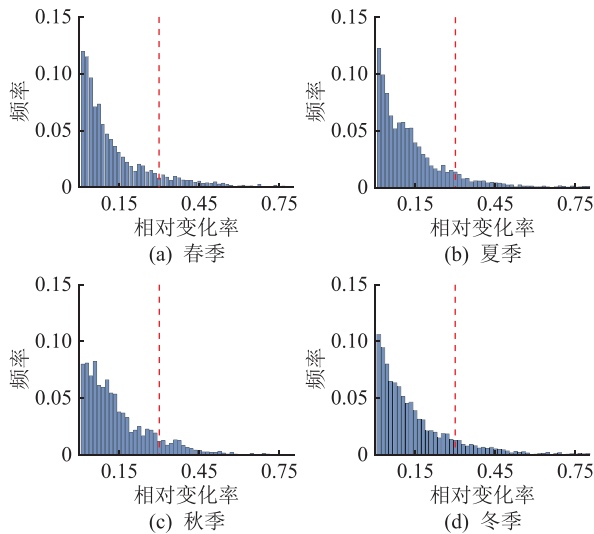


图2 不同季节下光伏功率相对变化率的频率分布

Fig.2 Frequency distribution of relative change rate of photovoltaic power under different seasons

由图2可知,该地区光伏功率的相对变化率主要集中在0.3以内。其中,冬季相对变化率最为集中,秋季相对变化率较为分散,说明秋季短时间内光伏功率波动较为剧烈。综上所述,光伏功率的输出和光伏功率的波动性均具有显著的季节性差异,有必要分不同季节构建光伏功率预测模型。

2 光伏功率预测组合模型

不同的单一光伏功率预测模型因自身的训练原理不同,导致在相同场景体现出不同的光伏功率预测性能。根据预测场景的特点选择合适的预测模型,充分利用各种模型的优点,提高光伏功率预测精准度。

2.1 组合模型构建

文中基于机器学习算法的训练原理选择了5种常用光伏功率预测模型,分别为KELM、LSTM、GRU、XGboost和支持向量机(support vector mach-

ine,SVM),各模型输入参数设置如表2所示。

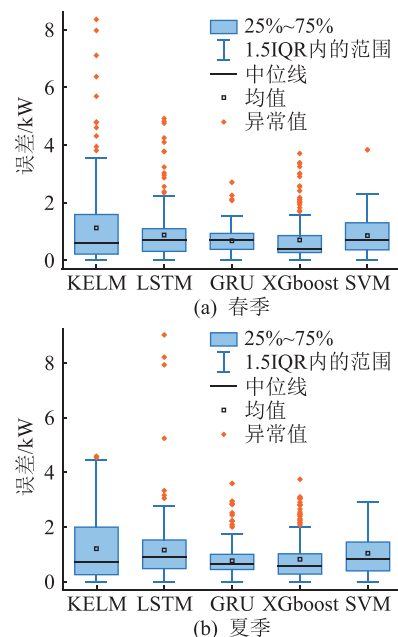
表2 各模型参数设置

Table 2 Parameter settings for each model

模型	参数	数值
KELM	核函数	RBF
	核参数	1
	正则化系数	10
LSTM	神经元个数	150
	学习率	0.01
GRU	最大训练轮次	200
	神经元个数	100
XGboost	学习率	0.01
	最大训练轮次	200
SVM	学习率	0.1
	节点的叶子最小分裂数	10
	最大深度	1
	训练数据占比	0.6
SVM	核函数	Gaussian
	核参数	1
	惩罚系数	10

在不同季节下,采用箱线图分析每个模型预测光伏功率的误差,如图3所示。其中,25%~75%表示每个模型预测误差的中间50%的数据范围;1.5IQR内的范围表示每个模型预测误差0~75%的数据范围,即正常数据范围;中位线表示每个模型预测误差的中位数;均值表示每个模型预测误差的均值;异常值表示每个模型预测误差76%~100%的数据范围。

根据模型预测误差,采用皮尔逊相关系数法分析各模型间预测误差的相关性,如图4所示。从图3、图4可以看出,GRU具有较好的泛化性,整体预



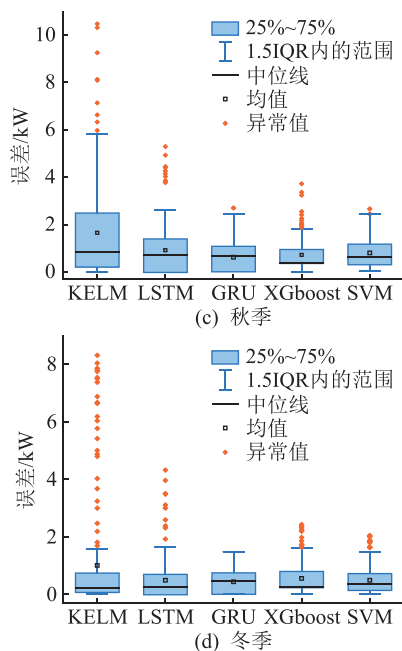


图3 不同季节下单模型的光伏功率预测误差分析
Fig.3 Prediction error analysis of single model photovoltaic power under different seasons

测误差最小; XGboost 在冬季的预测误差仅小于 KELM, KELM 预测效果最差, 且有大量异常值; SVM 异常值较少, 但其预测误差较高。在春季和冬季中 GRU 和 XGboost 的相关性最低, 二者在夏季和秋季中相关性也不高。XGboost 不仅与其他模型相关性低, 而且训练时间短。LSTM 与 GRU 相关性很高, 原因是二者属同一类型神经网络, 2 个模型的预测优缺点相似度高, 多样性低。但 GRU 需设置的超参数更少, 意味着训练时间更短。综上所述, 文中选择 GRU 和 XGboost 构建组合模型。

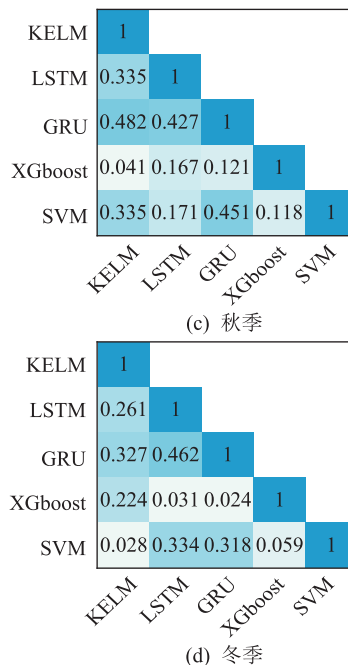
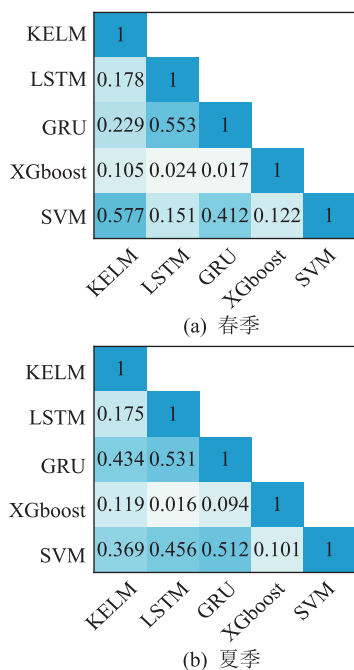


图4 不同季节下单模型间的皮尔逊相关系数
Fig.4 The Pearson correlation coefficient of single models under different seasons

组合模型的构建方式很多, 如 Stacking 集成学习、Bagging 集成学习和权重系数组合^[28-29]。文中选择权重系数组合的方法构建组合模型, 并利用误差倒数法^[30]确定模型的权值。误差倒数法依据每个模型的预测误差确定权重系数, 可以充分利用单模型的不同优点。相对于集成学习, 权重系数组合能够直观判断每个单模型的贡献度, 通过权值的变化实时调整模型, 极大地提高了预测模型构建的效率。

2.2 GRU-XGboost 组合模型预测流程

GRU-XGboost 组合模型流程如图 5 所示, 主要包含 5 个步骤。

(1) 步骤一: 原始数据集主要包含光伏功率和相关的气象数据, 对原始数据进行清洗、剔除坏数据。由于数据集内各类气象数据与光伏功率量纲不同、数值差异大, 故对数据进行归一化处理。

(2) 步骤二: 利用皮尔逊相关系数法, 选择 s 个与光伏功率高相关的气象因子 $M = \{m_1, m_2, \dots, m_s\}$, 构建模型输入特征、训练集、验证集和测试集。

(3) 步骤三: 设置 GRU 超参数和 XGboost 超参数, 训练模型并通过验证集调整超参数。

(4) 步骤四: 由误差倒数法确定组合模型权重系数。首先, 分别计算 GRU 和 XGboost 的验证集预测误差 e_1 和 e_2 , 如式 (3) 所示; 然后, 基于误差倒数法计算模型的权重系数, 如式 (4) 所示。GRU 的权重系数为 β , XGboost 的权重系数为 $1-\beta$ 。

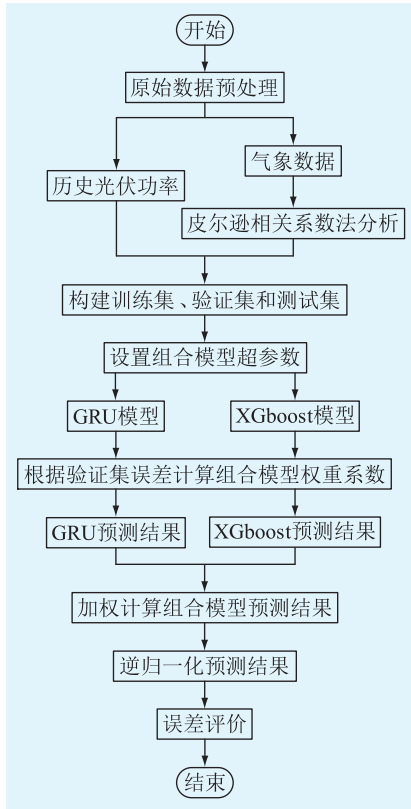


图5 组合模型预测流程

Fig.5 Flow chart of combined model prediction

$$\begin{cases} e_1 = \frac{1}{n} \sum_{t=1}^n |P_{GRU,t} - P_{V,t}| \\ e_2 = \frac{1}{n} \sum_{t=1}^n |P_{XGboost,t} - P_{V,t}| \end{cases} \quad (3)$$

$$\beta = \frac{e_2}{e_1 + e_2} \quad (4)$$

式中: n 为验证集的预测时刻总数; $P_{GRU,t}$ 、 $P_{XGboost,t}$ 分别为 t 时刻GRU和XGboost的光伏功率预测值; $P_{V,t}$ 为 t 时刻验证集的实际光伏功率。

(5) 步骤五:使用完成训练的GRU和XGboost算法分别进行光伏功率预测,将2个模型的预测结果加权计算得到组合模型预测结果并逆归一化。采用均方根误差 I_{RMSE} 、平均绝对误差 I_{MAE} 和平均绝对百分误差 I_{MAPE} 对光伏功率预测结果进行评估,如式(5)所示。

$$\begin{cases} I_{RMSE} = \sqrt{\frac{\sum_{t=1}^N (P_{Te,t} - P_{P,t})^2}{N}} \\ I_{MAE} = \frac{1}{N} \sum_{t=1}^N |P_{Te,t} - P_{P,t}| \\ I_{MAPE} = \frac{100\%}{N} \sum_{t=1}^N |(P_{Te,t} - P_{P,t})/P_{P,t}| \end{cases} \quad (5)$$

式中: N 为测试集的预测时刻总数; $P_{P,t}$ 、 $P_{Te,t}$ 分别为

t 时刻测试集的实际光伏功率和预测结果。

3 算例仿真与分析

文中以宁夏某光伏场站2019年的实测光伏输出功率与气象数据为样本进行仿真实验,预测未来24h光伏功率,其装机容量为130kW。在23:00—06:00,该地区光伏功率真实值接近0,故文中预测时间选定为07:00—22:00。数据采样时间间隔为15min,每日有64个数据窗口,春季和夏季都有5888个数据窗口,秋季有5824个数据窗口,冬季有5760个数据窗口。每个季节数据集划分如下:70%季节数据为训练集,15%季节数据为验证集,15%季节数据划分为晴天、多云和阴雨天3种天气的典型日,并从典型日中选择任意1d构建测试集,如表3所示。由于气候具有滞后性,如果某季节测试集中缺少某些类型天气,可从相邻后1个季节的前15%数据集中补充完整测试集。

表3 测试集

Table 3 Test set

季节	晴天日期	多云日期	阴雨天日期
春季	03.17	03.26	03.20
夏季	08.31	08.22	09.01
秋季	11.29	11.27	12.03
冬季	12.30	12.25	12.23

3.1 最优数据窗口

文中使用气象数据作为光伏功率预测模型的输入数据,不仅包含待预测时刻气象预报数据,还包含历史气象数据。历史气象数据对待预测时刻气象预报数据起补充作用,因此历史气象数据的输入窗口将会影响模型的预测性能。文中采用皮尔逊相关系数法计算某时刻气象数据与相邻前10个数据窗口之间的相关性,初步确定历史气象数据输入窗口的范围,如表4所示。由表4可以看出,随着数据窗口的增大,其与待测时刻气象数据相关性越来越小。

表4 待测时刻与相邻数据窗口的皮尔逊相关系数

Table 4 The Pearson correlation coefficient between the moment to be predicted and the neighboring data windows

数据窗口	S	数据窗口	S
1	0.978	6	0.747
2	0.940	7	0.705
3	0.896	8	0.633
4	0.841	9	0.548
5	0.788	10	0.474

选择 GRU-XGboost 组合模型对不同数据窗口进行预测,得到不同数据窗口对应的误差评价指标 I_{RMSE} 、 I_{MAE} 、 I_{MAPE} ,如图 6 所示。从图 6 可以看出,当历史气象数据的输入窗口为 1 时,整体预测精度最高,相较于数据窗口为 3 时, I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别下降了 9.96%、11.72% 和 10.21%。因此,在 GRU-XGboost 组合模型有效性验证实验中,光伏功率预测模型的数据窗口统一设置为 2,即输入待预测时刻气象数据与前一时刻气象数据。

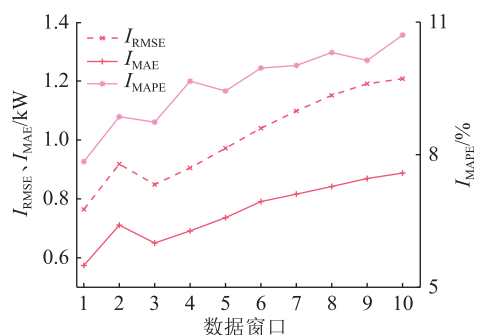


图 6 不同数据窗口的预测误差

Fig.6 Prediction errors for different data windows

3.2 GRU-XGboost 组合模型有效性验证

为验证 GRU-XGboost 组合模型的有效性,选取 GRU、XGboost、文献[31]所提出的 LSTM-XGboost 模型进行对比。采用表 3 的测试集进行测试,仿真结果为 30 次训练结果取均值。预测误差如表 5 所示。

各季节预测误差分析如下:

(1) 春季场景下,单模型预测精度大致相同,组

合算法预测效果更好。晴天和多云天气下,各模型预测误差接近;在阴雨天气下,GRU、XGboost 和 LSTM-XGboost 预测误差变大,GRU-XGboost 仍有较高的预测精度。相比其他 3 种模型,多云和阴雨天气下,GRU-XGboost 的 I_{MAPE} 明显更低,即对光伏功率趋势的拟合度更高。春季 3 种典型天气的光伏功率预测曲线如图 7 所示。

(2) 夏季日照时间和无云层遮挡时间长,所有模型的光伏功率趋势拟合度高。3 种典型天气下,GRU-XGboost 预测准确率明显最高,相较于 LSTM-XGboost,夏季中 3 种典型天气的平均 I_{RMSE} 和平均 I_{MAE} 分别降低了 16.85% 和 22.45%。夏季 3 种典型天气的光伏功率预测曲线如图 8 所示。

(3) 由于秋季的光伏功率在短时间内波动剧烈,单模型预测误差明显增高,原因是秋季早晚气温变化较大,天气变化频繁。GRU-XGboost 与 GRU 相比,平均 I_{RMSE} 、平均 I_{MAE} 和平均 I_{MAPE} 分别降低了 25.23%、25.54% 和 14.38%,预测精度明显提升。GRU-XGboost 的光伏功率预测精度仍高于 LSTM-XGboost。秋季 3 种典型天气的光伏功率预测曲线如图 9 所示。

(4) 冬季天气稳定,且晴天较多,各模型都能很好的预测光伏功率的趋势。单模型中 GRU 表现出优秀的预测性能,在晴天和多云天气下预测精度大多 LSTM-XGboost。在多云天气下,GRU-XGboost 的 I_{MAPE} 为 1.915%,GRU 的 I_{MAPE} 为 1.811%,略低于

表 5 不同季节下各模型的预测误差

Table 5 Prediction errors of each model under different seasons

季节	预测模型	晴天			多云			阴雨天		
		I_{RMSE}/kW	I_{MAE}/kW	$I_{MAPE}/%$	I_{RMSE}/kW	I_{MAE}/kW	$I_{MAPE}/%$	I_{RMSE}/kW	I_{MAE}/kW	$I_{MAPE}/%$
春季	GRU	0.581	0.441	2.152	0.620	0.507	2.442	0.662	0.503	3.624
	XGboost	0.531	0.429	2.187	0.679	0.526	2.722	0.705	0.535	3.614
	LSTM-XGboost	0.502	0.397	1.764	0.589	0.466	2.545	0.683	0.490	3.543
	GRU-XGboost	0.434	0.362	1.974	0.484	0.391	2.375	0.500	0.417	2.995
夏季	GRU	0.499	0.403	1.882	0.535	0.432	2.395	0.613	0.516	2.496
	XGboost	0.548	0.421	1.817	0.555	0.430	2.429	0.617	0.539	2.349
	LSTM-XGboost	0.518	0.429	1.879	0.489	0.402	2.484	0.548	0.501	2.279
	GRU-XGboost	0.413	0.332	1.739	0.424	0.334	2.105	0.456	0.367	2.153
秋季	GRU	0.619	0.511	2.192	0.810	0.658	2.839	0.933	0.726	2.901
	XGboost	0.641	0.529	1.949	0.744	0.594	2.683	1.163	0.793	3.508
	LSTM-XGboost	0.518	0.431	1.819	0.637	0.525	2.510	0.829	0.652	2.957
	GRU-XGboost	0.485	0.370	1.791	0.541	0.451	2.314	0.740	0.590	2.686
冬季	GRU	0.263	0.230	1.778	0.299	0.237	1.811	0.441	0.337	2.677
	XGboost	0.430	0.315	1.974	0.446	0.335	2.455	0.515	0.361	2.670
	LSTM-XGboost	0.338	0.270	1.471	0.361	0.291	1.926	0.424	0.332	2.769
	GRU-XGboost	0.234	0.216	1.467	0.281	0.233	1.915	0.371	0.279	2.259

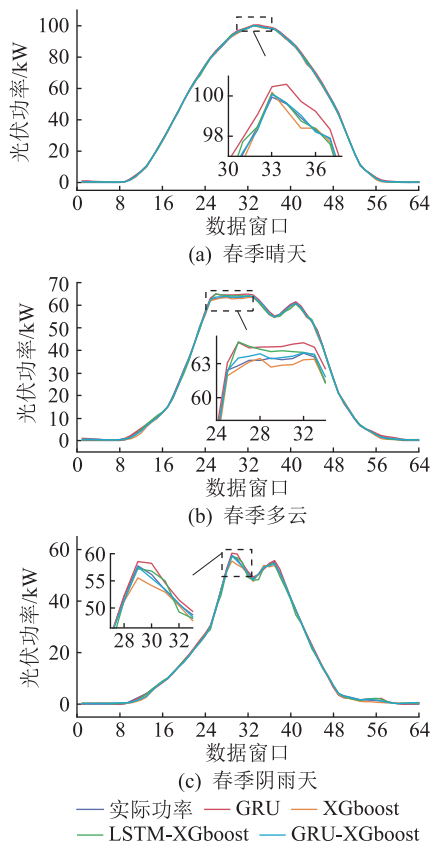


图7 春季3种典型天气类型的光伏功率预测曲线
Fig.7 Photovoltaic power predicted curves for three typical weather types in spring

GRU,因此,GRU 的光伏功率趋势拟合度略高于 GRU-XGboost。但是从整体来看,无论是预测精度还是光伏功率趋势拟合度,GRU-XGboost 表现都是最好的。冬季3种典型天气的光伏功率预测曲线如图10所示。

与单模型 GRU 相比,GRU-XGboost 的全年平均 I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别降低了 21.99%、21.07%和 11.70%。

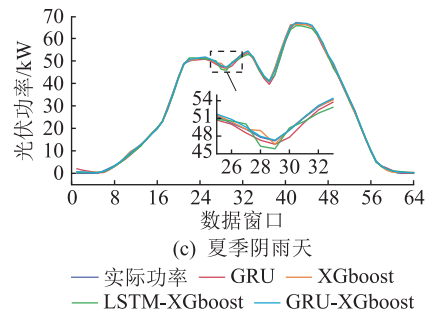
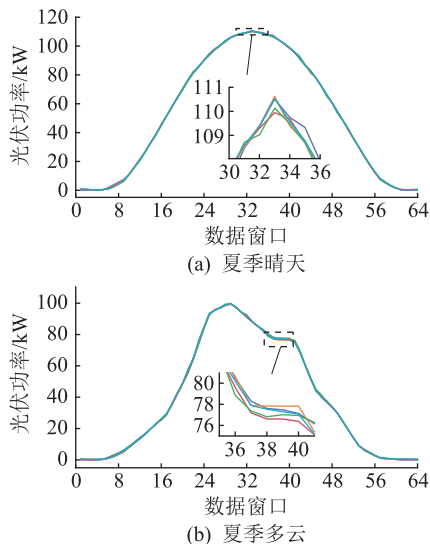


图8 夏季3种典型天气类型的光伏功率预测曲线
Fig.8 Photovoltaic power predicted curves for three typical weather types in summer

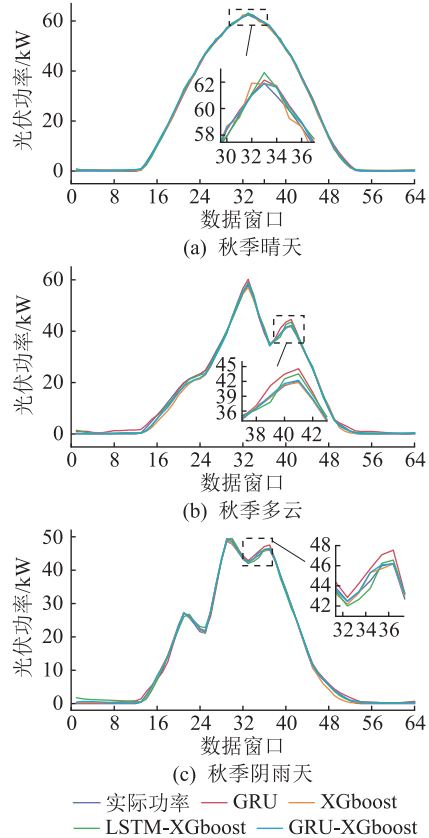


图9 秋季3种典型天气类型的光伏功率预测曲线
Fig.9 Photovoltaic power predicted curves for three typical weather types in autumn

3.3 考虑季节特性的精细化建模有效性验证

光伏功率的输出和波动均具有显著的季节性差异,考虑季节的精细化建模不仅可以提高模型泛化性,还可实现高精度的光伏功率预测。基于 GRU-XGboost 模型,分别构建季节预测模型和全年预测模型,并比较2种模型的预测精度。其中,季节预测模型根据不同季节的数据进行训练,全年模型以一整年的数据进行训练,测试集见表3,预测误差评价结果如表6所示。误差指标 I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 均为3种典型天气的平均值,模型训练时间为3种典型天气的平均训练时间。

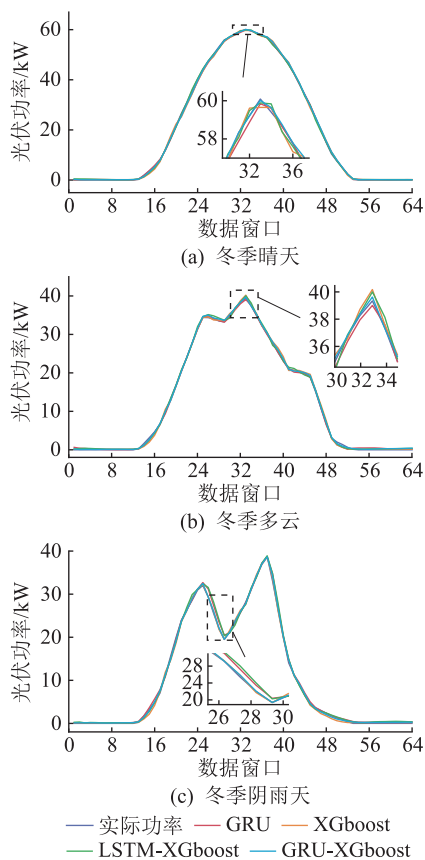


图 10 冬季 3 种典型天气类型的光伏功率预测曲线
Fig.10 Photovoltaic power predicted curves for three typical weather types in winter

表 6 季节模型和全年模型的预测误差评价
Table 6 Evaluation of prediction errors in seasonal and year-round models

季节	预测模型	I_{RMSE}/kW	I_{MAE}/kW	$I_{MAPE}/%$	训练时间/s
春季	季节模型	0.473	0.390	2.448	37
	全年模型	0.626	0.463	3.453	128
夏季	季节模型	0.456	0.371	2.059	36
	全年模型	0.581	0.437	3.193	126
秋季	季节模型	0.537	0.437	2.168	29
	全年模型	0.501	0.389	2.937	114
冬季	季节模型	0.375	0.294	1.991	23
	全年模型	0.551	0.431	2.309	91

可以看出,季节模型整体预测误差小于全年模型,全年模型仅在秋季的预测误差小于季节模型,其余季节均是季节模型预测误差更小。考虑季节的光伏功率短期预测建模相较全年气象数据建模,平均 I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别下降 18.50%、13.26% 和 27.13%,全年模型和季节模型的平均训练时间分别为 114.75 s 和 31.25 s,季节模型平均训练时间远小于全年模型。综上所述,考虑季节的精细化建模有更高的预测精度和更短的预测时长,这为大规模光伏消纳提供了高效、可靠的解决方案。

4 结论

气候的间歇性和波动性给短期光伏功率的预测带来了困难,以全年数据建模的传统方法效率低且难以得到满意的模型泛化能力。为了提高光伏短期输出功率预测的准确性,文中提出一种考虑季节特性的 GRU-XGboost 短期光伏功率预测组合模型,通过与实际光伏发电数据的仿真对比分析,得到以下结论:

(1) 考虑季节的光伏功率短期预测建模相较全年气象数据建模,平均 I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别下降 18.50%、13.26% 和 27.13%,模型的平均训练时间减少了 72.77%,有效提高了训练效率和预测精度。

(2) 选择皮尔逊相关系数法分析历史气象数据不同的输入窗口与待预测时刻气象数据之间的相关性,然后采用 GRU-XGboost 对不同数据窗口进行测试,数据窗口为 1 时预测误差最小,相较于数据窗口为 3 时, I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别下降了 9.96%、11.72% 和 10.21%。

(3) 引入 GRU-XGboost 组合模型,GRU 能充分挖掘时间序列的深层次特征规律,XGboost 并行处理能力能大大缩短模型运行时间,通过权重系数集成二者预测结果。与单模型 GRU 相比,GRU-XGboost 的全年平均 I_{RMSE} 、 I_{MAE} 和 I_{MAPE} 分别降低了 21.99%、21.07% 和 11.70%。

(4) 通过历史预测模型训练经验设置光伏功率预测模型超参数不能充分发挥预测模型的性能,未来工作将引入智能优化算法选择预测模型的超参数,提升模型预测精度。模型权重系数需要大量的数据来保证其具有足够的泛化性,未来工作将进一步优化组合模型的权重系数,使其能充分发挥各模型的预测性能。

参考文献:

[1] ZANG H X, CHENG L L, DING T, et al. Day-ahead photovoltaic power forecasting approach based on deep convolutional neural networks and meta learning[J]. International Journal of Electrical Power & Energy Systems, 2020, 118: 105790.

[2] 赵俊浩, 吴杰康, 张文杰, 等. 基于布谷鸟-Elman 算法的光伏发电预测[J]. 电力工程技术, 2020, 39(2): 81-88.

ZHAO Junhao, WU Jiekang, ZHANG Wenjie, et al. Photovoltaic power prediction based on Elman neural network with improved cuckoo algorithm[J]. Electric Power Engineering Technology, 2020, 39(2): 81-88.

[3] 董存, 王铮, 白捷予, 等. 光伏发电功率超短期预测方法综述[J]. 高电压技术, 2023, 49(7): 2938-2951.

DONG Cun, WANG Zheng, BAI Jieyu, et al. Summary of ultra-short-term forecasting methods for photovoltaic power generation

- [J]. High Voltage Engineering, 2023, 49(7): 2938-2951.
- [4] 孙新程, 万玥, 丁宏, 等. 基于小波去噪-KPCA 神经网络的光伏功率预测方法[J]. 电力工程技术, 2020, 39(2): 110-116.
SUN Xincheng, WAN Yue, DING Hong, et al. Forecasting method of photovoltaic output power based on wavelet denoising/KPCA/PSOBP[J]. Electric Power Engineering Technology, 2020, 39(2): 110-116.
- [5] 郑国权, 祝恩国, 张海龙, 等. 基于主从博弈的高比例光伏配电台区柔性互联规划[J]. 电力建设, 2024, 45(4): 100-110.
ZHENG Guoquan, ZHU Enguo, ZHANG Hailong, et al. Flexible interconnection planning for distribution station areas of high-ratio photovoltaic based on master-slave game[J]. Electric Power Construction, 2024, 45(4): 100-110.
- [6] 黄泽, 毕贵红, 谢旭, 等. 基于 MBI-PBI-ResNet 的超短期光伏功率预测[J]. 电力系统保护与控制, 2024, 52(2): 165-176.
HUANG Ze, BI Guihong, XIE Xu, et al. Ultra-short-term PV power prediction based on MBI-PBI-ResNet[J]. Power System Protection and Control, 2024, 52(2): 165-176.
- [7] 宋文乐, 张焯, 刘航旭, 等. 基于光伏组件老化特性解耦的渔光互补光伏短期出力预测方法[J]. 电力建设, 2024, 45(7): 25-33.
SONG Wenle, ZHANG Ye, LIU Hangxu, et al. Short-term output prediction method for complementary fishing and solar power decoupling the aging characteristics of photovoltaic modules[J]. Electric Power Construction, 2024, 45(7): 25-33.
- [8] 陈文进, 陈菁伟, 钱建国, 等. 气象特征频繁变化区域的光伏功率预测方法[J]. 浙江电力, 2023, 42(3): 37-46.
CHEN Wenjin, CHEN Jingwei, QIAN Jianguo, et al. A photovoltaic power prediction method for regions with frequent changes of meteorological characteristics[J]. Zhejiang Electric Power, 2023, 42(3): 37-46.
- [9] 王亚军, 王继东, 曹曼, 等. 光伏电站理论发电功率的优化计算模型与方法[J]. 电力系统保护与控制, 2024, 52(14): 144-153.
WANG Yajun, WANG Jidong, CAO Man, et al. Optimization calculation models and methods for the theoretical power of a photovoltaic power station[J]. Power System Protection and Control, 2024, 52(14): 144-153.
- [10] MAYER M J. Benefits of physical and machine learning hybridization for photovoltaic power forecasting[J]. Renewable and Sustainable Energy Reviews, 2022, 168: 112772.
- [11] DIMD B D, VÖLLER S, CALI U, et al. A review of machine learning-based photovoltaic output power forecasting: Nordic context[J]. IEEE Access, 2022, 10: 26404-26425.
- [12] 朱琼锋, 李家腾, 乔骥, 等. 人工智能技术在新能源功率预测的应用及展望[J]. 中国电机工程学报, 2023, 43(8): 3027-3047.
ZHU Qiongfeng, LI Jiateng, QIAO Ji, et al. Application and prospect of artificial intelligence technology in new energy power prediction[J]. Proceedings of the CSEE, 2023, 43(8): 3027-3047.
- [13] 朱玥, 顾洁, 孟璐. 基于 EMD-LSTM 的光伏发电预测模型[J]. 电力工程技术, 2020, 39(2): 51-58.
ZHU Yue, GU Jie, MENG Lu. Photovoltaic power generation prediction model based on EMD-LSTM[J]. Electric Power Engineering Technology, 2020, 39(2): 51-58.
- [14] 唐雅洁, 林达, 倪筹帷, 等. 基于 XGBoost 的双层协同实时校正超短期光伏预测[J]. 电力系统自动化, 2021, 45(7): 18-27.
TANG Yajie, LIN Da, NI Chouwei, et al. Ultra-short-term photovoltaic prediction based on XGBoost and double-layer collaborative real-time correction[J]. Automation of Electric Power Systems, 2021, 45(7): 18-27.
- [15] GRZEBYK D, ALCANIZ A, DONKER J C B, et al. Individual yield nowcasting for residential PV systems[J]. Solar Energy, 2023, 251: 325-336.
- [16] ZHANG N, REN Q, LIU G C, et al. Short-term PV output power forecasting based on CEEMDAN-AE-GRU[J]. Journal of Electrical Engineering & Technology, 2022, 17(2): 1183-1194.
- [17] 卞海红, 孙健硕. 基于典型气象周的 GRNN 光伏发电量预测模型[J]. 电力工程技术, 2021, 40(5): 94-99.
BIAN Haihong, SUN Jianshuo. Photovoltaic power generation prediction model based on optimized TMY method-GRNN[J]. Electric Power Engineering Technology, 2021, 40(5): 94-99.
- [18] 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): 4372-4384.
WANG Kaiyan, DU Haodong, JIA Rong, et al. Short-term interval probability prediction of photovoltaic power based on similar day clustering and QR-CNN-BiLSTM model[J]. High Voltage Engineering, 2022, 48(11): 4372-4384.
- [19] 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 Liqun, LI Hongbo, HOU Yadong, et al. Short-term photovoltaic power generation prediction based on VMD-ISSA-KELM[J]. Power System Protection and Control, 2022, 50(21): 138-148.
- [20] 赵耀, 高少炜, 李东东, 等. 基于天气相似聚类与 QRNN 的短期光伏功率区间概率预测[J]. 电力系统自动化, 2023, 47(23): 152-161.
ZHAO Yao, GAO Shaowei, LI Dongdong, et al. Short-term interval probability prediction of photovoltaic power based on weather similarity clustering and quantile regression neural network[J]. Automation of Electric Power Systems, 2023, 47(23): 152-161.
- [21] ZHENG L W, SU R, SUN X Y, et al. Historical PV-output characteristic extraction based weather-type classification strategy and its forecasting method for the day-ahead prediction of PV output[J]. Energy, 2023, 271: 127009.
- [22] 雷柯松, 吐松江·卡日, 伊力哈木·亚尔买买提, 等. 基于 WGAN-GP 和 CNN-LSTM-Attention 的短期光伏功率预测[J]. 电力系统保护与控制, 2023, 51(9): 108-118.

- LEI Kesong, TUSONGJIANG·Kari, YILIHAMU·Yaermainaiti, et al. Prediction of short-term photovoltaic power based on WGAN-GP and CNN-LSTM-Attention[J]. Power System Protection and Control, 2023, 51(9): 108-118.
- [23] 张云钦,程起泽,蒋文杰,等. 基于 EMD-PCA-LSTM 的光伏功率预测模型[J]. 太阳能学报, 2021, 42(9): 62-69.
ZHANG Yunqin, CHEN Qize, JIANG Wenjie, et al. Photovoltaic power prediction model based on EMD-PCA-LSTM[J]. Acta Energetica Sinica, 2021, 42(9): 62-69.
- [24] WANG Z, WANG Y S, CAO S L, et al. A robust spatial-temporal prediction model for photovoltaic power generation based on deep learning [J]. Computers and Electrical Engineering, 2023, 110: 108784.
- [25] 杨帆,申亚,李东东,等. 基于 GA-GNNM 的极地光伏发电功率预测方法[J]. 太阳能学报, 2022, 43(4): 167-174.
YANG Fan, SHEN Ya, LI Dongdong, et al. Power prediction method of polar photovoltaic generation based on GA-GNNM [J]. Acta Energetica Sinica, 2022, 43(4): 167-174.
- [26] BENESTY J, CHEN J D, HUANG Y T, et al. Pearson correlation coefficient [M]//Springer Topics in Signal Processing. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 1-4.
- [27] 王伟胜,车建峰,王勃,等. 光伏发电功率预测技术及应用[M]. 北京:中国电力出版社, 2019: 9-10.
WANG Weisheng, CHE Jianfeng, WANG Bo, et al. Photovoltaic power prediction technology and its application[M]. Beijing: China Electric Power Press, 2019: 9-10.
- [28] 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.
- TAN Xiaoyu, LIU Fang, MA Junjie, et al. Ultra-short-term forecasting model of photovoltaic power based on combination of DBN and T-S time-varying weights[J]. Acta Energetica Sinica, 2021, 42(10): 42-48.
- [29] 于涛,丁海旭,黄卫民,等. 面向复杂异质数据的集成学习研究综述[J]. 控制工程, 2023, 30(8): 1425-1435.
YU Tao, DING Haixu, HUANG Weimin, et al. Summary of integrated learning for complex heterogeneous data [J]. Control Engineering of China, 2023, 30(8): 1425-1435.
- [30] GAO H B, QIU S, FANG J, et al. Short-term prediction of PV power based on combined modal decomposition and NARX-LSTM-LightGBM[J]. Sustainability, 2023, 15(10): 8266.
- [31] 谭海旺,杨启亮,邢建春,等. 基于 XGBoost-LSTM 组合模型的光伏发电功率预测[J]. 太阳能学报, 2022, 43(8): 75-81.
TAN Haiwang, YANG Qiliang, XING Jiangchun, et al. Photovoltaic power prediction based on combined XGBoost-LSTM model[J]. Acta Energetica Sinica, 2022, 43(8): 75-81.

作者简介:



张静

张静(1996),男,硕士在读,研究方向为光伏预测、配电网优化运行(E-mail: 104909133-5@qq.com);

熊国江(1986),男,博士,教授,通信作者,研究方向为电力系统优化运行、电网故障诊断。

Short-term photovoltaic power prediction combination model considering seasonal characteristic and data window

ZHANG Jing¹, XIONG Guojiang^{1,2}

(1. College of Electrical Engineering, Guizhou University, Guiyang 550025, China;

2. Guizhou University Institute of Engineering Investigation & Design Co., Ltd., Guiyang 550025, China)

Abstract: The intermittency and randomness of photovoltaic power present different characteristics due to seasonal variations, so it is important to consider seasonal characteristics to improve the accuracy of photovoltaic power prediction. Therefore, a short-term photovoltaic power prediction combination model considering seasonal characteristic and data window is proposed in the paper. Firstly, the Pearson correlation coefficient method is adopted to determine suitable meteorological factors with high contribution to photovoltaic power and reduce the input feature dimensions of the prediction model. Secondly, the prediction error of different photovoltaic power models is compared, and the two models with the lowest photovoltaic power prediction error and the lowest correlation are selected to construct the combination model, i.e., gated recurrent unit (GRU) model and extreme gradient boosting (XGboost) model. Thirdly, the effects of different input windows in the historical meteorological data on the prediction accuracy of GRU-XGboost model are analyzed to determine the optimal data window. Finally, on this basis, GRU and XGboost predict the photovoltaic power respectively. The final prediction is obtained by weighted combination of the two predictions. Simulation results show that the proposed model has stronger adaptability and higher prediction accuracy than other models.

Keywords: short-term photovoltaic power prediction; seasonal characteristics; data window; gated recurrent unit (GRU); extreme gradient boosting (XGboost); combination model

(编辑 吴昊)