

基于气象特征和改进 Transformer 的光伏功率短期预测

张建辉¹, 滕 婕¹, 李秀慧¹, 谭庄熙²

(1. 国网甘肃省电力公司经济技术研究院, 甘肃 兰州 730030; 2. 湖南科技大学信息与电气工程学院, 湖南 湘潭 411201)

摘要: 光伏出力易受气象因素影响, 从而呈现间歇性和随机性。而准确可靠地预测光伏出力不仅可以缓解高比例光伏并网对电网的冲击, 还可以为电网的调度决策人员提供数据参考。本文提出一种基于气象特征和改进 Transformer 的光伏功率短期预测方法。首先针对光伏相关的气象因素提取增量特征、统计特征和时变特征; 然后将提取的特征和光伏出力数据输入 BOA-iTransformer 模型, 再将每个变量独立嵌入, 便于模型捕捉关键气象特征和多元数据的关联性; 随后采用贝叶斯优化调参进行特征选择, 得到最优特征组合, 以此建立 BOA-iTransformer 光伏预测模型; 最后采用中国某地区实际光伏电站数据进行对比实验。实验结果表明, 本文模型比 iTransformer、Transformer 和 LSTM 模型预测精度分别提高了 3.54%、7.24% 和 14.2%。

关键词: 光伏预测; 深度学习; 贝叶斯优化; 特征构造

中图分类号: TM615 文献标识码: A DOI: 10.16146/j.cnki.rndlgc.2024.08.017

[引用本文格式] 张建辉, 滕 婕, 李秀慧, 等. 基于气象特征和改进 Transformer 的光伏功率短期预测[J]. 热能动力工程, 2024, 39(8): 146-154. ZHANG Jianhui, TENG Jie, LI Xiuhui, et al. Short-term prediction of photovoltaic power based on meteorological features and improved transformer[J]. Journal of Engineering for Thermal Energy and Power, 2024, 39(8): 146-154.

Short-term Prediction of Photovoltaic Power based on Meteorological Features and Improved Transformer

ZHANG Jianhui¹, TENG Jie¹, LI Xiuhui¹, TAN Zhuangxi²

(1. Economic and Technological Research Institute of State Grid Gansu Electric Power Co., Ltd., Lanzhou, China, Post Code: 730030; 2. School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan, China, Post Code: 411201)

Abstract: Photovoltaic (PV) output is susceptible to meteorological factors, thus showing intermittency and randomness. Accurate and reliable prediction of PV power can not only alleviates the impact of high percentage of PV grid-connectedness on the power grid, but also provides data reference for grid scheduling decision makers. In this paper, we proposed a short-term prediction method of PV power based on meteorological features and improved Transformer. Firstly, incremental features, statistical features and time-varying features were extracted for PV-related meteorological factors; then, the extracted features and PV output data were input into the BOA-iTransformer model, and each variable was embedded independently, which was convenient for the model to capture the key meteorological features and the correlation of multivariate data; subsequently, Bayesian optimal tuning was used for feature selection to obtain the optimal feature combinations, which was used to build the BOA-iTransformer PV prediction model; finally, the actual data of photovoltaic power stations in a region of China were used for comparative experi-

收稿日期: 2023-12-17; 修订日期: 2024-02-22

基金项目: 国网总部科技项目(52272810005); 湖南省自然科学基金(2022JJ40150)

Fund-supported Project: State Grid Headquarters Science and Technology Project (52272810005); Natural Science Foundation of Hunan Province (2022JJ40150)

作者简介: 张建辉(1979-), 男, 国网甘肃省电力公司经济技术研究院高级工程师。

ments. The experimental results show that the prediction accuracy of this model can be improved by 3.54%, 7.24% and 14.2% compared with iTransformer, Transformer and LSTM models, respectively.

Key words: photovoltaic prediction, deep learning, Bayesian optimization, feature mining

引言

在双碳目标要求下,国际能源署 IEA 预计 2040 年新能源比例将达到 60%,其中光伏和风能占比超过 50%^[1]。中国每年新能源装机容量高达 75GW,其中 2022 年分布式光伏装机容量达到 51.1GW^[2]。由于光伏出力具有较强的波动和间歇特性,导致大比例光伏并网时给电网的安全运行带来巨大冲击。精准可靠的光伏短期出力预测不仅可以促进新能源消纳,还为电网安全稳定运行提供安全保障。

现有光伏出力预测可分为物理方法预测、直接预测和间接预测 3 类^[3-6]。其中物理方法预测基于卫星云图信息对分布式发电系统的出力进行预估,成本较高^[3]。直接预测法通过光伏出力曲线的周期性、趋势性等外在表征建立光伏出力历史数据与未来光伏出力数据之间的映射关系^[4]。但此类方法未考虑太阳辐照度、湿度、温度、压力、风速等与光伏出力相关的气象信息。间接预测通过对影响光伏出力大小的气象数据进行预测,再以气象数据与光伏出力的映射关系来预测对应时刻的光伏出力。此类常规方法如:人工神经网络^[5]、支持向量机^[6]、马尔科夫链^[7]。由于同类天气模式存在明显的相似性,因此有学者进行划分相似日天气类型的光伏出力预测研究。此方法相较于前两种方法虽然考虑了影响光伏出力的气象信息,但仅从气象数据的原始特征建模,未进一步对气象特征进行挖掘^[8]。

传统方法仅从气象数据和光伏出力数据的宏观规律进行建模,未能进一步对气象特征进行深度挖掘。随着近年来深度学习算法发展迅速,部分学者将深度学习算法应用于光伏出力预测^[9]。这些预测方法有长短时记忆 LSTM (Long Short Term Memory)^[10]、基于 LSTM 的自编码器^[11]、深度信念网络^[12]、基于小波分解 DWT 的 CNN-LSTM^[13] 等,在一定程度上提高了光伏出力预测的准确率。深度学习模型可较好拟合光伏出力曲线,而深度学习算法中 Transformer 模型^[14] 通过引入自注意力机制具有并行速度快、模型可解释性高等优点,并能有效捕捉长

时间序列依赖关系。需要指出的是,特征是模型预测的核心,而光伏发电功率与辐照度、气温、湿度、风速等气象因素有较强相关关系^[15]。将气象特征独立嵌入 Transformer 模型,利用自注意力机制处理气象特征与光伏出力之间的相关性,可对影响光伏出力的关键气象特征进行深层挖掘,进一步提高模型预测精度。

本文提出一种基于气象特征挖掘和改进 Transformer 的光伏出力短期预测方法。首先针对光伏相关的气象因素提取增量特征、统计特征和时变特征;然后建立 BOA-iTransformer 模型对经特征构造后更细致、多元的特征进行挖掘,从而实现光伏出力的预测。该模型输入嵌入部分将每个变量独立嵌入为一个标记,用于捕捉关键气象特征和多元数据关联性。贝叶斯优化调参 (Bayes Optimization, BOA) 用于找出最优时间窗和减少特征冗余度。

1 基于气象特征挖掘和改进 Transformer 模型理论

1.1 特征构造

光伏出力受众多环境因素的影响,其中天气条件是影响最大的因素之一。光辐照强度、温度、云层覆盖、风速和湿度等都会影响光伏出力的大小和效率。因此对天气特征进行特征构造,挖掘与光伏出力对应的天气条件下的细粒度特征,可提高光伏出力预测模型的精度。对于气象特征 $X_0 = [X_1, X_2, \dots, X_N]^T$, X_i 表示第 i ($i = 1, 2, \dots, N$) 时刻的特征, $X_i = [x_{i1}, x_{i2}, \dots, x_{iM}]$, x_{ik} ($k = 1, 2, \dots, M$) 表示 X_i 的第 k 个特征, M 表示总的特征数量。特征构造如下:

(1) 原始特征,即 $X_0 = [X_1, X_2, \dots, X_N]^T$ 。

(2) 增量特征,用于表示天气数据在偏移时间 t 内的变化趋势,计算公式如下:

$$\begin{cases} X_{i,z} = [X_i - X_{(i-t)}, \dots, X_i - X_{(i-t+n)}, \dots, X_i - X_{(i+n)}] \\ X_i - X_{(i-t+n)} = [x_{i1} - x_{(i-t+n)1}, \dots, x_{ik} - x_{(i-t+n)k}, \dots, \\ x_{iM} - x_{(i-t+n)M}] \end{cases}, n = 0, 1, \dots, 2t \quad (1)$$

式中: $X_{i,z}$ — 增量特征。

(3) 统计特征,用于衡量时间窗参数 t 内的气象

数据的波动程度,表达式如下:

$$\begin{cases} \bar{x}_{ikt} = \frac{1}{2t+1} \sum_{v=-t}^t x_{(i+v)k} \\ S_{ikt} = \left[\frac{1}{2t+1} \sum_{v=-t}^t (x_{(i+v)k} - \bar{x}_{ikt})^2 \right]^{\frac{1}{2}} \\ x_{\max,kt} = \max \{ x_{vk}, x_{v+1}, \dots, x_{v+tk} \} \end{cases} \quad (2)$$

式中: $x_{\max,kt}$ 、 S_{ikt} 、 \bar{x}_{ikt} —气象数据在 $[i-t, i+t]$ 内的最大值、标准差和均值。

(4) 时变特征,用于挖掘光伏出力数据在不同时间尺度(如日、月和季节)上的短期规律性,从而更精准地提取光伏出力的时变特性。假设光伏出力数据为 $Y = [y_1, y_2, \dots, y_n]^T$,则时变特征可表示为:

$$Y_{i,s} = [y_{i,\text{minute}}^{l_1}, \dots, y_{i,\text{minute}}^{l_K}, y_{i,\text{hour}}^{l_1}, \dots, y_{i,\text{hour}}^{l_K}] \quad (3)$$

式中: $Y_{i,s}$ —时变特征; $l = [l_1, \dots, l_K]$ —时间窗范围; K —时间窗种类; $y_{i,\text{minute}}^{l_K}$ 和 $y_{i,\text{hour}}^{l_K}$ —不同时窗 K 内同分钟/小时的平均光伏出力。

综上所述,通过构造增量特征、统计特征和时变特征共3类特征而成的细粒度特征,可实现对影响光伏出力大小的天气条件及时变特性更为精准、可靠的描述。

1.2 iTransformer 模型

Transformer 引入自注意力机制实现了并行计算,提高了计算速度,并能够有效地捕捉长时间序列依赖关系,且在自然语言处理、图像领域中取得了不错的效果^[16]。然而受多变量合并嵌入的影响导致模型难以保留多元数据之间的关联性,因此限制了模型捕捉关键气象特征。iTransformer 在传统 Transformer 的基础上,通过将每个变量的时间序列独立嵌入为一个标记,利用自注意力机制处理多元相关性。其整体框架结构可分为输入嵌入、编码器、投影层共3个模块,如图1所示^[17]。

(1) 输入嵌入模块,将天气特征、光伏出力时间序列输入 iTransformer 模型。与 Transformer 结构不同的是 iTransformer 将单个变量时间序列独立地嵌入为一个标记。由于输入天气特征和光伏出力序列的顺序隐含地存储在前馈神经网络的神经元排列中,因此不需要 Transformer 结构中的位置嵌入。其示意图如图2所示,计算公式为:

$$h_n^0 = \text{Embedding}(X) \quad (4)$$

式中: X —天气特征、光伏出力时间序列等输入矩阵; h_n^0 —天气特征、光伏出力时间序列的嵌入标记;Embedding—嵌入操作。

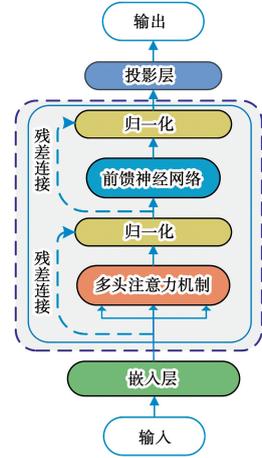


图1 iTransformer 的整体结构

Fig.1 Overall structure of iTransformer

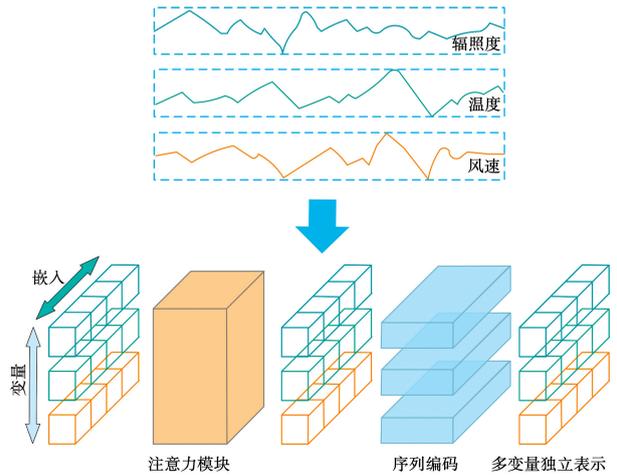


图2 输入嵌入示意图

Fig.2 Schematic diagram of input embedding

(2) 编码器模块,采用多头注意力机制对独立嵌入的多变量计算注意力分布值,对重要特征进行高权重分配,从而实现对影响光伏出力影响较大的天气数据非线性特征的深度挖掘,计算公式如下:

$$Q^c = W^Q X \quad (5)$$

$$K = W^K X \quad (6)$$

$$V^c = W^V X \quad (7)$$

$$T_A(Q^c, K, V^c) = \text{softmax}(Q^c K^T / \sqrt{d_k}) V^c \quad (8)$$

$$T_M(\mathbf{Q}^c, \mathbf{K}, \mathbf{V}^c) = \text{concat}(T_{\text{head}_1}, \dots, T_{\text{head}_i}) \mathbf{W}^o \quad (9)$$

$$T_{\text{head}_i} = T_A(\mathbf{Q}_i^c, \mathbf{K}_i, \mathbf{V}_i^c) \quad (10)$$

式中: \mathbf{Q}^c —查询矩阵; $\mathbf{W}^o \in \mathbf{R}^{d_{\text{model}} \times d_k}$; \mathbf{K} —键矩阵; $\mathbf{W}^k \in \mathbf{R}^{d_{\text{model}} \times d_k}$; \mathbf{V}^c —值矩阵; $\mathbf{W}^v \in \mathbf{R}^{d_{\text{model}} \times d_v}$; $\mathbf{W}^o \in \mathbf{R}^{d_v \times d_{\text{model}}}$ —参数向量; d_k 和 d_v —键和值的维度; d_{model} —输入的维度; softmax—激活函数; concat—连接函数; T_A —注意力分布值; T_M —多头注意力分布值; T_{head_i} —第 i 个注意力分布值。

利用前馈神经网络来降低长时间序列对模型过拟合的影响;归一化层则将数据规范化,可减少不同变量之间的差异对模型的影响。计算公式如下:

$$T_{\text{FFN}}(\mathbf{x}) = \max(0, \mathbf{x} \mathbf{W}_1 + \mathbf{b}_1) \mathbf{W}_2 + \mathbf{b}_2 \quad (11)$$

$$\mathbf{X}' = T_L(\mathbf{X} + T_{M(\mathbf{Q}, \mathbf{K}, \mathbf{V})}) \quad (12)$$

$$\mathbf{X}_{\text{out}} = T_L(\mathbf{X}' + T_{\text{FFN}}(\mathbf{X})) \quad (13)$$

式中: \mathbf{X}' —归一化后的天气特征输入; \mathbf{x} — \mathbf{X} 中的元素; max—激活函数; T_L —归一化处理; T_{FFN} —前馈神经网络的线性处理; \mathbf{X}_{out} —输出矩阵; \mathbf{W}_1 和 \mathbf{W}_2 —模型的权重; \mathbf{b}_1 和 \mathbf{b}_2 —模型的偏置。

(3) 投影层模块,由多层感知器(Multi-Layer Perceptron, MLP)组成。该模块将通过编码器模块独立处理后的多变量标记进行非线性映射输出。

综上所述, iTransformer 模型将每个时间步嵌入时间标记,这种将单一变量序列独立嵌入变量标记,有利于注意力模块捕捉多变量间的相关性,再通过前馈神经网络来进行序列编码,旨在提高模型对多元时间序列的表示学习和关联性建模,从而提高光伏出力预测模型的精度。

2 基于气象特征挖掘和改进 Transformer 的光伏出力预测模型

2.1 贝叶斯优化调参

为了提高光伏预测模型的精确度,本文采用了机器学习中常用的贝叶斯优化方法^[18]来进行参数优化。为了减少模型的计算量并找到最佳的特征组合,将 1.1 节提到的可能的特征组合作为参数输入到贝叶斯优化中进行选择。

本文对时间窗参数 t 进行贝叶斯优化调参,从而找到最优特征组合,减少特征冗余度。构建的气象特征参数组合为 $[p_1, p_2, \dots, p_q]$, 共有 q 类特征组

合,贝叶斯优化调参步骤如下:

步骤(1) 利用参数组合来对损失函数 $L(p)$ 进行高斯过程预测:

挑选若干参数组合 $[p_1, p_2, \dots, p_r]$, 计算相应的损失函数值,并将参数与损失函数值配对 D 。

$$D = \{(p_s, L(p_s)), s = 1, \dots, r\} \quad (14)$$

假设损失函数 $L(p_s)$ 遵循高斯分布,则:

$$\begin{cases} L(p_s) \sim N(0, U) \\ U = \begin{bmatrix} k(p_1, p_1) & \dots & k(p_1, p_r) \\ \vdots & \ddots & \vdots \\ k(p_r, p_1) & \dots & k(p_r, p_r) \end{bmatrix} \end{cases} \quad (15)$$

式中: $N(0, U)$ —高斯分布;核函数 $k(\cdot)$ — p 的协方差; U —协方差。

步骤(2) 采用预期改进 EI (Expected Improvement) 函数来进行采样选择。采用 TPE (Tree-structured Parzen Estimator) 算法^[19]来构建 EI 函数。首先选取当前的 p^+ 作为最优参数组合^[11],然后将大于 $L(p^+)$ 的 y^* 值作为阈值,通过式(16)得到对应 EI 函数:

$$\begin{cases} \text{EI}_{y^*}(p) = \int_{-\infty}^{y^*} (y^* - y) p(y | p) dy \\ = \frac{\gamma y^* l(p) - l(p) \int_{-\infty}^{y^*} p(y) dy}{\gamma l(p) + (1 - \gamma) g(p)} \\ p(p | y) = \begin{cases} l(p), & \text{if } y \leq y^* \\ g(p), & \text{if } y > y^* \end{cases} \end{cases} \quad (16)$$

式中: $l(p)$ —小于 y^* 的分布函数; $g(p)$ —大于 y^* 的分布函数; γ —分布比重。

选择使得 EI 函数最大的点作为采样点。

$$p_{\text{new}} = \text{argmax EI}(p) \quad (17)$$

式中: p_{new} —EI 函数最大的点。

步骤(3) 更新 $L(p_s)$ 的高斯分布:

在增加了新样本后,新的高斯分布表示如下:

$$\begin{cases} \begin{bmatrix} L(p_s) \\ L(p_{\text{new}}) \end{bmatrix} \sim N\left(0, \begin{bmatrix} U & k' \\ k'^T & k(p_{\text{new}}, p_{\text{new}}) \end{bmatrix}\right) \\ k' = [k(p_{\text{new}}, p_1), \dots, k(p_{\text{new}}, p_r)] \end{cases} \quad (18)$$

根据文献[18]中的方法,计算得到 $L(p_{s+1})$ 的新分布 $P(F(p_{\text{new}}) | D, p_{\text{new}})$, 具体表示如下:

$$\begin{cases} P(L(p_{new} | D, p_{new})) \sim N(\mu, \sigma^2) \\ \mu = k'^T U^{-1} L(p_s) \\ \sigma^2 = k(p_{new}, p_{new}) - k'^T U^{-1} k' \end{cases} \quad (19)$$

式中： μ —均值； σ —标准差。

步骤(4)持续进行步骤2和步骤3的循环操作，直到完成设定的最大迭代次数为止，在此过程中确定最佳输出点。

2.2 光伏出力预测模型框架

本文提出的 BOA-iTransformer 光伏预测模型其流程详见图3。首先对气象特征挖掘时间窗口 t 下的增量特征、统计特征和时变特征进行特征构造。然后输入 iTransformer 模块，将每个变量独立嵌入为一个标记，用于捕捉关键细粒度特征和多元数据的关联性；再采用贝叶斯优化调参，直到完成设定的最大迭代次数为止，得到光伏出力预测模型。

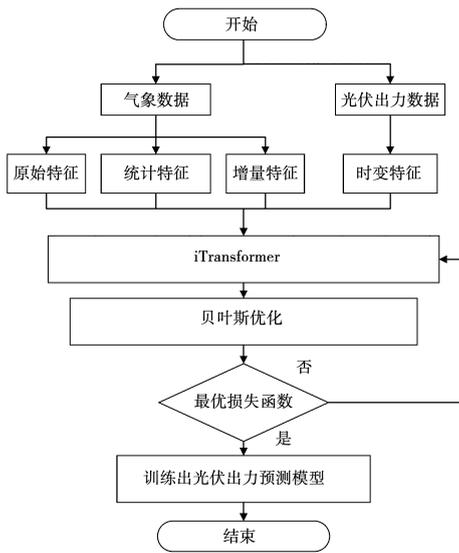


图3 模型流程图

Fig.3 Model flow chart

2.3 评价指标

为更好地评价本文所提模型和对比方法的预测结果，采用文献[9]的评价指标对模型进行评价：

(1) 可决系数 R^2 ：

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2}{\sum_{i=0}^{n-1} (y_i - \bar{y})^2} \quad (20)$$

(2) 绝对平均误差 MAE：

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \quad (21)$$

(3) 均方误差 MSE：

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (22)$$

式中： n —待预测光伏点的数量； y_i 和 \hat{y}_i —第 i 个光伏出力的实际值和预测值； \bar{y} —光伏实际出力样本的均值。

3 算例分析

3.1 仿真条件

本文的光伏数据选取中国某地光伏电站的实际光伏发电数据，数据的采样频率为 15 min。光伏电站样本数据包括气温、辐照度、湿度、风速等以及光伏实际出力。训练集数据条数为 62 042 条，时间跨度为 2016 年 4 月 1 日 - 2017 年 12 月 12 日。测试集数据条数为 13 371 条，时间跨度为 2017 年 12 月 13 日 - 2018 年 4 月 30 日。仿真实验使用的硬件环境为：CPU, i7 - 13700K; GPU, 3090 24G 显存; 内存 128 G。

3.2 参数设定

本文所采用的改进深度学习模型的具体训练参数设定如下：学习率设定为 0.001，批量大小设定为 128，轮次设定为 100，损失函数设定为 MSE，优化方法设定为 Adam Optimizer。

3.3 实验结果分析

根据 1.1 节对气象数据进行特征构造，原始特征包括辐照度、湿度、温度、压力、风速 5 类气象特征，每个特征提取 $2 \times 24 \times 5$ 个增量特征、 3×5 个统计特征和 2×9 个时变特征，共计 273 个特征。由于特征数量过多，导致模型拟合的计算量显著增加。故本文采用 2.1 节所提贝叶斯优化方法来优化参数，减少特征冗余度。

采用本文所提模型进行建模验证，损失函数选用均方误差 MSE。将光伏电站的训练集数据按时间顺序以 8:2 比例划分为训练集和验证集，以验证集的误差为目标进行贝叶斯优化调参。模型选择 200 步作为迭代步数，最优参数出现在 168 次， $t = 36$ 时，MSE 为 0.875。相比没有优化的 $t = 48$ 模型 MSE 减少了 0.169，优化过程中的迭代次数 - 误差指标联合分布如图 4 所示。

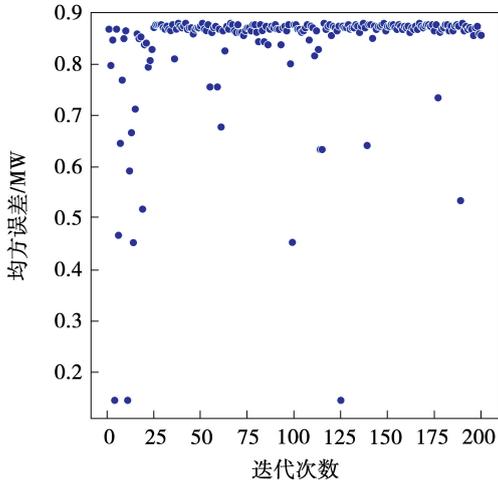


图 4 迭代次数 - 误差指标联合分布图

Fig. 4 Joint distribution of iteration count-error metrics plot

3.4 预测结果分析

分别构建以下几种模型和本文所提模型进行对比实验:

第 1 种模型:采用 1.1 节所述方法构造特征作为输入,本文所提模型作为训练模型,完成对光伏出力预测模型的训练和预测。

第 2 种模型:原始特征作为输入,BOA-iTransformer 模型作为训练模型。

第 3 种模型:原始特征作为输入,iTransformer 模型^[17]作为训练模型。

第 4 种模型:原始特征作为输入,Transformer 神经网络模型^[16]作为训练模型。

第 5 种模型:原始特征作为输入,LSTM 模型^[10]作为训练模型。

光伏电站的预测结果如表 1 所示。由表 1 中评价指标可知,本文所提方法预测模型的 R^2 值最大,MSE 和 MAE 值最小。其中,MSE 值相较未进行气象特征构造的 BOA-iTransformer、iTransformer、Transformer 和 LSTM 分别提高 1.73%,3.54%,7.24% 和 14.2%,因此效果最佳。综合来看,经过了特征构造的本文所提模型相较于未经过特征构造的 BOA-iTransformer、iTransformer、Transformer 和 LSTM 模型预测准确率有明显提高,证明了本文所提方法的有效性。

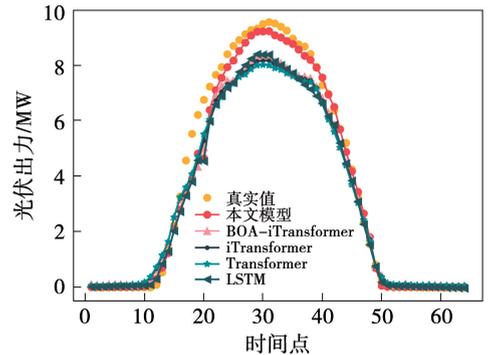
表 1 光伏电站预测结果

Tab. 1 Forecast result of photovoltaic power plant

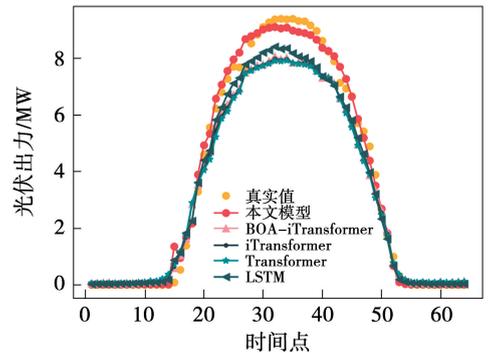
模型	MSE	MAE	R^2
本文模型	1.269	0.624	0.875
BOA-iTransformer	1.291	0.651	0.874
iTransformer	1.314	0.664	0.872
Transformer	1.361	0.682	0.868
LSTM	1.450	0.663	0.859

3.4.1 晴天天气下光伏出力预测情况

以晴天天气类型下某两日为预测日样本,采用上述 5 种模型对光伏出力进行预测,结果如图 5 和表 2 所示。



(a) 晴天1预测结果



(b) 晴天2预测结果

图 5 光伏出力晴天预测情况

Fig. 5 Photovoltaic output forecast on sunny days

表 2 晴天天气下光伏出力预测误差

Tab. 2 PV output forecast error in sunny weather

模型	晴天 1			晴天 2		
	MSE	MAE	R^2	MSE	MAE	R^2
本文模型	0.336	0.381	0.970	0.128	0.204	0.983
BOA-iTransformer	0.883	0.803	0.922	0.385	0.449	0.974
iTransformer	0.969	0.851	0.915	0.542	0.543	0.963
Transformer	1.074	0.898	0.906	0.595	0.574	0.960
LSTM	1.027	0.843	0.910	0.655	0.606	0.956

由图 5 和表 2 可知:(1) 在晴天天气类型下,由于光伏出力波动较小,光伏出力曲线变化趋势较为平稳。5 种模型对光伏出力预测都相对较准确,但本文所提模型相比其他模型有较明显优势。(2) 相较其他 3 种模型,未采用特征挖掘的 BOA-iTransformer 模型预测精度有略微提高,相较 iTransformer、Transformer 和 LSTM,本文模型仅降低了 0.07,0.11 和 0.09;而经过特征挖掘的本文所提方法模型精度显著提高,其中 MAE 相对对比方法依次降低 0.33,0.40,0.44 和 0.43,因此本文所提方法能提高对晴天天气下的光伏出力预测精度。

3.4.2 多云天气下光伏出力预测情况

以多云天气类型下某两日为预测日样本,采用上述 5 种模型对光伏出力进行预测,结果如图 6 和表 3 所示。

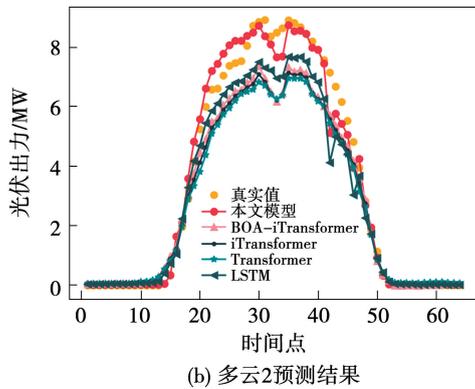
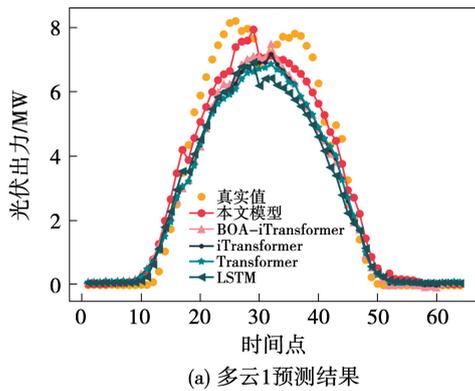


图 6 光伏出力多云天预测情况

Fig. 6 Photovoltaic output forecast on cloudy days

表 3 多云天气下光伏出力预测误差

Tab. 3 PV output forecast error in cloudy weather

模型	多云 1			多云 2		
	MSE	MAE	R ²	MSE	MAE	R ²
本文模型	0.346	0.435	0.963	0.558	0.663	0.941
BOA-iTransformer	1.424	1.014	0.847	1.103	0.859	0.883
iTransformer	1.669	1.105	0.821	1.191	0.879	0.874
Transformer	1.954	1.202	0.791	1.306	0.921	0.861
LSTM	1.152	0.871	0.877	1.479	1.022	0.843

由图 6 和表 3 可知:(1) 相比晴天天气类型,多云天气类型下各模型的预测曲线对真实值的拟合度优度都有所降低,但本文模型和 BOA-iTransformer 模型 R² 值分别达到 0.963 和 0.941,更加贴近真实值变化趋势。(2) 在多云天气类型下,未采用特征构造的 BOA-iTransformer 模型相较其他 3 种模型预测精度有略微提高;经特征挖掘后的本文模型相较 BOA-iTransformer、iTransformer、Transformer、LSTM 模型在 MSE 值上分别提高 0.811,0.978,1.178 和 0.863,因此本文模型能更好地对多云天气类型下的气象特征与光伏出力的映射关系进行拟合,从而提高预测精度。

3.4.3 雨天天气下光伏出力预测情况

以雨天天气类型下某两日为预测日样本,采用上述 5 种模型对光伏出力进行预测,结果如图 7 和表 4 所示。

由图 7 和表 4 可知:(1) 在雨天天气类型下,受气象影响引起辐照度波动从而导致光伏出力波动较大,5 种模型的预测效果均有所下降。对于雨天 1,本文模型和 BOA-iTransformer 模型也未能很好的拟合真实值变化趋势。但相较于对比的 3 种模型,预测精度仍有所提高。(2) 对于雨天 2,光伏出力真实值波动相较于多云天气条件波动更大,但本文模型和 BOA-iTransformer 模型相较其他 3 种模型拟合优度明显提升,分别为 0.848 和 0.812;误差 MSE 和 MAE 均明显下降,表明本文模型能更细致地挖掘影响光伏出力的气象因素变化特征,从而减少预测误差。

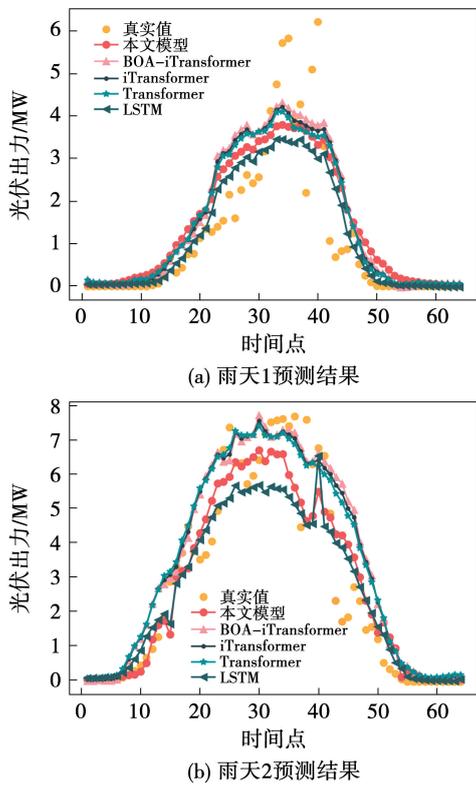


图 7 光伏出力雨天预测情况

Fig. 7 Photovoltaic output forecast on rainy days

表 4 雨天天气下光伏出力预测误差

Tab. 4 PV output forecast error in rainy weather

模型	雨天 1			雨天 2		
	MSE	MAE	R ²	MSE	MAE	R ²
本文模型	1.096	0.793	0.625	1.022	0.747	0.848
BOA-iTransformer	1.037	0.694	0.675	1.388	0.868	0.812
iTransformer	1.249	0.888	0.608	1.395	0.906	0.594
Transformer	1.218	0.878	0.618	1.349	0.888	0.608
LSTM	1.295	0.906	0.594	1.318	0.878	0.618

4 结 论

(1) 基于光伏发电与辐照度、气温、湿度、风速等气象因素有较强相关关系,提出一种气象特征构造方法。该方法构造增量特征、统计特征和时变特征共 3 类特征,从而实现对影响光伏出力的气象因素和时变特性更为精准、可靠的描述。

(2) 所提方法在光伏相关的气象特征上进行深层挖掘,算例通过对辐照度、湿度、温度、压力、风速 5 类气象特征构造增量特征、统计特征和时变特征共 273 个,考虑气象条件的相似性和短时变化性以

及光伏出力的时变特性,同时利用贝叶斯优化调参,剔除掉冗余特征,结果表明辐照度与湿度、辐照度与压强等细粒度特征组合可实现对光伏数据更细致、精准的描述,从而提高模型预测精度。

(3) 以光伏电站的实际光伏发电数据进行算例分析,验证了本文所提方法的有效性和优势。未经过特征构造的 BOA-iTransformer 模型相比其他模型可有效提高光伏出力预测的精度,但经过特征构造的本文所提模型的由于将每个输入变量独立嵌入,可更好地捕捉关键气象特征和多元数据关联性,相比对比方法 iTransformer、Transformer 和 LSTM 本文模型预测精度可分别提高 3.54%、7.24% 和 14.2%。

参考文献:

[1] 韩 坤. 基于 ALSTM 的光伏出力短期预测研究[J]. 仪器仪表用户, 2022, 29(2): 95 - 100.
HAN Kun. Study on short-term forecast of photovoltaic output based on ALSTM[J]. Instrumentation, 2022, 29(2): 95 - 100.

[2] 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.

[3] 李 斌, 张一凡, 颜世焯, 等. 基于改进极限学习机 ELM 的光伏发电预测方法研究[J]. 热能动力工程, 2022, 37(10): 207 - 214.
LI Bin, ZHANG Yifan, YAN Shiye, et al. Research on photovoltaic power generation prediction method based on improved extreme learning machine (ELM)[J]. Journal of Engineering for Thermal Energy and Power, 2022, 37(10): 207 - 214.

[4] 李立刚, 孙立群, 戴永寿, 等. 基于时空图卷积神经网络的光伏发电功率超短期预测方法[J]. 热能动力工程, 2023, 38(9): 152 - 157, 173.
LI Ligang, SUN Liqun, DAI Yongshou, et al. Ultra-short-term prediction method of photovoltaic power based on spatio-temporal graph convolutional neural network[J]. Journal of Engineering for Thermal Energy and Power, 2023, 38(9): 152 - 157, 173.

[5] 孟祥剑. 基于数据驱动的光伏发电系统最大功率跟踪及功率预测方法[D]. 济南: 山东大学, 2021.
MENG Xiangjian. Data-driven maximum power point tracking and power forecasting methods of photovoltaic power generation systems [D]. Jinan: Shandong University, 2021.

- [6] 魏鹏飞,石新聪,朱咏明,等.基于CEEMD-SSA-SVM的短期光伏发电功率预测[J].水力发电,2024,50(4):87-94.
WEI Pengfei,SHI Xincong,ZHU Yongming,et al. Short-term photovoltaic power prediction based on CEEMD-SSA-SVM[J]. Water Power,2024,50(4):87-94.
- [7] SHAKYA A, MICHAEL S, SAUNDERS C, et al. Using Markov switching model for solar irradiance forecasting in remote microgrids[C]//2016 IEEE Energy Conversion Congress and Exposition, Milwaukee, WI, USA, 2016.
- [8] 徐鹤勇,张倩.基于数字孪生和改进LSTM的光伏发电预测技术[J].热能动力工程,2023,38(2):84-91,100.
XU Heyong,ZHANG Qian. Photovoltaic power generation prediction technology based on digital twin and improved LSTM[J]. Journal of Engineering for Thermal Energy and Power, 2023, 38(2):84-91,100.
- [9] 杨京渝,罗隆福,阳同光,等.基于气象特征挖掘和改进深度学习模型的风电功率短期预测[J].电力自动化设备,2023,43(3):110-116.
YANG Jingyu,LUO Longfu,YANG Tongguang,et al. Wind power short-term forecasting based on meteorological feature exploring and improved deep learning model[J]. Electric Power Automation Equipment,2023,43(3):110-116.
- [10] 刘康,刘鑫,张蓬鹤,等.基于负荷尖峰特征LSTM自编码器的窃电识别方法[J].电力系统自动化,2023,47(2):96-104.
LIU Kang,LIU Xin,ZHANG Penghe,et al. Identification method of electricity theft based on long short-term memory autoencoder with load peak features[J]. Automation of Electric Power Systems,2023,47(2):96-104.
- [11] 谢从珍,王江储,谢心昊,等.基于细粒度特征的BOA-GBDT光伏出力预测[J].电网技术,2020,44(2):689-696.
XIE Congzhen,WANG Jiangchu,XIE Xinhao,et al. PV output prediction based on gradient boosting decision tree model with Bayesian optimization algorithm and fine-grained features[J]. Power System Technology,2020,44(2):689-696.
- [12] 宋煜,郭军红,袁荔,等.基于逐步聚类分析的短期光伏发电预测方法[J].热能动力工程,2023,38(10):158-166.
SONG Yu,GUO Junhong,YUAN Li,et al. Short-term photovoltaic power generation forecasting method based on stepwise clustering analysis[J]. Journal of Engineering for Thermal Energy and Power,2023,38(10):158-166.
- [13] HUANG C,KUO P. Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting[J]. IEEE Access,2019,7:74822-74834.
- [14] 龚小梅,张轶,胡术.引入特征融合和Transformer模型预测器的目标跟踪算法[J/OL].计算机工程与应用:1-11[2024-01-29].<http://kns.cnki.net/kcms/detail/11.2127.TP.20240118.1018.011.html>.
GONG Xiaomei,ZHANG Yi,HU Shu. Target tracking algorithm introducing feature fusion and Transformer model predictor[J/OL]. Computer Engineering and Applications:1-11[2024-01-29]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20240118.1018.011.html>.
- [15] 董明,李晓枫,杨章,等.基于数据驱动的分布式光伏发电功率预测方法研究进展[J].电网与清洁能源,2024,40(1):8-17,28.
DONG Ming,LI Xiaofeng,YANG Zhang,et al. Research progress on data-driven prediction methods for distributed photovoltaic power generation[J]. Power System and Clean Energy, 2024, 40(1):8-17,28.
- [16] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems,2017,30.
- [17] LIU Y,HU T,ZHANG H,et al. Itransformer: Inverted transformers are effective for time series forecasting[J]. arXiv Preprint arXiv,2023,12:2310.06625v2.
- [18] 刘哲.基于模糊聚类和高斯过程回归的光伏功率预测[D].广州:广东工业大学,2019.
LIU Zhe. Short-term solar power forecasting based on fuzzy C-means and Gaussian process regression[D]. Guangzhou: Guangdong University of Technology,2019.
- [19] 陈德余.基于改进多元组合算法的光伏发电功率预测方法研究[D].济南:齐鲁工业大学,2023.
CHEN Deyu. Research on photovoltaic power prediction method based on improved multivariate combination algorithm[D]. Jinan: Qilu University of Technology,2023.

(姜雪梅 编辑)