

DOI:10.13296/j.1001-1609.hva.2026.08.022

结合时空局部注意力网络与LSTM的日前光伏功率预测模型

宿连超¹, 樊静雨², 只德天³, 吴凯⁴

(1. 国网潍坊供电公司, 山东 潍坊 261000; 2. 国网山东省电力公司, 济南 250013; 3. 天津国投津能发电有限公司, 天津 300450; 4. 华北电力大学新能源电力系统国家重点实验室, 北京 102206)

摘要: 光伏功率随时间的波动具有不确定性和随机性, 准确可靠的光伏功率预测对电网的实时调度和频率调节具有重要意义。为了提高光伏功率预测的准确性, 提出一种结合时空局部注意力网络(spatio-temporal local attention network, STLAN)和长短时记忆网络(long short term memory, LSTM)的日前光伏发电功率预测模型(STLAN-LSTM)。所提STLAN-LSTM由双自注意力机制组成, 通过时序自注意力机制(temporal self-attention, TSA)动态定位相似的历史光伏功率序列, 同时采用空间自注意力机制(spatial self-attention, SSA)自动捕获相邻站点匹配的光伏功率序列。然后, 通过改进的相似序列分数估计策略统计两个注意力通道的输出, 使得模型侧重于日常特征中相似局部序列。最后, 使用极限学习机(extreme learning machine, ELM)将注意力通道的输出转换为非线性空间, 并输入至LSTM网络实现功率预测。在光伏电站测量数据上的实验结果表明, STLAN-LSTM的预测指标MAE、RMSE、MAPE值分别为1.26%、1.61%、23.21%, 具有较高的预测准确性, 为日前光伏发电功率预测提供有效的技术支持。

关键词: 光伏功率预测; 注意力机制; 解码; 时空相关性; LSTM网络

Day-ahead Photovoltaic Power Forecasting Model Based on Spatio-temporal Local Attention Network and LSTM

SU Lianchao¹, FAN Jingyu², ZHI Detian³, WU Kai⁴

(1. State Grid Weifang Power Supply Company, Shandong Weifang 261000, China; 2. State Grid Shandong Electric Power Company, Jinan 250013, China; 3. Tianjin Guotou Jinneng Power Generation Co., Ltd., Tianjin 300450, China; 4. State Key Laboratory of Alternate Electrical Power System With Renewable Energy Sources, North China Electric Power University, Beijing 102206, China)

Abstract: The fluctuation of photovoltaic(PV) power over time is characterized by uncertainty and randomness. Accurate and reliable PV power forecasting is of great significance for real-time grid dispatch and frequency regulation. For improving the accuracy of PV power forecasting, a day-ahead PV power forecasting model STLAN-LSTM based on Spatiotemporal Local Attention Network and Long Short-Term Memory (STLAN-LSTM) is proposed. The proposed STLAN-LSTM is composed of a dual self-attention mechanism. First, the similar historical PV power sequences is dynamically located through the Temporal Self-Attention(TSA) and, at the same time, the spatial self-attention(SSA) mechanism is used to automatically capture matching photovoltaic power sequence from adjacent stations. Then, an improved similarity sequence score estimation strategy is used to aggregate the output of the two attention channels, enabling the model to focus on similar local sequence within the daily features. Finally, the outputs of the attention channels are transformed into a nonlinear space using Extreme Learning Machine (ELM), which are then fed into an LSTM network to achieve power forecasting. Experimental results on measured data from PV power plants show that the forecasting indicators MAE, RMSE, and MAPE of STLAN-LSTM are 1.26%, 1.61%, and 23.21% respectively, indicating high forecasting accuracy and providing effective technical support and day-ahead PV power forecasting.

Key words: photovoltaic power forecasting; attention mechanism; decoder; spatio-temporal correlation; LSTM network

收稿日期:2026-02-22; 修回日期:2026-04-26

基金项目:国家自然科学基金资助项目(52007174)。

Project Supported by National Natural Science Foundation of China(52007174).

0 引言

为解决能源短缺和环境危机,光伏发电逐渐成为电力系统中不可或缺的部分^[1-5]。然而,由于各种气象因素的影响,光伏发电量的波动性和随机性给电力系统稳定运行带来巨大挑战^[6-7]。准确的光伏发电量预测是解决这个问题的重要方法之一。

光伏发电具有每日的时间周期性,还容易受到季节、天气类型、云运动等影响,具有一定的相似性和随机性。鉴于此,研究人员通常基于时间数据和空间相关性来提高光伏功率预测的准确性。基于时间数据的方法是将光伏发电预测视为时间序列预测问题,并将相似的天气日等信息纳入预测中,以提高光伏发电预测的准确性^[8-10]。例如,Chen等^[11]根据辐照度范围手动分类各种天气类型,Li等^[12]使用每周的天气过程将光伏发电分类为六类。但是这些手动分区方法效率低下且严重依赖经验。此外,研究人员还通过分类和聚类模型利用类似的气象条件构建光伏发电预测模型。例如,Feng等^[13]利用支持向量机模式识别分类模型自动分类预测的天气数据。Zhou等^[14]利用自组织映射将输入的天气日分组为多个簇,估计不同集合的预测模型参数。这些方法使用每日的粒度信息进行预测,忽略了一天中可能出现的多种不同天气类型,使得模型无法整合不同天数上多个相似序列的发电数据,存在历史信息利用不足的问题,因此仍需利用不同天数上的细粒度相似序列数据来改善光伏发电预测。

除了时间数据,研究人员还利用空间相关性来提高光伏预测的准确性^[15]。由于光伏输出与云的存在直接相关,在一定地理范围内云的运动使得相邻距离的光伏电站具有相似性^[16-17]。基于此,Zhang等^[18]通过计算欧氏距离、动态时间规整和Pearson相关系数来选择具有空间相似性的候选站点,并将这些站点作为贝叶斯网络的输入,实现光伏功率预测。Agoua等^[19]分析相邻站点的输出序列以及滞后序列之间的互相关关系,并选择具有一定空间相关性的序列作为预测模型的输入。这些提取空间相关性的方法基于相邻站点的长时间序列或滞后序列,而云的运动持续时间较短,将长时间的时间序列作为预测输入特征会引入额外的噪音并影响预测准确性。因此,仍需要一种方法,可以呈现由于云运动而产生的功率序列波动,以计算空间相关性,从而提高空间相关性信息的利用率。

与此同时,一些研究将时空相关性引入光伏发

电预测模型中^[20-21]。该类方法使用不同的相似性指标测量时空相关性,并将相关序列选作预测模型的输入。这种方法通过逐步优化特征提取和预测过程,例如,Simeunovic等^[22]构建了一个端到端的GCLSTM深度网络结构来解决模型融合的问题。王海军等^[23]结合贝叶斯神经网络和LSTM,基于时空关联特征实现分布式光伏功率的预测。宋绍剑等^[24]通过主成分分析法对光伏发电功率中的多元数据序列进行降维,并基于LSTM网络对短期光伏发电功率进行预测。然而,这些方法提取空间和时间相关性的方法相对局部化,并未从全局角度解决空间和时间相关性问题,且寻找最优解仍是一个挑战性问题。

上述预测算法引入了时空信息,但对时空相关性的利用不足,导致预测准确性低下。近期基于注意力机制的Transformer^[25]在时间序列预测方面表现出很强的性能,其关键技术自注意力可以有选择性地关注序列之间的相似信息^[26-27]。Transformer是一个端到端的框架,可以将光伏发电的空间—时间预测模块整合到整个框架中,进而统一整合时空相关性的提取和预测功能,在此基础上,搭配LSTM神经网络可以充分利用细粒度的时空相关性以提高预测准确性。

基于以上研究,文中提出一种基于双自注意力机制的LSTM日前光伏功率概率预测方法,称为STLAN-LSTM。该方法集成时序自注意力机制和空间自注意力机制到LSTM神经网络,充分考虑光伏数据的时空相关性。实验结果表明,STLAN-LSTM具有较高的预测准确性和鲁棒性。

1 光伏发电功率预测模型

1.1 模型框架

模型充分利用光伏发电的时空特征,基于时空局部注意力网络(spatio-temporal local attention network, STLAN),通过时序自注意力机制(temporal self-attention, TSA)计算注意力得分来动态地选择和利用历史数据。同时,采用空间自注意力机制(spatial self-attention, SSA)来自动捕捉邻近站点的相似光伏发电序列。为了简化和统一注意分数的计算,提出一种计算相似序列注意分数的新方法。该方法侧重于日常特征中相似局部序列的自动注意,改进原有的注意评分方法,使其更适合于光伏发电预测。最后通过ELM隐层的非线性扩展继承两个注意力分支的输出,将数据输入至LSTM实现功率预测。

1.2 时空局部注意力网络

文中提出时空局部注意力网络融合时空信息,其中包含时序自注意力机制和空间自注意力机制,

其结构见图1。为了防止多个数据输入和额外的时空信息融合引起的错误,采用并行结构连接多个注意力模块。

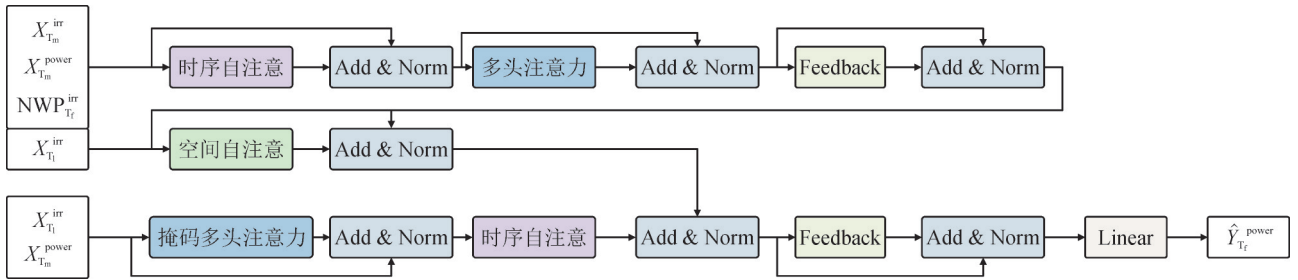


图1 时空局部注意力网络框架

Fig. 1 Spatio-temporal local attention network framework

在时间维度上,光伏输出的相似性通常体现在辐照度和天气变化的周期性上。因此,提出基于多头自注意力的时序自注意力机制,具体如下

$$TSA = SSAS \left[NWP_{T_f}^{irr}, \text{Concat} \left(X_{T_m}^{irr}, X_{T_m}^{power} \right), N_{Flag} \right] \quad (1)$$

式(1)中:SSAS(·)为相似性序列注意力评分策略; N_{Flag} 为超参数; $NWP_{T_f}^{irr}$ 为未来可获得的太阳辐照度; $X_{T_m}^{irr}$ 、 $X_{T_m}^{power}$ 分别为测量的辐照度、实际功率输出。细粒度注意力主要计算来自 $NWP_{T_f}^{irr}$ 和历史多日序列数据 $X_{T_m}^{irr}$ 、 $X_{T_m}^{power}$ 之间的相似序列得分。在STLAN中,TSA实现了对预测序列中每个周期的不同天气类型和周期数据的选择性注意。

STLAN包含两个时序细粒度时序注意力模块,利用编码器中的该模块来建模时间相关性,通过多头自注意力对周期进行进一步建模后,可以得到输出编码。在光伏数据的日季节模式中,时间维度的周期性占较大比例。因此,在解码器中,添加额外时序自注意力来强调时间维度的周期性。

此外,目标站点的光伏输出可能与邻近站点的任何历史时期具有一定的相似性,因此,利用邻近站点的历史数据作为额外的特征输入,可以帮助提高光伏功率预测模型的准确性和稳定性。具体来说,根据风速计算云移动的近似速度后,选择邻近站点一天的历史辐照度数据 $X_{T_l}^{irr}$ 作为输入。SSA计算待预测序列和邻近站点每个周期的相似序列注意力分数,从而学习空间相关性,具体表示为

$$SSA = SSAS \left(NWP_{T_f}^{irr}, X_{T_l}^{irr}, N_{Flag} \right) \quad (2)$$

编码器包含一个空间自注意力模块,主要用于空间相关建模。该模块以细粒度为标准在 $X_{T_l}^{irr}$ 中找到相似的云覆盖事件,并将提取的空间相关性嵌入到输出编码中。

1.3 相似序列注意力分数计算

STLAN结合了历史数据的时序特性和邻近站点数据的空间特性,而无论是从历史输出序列中还是从邻近站点的发电情况中寻找相似的时期,估计序列的相似性是至关重要的。因此,提出一种新的光伏数据相似序列注意力计算方法,相似序列注意力分数计算(similar sequence attention score, SSAS)。该方法采用皮尔逊相关系数 $PCC(\cdot)$ 作为相似性度量,皮尔逊相关系数是一种统计指标,用于量化两个变量之间的线性关系,在文中即光伏发电序列。通过计算不同序列之间的皮尔逊相关系数,可以确定他们的相似程度,从而识别数据中的相似模式或趋势, $PCC(\cdot)$ 表示为

$$PCC(X_a, X_b) = \frac{COV(X_a, X_b)}{\sigma_{X_a} \sigma_{X_b}} \quad (3)$$

式(3)中: X_a 、 X_b 为特征矩阵中的任意两个向量; COV 为协方差操作; σ 为标准差。

光伏数据多为非线性分布,皮尔逊相关系数可以度量两组数据之间的线性相关性,因此利用它来计算细粒度的序列相似度。

自注意力机制是Transformer中提高记忆能力的关键的模块,适用于在时间序列预测中分配特征权重。鉴于此,文中提出一种注意力分数计算模块,通过将相似日与自注意力机制相结合,自动关注相似序列见图2,该模块主要由两部分组成:相似序列的筛选和注意力分数的计算。输入是张量Query(Q)和Key(K),其中 Q 是未知事件类别的每日数据,而 K 为包含多天各种事件的长序列。

在相似序列筛选的部分,将 Q 、 K 中的每日数据分成 T 个不同的时间段,针对 Q 中的每个时间段和 K 中对应的时间段,计算每天的皮尔逊相关系数,并得到 U 组中最相似序列的索引,得

$$\text{Index}_U(Q, K, N_{Flag}) = \arg \text{Top}_U \left[PCC(Q_i, K_{jk}) \right] \quad (4)$$

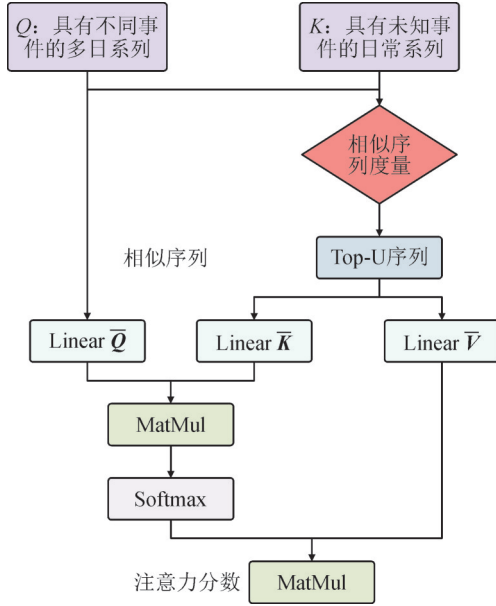


图2 相似序列注意力分数计算方法

Fig. 2 Method for calculating attention scores for similar sequences

式(4)中: U 为超参数, 表示相似序列的数量; N_{Flag} 为超参数, 用来表示时空信息, 当 N_{Flag} 表示时序时, STLAN-LSTM 提取时间相关性, 并仅计算处于相同时间段的序列之间的相关性。当 N_{Flag} 表示空间时, 提取空间相关性, 并计算所有时间段之间的相关性; Index_U 是计算后得到的时间索引; i, j 为不同的时间段; k 为第 k 天。根据索引可获得相似序列 SS

$$SS(Q, K, N_{\text{Flag}}) = \text{Linear} \left[K_{\text{Index}_U(Q, K, N_{\text{Flag}})} \right] \quad (5)$$

在注意力分数计算的部分, 使用点积注意力来实现类似序列之间关联规律的建模。为了能够无缝替代自注意力, 相似序列的注意力模块的输入包括查询向量 $\bar{Q}$ 、键向量 $\bar{K}$ 和值向量 $\bar{V}$, $\bar{K}$ 和 $\bar{V}$ 是对 K_{Index_U} 进行线性映射得到的结果, $\bar{Q}$ 是对 Q 进行线性映射得到的结果。首先, 通过对 $\bar{Q}$ 和 $\bar{K}$ 进行点积操作, 可以得到 $\bar{K}$ 中每个相似序列与 $\bar{Q}$ 的相关性。所得到的相关性经由缩放因子 $1/\sqrt{d_K}$ 和 Softmax 函数计算, 以获得 $\bar{K}$ 中每个序列的权重, 然后将权重与 $\bar{V}$ 相结合, 得到 Q 对于 K_{Index_U} 中相似序列的注意力分数

$$AS(\bar{Q}, \bar{K}, \bar{V}) = \text{Softmax} \left(\frac{\bar{Q} \bar{K}^T}{\sqrt{d_K}} \right) \bar{V} \quad (6)$$

最终, 计算相似序列注意力分数为

$$SSAS(Q, K, Flag) = \text{Linear} [\text{Concat}(\text{head}_1, \dots, \text{head}_h)]$$

where $\text{head}_i =$ (7)

$$AS[\text{Linear}(Q), SS(Q, K, N_{\text{Flag}}), SS(Q, K, N_{\text{Flag}})]$$

式(7)中, h 为超参数, 表示头部的数量, 而 head_i

对应不同头部的输出, 其计算方式为

$$\text{head}_i = AS[\text{Linear}(Q), SS(Q, K, N_{\text{Flag}}), SS(Q, K, N_{\text{Flag}})] \quad (8)$$

并行计算多个注意点, 在不同的子空间中学习相似的信息, 在输入 Q, K 之间实现相似的序列相关建模。模型最终利用 SSAS 实现简化和统一注意分数计算, 使得模型侧重于日常特征中相似局部序列的自动注意, 增强了 Transformer 中原有的注意力评分方法, 使其更适合于光伏发电预测。其具体过程见图 3。

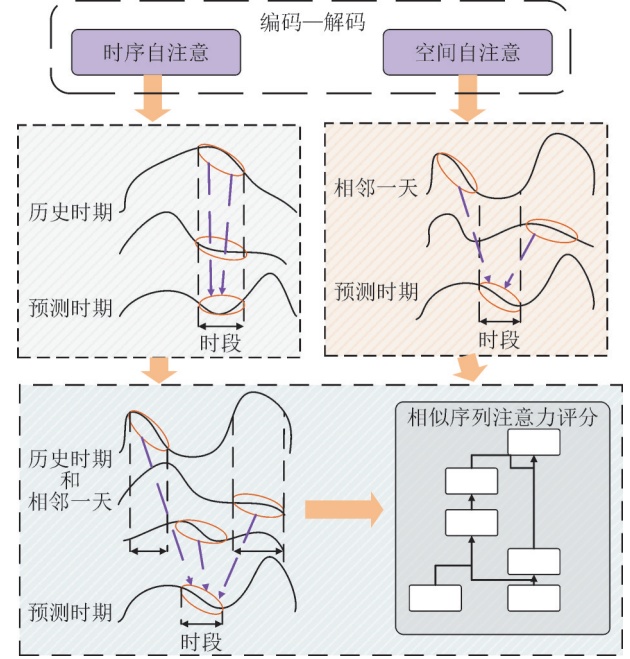


图3 双自注意力机制执行过程

Fig. 3 Dual self-attention mechanism execution process

1.4 非线性扩展

时空双自注意力机制能够有效捕捉时间和空间上的关联信息, 有助于模型更好地理解光伏功率预测中的时序和空间特征, 这可以提高模型对复杂关系的建模能力。但是这种对时空相关性的高度关注可能增加模型的复杂度和计算成本。ELM (extreme learning machine) 是一种具有单隐藏层的前馈神经网络。ELM 的独特特点在于, 从输入层到隐藏层的连接权重和偏置是随机初始化的。随后, 通过线性代数直接求解从隐藏层到输出层的连接权重。

设输入数据为 X , 输出数据为 Y , 隐藏层输出为 H , 则有以下关系成立

$$\begin{cases} Y = \beta \cdot g(\omega \cdot X + b) \\ H = g(\omega \cdot X + b) \\ \beta = H^+ Y \end{cases} \quad (9)$$

式(9)中: ω, b 为隐含层的连接权重和偏置向量。输出层的权重 β 通过求解 H 的 Moore-Penrose 伪

逆来确定。在STLAN-LSTM模型中,ELM的输入 X 是来自SSAS输出的相似序列,输出 Y 为经过非线性空间转换后的序列数据。这种直接的方法使得ELM具有较高的训练速度,并减少了对调参数的依赖,从而降低了ELM的复杂性。

此外,ELM通过隐层的非线性变换,可以更好地捕捉光伏功率预测中复杂的非线性关系。模型最后将原始数据和编码解码数据通过ELM的隐层转化为非线性空间,提取双通道特征,输入至LSTM网络进行回归预测,得到预测结果。

1.5 LSTM网络

LSTM神经网络由输入层、隐藏层和输出层构成,隐藏层包含一定数量的带有输入门和输出门的记忆单元。Gers等人^[28]通过引入遗忘门,对LSTM神经网络进行了改进。这些门控制了信息在记忆单元中的流动。隐藏层的核心组件被称为记忆块,其中一些记忆单元共享相同的门单元,记忆块的结构见图4。

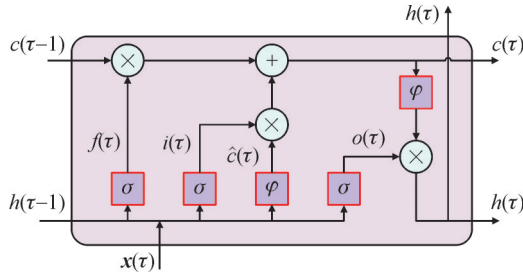


图4 LSTM记忆块结构

Fig. 4 Member module structure of LSTM

时间帧中的输入序列表示为 $\{x(1), x(2), \dots, x(M)\} \in R^{K \times M}$, 其中 $x(\tau) \in R^K$ 是时间戳 τ 处的特征向量。一个记忆块具有 J 个记忆单元和一个输入特征向量 $x(\tau) \in R^K$, 它会进行 M 次更新, 每个输入序列中的特征向量进行一次更新。记忆块的每次更新会生成当前状态向量 $c(\tau) \in R^J$ 。来自前一个时间戳的单元状态向量 $c(\tau-1) \in R^J$ 和单元输出向量 $h(\tau-1) \in R^J$ 被用于计算当前的输出向量 $h(\tau) \in R^J$ 。输入激活向量 $i(\tau) \in R^J$ 、遗忘激活向量 $f(\tau) \in R^J$ 和输出激活向量 $o(\tau) \in R^J$ 在输入序列的每个时间步更新一次, 使用Sigmoid激活函数(σ), 见式(10)、(11)、(13)。双曲正切激活函数(φ)用于计算中间单元状态, 称为单元候选状态 $\hat{c}(\tau) \in R^J$, 见式(13)。相应的 $(J \times K)$ 输入权重矩阵被连接成 $\{W_i^T, W_f^T, W_c^T, W_o^T\}^T$, 得到一个 $(4J \times K)$ 矩阵。类似地, 相应的 $(J \times J)$ 循环权重矩阵被连接成 $\{U_i^T, U_f^T, U_c^T, U_o^T\}^T$, 得到一个 $(4J \times J)$ 矩阵。相应的 $(J \times 1)$ 偏差被连接成 $\{b_i^T, b_f^T, b_c^T, b_o^T\}^T$, 得到一个 $(4J \times 1)$ 矩阵。记忆块在时间步 τ 的更新如下:

$$f(\tau) = \sigma[W_f x(\tau) + U_f h(\tau-1) + b_f] \quad (10)$$

$$i(\tau) = \sigma[W_i x(\tau) + U_i h(\tau-1) + b_i] \quad (11)$$

$$\hat{c}(\tau) = \varphi[W_c x(\tau) + U_c h(\tau-1) + b_c] \quad (12)$$

$$o(\tau) = \sigma[W_o x(\tau) + U_o h(\tau-1) + b_o] \quad (13)$$

$$c(\tau) = f(\tau) \otimes c(\tau-1) + i(\tau) \otimes \hat{c}(\tau) \quad (14)$$

$$h(\tau) = o(\tau) \otimes \varphi[c(\tau)] \quad (15)$$

式(14)、(15)中, $\otimes$ 为逐元素乘法。

1.6 预测步骤

综合上述方法,STLAN-LSTM模型训练和预测流程见图5。

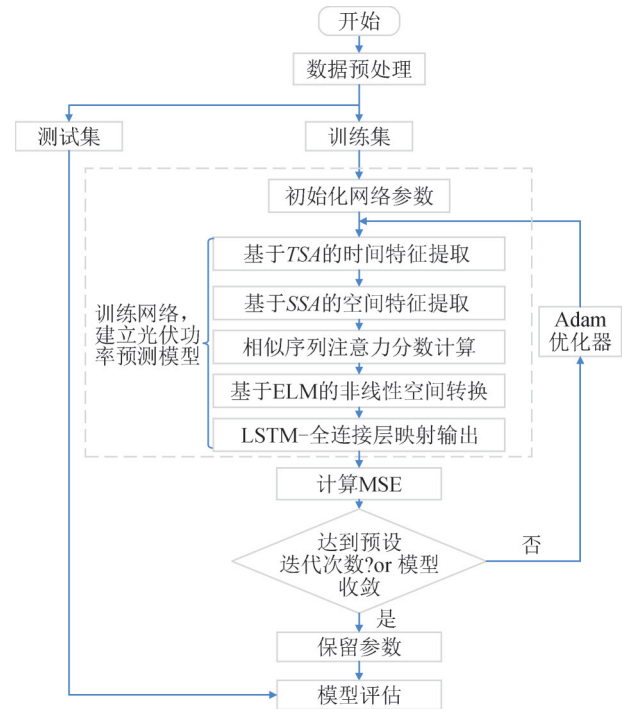


图5 STLAN-LSTM模型训练和预测流程

Fig. 5 Training and prediction process of STLAN-LSTM

根据图5,STLAN-LSTM模型的预测步骤如下。

1.6.1 模型训练

- 1)收集和整理用于训练和预测的序列数据,确保数据按时间顺序划分,并分为训练集和测试集。
- 2)将原始数据归一化为区间 $(0,1]$,具体有

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (16)$$

式(16)中: X 为原始数据; $X_{\min}$ 、 $X_{\max}$ 分别为最小值、最大值; X' 为归一化后的输入。

- 3)基于训练集训练模型,其中基于双自注意力的网络架构用于捕捉序列之间的全局依赖关系,LSTM用于处理序列内部的局部依赖关系。
- 4)通过Adam优化器和MSE损失函数对模型进行参数优化和训练,重复迭代训练过程,直到模型收敛或达到预设迭代次数。

5)使用训练好的模型对测试集进行预测,将测试集输入网络,获取预测值。

1.6.2 模型预测

1)使用时空局部注意力网络对 X 进行编码—解码,其中TSA用来计算被预测时期和历史时期之间的注意力得分,系统会自动选择最相似的历史周期并用于预测。SSA负责度量当前临近站点与目标站点之间的相似性,并计算空间注意力得分。

2)采用相似序列注意力计算方法来简化计算,充分利用细粒度的时空信息。

3)考虑到数据的非线性特征,将编码器的输出通过ELM隐层进行非线性展开,提取对应通道特征。并将ELM扩展的非线性特征输入LSTM中进行回归预测。

2 算例分析

2.1 数据准备

文中使用的数据样本来自中国中部的多个光伏发电厂的实测数据(时间跨度:从2022年1月1日的00:00-12月31日的23:45)。由于光伏功率输出具有基于太阳辐射和动力学的年周期性特征,文中选择一年的数据集。白天持续时间为15 h,数据的时间步长为15 min,电厂的装机容量为20 MW,此外,还包括总辐射、直接辐射、散射辐射、温度、环境温度、气压、环境湿度和实测功率数据。

在数据集中的13个备选电厂中,根据云移动的近似速度,选择了目标站点50 km范围内的13个光伏电站进行分析,见图6。如图6所示(目标站点标记为 N_0),选择的相邻站点与目标站点的历史辐照度数据的相关性更高,可以为目标站点的PV输出预测提供更多的可用信息。

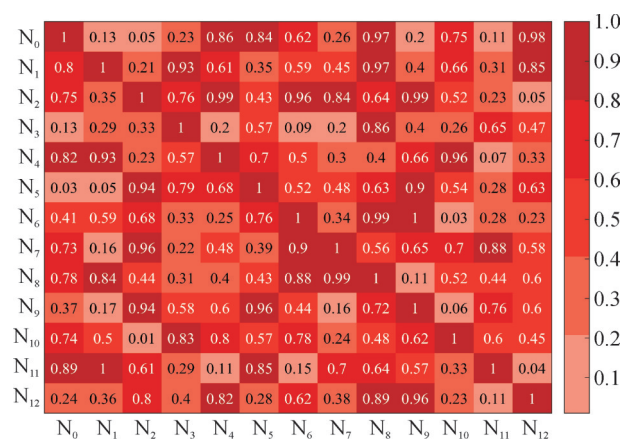


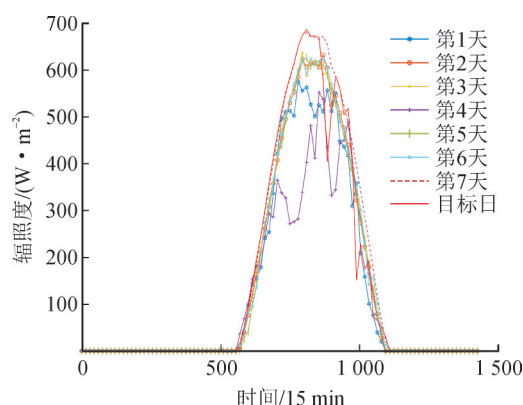
图6 13个站点相关性热力图

Fig. 6 Correlation heat maps of 13 stations

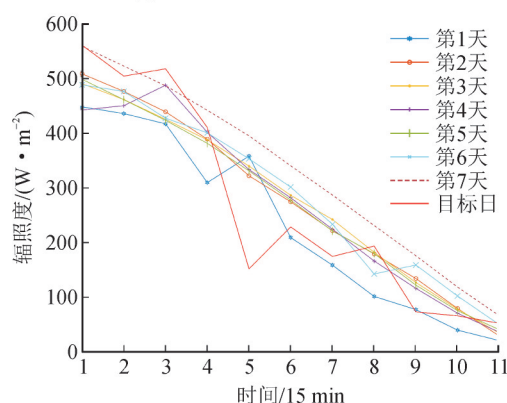
光伏发电受到太阳辐射、云运动等因素的影

响。类似的光伏发电也将出现在类似的季节和天气条件下。现有的基于相似天数的光伏功率预测方法利用这一特性提高了预测精度。但是,根据相似日的方法选择的序列在局部时间段可能不是最相似的,这将导致信息利用率较低。为了验证这一想法,随机选择特定某一天目标地点的NWP辐照度数据和前7天的辐照度数据。

目标日和最后7天的辐照度数据见图7(a),即全天数据。同时,放大目标日到特定时间段的辐照度曲线,也放大前7天到同一时间段,得到聚焦于局部序列的曲线见图7(b)。



(a) 目标日和最后7天的辐照度数据



(b) 目标日特定时间段的辐照度曲线

图7 历史7天和目标日辐照度数据

Fig. 7 Historical seven-day and target-day irradiance data

如图7(a)所示,当比较全天的曲线时,在历史记录的第7天与目标日之间存在相似的辐照度模式。然而,当在图7(b)中关注局部时期时,历史上的第1天比第7天更相似。这说明,在引入历史数据时,从局部序列的角度分析相似性比从一整天的角度分析更合理。

此外,将预测数据按季节划分,以严格验证预测模型的性能,2022年气温变化曲线见图8。

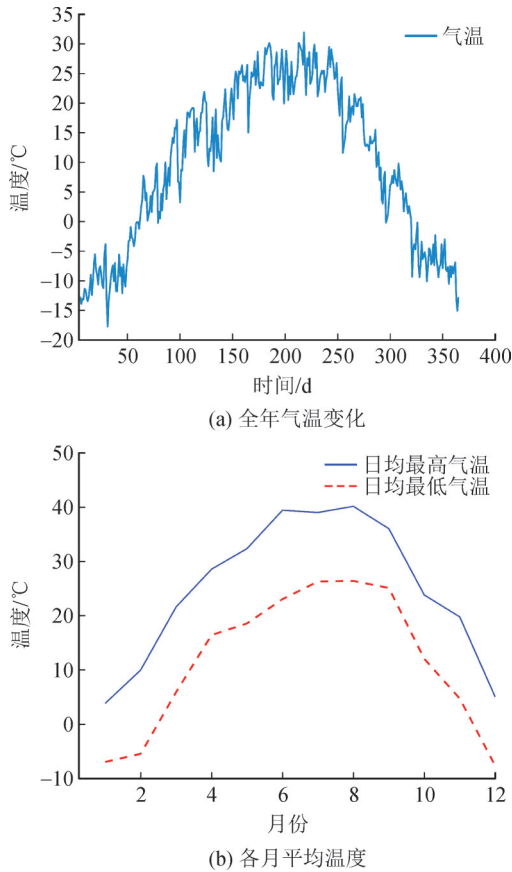


图8 2022年发电站气温变化曲线

Fig. 8 Temperature change curve of power stations in 2022

由图8中可知,1、12月时气温最低,7、8月气温最高,2~6月和9~11月气温呈现过渡趋势,结合气候特征季节划分法,定义2~5月为春季、6~8月为夏季、9~11月为秋季、12至次年1月为冬季。

模型编码器中时序自注意力的最相似序列的数量 $U=4$,其余相似序列注意中 $U=3$,并行注意头数量 $h=8$ 。采用Adam优化器和MSE损失函数用于模型训练,初始学习率0.001,批处理大小16。相邻站点数量设置为13,在相似序列注意力机制中的每日周期数 T 设置为6。此外,由于白天有15 h,采样频率为15 min/次,所以每日数据序列的长度设置为60。最后,在训练阶段,模型使用气象特征、时间特征、光伏电站特征等历史数据来学习模式和关联,在预测阶段,同样使用这些特征数据,用来预测未来某时刻的光伏功率。

2.2 评估指标

文中采用平均绝对误差(MAE)、均方根误差(RMSE)、平均绝对百分比误差(MAPE) 3种统计指标来评价该模型的预测性能。

$$e_{MAE} = \frac{1}{n} \sum_{i=1}^n |x_{\text{predicted}}[i] - x_{\text{actual}}[i]| \quad (17)$$

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{\text{predicted}}[i] - x_{\text{actual}}[i])^2} \quad (18)$$

$$e_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_{\text{predicted}}[i] - x_{\text{actual}}[i]}{x_{\text{actual}}[i]} \right| \times 100\% \quad (19)$$

式(17)-(19)中: n 为样本数; $x_{\text{predicted}}$ 、 x_{actual} 为预测值和真实值。

2.3 与其他预测方法的对比

选择具有代表性的预测模型与文中所提方法进行比较。这4个基线模型分别是时间序列预测模型ARIMA^[29]、机器学习模型SVR^[30]、深度神经网络LSTM、时空方法ConvLSTM^[31],结果见表1。

表1 与不同预测模型的对比结果

Table 1 Comparison results with different prediction models

模型	e_{MAE}	e_{RMSE}	$e_{MAPE}/\%$
ARIMA	3.29	4.46	51.25
SVR	1.99	2.83	34.96
LSTM	1.98	2.85	29.69
ConvLSTM	1.86	2.68	26.33
STLAN-LSTM	1.56	2.47	23.21

从表1可以看出,时间序列模型的预测效果最差。SVR和LSTM预测效果优于ARIMA。同时,LSTM不利用时空信息,缺乏对不同天气类型的学习能力,因此性能不佳。而ConvLSTM可以提取时空信息,相当于在LSTM中添加了空间信息,以提高预测效果。然而,由于ConvLSTM使用不同的方法来提取时空相关性,从较长序列的角度出发,其预测效果不如STLAN-LSTM,STLAN-LSTM更关注局部相似的信息。

为了进一步说明所提模型所采用的相似序列注意力方法优于先前的基于天数的分类方法,使用K-means模型对基于NWP的每日数据进行天气类型分类,并针对每种天气类型训练不同的预测模型。在这里,选择3种代表性的预测模型:SVR(机器学习模型)、Transformer(长序列时间序列预测模型)和ConvLstm(时空模型),结果见表2。

表2 与基于K-means模型的对比结果

Table 2 Comparison results with K-means-based models

模型	e_{MAE}	e_{RMSE}	$e_{MAPE}/\%$
K-means+SVR	1.92	2.74	30.42
K-means+Transformer	1.84	2.73	28.67
K-means+ConvLSTM	1.73	2.51	25.61
STLAN-LSTM	1.56	2.27	23.33

表2的结果显示,使用K-means进行天气分类后,3种模型的预测结果都得到了改善,但均差于

STLAN-LSTM。证明了所提局部相似序列自动注意方法能以比相似日方法更细的粒度提取相似天气类型的历史信息,使模型具有更好的预测结果。

2.4 不同季节预测性能分析

实验分析所提模型不同季节的光伏功率预测性能,结果见表3。

表3 STLAN-LSTM的四季预测结果

Table 3 Forecasting results of four seasons for STLAN-LSTM

季节	e_{MAE}	e_{RMSE}	$e_{MAPE}/\%$
春季	1.53	2.06	25.63
夏季	1.81	2.43	27.43
秋季	1.26	1.61	23.21
冬季	2.79	3.32	30.17

根据表3数据,秋季的功率预测性能整体优于

其他季节,MAE、RMSE、MAPE值分别为1.26%、1.61%、23.21%。这是由于秋季的天气相对晴朗、且湿度较低,便于功率预测模型更精准的预测。4个季节模型功率预测值和真实值对比曲线见图9。

2.5 长序列预测性能分析

实验分析日前预测中每个时间索引的MAE,以验证所提模型的性能。不同模型的MAE曲线见图10,其中虚线表示平均预测误差。

在提前2 h的预测中,所提模型并没有达到最好的预测精度。ARIMA和SVR的预测精度较高。然而,随着预测序列的增长,所提模型的误差增加较小,更适合于长序列的输出预测。平均总体预测结果最准确。此外,从图10中可以看出,ARIMA的预测结果在中午达到峰值,具体数据结果见表4。

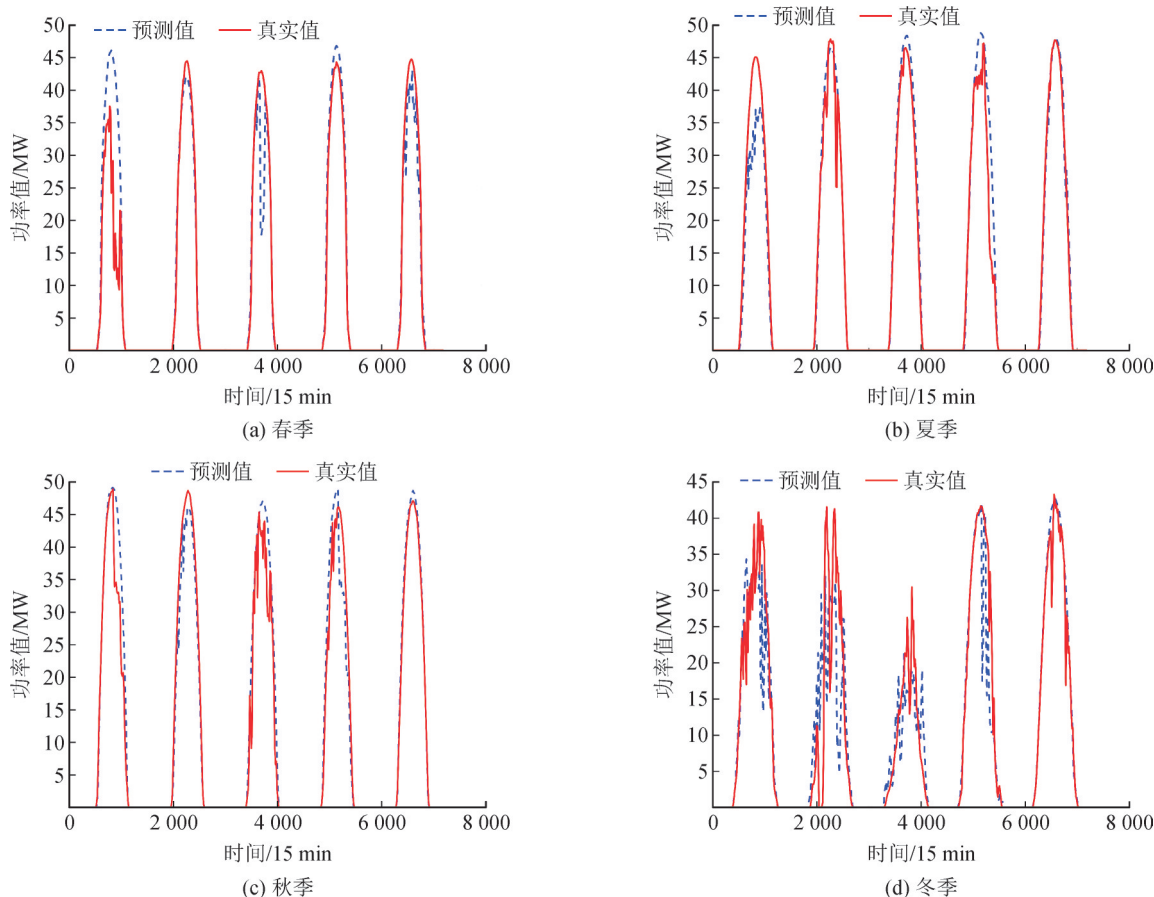


图9 四季预测值和真实值对比

Fig. 9 Comparison of forecasting value and real value in four seasons

从表4可以看出,在提前0.5 h的预测中,SVR和ARIMA的误差较小,MAE、RMSE值分别达到了1.02、1.52和1.17、1.62,而此时STLAN-LSTM为1.46和2.32,LSTM的误差相对较大,这是由于NWP数据本身的不确定性引入的显著误差。在提前1 h的预

测中,ARIMA的误差随着序列的增加而增加,而其他模型的性能与提前半小时的预测结果相似,此时SVR模型达到了最佳预测精度,MAE、RMSE值分别为1.33、1.86,相较于STLAN-LSTM分别降低了10.14%、21.19%,这表明注意力机制对这种短期内的

预测性能提高不显著。

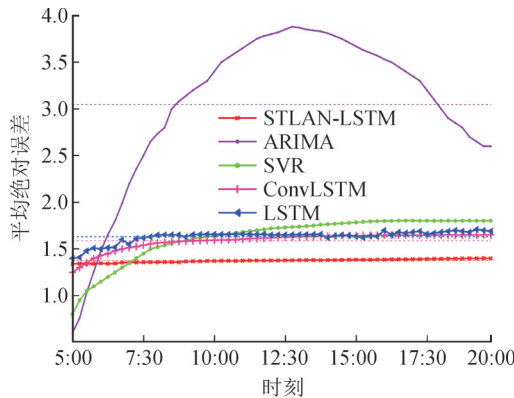


图10 提前一日预测误差比较曲线

Fig. 10 One-day ahead forecasting error comparison curve

随着预测时间的增加,预测误差从提前3 h增加到提前一天,STLAN大部分时间都保持着最高的预测精度,其平均预测结果MAE、RMSE分别达到了

表4 长序列预测的性能分析

Table 4 Performance analysis of long sequence forecasting

模型	提前 30 min		提前 1 h		提前 3 h		提前 6 h		提前 15 h		平均	
	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}
ARIMA	1.02	1.52	1.44	2.03	2.82	3.80	3.97	5.32	2.92	3.86	3.29	4.46
SVR	1.17	1.62	1.33	1.86	1.70	2.44	2.07	2.96	2.22	3.16	1.99	2.83
LSTM	1.59	2.19	1.76	2.46	1.92	2.77	2.01	2.90	2.09	3.02	1.98	2.85
ConvLSTM	1.53	2.15	1.60	2.26	1.77	2.52	1.88	2.72	1.97	2.86	1.86	2.68
STLAN-LSTM	1.46	2.32	1.48	2.36	1.52	2.44	1.55	2.47	1.60	2.50	1.56	2.47

2.6 天气类型预测性能分析

实验比较了基线模型和所提模型在不同天气类型下的性能。对测试集中理论全球水平辐照度(GHI)与实测辐照度之间的差异进行了K-means聚类分析。根据天气类型,将这些天分为晴天和非晴天两类见表5。

表5 天气类型的性能分析

Table 5 Performance analysis of weather type

模型	晴天		非晴天	
	e_{MAE}	e_{RMSE}	e_{MAE}	e_{RMSE}
ARIMA	3.87	5.40	3.06	3.93
SVR	2.06	2.72	3.15	4.37
LSTM	1.83	2.58	2.82	4.05
ConvLSTM	1.85	2.62	2.71	3.98
STLAN-LSTM	1.40	2.03	2.22	3.42

如表5所示,除了ARIMA外,其他模型在晴天的预测误差小于非晴天。这是因为太阳辐照度的波动性较大,非晴天的辐照度较低,导致普遍存在较大的预测误差。在晴天中,LSTM的预测比ConvLSTM更准确,MAE、RMSE值分别降低了1.08%、1.53%。但在非晴天中,ConvLSTM表现更

好,预测误差相较于LSTM分别降低了3.90%、1.73%,说明添加空间信息能够更好地捕捉不确定的云运动模式,从而提高非晴天光伏输出的预测准确性。STLAN-LSTM提高了晴天的输出预测和非晴天的低不确定性预测精度,在晴天的预测中减少了误差,其MAE、RMSE值分别达到了1.41、2.03,优于对比模型中在晴天精度最高的LSTM。此外,细粒度的时间注意力和云覆盖的空间注意力可以合理地将时空相关信息引入预测模型,从而使对波动较大的非晴天的输出预测更加准确,在非晴天也展现出了最低的预测误差,其MAE、RMSE值分别达到了2.22、3.42,优于在非晴天精度最高的ConvLSTM。以上结果进一步证明,在不晴朗的天气条件下,STLAN-LSTM更有利于预测光伏产量。

2.7 时空相关学习性能分析

为了验证基于相似序列注意力的空间—时间相关性提取的有效性,通过移除和保留时序自注意力或空间自注意力来评估其对预测误差的影响。此外,还通过原始的多头注意力机制来提取时间和空间相关趋势,相对地展示相似序列注意力的优

势,结果见表6。

表6 时空相关学习性能

Table 6 Spatio-temporal learning performance

模型	e_{MAE}	e_{RMSE}
STLAN-LSTM	1.26	1.61
Transformer	1.58	2.12
仅 TSA	1.33	1.82
仅 SSA	1.64	2.98
时序多头注意力+SSA	1.54	2.05
TSA+空间多头注意力	1.38	1.89

从表6可知,与Transformer模型的预测误差1.58、2.12相比,STLAN-LSTM的预测误差分别降低了20.25%、24.06%,与仅时序自注意力模型(TSA)相比,STLAN-LSTM的预测误差分别降低了6.02%、11.54%,与仅空间自注意力模型(SSA)相比,STLAN-LSTM的预测误差分别降低了23.17%、45.97%。这些结果验证了时间自注意力和空间自注意力对提高光伏功率预测的准确性的有效性。同时,与原始Transformer模型相比,STLAN-LSTM的误差降低程度超过了仅时间注意力和仅空间注意力的降低之和,这是由于时间和空间维度之间存在特定的耦合关系,以及提取空间—时间相关性的算法增强了预测准确性。

此外,实验采用原始的多头注意力机制替换了提出的时空注意力来提取空间—时间信息,结果显示,相较于原始时序多头注意力机制,STLAN-LSTM的MAE、RMSE值分别降低了18.18%、21.46%,相较于原始空间多头注意力机制,STLAN-LSTM的MAE、RMSE值分别降低了8.70%、14.81%。值得注意的是,仅具有空间注意力时,预测误差MAE、RMSE值分别达到了1.64、2.98,比原始Transformer模型高3.80%、40.57%,原因是空间相关性信息的复杂性较高,如果不能正确筛选出相似序列的注意力矩阵,则很难改善预测结果,而本文提出的时空注意力能够自动精确地完成这一任务。

3 结论

为提高日前光伏功率预测的准确性和可靠性,提出了一种基于STLAN-LSTM的光伏功率预测模型。该方法充分考虑光伏功率的时空特性和相关性,采用时序自注意力和空间自注意力加强模型时空特征的提取,通过一种新的相似序列分数计算方法简化注意力通道输出。所提出的神经网络在端到端训练过程中进行了优化,而不需要进行特征工程。主要结论如下:

1)在光伏发电实测数据上的实验结果表明,STLAN-LSTM在秋季实现最佳功率预测准确度,MAE、RMSE、MAPE指标分别达到了1.26%、1.61%、23.21%,且在晴天和非晴天预测精度对比结果中,误差相较于对比模型中晴天精度最高的LSTM降低了23.50%、22.52,相较于在非晴天精度最高的ConvLSTM降低了18.08%、14.07%,具有较高的准确性和鲁棒性。

2)长序列预测性能比较实验结果显示,在提前3 h到提前一天的预测中,STLAN-LSTM具有最高预测精度,误差平均值分别达到了1.56、2.47,相较于时空模型ConvLSTM降低了16.13%、7.84%,相较于时序模型LSTM降低了21.21%、13.33%,在时空相关性和长序列预测方面展现出了优越性能。

3)时空相关性分析实验结果显示,STLAN-LSTM的预测误差MAE和RMSE相较于原始多头注意力模型Transformer分别提高了20.25%、24.06%,表明所提方法可以精准地将时空相关信息引入预测模型,从细粒度信息和云覆盖信息中更有效地提取时空相关性。

综上所述,所提模型在长序列预测性能、天气类型预测性能以及时空相关性学习性能3个方面均具有优越表现,为日前光伏发电功率预测提供了可行的方案。然而,STLAN-LSTM采用了双自注意力机制,虽然有效地提高了模型的预测精度,但模型的可解释性仍有较大提升空间,未来工作将把主流的方法与编码—解码器相结合,以探索更合理的模型解释方法。

参考文献:

- [1] Yao Tiechui, Wang Jue, Wu Haoyan, et al. Intra-hour photovoltaic generation forecasting based on multi-source data and deep learning methods[J]. IEEE Transactions on Sustainable Energy, 2022, 13 (1): 607-618.
- [2] 徐其春, 李良杰, 卢泽汉. 电网接入光伏发电功率的实时调度与预测研究[J]. 电网与清洁能源, 2023, 39(12): 141-147.
Xu Qichun, Li Liangjie, Lu Zehan. A study on the real-time scheduling and prediction of photovoltaic power connected to the power grid[J]. Advances of Power System & Hydroelectric Engineering, 2023, 39(12): 141-147.
- [3] 冯俊琨, 王黎光. 基于马尔可夫链的智能微电网光伏发电预测[J]. 计算机应用与软件, 2023, 40(4): 343-349.
Feng Junkun, Wang Liguang. Photovoltaic power generation prediction of smart microgrid based on markov chain[J]. Computer Applications and Software, 2023, 40(4): 343-349.
- [4] 仇晨光, 倪杰, 王博仑, 等. 基于相似日与数据机理混合驱动的光伏发电功率预测[J]. 电网与清洁能源, 2025, 41(9): 101-109.
Qiu Chenguang, Ni Jie, Wang Bolun, et al. Photovoltaic power

- generation forecasting based on hybrid-driven approach combining similar days and data mechanism[J]. *Power System and Clean Energy*, 2025, 41(9): 101-109.
- [5] 刘庆, 王有军, 张垚, 等. 基于注意力机制的融合式NCP-DCNN短期光伏功率预测方法[J]. *智慧电力*, 2025, 53(8): 62-69.
Liu Qing, Wang Youjun, Zhang Yao, et al. Attention mechanism-based integrated NCP-DCNN method for short-term photovoltaic power forecasting[J]. *Smart Power*, 2025, 53(8): 62-69.
- [6] Antonanzas J, Osorio N, Escobar R, et al. Review of photovoltaic power forecasting[J]. *Solar Energy*, 2016(136): 78-111.
- [7] 吐松江·卡日, 吴现, 马小晶, 等. 基于地基云图数据多维特征融合的光伏功率预测算法[J]. *电力系统保护与控制*, 2025, 53(11): 84-94.
Tusongjiang Kari, Wu Xian, Ma Xiaojing, et al. Photovoltaic power prediction algorithm based on multidimensional features fusion of ground-based cloud images[J]. *Power System Protection and Control*, 2025, 53(11): 84-94.
- [8] Kim H, Lee D. Probabilistic solar power forecasting based on bivariate conditional solar irradiation distributions[J]. *IEEE Transactions on Sustainable Energy*, 2021, 12(4): 2031-2041.
- [9] 王健, 刘汇源, 张占喜, 等. 基于时空特征提取与跨模态融合的光伏集群功率预测[J]. *电力建设*, 2025, 46(11): 121-129.
Wang Jian, Liu Huiyuan, Zhang Zhanxi, et al. Power prediction of photovoltaic clusters based on spatio-temporal feature extraction and cross-modal fusion[J]. *Electric Power Construction*, 2025, 46(11): 121-129.
- [10] 史文瑜, 张珍翼, 杨德昌. 基于时间序列大模型TimeGPT光伏功率预测方法研究[J]. *电力科学与技术学报*, 2025, 40(4): 150-160.
Shi Wenyu, Zhang Zhenyi, Yang Dechang. Research on photovoltaic power forecasting method based on time series large model TimeGPT[J]. *Journal of Electric Power Science and Technology*, 2025, 40(4): 150-160.
- [11] Chen Changsong, Duan Shanxu, Cai Tao, et al. Online 24 h solar power forecasting based on weather type classification using artificial neural network[J]. *Solar Energy*, 2011, 85(11): 2856-2870.
- [12] Li Hui, Ren Zhouyang, Xu Yan, et al. A multi-data driven hybrid learning method for weekly photovoltaic power scenario forecast[J]. *IEEE Transactions on Sustainable Energy*, 2022, 13(1): 91-100.
- [13] Feng Cong, Cui Mingjian, Hodge B M, et al. Unsupervised clustering-based short-term solar forecasting[J]. *IEEE Transactions on Sustainable Energy*, 2019, 10(4): 2174-2185.
- [14] Zhou Nan, Xu Xiaoyuan, Yan Zheng, et al. Spatio-temporal probabilistic forecasting of photovoltaic power based on monotone broad learning system and copula theory[J]. *IEEE Transactions on Sustainable Energy*, 2022, 13(4): 1874-1885.
- [15] Zhang Ruiyuan, Ma Hui, Saha T K, et al. Photovoltaic nowcasting with bi-level spatio-temporal analysis incorporating sky images[J]. *IEEE Transactions on Sustainable Energy*, 2021, 12(3): 1766-1776.
- [16] Yang Dazhi, Gueymard C. Probabilistic post-processing of gridded atmospheric variables and its application to site adaptation of shortwave solar radiation[J]. *Solar Energy*, 2021, 225(1): 427-443.
- [17] Yagli G M, Yang Dazhi, Srinivasan D. Ensemble solar forecasting and post-processing using dropout neural network and information from neighboring satellite pixels[J]. *Renewable and Sustainable Energy Reviews*, 2022 (155): 111909.
- [18] Zhang Ruiyuan, Ma Hui, Hua Wen, et al. Data-driven photovoltaic generation forecasting based on a bayesian network with spatial-temporal correlation analysis[J]. *IEEE Transactions on Industrial Informatics*, 2020, 16(3): 1635-1644.
- [19] Agoua X, Girard R, Kariniotakis G. Probabilistic models for spatio-temporal photovoltaic power forecasting[J]. *IEEE Transactions on Sustainable Energy*, 2019, 10(2): 780-789.
- [20] Bessa R, Trindade A, Miranda V. Spatial-temporal solar power forecasting for smart grids[J]. *IEEE Transactions on Industrial Informatics*, 2015, 11(1): 232-241.
- [21] Qin Jun, Jiang Hou, Lu Ning, et al. Enhancing solar PV output forecast by integrating ground and satellite observations with deep learning[J]. *Renewable and Sustainable Energy Reviews*, 2022 (167): 112680.
- [22] Simeunović J, Schubnel B, Alet P J, et al. Spatio-temporal graph neural networks for multi-site PV power forecasting[J]. *IEEE Transactions on Sustainable Energy*, 2022, 13(2): 1210-1220.
- [23] 王海军, 居蓉蓉, 董颖华. 基于时空关联特征与B-LSTM模型的分布式光伏功率区间预测[J]. *中国电力*, 2024, 57(7): 74-80.
Wang Haijun, Ju Rongrong, Dong Yinghua. Distributed photovoltaic power interval prediction based on spatio-temporal correlation feature and B-LSTM model[J]. *Electric Power*, 2024, 57(7): 74-80.
- [24] 宋绍剑, 李博涵. 基于LSTM网络的光伏发电功率短期预测方法的研究[J]. *可再生能源*, 2021, 39(5): 594-602.
Song Shaojian, Li Bohan. Short-term forecasting method of photovoltaic power based on LSTM[J]. *Renewable Energy Resources*, 2021, 39(5): 594-602.
- [25] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need [EB/OL]. [2026-02-22]. <https://arxiv.org/abs/1706.03762>.
- [26] Zhou Hangxia, Zhang Yujin, Yang Lingfan, et al. Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism[J]. *IEEE Access*, 2019 (7): 78063-78074.
- [27] Shih S Y, Sun Fankeng, Lee H Y. Temporal pattern attention for multivariate time series forecasting[J]. *Machine Learning*, 2019, 108(8): 1421-1441.
- [28] Gers F A, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM[J]. *Neural Computation*, 2000, 12(10): 2451-2471.
- [29] 王瑞, 高强, 逮静. 基于CEEMDAN-LSSVM-ARIMA模型的短期光伏功率预测[J]. *传感器与微系统*, 2022, 41(5): 118-122.
Wang Rui, Gao Qiang, Lu Jing. Short-term photovoltaic power prediction based on CEEMDAN-LSSVM-ARIMA model[J]. *Transducer and Microsystem Technologies*, 2022, 41(5): 118-122.
- [30] 周鑫, 李燕, 曾永辉, 等. 基于SARIMAX-SVR的光伏发电功率预测[J]. *电力系统及其自动化学报*, 2024, 36(5): 1-8.
Zhou Xin, Li Yan, Zeng Yonghui, et al. Forecasting of photovoltaic power generation based on SARIMAX-SVR[J]. *Proceedings of the CSU-EPSA*, 2024, 36(5): 1-8.
- [31] Lin Ziwei, Li Mengting, Zheng Zihan, et al. Self-attention ConvLSTM for spatiotemporal prediction[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*, 2020: 11531-11538.

宿连超(1986—), 男, 高级工程师, 研究方向为基于大数据和人工智能的电力系统应用(通信作者)(E-mail: wssg878@sina.com)。