

基于 JSD-Informer 模型的分布式光伏发电功率短期预测方法

谭晶¹, 马洲俊¹, 蒋承伶¹, 郭金坤², 徐溯², 魏训虎³, 富思³

(1. 国网江苏省电力有限公司, 南京 210000;

2. 国网江苏省电力有限公司 镇江供电分公司, 江苏 镇江 212000;

3. 南京南瑞信息通信科技有限公司, 南京 210000)

摘要: 精确的气象信息是实现光伏发电功率精准预测的基础; 然而, 分布式光伏电站通常布局分散、装机容量小, 且普遍缺乏专用气象站; 基于此, 提出了基于 JSD-Informer 模型的分布式光伏发电功率短期预测方法; 通过分析光伏电站周边的气象站点的数值天气预报、光伏电站的发电功率及其地理信息之间的映射关系; 精准量化多源气象数据的分布差异, 融合多源气象站数值天气预报, 生成分布式光伏电站所在地的气象信息数据; 结合融合的气象信息, 利用詹森-香农散度来优化 Informer 网络的稀疏自注意力机制, 构建分布式光伏发电功率短期预测模型; 通过对比实验分析表明, 所提方法具有较高的预测精度和较好的鲁棒性, 分布式光伏功率预测计算的均方根误差和平均绝对误差分别低至 0.10 和 0.06。

关键词: 气象信息; 光伏发电; 融合; 詹森-香农散度; 自注意力机制; 鲁棒性

Short-Term Prediction Method for Distributed Photovoltaic Power Generation Based on JSD-Informer Model

TAN Jing¹, MA Zhoujun¹, JIANG Chengling¹, GUO Jinkun², XU Su², WEI Xunhu³, FU Si³

(1. State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210000, China;

2. Zhenjiang Power Supply Branch, State Grid Jiangsu Electric Power Co., Ltd., Zhenjiang 212000, China;

3. Nanjing Nanrui Information and Communication Technology Co., Ltd., Nanjing 210000, China)

Abstract: It is a foundation for accurate meteorological information to precisely predict photovoltaic power generation; However, distributed photovoltaic power plants often have the characteristics of decentralized layout, small installed capacity, and generally lacking dedicated meteorological stations; Based on this, a distributed photovoltaic power short-term prediction method using the JSD Informer model is proposed; By analyzing the numerical weather forecast of meteorological stations around the photovoltaic power station, the mapping relationship between the power generation of the photovoltaic power station and its geographic information, the distribution differences of multi-source meteorological data are accurately quantified. The numerical weather forecasts from multi-source meteorological stations are integrated to generate the meteorological information data in distributed photovoltaic power stations; By combining with the fused meteorological information and utilizing the Jensen-Shannon divergence to optimize the sparse self-attention mechanism of the Informer network, a distributed photovoltaic power generation short-term prediction model is constructed; Comparative experimental analysis shows that the proposed method has a high prediction accuracy and good robustness, and the root mean square (RMS) and average absolute errors of the distributed photovoltaic power prediction calculation are as low as 0.10 and 0.06, respectively.

Keywords: meteorological information; photovoltaic power generation; fusion; Jensen-Shannon divergence; self-attention; robustness

收稿日期:2025-07-08; 修回日期:2025-08-20。

作者简介:谭晶(1984-),男,博士,高级工程师。

引用格式:谭晶,马洲俊,蒋承伶,等. 基于 JSD-Informer 模型的分布式光伏发电功率短期预测方法[J]. 计算机测量与控制, 2026,34(6):208-215.

0 引言

当代社会正处于新能源兴起的浪潮之中, 太阳能资源由于可持续、无污染、分布广等特点受到了广泛的关注, 使得以太能为基础的分布式光伏发电技术得到了迅速推广和应用。截至 2024 年底, 我国分布式光伏累计装机量达 370 GW, 占全国光伏总装机量的 42%。2024 年, 我国全年新增分布式光伏装机量 120 GW, 占光伏新增装机总量 43% (国家能源局)。随着光伏装机容量的不断增加, 当大规模光伏电站并网时, 光伏发电具有的间歇性、波动性等特性对电力系统的安全稳定运行造成的隐患日渐显现。因此, 分布式光伏发电功率的精准预测对电网的稳定运行和智能运维变得愈发重要。

在中国, 分布式光伏电站具有数量庞大、分布广泛、装机容量低的特点 (全国超 120 万座, 98% 县域覆盖, 单体中位容量仅 186 kW), 精准的气象信息是实现分布式光伏发电功率预测的基础。气象站是气象信息的主要来源, 但其高昂的安装和运维成本限制在分布式光伏电站的部署, 配备专用气象监测设备的分布式光伏电站占比不足 5% (《中国光伏行业气象监测白皮书 (2023)》)。因此, 开发高效可靠、经济实惠的气象信息计算方法, 对提高分布式光伏电站发电功率预测精度, 实现光伏电站智能化运营, 保障光伏电站安全并网具有重要意义。

根据光伏发电功率预测方法的差异性, 目前常见的功率预测方法可以分为物理方法, 统计方法和深度学习方法^[1]。

物理方法是通过模拟太阳辐照强度的能量传递和光伏组件发电的物理过程实现功率预测, 其预测原理简单, 不需要大量的历史数据, 但在建模过程中需要详细的光伏电站位置信息、光伏组件的特性参数及安装信息等, 针对不同的信息参数需要建立不同的预测模型^[2-4], 推广性较差。

统计方法则无需光伏电站各种元器件的具体物理参数和安装信息, 而是通过分析数值天气预报与历史发电功率数据, 挖掘出数值天气预报与历史发电功率数据内在的规律和映射关系, 进而预测下一时段光伏电站的发电功率^[5-6]。该类方法需要长期的数据积累, 对异常值敏感, 难以拟合复杂的非线性关系, 尤其在复杂多变的天气下, 预测精度较低。

不同于传统的统计方法使用数据分析来建立数据特征映射关系, 深度学习是一种基于样本数据驱动的方法, 具有强大的数据关系拟合能力, 它可以利用自身的学习能力建立多元非线性映射关系, 充分发掘数据之间的潜在联系和数据分布特性, 在光伏发电功率预测中具

有较强的应用潜力。文献 [7] 采用长短期记忆力网络 (LSTM, long short-term memory) 模型对光伏电站进行了光伏发电功率超短期预测, 相比于基于统计的预测方法, 预测精度得到了较大的提升。文献 [8-10] 从不同的方向对 LSTM 模型进行了优化, 并用于光伏电站发电功率预测, 由于 LSTM 模型在捕获时间序列依赖关系的能力较弱, 用于光伏发电功率预测存在明显的缺陷。文献 [11] 采用 Transformer 模型进行光伏功率预测, 文献 [12] 使用时序卷积网络和 Transformer 模型相结合的方法, 提取数据的空间特征和时序特征进行光伏功率预测, 文献 [13] 从相似时段的角度出发, 结合 Transformer 模型进行功率预测。相比基于 LSTM 模型的光伏功率预测, Transformer 模型采用注意力机制能够很好地捕获时间序列中的长期依赖关系, 但其复杂度高, 内存占用大, 运行速度慢, 硬件资源消耗大等缺陷一直是困扰着工程应用的难题。文献 [14-15] 提出使用 Informer 模型来实现中期的光伏发电功率预测, 相比较 Transformer 模型而言, Informer 模型使用概率自注意力机制, 大大降低了模型复杂度和内存占用, 是一种高效的长时间序列预测模型。

上述研究成果很好地证明了深度学习方法在光伏发电功率预测任务上的有效性, 然而, 以上各种方法均依赖于精确的气象信息。而分布式光伏电站具有分布广、数量多、装机容量小等特点, 对于大多数分布式光伏电站而言均缺乏专用的气象站, 获得精准的气象信息较为困难, 这对分布式光伏发电功率预测精度的提升带来了挑战。基于此, 本文提出了基于 JSD-Informer 模型的分布式光伏发电功率短期预测方法, 该方法首先精准量化多源气象数据的分布差异, 融合多源气象站数值天气预报, 生成分布式光伏电站所在地的气象信息数据; 结合融合的气象信息, 利用詹森-香农散度 (JSD, Jensen-Shannon divergence) 来优化 Informer 网络的稀疏自注意力机制, 构建分布式光伏发电功率短期预测模型; 增强模型的稳定性和鲁棒性, 提高模型的预测性能。主要工作分为以下 3 个部分:

1) 为解决由计量偏差、人为误操作及环境扰动导致的数据噪声与缺失问题, 本文采用孤立森林 (Isolation Forest) 算法进行异常值识别, 结合 K 近邻填补实现数据修复。

2) 分析在同一气象站观测条件下, 区域差异性对气象信息与分布式光伏电站输出功率关联关系的影响。以分布式光伏电站为目标点, 分析融合周围不同气象站点的气象信息, 提出基于地理位置信息计算光伏电站所在地的气象信息方法。

3) 提出 JSD-Informer 模型来实现光伏发电功率的预测, 使用 JSD 的对称性优化 Informer 的稀疏自注意

力机制,增强模型的稳定性和鲁棒性,提高模型的预测精度。

1 光伏发电输出功率影响因素分析

1.1 数据异常值处理

数据噪声是指在实际数据采集、传输和存储过程中,由于各种不可控因素导致的数据失真现象。这些噪声可能来源于设备精度、人为操作失误、环境干扰等多个方面。数据噪声的存在会破坏数据的真实性和完整性,对数据进行预处理,减少噪声数据影响,是数据挖掘和数学建模的重要步骤之一。

本文使用孤立森林算法^[16]发现异常值,使用 K 近邻填补法填补缺失值和异常值,确保数据的准确性和可靠性。孤立森林算法的核心优势在于其通过随机分割策略高效识别异常点。该算法具有近线性的时间复杂度和低内存占用特性,能快速处理海量数据;其完全无监督机制不依赖先验标签或数据分布假设,可自适应捕捉复杂数据模式;针对高维数据,算法通过随机特征选择策略,有效规避维度灾难,并对冗余特征保持强鲁棒性。这些特性使其尤为适配光伏数据噪声处理场景,孤立森林算法异常值的判断主要分为两个步骤:生成孤立树和计算异常分数。

孤立森林算法生成孤立树具体步骤如下:

- 1) 从光伏发电输出功率训练数据中随机选择 n 个样本点作为样本子集,放入树的根节点;
- 2) 随机指定一个样本点作为新节点,在当前节点数据中随机产生一个切割点 P ;
- 3) 以此切割点生成一个超平面,然后将当前节点数据空间分为 2 个子空间:把指定样本点小于 P 的数据放在当前节点的左子节点,把大于等于 P 的数据放在当前节点的右子节点;
- 4) 循环执行 2 和 3),不断构造新节点,在无法开展数据划分或满足终止条件时,停止行为;
- 5) 循环执行 1)~4),生成 t 个孤立树 $iTree$ 。

孤立森林算法计算孤立树异常分数 s 的计算公式为:

$$s(x, n) = 2^{-E[h(x)]/c(n)} \quad (1)$$

式中, $h(x)$ 为样本 x 的平均路径长度, $E[h(x)]$ 为 $h(x)$ 的期望值, $c(n)$ 为平均路径长度的归一化因子,如式 (2)~(4) 所示:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n} \quad (2)$$

$$H(i) = \ln(i) + 0.5772156649 \quad (3)$$

$$h(x) = e + c(n) \quad (4)$$

式中, e 为样本 x 从树的根节点到叶子节点过程中经历的边的个数, n 表示和样本 x 在同一个叶子节点的样本个数。

基于孤立森林算法的异常评分机制,样本的异常概率随评分 s 增大而增大。根据实验分析,本文设定阈值 $s=0.6$ 作为异常判定边界:

1) 当 $s > 0.6$ 时,判定样本为异常点,执行异常剔除操作;

2) 当 $s \leq 0.6$ 时,保留样本作为有效观测值。

对于被剔除的异常点,采用时间窗口约束下的 K 近邻填补法进行数据重建,确保时间序列的连续性。

1.2 数据归一化处理

影响光伏发电功率的特征因素较多,例如,辐照度、温度、功率等,它们具有不同物理量纲的特征。这些不同的量纲会对数据特征分析造成不利的影响,例如,高幅值参数(辐照度 $1\,000\text{ W/m}^2$)掩盖低幅值关键信号($10\text{ }^\circ\text{C}$)。因此,需要对实验数据进行归一化处理^[17]。

本文选择使用 Min-Max 方法对数据进行归一化处理。Min-Max 方法核心作用在于通过消除不同特征的物理量纲差异,显著提升机器学习模型的性能与鲁棒性。具体体现为:

- 1) 将不同物理量纲的特征(如辐照度、温度、功率等)映射到统一数值范围,避免模型因特征尺度差异而丢失有效信息;
- 2) 加速模型收敛,降低模型训练迭代次数,使损失函数优化轨迹更稳定;
- 3) 适配多源异构数据融合需求,为光伏发电功率预测模型提供标准化输入;

归一化公式如式 (5) 所示:

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

式中, $\bar{x}$ 为归一化后的值, x 为原始数据值, $x_{\min}$ 为原始数据集的最小值, $x_{\max}$ 为原始数据集的最大值。

1.3 影响因素分析

光伏发电功率的大小很大程度上取决于气象数据的变化。常见的气象数据有温度、相对湿度、降雨量、风向、太阳总水平辐射、水平散射辐射等。本文使用 Spearman 相关系数^[18]度量光伏发电功率和气象数据之间的相关关系。Spearman 系数是一种秩相关系数,它不要求数据服从正态分布,对数据噪声不敏感,既可以度量线性关系,也可以度量非线性关系,相比于皮尔逊相关系数,适用范围更广泛。

设观测数据集包含 n 个独立样本对 $(X_i, Y_i), i = 1, 2, \dots, n$ 。Spearman 数学表达式如公式 (6) 所示:

$$r = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (6)$$

式中, d 为观测数据集 $(X_i, Y_i), i = 1, 2, \dots, n$ 的秩次差

集合, 其计算过程如下步骤所示:

- 1) 将观测值 $X_1, X_2, \dots, X_n$ 按照升序排列, 其秩次 R_i 表示: 升序排列中 X_i 的位置序号;
- 2) 将观测值 $Y_1, Y_2, \dots, Y_n$ 按照升序排列, 其秩次 S_i 表示: 升序排列中 Y_i 的位置序号;
- 3) 对每个样本点 i , 计算 X 与 Y 秩次的差值: $d_i = R_i - S_i$, 秩次差集合定义为所有 d_i 的集合: $d = \{d_1, d_2, \dots, d_n\}$ 。

Spearman 相关系数的取值范围为 $[-1, 1]$, 绝对值越大, 相关性越强。当 $r > 0$ 时, 两个变量集合正相关, 反之, 两个变量集合负相关。

为准确反映气象因素与不同地理位置分布式光伏电站的相关性, 本文选取华东地区某一气象站 Q 下的 A、B、C 三个分布式光伏电站, 其中 A 光伏电站距离 Q 气象站直线距离为 5.3 km, B 光伏电站距离 Q 气象站直线距离为 37.1 km, C 光伏电站距离 Q 气象站直线距离为 85.3 km。气象站与光伏电站的地理位置分布如图 1 所示。

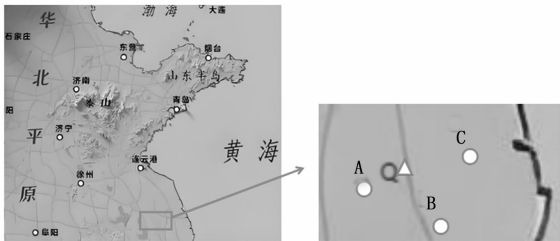


图 1 气象站和分布式光伏电站分布示意图

基于公式 (6) 计算图 1 中 A、B、C 三个分布式光伏电站输出功率与 Q 气象站监测的气象数据的 Spearman 相关系数如表 1 所示。

表 1 气象因素与光伏发电功率的 Spearman 相关系数

气象因素	A 光伏电站	B 光伏电站	C 光伏电站
温度	0.429	0.321	0.421
湿度	0.522	0.502	0.431
太阳总辐射	0.977	0.873	0.542
太阳直接辐射	0.833	0.683	0.432
风速	0.092	0.121	0.091
云量	0.446	0.391	0.562

由表 1 可以看出, A 光伏电站发电功率与太阳直接辐射、太阳总辐射呈高度正相关, 与温度、云量呈正相关, 而与湿度、风速不相关。而距离气象站较远的 C 光伏电站发电功率与太阳直接辐射、太阳总辐射所呈现相关性与 A 站点相比, 明显变弱, 这是气象站距离 C 光伏电站距离较远, 监测到的气象信息不准确造成的。基于此, 本文选择太阳直接辐射、太阳总辐射、温度、云量作为影响光伏发电的主要因素。

2 光伏发电功率预测模型

2.1 多源气象信息融合计算

由于气象监测站设备庞大、安装复杂、维护成本高, 难以有效覆盖分布式光伏电站的气象数据监测, 导致光伏电站气象信息不准确。本文使用地理位置信息数据融合多气象站数据计算分布式光伏站点数据的方法, 通过空间插值技术生成光伏站点气象数据。具体而言, 以气象站与光伏站点间直线距离的倒数作为加权重, 量化地理空间的关联性并驱动气象数据传递^[19]。

设区域内有 m 个气象观测站 (位置集 $S = \{s_1, s_2, \dots, s_m\}$) 和 n 个分布式光伏站点 (位置集 $P = \{p_1, p_2, \dots, p_n\}$), 则构建映射函数为:

$$f: R^m \times R^2 \rightarrow R, \hat{v}_k = f(v_1, v_2, \dots, v_m; x_k) \quad (7)$$

式中, v_j 为气象站 s_j 的观测值 (如辐照度、温度), x_k 为光伏站点 p_k 的地理坐标 (经度和纬度), $\hat{v}_k$ 为 p_k 光伏电站点的气象信息估计值。

为方便计算, 假设西经为负值, 东经为正值, 北纬为负值, 南纬为正值, 则纬度的取值范围可以表示为 $[-\frac{\pi}{2}, \frac{\pi}{2}]$, 经度的取值范围表示为 $[-\pi, \pi]$ 。对于地球表面上任意一点 P , 设其纬度和经度分别为 α 和 β , 则将该点转化到三维空间直角坐标系中, 如公式 (8) 所示:

$$\begin{cases} x = r \cos \alpha \cos \beta \\ y = r \cos \alpha \sin \beta \\ z = r \sin \alpha \end{cases} \quad (8)$$

式中, r 为地球的半径。

如图 2 所示, 设地心点为 O , 地球表面上 A 点的纬度和经度分别为 α_1 和 β_1 , B 点的纬度和经度分别为 α_2 和 β_2 , 在空间直角坐标系中 A 点和 B 点可以表示为 (x_1, y_1, z_1) 和 (x_2, y_2, z_2) , 则有 $\angle AOB$ 的余弦值为:

$$\cos \angle AOB = \frac{x_1 x_2 + y_1 y_2 + z_1 z_2}{r^2} = \sin \alpha_1 \sin \alpha_2 + \cos \alpha_1 \cos \alpha_2 \cos (\beta_2 - \beta_1) \quad (9)$$

根据公式 (9), 地球表面两点之间的距离可以表示为:

$$d = r \arccos [\sin \alpha_1 \sin \alpha_2 + \cos \alpha_1 \cos \alpha_2 \cos (\beta_2 - \beta_1)] \quad (10)$$

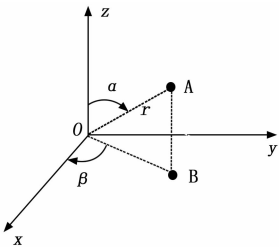


图 2 球面三维坐标示意图

取分布式光伏电站与气象站直线距离的倒数为气象数据加权重,即距离光伏电站越远的气象监测站,监测到的气象数据对光伏发电功率预测精度的影响越小;距离光伏电站越近的气象监测站,监测到的气象数据对光伏发电功率预测精度的影响越大。基于以上描述,定义多气象信息融合计算方式如公式(11)所示:

$$\hat{v}_k = \frac{\sum_{j=1}^m w_{kj} v_j}{\sum_{j=1}^m w_{kj}} \quad (11)$$

$$w_{kj} = \frac{1}{d(s_j, p_k)^\beta} \quad (12)$$

式中, β 为距离衰减因子,经实测标定取 $\beta=1.8$,引入距离阈值上限 $d_{\max}=100$ km 和下限 $d_{\min}=10$ km,忽略 $d>d_{\max}$ 的站点以避免噪声,对于光伏站点 p_k 在距离下限 $d_{\min}$ 范围内存在多个气象站的情况,取该范围内的气象数据均值作为该空间领域内的聚合气象站点。

2.2 Informer 模型原理

Informer 模型^[20]通过自回归转换机制、编码器—解码器结构、自注意力机制和跨时间交互技术等,实现了对长期依赖数据的处理。其模型结构如图3所示。

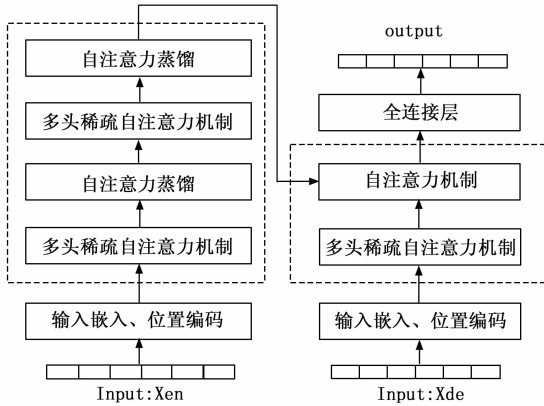


图3 Informer 模型结构

如图3所示,Informer 模型采用了一种经典的编码器—解码器结构。编码器和解码器部分都是由多层自注意力机制、全连接层和正则化层组成。编码器的任务是对输入的时间序列数据进行编码,提取数据特征。解码器的任务是根据编码器的输出,预测下一个时间段的数据,这种结构使模型能够同时处理序列中的长期和短期依赖关系。但这种结构对于光伏发电功率中期预测任务来说,未能考虑数据噪声和天气突变引起的训练参数抖动和梯度爆炸等问题。因此,本文提出在 Informer 模型的基础上引入 JSD-ProbSparse 自注意力机制,提升模型的稳定性和鲁棒性。

2.3 JSD-ProbSparse 自注意力机制

ProbSparse 自注意力机制^[20]基于三元组输入,即查询矩阵 (Q , Query)、键矩阵 (K , Key) 和值矩阵

(V , Value), 其计算定义如下:

$$A(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \quad (13)$$

式中, $\text{softmax}(\cdot)$ 为激活函数, d 为 K 的维度。

QK^T 是为了计算查询矩阵 Q 中每个行向量 q 与键矩阵 K 中每个行向量 k 之间的相关性,从而获得每个查询向量 q 与每个键向量 k 之间的注意力系数。

原始的 Informer 模型使用 KL 散度作为度量来自适应的选择具有代表性的查询。如式(14)所示:

$$S_{KL}(q_i, K) = \ln L - H(P_i) \quad (14)$$

其中: $H(P_i) = -\sum_j p_i(j) \log p_i(j)$ 为香农熵:

$$P_i = \text{softmax}\left(\frac{q_i K^T}{\sqrt{d_k}}\right) \quad (15)$$

这一方式存在两点局限性:

- 1) KL 散度非对称性导致分布差异度量偏差;
- 2) 边界敏感性在长序列中加剧信息损失。

JSD 是概率分布之间的一种相似性度量,在衡量两个分布之间的相似性和区分度时,它对称且有界。给定两个序列 M 和 N ,则 JSD 的数学表达式^[21],如公式(16)~(19)所示:

$$JS(M\|N) = \frac{1}{2}KL(M\|W) + \frac{1}{2}KL(N\|W) \quad (16)$$

$$W = \frac{1}{2}(M+N) \quad (17)$$

$$KL(M\|W) = \sum_x M(x) \log \frac{M(x)}{W(x)} \quad (18)$$

$$KL(N\|W) = \sum_x N(x) \log \frac{N(x)}{W(x)} \quad (19)$$

在序列长度为 L 的注意力机制中,均分分布 U 定义为一个离散均分分布:

$$U = \left(\frac{1}{L}, \frac{1}{L}, \dots, \frac{1}{L}\right) \quad (20)$$

它表示一个无信息的参考分布,即所有位置被平等对待的基准。

基于以上描述,对于每一个查询 q_i ,定义 JSD 稀疏性度量为:

$$S_{JS}(q_i, K) = JS(P_i\|U_L) \quad (21)$$

将公式(21)展开,可得:

$$\begin{aligned} S_{JS} &= \frac{1}{2}[KL(P_i\|W) + KL(U_L\|M)] = \\ &= \frac{1}{2}\left[\sum_j P_i(j) \ln \frac{P_i(j)}{M(j)} + \sum_j U_L(j) \ln \frac{U_L(j)}{M(j)}\right] = \\ &= \frac{1}{2}\left[\sum_j P_i(j) \ln \frac{2P_i(j)}{P_i(j) + U_L(j)} + \sum_j \frac{1}{L} \ln \frac{2U_L(j)}{P_i(j) + U_L(j)}\right] = \\ &= \frac{1}{2}\left[\sum_j P_i(j) + \ln \frac{2P_i(j)}{P_i(j) + 1/L} + \frac{1}{L} \sum_j \ln \frac{2/L}{P_i(j) + 1/L}\right] \end{aligned} \quad (22)$$

2.4 JSD 性能分析

2.4.1 稳定性

当查询向量 $\mathbf{P}$ 接近均分分布 U 时, KL 散度的梯度发散, 如公式 (23) 所示, JS 散度的梯度趋于 0^[22], 如公式 (24) 所示:

$$\nabla KL = (P_i \| U_L) \propto \ln(P_i) \rightarrow \infty \tag{23}$$

$$\lim_{P_i \rightarrow U_i} \nabla S_{JS} \tag{24}$$

JSD 的零梯度特性有效避免训练过程中的梯度爆炸, KL 散度在边界区域梯度会趋于无穷大, 导致训练不稳定。这一特性在光伏发电功率预测任务中具有重要的工程意义: 在概率边界区域 (如晴空天气突现云层), JSD 避免了因微小概率值导致的数值不稳定, 保障训练收敛性。

2.4.2 鲁棒性

数据采集过程中, 噪声的存在是不可避免的。这一现象会带来伪概率偏移现象:

$$P_i^{noise} = P_i + \epsilon, \epsilon \in (0, \sigma^2) \tag{25}$$

当 P 接近均分分布 U 时, JSD 的抗干扰能力:

$$E_{\epsilon} [\| \nabla S_{JS}(P_i^{noise}) \|] \approx 0 \tag{26}$$

KL 散度的敏感性:

$$E_{\epsilon} [\| S_{KL}(P_i^{noise}) \|] \propto \sigma \tag{27}$$

结合式 (26) 和 (27) 可以看出, JSD 对噪声数据不敏感, JS 散度相当于给损失函数加了正则项, 减少对微小扰动的过拟合^[23], 尤其适合光伏预测中气象噪声多的场景, 可以有效提升模型泛化性。

2.5 光伏发电功率短期预测方法流程

JSD-Informer 的光伏发电功率中期预测具体步骤为:

- 1) 使用历史光伏发电功率和对应的气象数据计算斯皮尔曼相关系数, 选择出对光伏发电功率影响比较大的气象数据。
- 2) 根据气象信息融合计算方法, 对以目标光伏电站为中心, 以 r 为半径的区域内, 所有气象站采集数据进行加权计算融合, 得到新的气象信息作为该光伏电站的气象信息。
- 3) 对数据集去除噪声及归一化处理, 按 7 : 2 : 1 的比例, 划分训练集、验证集和测试集。
- 4) 初始化模型参数, 包括损失函数、迭代次数、学习率、预测时间长度等参数。
- 5) 使用训练集分别训练对照模型, 保存最优的模型参数。
- 6) 将测试集数据输入到训练好的预测模型中进行预测, 并进行结果分析。

3 算例分析

3.1 实验数据

本研究采用分层随机抽样策略, 选取华中地区 (湖

北省、湖南省、河南省) 150 个分布式光伏电站和 42 个专业气象站作为研究对象。数据采集时间跨度为 2020 年 1 月 1 日至 2025 年 3 月 31 日, 采样时间分辨率为 5 分钟级高精度采样 (限定于有效日照时间段 07:00—18:00)。

本研究分层随机抽样选取的 150 个分布式光伏电站装机容量、地理分布、组件材料统计情况如表 2~4 所示。

表 2 光伏电站装机容量表

装机容量/kW	站数	占总量的百分比/%
<5	32	21.3
5~20	67	44.7
>20	51	34.0

在抽样的 150 个分布式光伏电站中, 装机容量小于 5 kW 的电站占比 21.3%, 装机容量在 5~20 kW 的电站占比 44.7%, 装机容量大于 20 kW 的电站占比 34.0%。样本覆盖分布式光伏典型容量谱系 (户用小微→工商业中型→农光互补大型), 满足《分布式光伏资源评估技术规范》(QX/T 556—2020) 的抽样要求。

表 3 光伏电站地理分布表

地理分布	站数	占总量的百分比/%
湖北	68	45.3
湖南	45	30.0
河南	37	24.7

样本电站覆盖经度为 112.3~115.8°E, 纬度为 27.4~31.6°N, 经度标准差为 0.82°, 纬度标准差为 0.76°, 与华中地区总体分布 (经度标准差为 0.85°, 纬度标准差为 0.79°) 一致。

表 4 光伏电站组件材料统计表

组件材料	站数	占总量的百分比/%
多晶硅	87	58.3
单晶硅	48	32.0
薄膜	15	10.7

本研究选取的 150 个电站, 其中多晶硅光伏组件电站占比 58.3%, 符合 2018—2022 年华中地区装机高峰期特征, 单晶硅光伏组件电站占比 32.0%, 符合 2023 年后技术升级浪潮的安装组件趋势。

本研究采用的 42 个专业气象站点在华中地区构成多层次监测网络, 典型地形分布统计表如表 5 所示。

表 5 气象站典型地形分布表

地形类型	站数	占总量的百分比/%
平原	18	42.9
丘陵	16	38.1
山地	8	19.0

42 个气象站分布于平原（占比 42.9%）、丘陵（38.1%）和山地（19.0%），基本覆盖华中地区的地形情况。

3.2 实验环境和评价指标

本实验所用训练设备系统为 Windows11。相关硬件设备型号如表 6 所示。

表 6 硬件配置统表		
组件	型号/参数	数量
CPU	Intel U9 285k 24 核 主频 3.2 G	1
GPU	RTX4080s	2 块(共 32 G)
主板	Intel Z890 DDR5	1
内存	DDR5	2 根(共 32 G)
存储	SSD/SAMSUNG. 870EVO	1T

本实验所用主要硬件设备如表 6 所示，主要包括 RTX4080s 的独立显卡两块，共 32 G；24 核 CPU 一个；内存条两根共 32 G；SSD 硬盘 1 T。

为评估本文所采用的模型预测结果的精确度，采用均方根误差（RMSE）、平均绝对误差（MAE）和回归系数（ R^2 ）3 种评价指标对模型进行评估。公式如（28）~（30）所示：

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \bar{y}_i)^2}$$

(28)

$$MAE = \frac{1}{m} \sum_{i=1}^m [y_i - \bar{y}_i]$$

(29)

$$R^2 = \frac{\sum_{i=1}^n \left(\bar{y}_i - \frac{1}{n} \sum_{i=1}^n y_i \right)^2}{\sum_{i=1}^n \left(y_i - \frac{1}{n} \sum_{i=1}^n y_i \right)^2}$$

(30)

式中， y_i 为光伏发电功率的真实值， $\bar{y}_i$ 为光伏发电功率的预测值。

3.3 实验结果分析

为充分验证所提方法在光伏发电功率短期预测任务中的有效性，搭建了 LSTM、Transformer、Informer 及 JSD-informer 模型进行实验对照。将采集的数据分为训练集、验证集和测试集 3 个部分，其中训练集 191 730 条，验证集 54 780 条，测试集 27 390 条。

表 7 为训练过程中 4 种模型的收敛时间对比统计表。

表 7 训练效率对比表			
模型名称	收敛时 epoch/次	每个 epoch 耗时的均值/s	收敛总耗时/s
LSTM	126	36	4 536
Transformer	102	40	4 080
Informer	86	28	2 408
JSD-Informer	71	30	2 130

从表 7 可知，在训练效率方面，JSD-Informer 训练效率表现最佳，这得益于其较快的收敛速度和较高的计

算效率。Informer 在单次迭代上效率最高，但由于收敛速度略慢于 JSD-Informer，所以总耗时稍高。Informer 因其概率注意机制，保持单轮 28 s 的计算效率优势。Transformer 和 LSTM 在训练效率上明显落后。JSD-Informer 通过算法层（JSD 对称性加速收敛）和架构层（概率注意机制降低单轮耗时）实现综合训练效率最优。

为有效验证模型在不同天气状况下的效果，本实验分别对晴天、雨天及多云 3 种常见的天气状况进行光伏发电功率预测实验。3 种天气下预测的功率曲线如图 4~6 所示。

图 4 为 2025 年 3 月 16 日样本验证集中 4 个模型晴天功率预测曲线图，由图中可以看出，LSTM 模型的功率预测曲线和 Transformer 模型的功率预测曲线，和真实值相比差别较大，而 Informer 模型的功率预测曲线和 JSD-Informer 模型的功率预测的曲线较接近真实值。

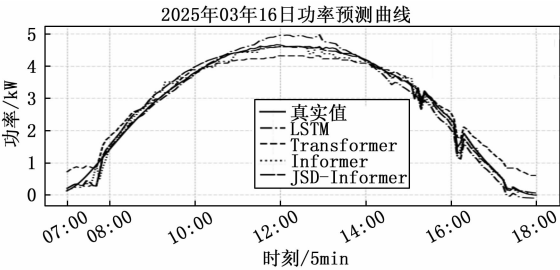


图 4 晴天功率预测曲线

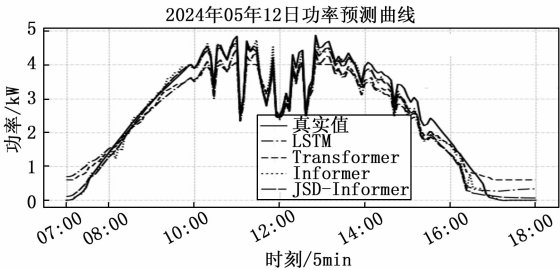


图 5 多云天气功率预测曲线

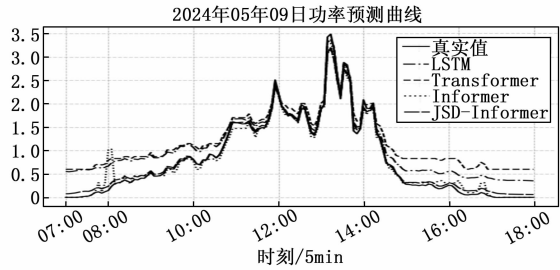


图 6 雨天功率预测曲线

图 5 为 2024 年 5 月 12 日样本验证集中 4 个模型多云功率预测曲线图。图 6 为 2024 年 5 月 9 日样本验证集中 4 个模型雨天功率预测曲线图。从图中可以看出，雨天状态下，LSTM 和 Transformer 预测的功率明显偏离功率真实值。

从表 8 中可以看出，LSTM 作为基础时序模型，取

得了 $RMSE=0.22$ 、 $MAE=0.12$ 和 $R^2=0.95$ 的预测性能。Transformer 模型通过自注意力机制有效捕捉长距离依赖, 将 $RMSE$ 降低至 0.19, MAE 降低至 0.10, R^2 提升至 0.958。而 Informer 模型针对长序列预测优化, 进一步将 $RMSE$ 降至 0.15, MAE 降至 0.08, R^2 提升至 0.971。最终, JSD-Informer 通过引入 JSD 优化注意力机制, 取得了最佳性能, $RMSE$ 低至 0.10, MAE 为 0.06, R^2 高达 0.976。实验结果表明, 模型架构的改进, 尤其是注意力机制的优化, 对提升光伏功率预测精度具有显著效果。

表 8 功率预测评价指标对比

模型名称	$RMSE$	MAE	R^2
LSTM	0.22	0.12	0.95
Transformer	0.19	0.10	0.958
Informer	0.15	0.08	0.971
JSD-Informer	0.10	0.06	0.976

4 结束语

本文提出的 JSD-Informer 框架在气象数据融合与 JSD 优化注意力机制上构建光伏发电功率预测模型, 实验证明对提升光伏功率预测精度具有显著效果, 其已被应用到华中部分地区智能电网调度、电站精益运营等领域, 将为新型电力系统建设提供核心算法支撑。但不可忽视的是, 模型还需在极端场景覆盖 (例如: 雾霾、沙尘暴)、模型轻量化及跨区泛化 (例如: 西北的高辐照区) 等方向持续优化, 为适应更多不同工况下的应用提供支撑。

参考文献:

[1] 丁 明, 王伟胜. 光伏发电系统功率预测物理方法研究 [J]. 中国电机工程学报, 2012, 32 (14): 1-8.

[2] SAID Y, BENDAOU R, ABIDI A E, et al. Evolution of the physical parameters of photovoltaic generators as a function of temperature and irradiance: new method of prediction based on the manufacturer's datasheet [J]. Energy Conversion and Management, 2020, 203: 112141.

[3] 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.

[4] 丁 坤, 刘增泉, 张经纬, 等. 基于图像奇异值分解的局部遮挡光伏阵列输出特性建模研究 [J]. 综合智慧能源, 2023, 45 (2): 53-60.

[5] 李燕斌, 张久菊, 肖俊明. 灰色-马尔可夫链统计组合模型在光伏发电功率短期预测中的应用 [J]. 电测与仪表, 2015, 52 (23): 111-116.

[6] 张雪松, 李 鹏, 周亦尧. 基于贝叶斯概率的光伏出力组合预测方法 [J]. 太阳能学报, 2021, 42 (10): 80-86.

[7] 周浩杰, 杨建卫, 王 尊, 等. 基于 LSTM 光伏发电功率超短期预测模型研究 [J]. 电源技术, 2023, 47 (6): 785-789.

[8] 赵佳蕊, 王玲芝, 李晨阳. 基于 LEA-LSTM 的光伏发电功率短期预测方法 [J]. 电力信息与通信技术, 2024, 22 (11): 34-42.

[9] 袁建华, 谈 顺, 刘 闯. 基于改进灰狼算法优化 LSTM 的光伏发电功率短期预测 [J]. 电力学报, 2024, 39 (2): 111-118.

[10] 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.

[11] 黄 莉, 甘恒玉, 刘兴举, 等. 基于 Transformer 编码器的超短期光伏发电功率预测 [J]. 智慧电力, 2024, 52 (5): 16-22.

[12] 盛瑞祥, 张啸宇. 基于概率 TCN-Transformer 的短期光伏功率预测模型 [J]. 综合智慧能源, 2024, 46 (11): 10-18.

[13] 杨鹏伟, 赵丽萍, 陈军法, 等. 基于相似时段匹配与 Transformer 网络建模的分布式光伏超短期功率预测方法 [J]. 中国电力, 2024, 57 (12): 60-70.

[14] 杨 晨. 基于 Informer 模型的区域光伏发电功率预测 [D]. 湘潭: 湘潭大学, 2023.

[15] 李之恒. 基于 Informer 模型的中期光伏发电量预测研究 [D]. 南京: 南京信息工程大学, 2024.

[16] 李新鹏, 高 欣, 阎 博, 等. 基于孤立森林算法的电力调度流数据异常检测方法 [J]. 电网技术, 2019, 43 (4): 1447-1456.

[17] 戴长春, 郑晓亮, 杨 铖, 等. 基于改进 DBO-BiLSTM-GRU 的光伏发电功率预测方法研究 [J]. 自动化应用, 2024, 65 (19): 89-93.

[18] 赵宇庆, 腾志军. 基于 Spearman-IPSO-LSSVM 的短期电力负荷预测方法研究 [J]. 电气自动化, 2025, 47 (1): 102-104.

[19] TOBLER W R. A computer movie simulating urban growth in the detroit region [J]. Economic Geography, 1970, 46 (2): 234-240.

[20] ZHOU H, ZHANG S, PENG J, et al. Informer: beyond efficient transformer for long sequence time-series forecasting [J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2022: 11106-11115.

[21] THIAGARAJAN P, GHOSH S. Jensen-Shannon divergence based novel loss functions for Bayesian neural networks [J]. Neurocomputing, 2025, 618: 129115.

[22] KOLCHINSKY A, TRUPECHIO M, KIM E, et al. Nonlinear information bottleneck [J]. Entropy, 2019, 21 (12): 1181.

[23] ZHANG H, ZHANG Y, LIU Y, et al. Jensen-Shannon divergence regularization for robust deep learning [J]. Baltimore: PMLR, 2022: 23047-23061.