

DOI: 10. 3969/j. issn. 2097-0706. 2025. 09. 006

基于 TimeGAN 的极端天气光伏功率预测方法

TimeGAN-based photovoltaic power prediction method under extreme weather events

孙师奇, 马刚*, 许文俊, 李豪, 马健

SUN Shiqi, MA Gang*, XU Wenjun, LI Hao, MA Jian

(南京师范大学 电气与自动化工程学院, 南京 210023)

(School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China)

摘要: 极端天气下准确预测光伏发电量对保障能源供应和电网稳定至关重要,但此类天气的突发性导致光伏电站历史数据稀缺,难以有效预测极端天气场景下的光伏功率。针对上述问题,提出一种基于时间序列生成对抗网络(TimeGAN)的少量历史数据扩充预测方法,捕捉光伏功率和天气条件中的复杂时间依赖关系,根据光伏电站已有少量历史数据,生成逼真的时间序列数据,模拟极端天气发生的过程,进而展开光伏功率预测。试验结果显示,相较于采用传统生成对抗网络(GAN)扩充小样本数据,采用TimeGAN扩充小样本后的预测结果有较好的拟合性,数据扩增25%后平均绝对误差(MAE)降低了1.14 MW,均方根误差(RMSE)降低了1.09 MW,数据扩增50%后MAE降低了1.08 MW, RMSE降低了0.99 MW,精确度得到了明显提高。

关键词: TimeGAN; 极端天气; 小样本扩充; 光伏功率预测; 时间序列

中图分类号: TK 51: TM 761 **文献标志码:** A **文章编号:** 2097-0706(2025)09-0051-09

Abstract: Accurate prediction of photovoltaic power generation under extreme weather events is crucial for ensuring energy supply and grid stability. However, the suddenness of such weather events leads to scarce historical data from photovoltaic power stations, making it difficult to effectively predict photovoltaic power under extreme weather conditions. To address this issue, a prediction method based on Time-series Generative Adversarial Networks (TimeGAN) was proposed to augment limited historical data. The method captured the complex temporal dependencies between photovoltaic power and weather conditions. Based on the limited historical data from photovoltaic power stations, the TimeGAN model generated realistic time-series data to simulate the occurrence of extreme weather events, and subsequently conducted photovoltaic power prediction. The experimental results showed that compared to traditional GAN for small sample augmentation, the TimeGAN-augmented prediction results demonstrated better fitting performance. After 25% data augmentation, the Mean Absolute Error (MAE) decreased by 1.14 MW, and the Root Mean Square Error (RMSE) decreased by 1.09 MW. After 50% data augmentation, the MAE decreased by 1.08 MW, and the RMSE decreased by 0.99 MW. These results indicated significant improvements in prediction accuracy.

Keywords: TimeGAN; extreme weather; small sample augmentation; photovoltaic power prediction; time series

0 引言

在全球能源结构加速低碳转型的背景下^[1-2],光伏产业凭借其技术突破与规模化应用,已成为推动可再生能源发展的核心力量^[3]。然而光伏出力具有间歇性和不确定性^[4-5],受天气影响较大,寒潮、雾霾、冰雹、沙尘暴等气象灾害会严重影响新能源电力系统的稳定输出,因此,在极端天气气象条件下

准确预测光伏出力,对提高现有电力系统的稳定性和可靠性至关重要^[6]。

文献[7]聚焦数据驱动方法在太阳辐照度建模领域的应用,通过对比分析其理论框架与技术路线的差异性特征,从技术演进视角系统评述了太阳辐射预测方法的创新进展,揭示了模型泛化能力不足等核心问题并提出面向复杂气象场景的跨模态优化路径;文献[8]构建了低参数依赖的晴空辐射计算新架构,融合动态参数自适应机制,实现弱气象扰动场景下光伏电站短时功率的超前预测;文献[9]提出了一种基于t-分布随机邻域嵌入(t-distributed Stochastic Neighbor Embedding, t-SNE)流

基金项目:江苏省碳达峰碳中和科技创新专项资金重点项目(BE2022003)

Jiangsu Province Carbon Peak and Carbon Neutrality Science and Technology Innovation Special Fund Project(BE2022003)

形学习与快速聚类算法的光伏逆变器故障预测技术,将历史信号作为原始数据提取特征值,利用 t-SNE 进行降维处理,提取特征矩阵,最终实现对故障的准确预测;文献[10]对 2020 年 6 月 21 日中国地区发生的日环食过程进行剖析,采用数值天气预报叠加事件影响的方法,提出了极端天文、天气事件对新能源间歇性与电网运行影响的预测方法。

文献[11]采用图像神经网络方法,解决了以往研究中多光伏场站间光伏出力的时空关联难以进行精准建模导致预测结果精度不够的问题;文献[12]针对天气状况不确定性限制光伏功率预测性能的问题建立混合模型,从多维时间序列数据中提取关键特征,缓解天气变化对光伏功率预测精度的不利影响;文献[13]提出了一种元学习方法,通过动态混合多个模型来学习每个模型在什么条件下表现最好,在不使用数值天气预报的情况下,使用 4 种不同长短期记忆(Long Short-Term Memory, LSTM)神经网络实现逐小时光伏发电量预测。

目前,大部分研究集中于算法的改进而忽视了原始数据量的重要性,由于寒潮等极端天气发生频次低且不可预测,造成光伏电站缺乏足够的极端工况运行数据,影响后续建模分析。为了能获取足够的训练集训练模型,文献[14]提出使用 Wasserstein 生成对抗网络(Wasserstein Generative Adversarial Network, WGAN)模型对光伏电站历史数据进行扩充,再将数据增强后的样本用于训练预测模型。文献[15]采用生成对抗网络(Generative Adversarial Network, GAN)模型对有限的历史负荷数据进行扩充,再将数据增强后的样本数据用于训练模型,使得预测模型对于负荷预测的精度变高。然而,目前几乎没有研究将 GAN 用于生成光伏时间序列数据,因此本文提出时间序列生成对抗网络(Time-series Generative Adversarial Network, TimeGAN)算法来扩充极端天气的小样本数据,解决极端天气的历史训练数据不足而导致的光伏发电功率预测精度较低问题,进而在时间和空间维度上捕捉光伏功率的发电规律,以提高光伏功率预测精度。

1 模型理论基础

1.1 极端天气特性分析

本文研究的江苏某光伏电站所处区域寒潮以强降温、大风及雨雪天气为特征,高发于 3 月、10 月及 11 月。依据中央气象台标准及区域实际情况,寒潮判定需满足:冷空气过境后 24 h 内降温 $\geq 8^\circ\text{C}$ 或 48 h 内 $\geq 10^\circ\text{C}$;过程最低气温 $\leq 6^\circ\text{C}$,两项条件同时

触发即构成寒潮事件^[16]。

云量、相对湿度、环境温度、风速、气压、辐照度等气象因素均会对光伏发电功率造成一定影响^[17],由于这些因素对光伏发电功率的影响程度不同,因此选取合适的环境因素作为预测模型的参数尤为重要。本文对各影响因素进行灰色关联度分析(见表 1),选取典型气象要素,以更好地进行气象分型聚类。

表 1 气象因素灰色关联度结果

Table 1 Grey relational analysis results of meteorological factors

区间	云量	相对湿度	环境温度	风速	气压	辐照度
0.1	0.312 4	0.225 2	0.344 3	0.381 4	0.183 1	0.684 2
0.2	0.363 3	0.258 1	0.471 4	0.502 1	0.208 7	0.736 4
0.3	0.495 2	0.287 4	0.572 3	0.592 1	0.216 4	0.771 9
0.4	0.567 1	0.291 5	0.593 6	0.645 2	0.263 2	0.819 2
0.5	0.600 4	0.345 3	0.645 3	0.697 2	0.298 8	0.832 3
0.6	0.628 4	0.378 9	0.691 4	0.739 3	0.313 1	0.858 2
0.7	0.661 1	0.482 1	0.725 6	0.754 5	0.328 4	0.890 1
0.8	0.699 8	0.531 4	0.748 4	0.788 2	0.357 9	0.901 3
0.9	0.725 3	0.559 0	0.761 2	0.791 7	0.389 1	0.910 4
1.0	0.746 7	0.574 1	0.777 5	0.813 1	0.393 5	0.923 7

从表 1 可以看出,辐照度对光伏发电效率的影响最大,然后依次为风速、环境温度、云量、相对湿度和气压。当环境温度较高时,太阳辐射也较强,因此在常规天气下,影响光伏发电效率的因素中,辐照度起着决定性的作用^[18]。

基于极端天气对光伏功率的影响机理,经过对比研究和分析,本文对寒潮事件的定义为:起始于剧烈降温开始的最高温度,中止于气温回升至低温阈值 6°C 以上。对寒潮天气做进一步拆分,其中雨雪、大风以及强降温为主要特征。

1.2 时间序列生成模型

原始的 GAN 容易产生两种问题:一是在训练过程中生成器出现梯度消失问题;二是生成器优化目标混乱导致梯度不稳定,样本多样性和准确性下降,引发模式崩溃。

对极端天气下收集到的光伏电站气象、功率数据样本集合 $x_{1:T} = \{(W_t, P_t) | W_t, P_t \in \Omega_{\text{ext}}\}$ (其中: W_t 为数值天气预报(Numerical Weather Prediction, NWP)气象预报数据; P_t 为功率; Ω_{ext} 为极端天气条件下整体的样本空间; $t = 1, \dots, T$) 进行扩充处理。样本的时间序列做如下分布 $p(x_t | x_{1:t-1})$ (表示在给定历史数据 $x_{1:t-1}$ 的条件下, t 时刻数据 x_t 的概率分布)。自编码器和 GAN 的联合学习过程如下。

重构损失函数为

$$L_R = E_{x_{i,r-p}} \left[\sum_t \|x_t - \tilde{x}_t\|_2 \right], \quad (1)$$

式中: $E_{x_{i,r-p}}$ 为从概率分布 p 中采样的时间序列 $x_{i,t}$ 的期望值; x_t 为整体时间序列的真实值; \tilde{x}_t 为自编码器学习映射后得到的数值; $\|x_t - \tilde{x}_t\|_2$ 为 x_t 与 \tilde{x}_t 之间的欧氏距离。

监督损失函数为

$$L_S = E_{x_{i,r-p}} \left[\sum_t \|h_t - g_X(h_{t-1}, z_t)\|_2 \right], \quad (2)$$

式中: h_t 为隐藏层向量,映射在潜在空间; g_X 为循环网络内置的生成器函数; z_t 为生成的自由向量。

无监督损失函数为

$$L_U = E_{x_{i,r-p}} \left[\sum_t \log y_t \right] + E_{x_{i,r-\hat{p}}} \left[\sum_t \log(1 - \hat{y}_t) \right], \quad (3)$$

式中: y_t 和 \hat{y}_t 分别为真实样本和生成样本,通过判别器进行判别后得到; \hat{p} 为条件分布,表示时序样本的合成值。

自动优化参数的自编码器过程为

$$\min_{\theta_e, \theta_r} (\lambda L_S + L_R), \quad (4)$$

参数优化GAN的过程为

$$\min_{\theta_g} (\eta L_S + \max_{\theta_d} L_U), \quad (5)$$

式中: θ_e 、 θ_r 、 θ_g 和 θ_d 分别为嵌入、恢复、生成器和判别器中的参数; λ 和 η 分别为平衡自编码器和GAN中损失函数的参数,均大于0。

对抗学习后,判别器 L_S 和生成器 L_U 在训练过程中达到纳什均衡,生成的样本结果更接近真实值。TimeGAN模型结构如图1所示。

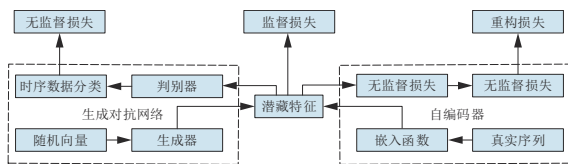


图1 TimeGAN模型结构

Fig. 1 Structure of TimeGAN model

1.3 功率预测模型

LSTM网络具备的3个门控机制有效控制了信息的流动,并在时间序列的长期依赖特性捕捉方面表现出色。由于门控机制允许并行计算,因此大大提高了模型的运行速度和效率^[19]。

数据中的局部特征提取、捕捉主要依靠卷积神经网络(Convolutional Neural Networks, CNN),将LSTM和CNN结合,可以实现对时空信息的综合利用,提高模型的性能。

因此本文构建CNN-LSTM神经网络模型,用于预测光伏场群的功率。应用CNN对时间序列的分

析,提取空间结构中多维时间序列数据之间的关系。通过CNN模块实现数据特征提取,CNN模块部分的输出作为LSTM模块的输入,由LSTM模块实现高精度预测。CNN-LSTM模型结构如图2所示。

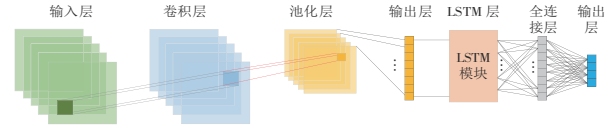


图2 CNN-LSTM模型

Fig. 2 CNN-LSTM model

为了避免陷入局部最优和梯度爆炸,达到优化模型的目的,需要设定合理的学习率和LSTM单元数量,以提高模型的训练速度^[20]。

遗传-粒子群优化(Genetic Algorithm-Particle Swarm Optimization, GA-PSO)算法在复杂多峰参数空间中具有更高的全局收敛效率,将其应用于LSTM网络的超参数优化,可以有效解决人工选择参数造成的模型精度不佳问题,减少人工因素对模型精度的影响。

GA-PSO算法分为GA部分和PSO算法部分,其中PSO算法用于优化GA选择步骤中未被选中的个体。为避免人工选择参数对功率预测结果造成干扰,本文采用GA-PSO算法对学习率、LSTM单元数量的设定进行优化,具体流程如图3所示。

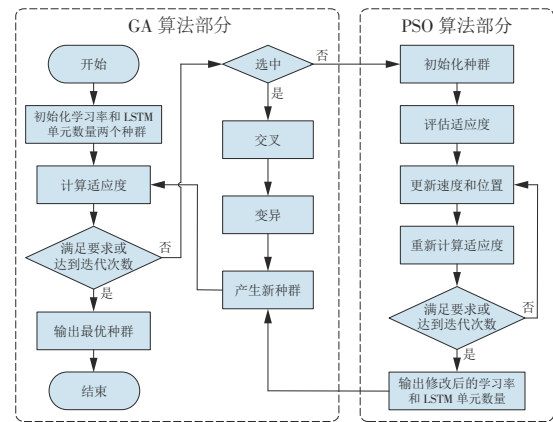


图3 GA-PSO算法优化LSTM模型超参数流程

Fig. 3 Flowchart of hyperparameter optimization for LSTM model using GA-PSO algorithm

模型的计算过程如下。

步骤1:将输入层的数据输入卷积层提取特征值,定义卷积核函数在卷积层中,进而输入数据进行卷积计算,然后通过激活函数并添加偏置值,得到该层的特征图。卷积计算过程为

$$M_i^l = f \left(\sum_{i \in N_l} M_i^{l-1} \otimes w_{i,j}^l + b_j^l \right), \quad (6)$$

式中: M_i^l 、 M_i^{l-1} 分别为第 l 、第 $l-1$ 个卷积层的第 i 个特征图; $w_{i,j}^l$ 为第 $l-1$ 个卷积层第 i 个特征图和第 l 个

卷积层第 j 个特征图间的权重; b_j^l 为第 l 个卷积层的第 j 个特征图偏置; N_i 为输入层的特征图集; \otimes 为卷积操作; $f(\cdot)$ 为激活函数, 此处选择 Relu 函数^[21-22]。

步骤 2: 进入池化层进一步获取隐藏信息, 降低特征维度。池化运算可以降低对输入变形的敏感度, 从而得到具有平移不变性的单元。平均池化是池化区域中所有值都被池化区域内的均值替代, 计算公式为

$$P_j^l = f(\beta_j^l \text{down}(M_i^{l-1}) + b_j^l), \quad (7)$$

式中: P_j^l 为第 l 个池化层的第 j 个特征图; β_j^l 和 b_j^l 分别为特征图的乘性偏置和加性偏置; $\text{down}()$ 为池化函数^[23]。

步骤 3: 数据进入 Flatten(展平层), 通过将卷积层输出的多维特征图转换为一维向量, 完成卷积层到全连接层的平稳过渡。

步骤 4: 将提取完特征的数据送入 LSTM 模块进行预测训练。

步骤 5: 全连接层与 LSTM 隐藏层的神经元之间以全连接方式连接, 作用是把这些特征进行综合, 生成最终的输出, 计算公式为^[24]

$$F_j^l = f\left(\sum_{i=1}^{l-1} w_{i,j}^l N_i^{l-1} + b_j^l\right), \quad (8)$$

式中: F_j^l 为第 l 层的第 j 个神经元。

2 基于 TimeGAN 的寒潮天气下光伏功率预测模型

2.1 预测模型框架

本文构建了寒潮天气下光伏功率预测模型, 如图 4 所示。

步骤 1: 对输入的数据进行预处理。先将 NWP 气象数据和功率数据在时间尺度上统一, 然后使用具有噪声的基于密度的空间聚类应用 (Density-Based Spatial Clustering of Applications with Noise, DBSCAN) 算法识别风光电站数据集中的分散型和堆积型异常数据并将其剔除后做归一化处理, 最后对寒潮天气下的气象和功率数据进行扩充。

步骤 2: 建立和训练预测模型并预测结果。首先, 将每个场站相关性最高的两个因素作为输入变量, 组成多任务学习 (Multi-Task Learning, MTL) 的硬参数共享模式; 其次, 采用 TimeGAN 对原始小样本数据进行扩增, 生成大量逼真数据作为训练集, 由共享层连接 CNN 模块进行特征提取; 然后, 通过 LSTM 部分进行预测; 最后, 输入测试集的数据, 得到预测结果并将其反归一化处理输出对应曲线图。

步骤 3: 误差校验。用主成分分析 (Principle

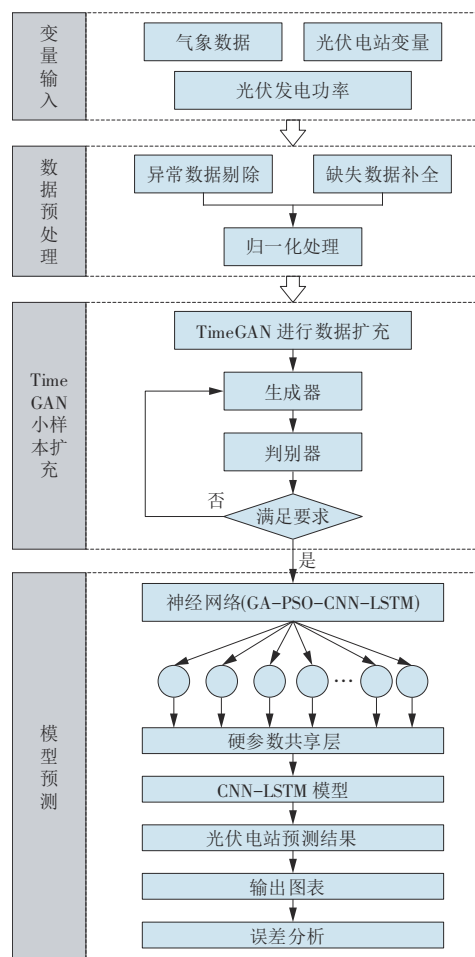


图 4 寒潮天气下光伏功率预测模型

Fig. 4 Photovoltaic power prediction model for cold wave weather

Component Analysis, PCA) 和 t-SNE 校验扩充数据的质量, 用平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Square Error, RMSE)、平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE) 和决定系数 R^2 判定模型的误差, 对结果进行分析, 检验模型在极端天气下光伏功率预测的效果。

2.2 误差校验

模型的准确性通常会通过预测结果和实际值的误差来反映, 本文采用 MAE, RMSE, MAPE 等指标来评定。其中 MAE 适合评估光伏功率在稳定天气下的平均误差, RMSE 适用于评估极端天气事件, MAPE 适合评估小功率时段预测精度的预测风险, 计算公式为

$$\delta_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \quad (9)$$

$$\delta_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (10)$$

$$\delta_{\text{MAPE}} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \quad (11)$$

式中: \hat{y}_i 为预测值; y_i 为实际值; n 为预测样本数。

预测结果和实际值的拟合水平一般可直观地由图表看出,决定系数 R^2 则可以量化拟合度,计算公式为

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2}, \quad (12)$$

式中: \bar{y}_i 为实际值的平均值。

MAE, RMSE 和 MAPE 越小表示误差越小; R^2 取值为 0—1, 数值越接近 1 表示预测曲线和实际曲线拟合度越高。

针对时间序列模型,而非持久性模型, R^2 可能高估模型性能,因此引入准确度 A_{cc} (accuracy)、精度 P_{re} (precision) 和召回率 R_{ec} (recall) 进行综合评价,计算公式为

$$A_{cc} = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{FP} + n_{FN} + n_{TN}}, \quad (13)$$

$$P_{re} = \frac{n_{TP}}{n_{TP} + n_{FP}}, \quad (14)$$

$$R_{ec} = \frac{n_{TP}}{n_{TP} + n_{FN}}, \quad (15)$$

式中: n_{TP} 和 n_{TN} 分别为被正确地分类为正例和负例的样本数量; n_{FP} 和 n_{FN} 分别为被错误地分类为正例和负例的样本数量。

A_{cc} 可以判断分类的正确率。 P_{re} 和 R_{ec} 是有关联的两个指标, R_{ec} 越高, P_{re} 就越低。

3 试验与分析

3.1 算例概况

本文以江苏某地 4 个光伏电站为研究对象,使用这 4 个光伏电站 2022 年 12 月—2023 年 1 月以寒潮、大风、降雨为主要气象因子的 NWP 气象数据和光伏电站输出功率数据。气象数据时间分辨率、光伏输出功率数据时间分辨率均为 15 min。结合中国气象网及当地气象局发出的寒潮预警及数据标准,统计了发生的 4 次寒潮事件,伴有大风 6 次、降雨 2 次,共计 2 000 条样本气象数据。

本文所采用模型的训练与验证均在 Tensorflow 平台进行,具体参数见表 2。

3.2 数据生成结果分析

数据集的数量和质量是影响开发数据驱动预测模型预测结果的最基本因素,本文使用 PCA 和 t-SNE 降维处理生成的数据,检验其分布情况是否合理且具有代表性特征。将 2 000 条样本数据的生成数据与原始数据进行对比,如图 5 所示。

高维数据降维技术中,PCA 与 t-SNE 各具特点,两者的协同应用为评估 TimeGAN 生成数据质量提

表 2 算法参数设置

Table 2 Parameter settings for algorithm

模型	参数	数值
TimeGAN	θ	2
	λ	0.1
	逻辑框架数量	3
	学习率 1	0.001
	学习率 2	0.005
	批量大小	56
CNN-LSTM	学习率	0.009
	隐藏层	134
	迭代次数	100
	批量大小	128
	梯度阈值	1

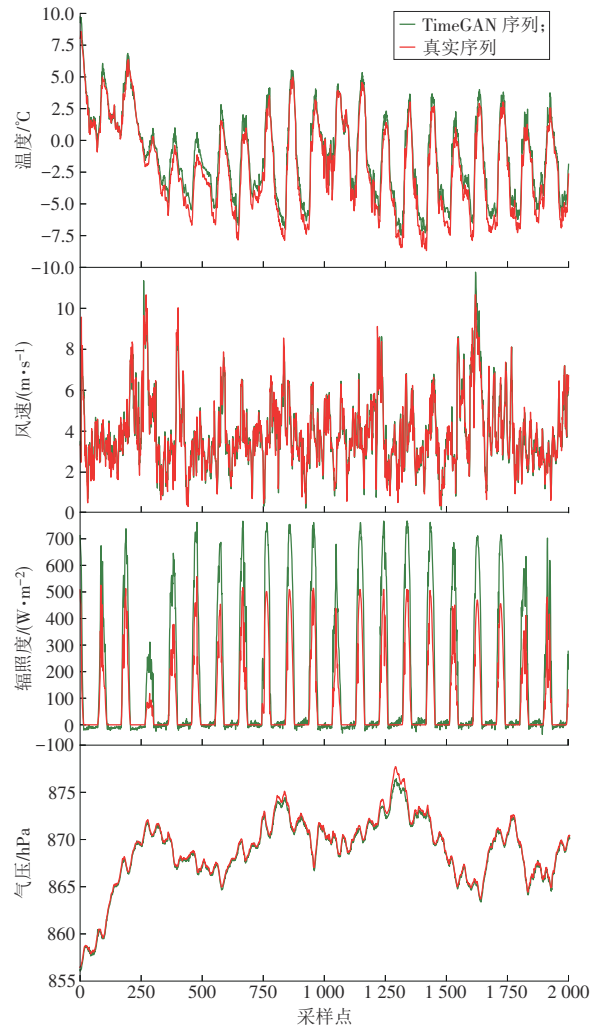


图 5 生成数据与原始数据对比

Fig. 5 Comparison between generated and original data

供了多维度的分析方法。PCA 基于协方差矩阵特征分解,通过构造正交基向量构建主成分空间,最大限度维持数据的总体变异量,首主成分承载最大方差,后续成分依次递减。这种线性投影机制对全局结构特征的保持具有显著效果。t-SNE 采用概率方法构建高维与低维空间的相似性分布,通过

Kullback-Leibler(KL)散度优化实现非线性映射。学生 t 分布的低维空间建模对数据点间的微观关联具有敏锐的捕捉能力,能有效放大局部邻域关系,因此,它对识别生成数据中的异常聚集、虚假聚类 etc 局部结构失真具有良好的效果。生成数据与原始数据之间的 PCA 和 t-SNE 分布如图 6 所示。

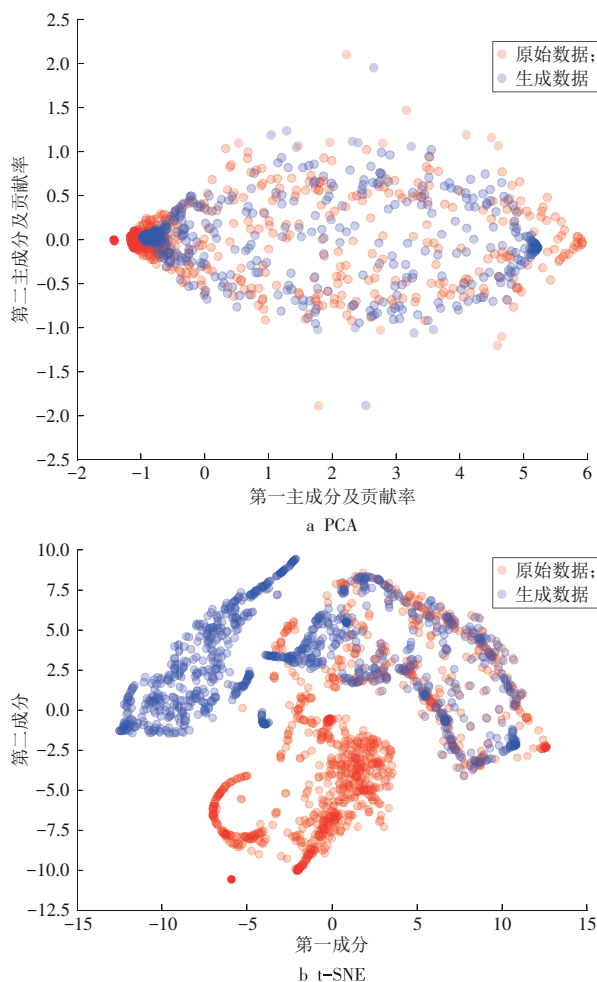


图 6 PCA 和 t-SNE 分析结果
Fig. 6 PCA and t-SNE analysis results

由图 6 可见:TimeGAN 生成的数据与原始数据分布规律相似,说明 TimeGAN 能较好地学习极端天气下的气象和出力特性并生成符合真实数据特征的样本数据;在降维空间中,生成数据复现了极端天气特有的气温骤降和剧烈波动等特征,生成数据与真实数据在 PCA 空间中全局分布重叠且在 t-SNE 中局部结构一致,说明 TimeGAN 成功捕捉了原始数据的时空特征。PCA 和 t-SNE 的结果能够有效论证 TimeGAN 生成数据在极端天气场景下的质量,为其在光伏功率预测中的应用提供可靠性支撑。

3.3 预测出力结果分析

利用 GAN 算法扩充的数据进行建模预测,分别对原始数据和扩增后的数据进行预测,预测结果如图 7 所示。

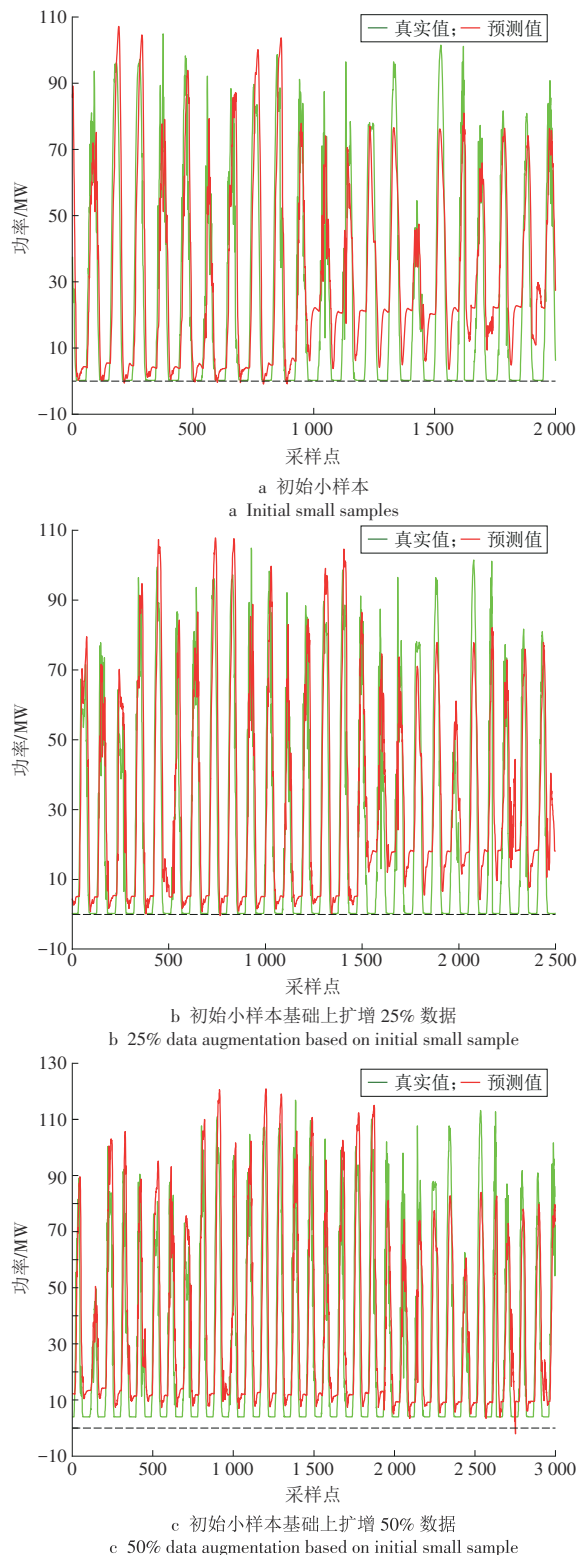


图 7 GAN 扩增方式下不同数据量的光伏出力预测结果
Fig. 7 Photovoltaic output prediction results with different data volumes under GAN-based augmentation

通过图 7 可以直观感受到,在原始历史数据不足的情况下,利用 GAN 扩增历史数据可以提高光伏功率预测的精准度,但对时间序列数据的处理仍存在局限性,误差对比见表 3。

为了使试验更具说服力,本文利用相同的数据样本采用 TimeGAN 扩增样本数据,进一步验证

表 3 GAN 扩增方式下 3 种数据量误差对比

Table 3 Error comparison of three data volumes under GAN-based augmentation

数据量	MAE/MW	RMSE/MW	R^2
无扩增	7.82	10.96	0.763
扩增 25%	5.40	5.81	0.873
扩增 50%	4.09	4.98	0.891

TimeGAN 处理时间序列的能力以及对于极端天气下光伏预测小样本扩充的可行性,预测结果和实际值的对比如图 8 所示。

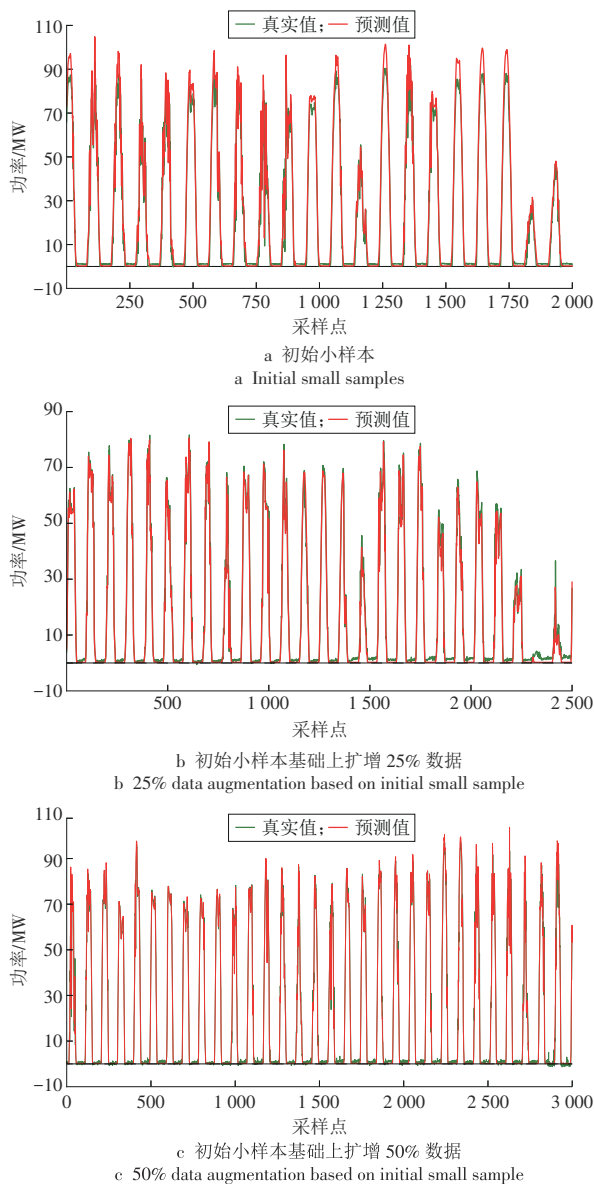


图 8 TimeGAN 扩增方式下不同数据量的光伏出力预测结果

Fig. 8 Photovoltaic output prediction results with different data volumes under TimeGAN-based augmentation

图 8 显示,采用 TimeGAN 捕捉时间特征,扩增数据量后的预测结果明显优于小样本历史数据的预测结果。

误差指标可以量化模型的准确性,夜间会存在

光伏发电功率为 0 的点,无法计算 MAPE 值,所以不考虑该指标。

光伏电站 3 种数据量的误差对比见表 4。GAN 与 TimeGAN 的综合评价对比见表 5。

表 4 TimeGAN 扩增方式下 3 种数据量误差对比

Table 4 Error comparison of three data volumes under TimeGAN-based augmentation

数据扩增	MAE/MW	RMSE/MW	R^2
无扩增	7.15	10.60	0.796
25%	4.26	4.72	0.878
50%	3.11	3.99	0.912

表 5 不同数据扩增方式的综合评价

Table 5 Comprehensive evaluation of different data augmentation methods

数据扩增方式	准确度	精度	召回率
GAN	96.09±0.79	92.93±2.76	74.75±0.63
TimeGAN	98.16±0.27	98.93±0.46	66.63±1.50

4 结论

针对极端天气下光伏电站历史数据获取困难导致历史样本数据过少的问题,本文提出了一种基于 TimeGAN 的极端天气光伏功率预测方法,得出以下结论。

(1) 利用 TimeGAN 对光伏电站历史数据进行扩充,解决了极端天气的样本原始数据量过小而导致训练集样本不足,进而影响预测精度的问题,通过扩充小样本数据,可为后续功率预测做好准备。通过对比生成数据和原始数据可以看出,扩增的数据与原始样本数据相似度极高,增强了样本的可靠性。

(2) 采用 TimeGAN 扩增小样本后的预测结果优于 GAN 扩增的小样本数据预测结果,有较好的拟合性,数据扩增 25% 后,MAE 降低了 1.14 MW,下降率为 21%,RMSE 降低了 1.09 MW,下降率为 18%;数据扩增 50% 后,MAE 降低了 1.08 MW,下降率为 23%,RMSE 降低了 0.99 MW,下降率为 20%。由此可见,在处理时间序列数据时,TimeGAN 扩增数据后的预测精确度得到了明显提高。

从预测结果可以看出,TimeGAN 对极端天气气象数据和光伏出力数据的扩充,可以较好地应用于极端天气光伏功率预测,有效弥补了原始数据量不足的问题,提高了预测模型的精准度,增强了结果的可靠性,当极端天气来临时,可以有效提高电力系统的应变能力,应用价值较高。

随着新能源技术的飞速发展,新能源发电场站数量不断增加,新建场站的历史数据较少,采用

TimeGAN 数据扩增的光伏功率预测技术可提升极端天气下的抗风险能力,通过精准预测新能源场站出力等方式全方位保障电力系统稳定运行。

参考文献:

- [1] 张冬冬, 单琳珂, 刘天皓. 人工智能技术在风力与光伏发电数据挖掘及功率预测中的应用综述[J]. 综合智慧能源, 2025, 47(3): 32-46.
ZHANG Dongdong, SHAN Linke, LIU Tianhao. Review on the application of artificial intelligence in data mining and wind and photovoltaic power forecasting [J]. Integrated Intelligent Energy, 2025, 47(3): 32-46.
- [2] 龙小慧, 秦际赞, 张青雷, 等. 基于相似日聚类及模态分解的短期光伏发电功率组合预测研究[J]. 电网技术, 2024, 48(7): 2948-2957.
LONG Xiaohui, QIN Jiyun, ZHANG Qinglei, et al. Short-term photovoltaic power prediction study based on similar day clustering and modal decomposition [J]. Power System Technology, 2024, 48(7): 2948-2957.
- [3] 卓振宇, 张宁, 谢小荣, 等. 高比例可再生能源电力系统关键技术及发展挑战[J]. 电力系统自动化, 2021, 45(9): 171-191.
ZHUO Zhenyu, ZHANG Ning, XIE Xiaorong, et al. Key technologies and developing challenges of power system with high proportion of renewable energy [J]. Automation of Electric Power Systems, 2021, 45(9): 171-191.
- [4] 邓芳明, 刘涛, 王锦波, 等. 基于地基云图与气象因素多模态融合的光伏功率预测方法[J/OL]. 中国电机工程学报, 2025: 1-14(2025-02-21)[2025-03-20]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=ZGDC20250220002&dbname=CJFD&dbcode=CJFQ>.
DENG Fangming, LIU Tao, WANG Jinbo, et al. Research on photovoltaic power prediction based on multimodal fusion of ground cloud map and meteorological factors [J/OL]. Proceedings of the CSEE, 2025: 1-14 (2025-02-21) [2025-03-20]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=ZGDC20250220002&dbname=CJFD&dbcode=CJFQ>.
- [5] 张元曦, 杨国华, 杨娜, 等. 基于 K-means 聚类的 LSTM-SVR-DE 光伏功率组合预测[J]. 综合智慧能源, 2025, 47(2): 71-78.
ZHANG Yuanxi, YANG Guohua, YANG Na, et al. Photovoltaic power prediction based on K-means clustering and the LSTM-SVR-DE model [J]. Integrated Intelligent Energy, 2025, 47(2): 71-78.
- [6] 王玉庆, 徐飞, 刘志坚, 等. 基于动态关联表征与图网络建模的分布式光伏超短期功率预测[J]. 电力系统自动化, 2023, 47(20): 72-82.
WANG Yuqing, XU Fei, LIU Zhijian, et al. Ultra-short-term power forecasting of distributed photovoltaic based on dynamic correlation characterization and graph network modeling [J]. Automation of Electric Power Systems, 2023, 47(20): 72-82.
- [7] 臧海祥, 程礼临, 刘玲, 等. 基于数据驱动的太阳辐射估计和预测研究与展望[J]. 电力系统自动化, 2021, 45(11): 170-183.
ZANG Haixiang, CHENG Lilin, LIU Ling, et al. Research and prospect for data-driven estimation and prediction of solar radiation [J]. Automation of Electric Power Systems, 2021, 45(11): 170-183.
- [8] 马原, 张雪敏, 甄钊, 等. 基于修正晴空模型的超短期光伏功率预测方法[J]. 电力系统自动化, 2021, 45(11): 44-51.
MA Yuan, ZHANG Xuemin, ZHEN Zhao, et al. Ultra-short-term photovoltaic power prediction method based on modified clear-sky model [J]. Automation of Electric Power Systems, 2021, 45(11): 44-51.
- [9] 张筱辰, 朱金大, 杨冬梅, 等. 基于 t-SNE 流形学习与快速聚类算法的光伏逆变器故障预测技术[J]. 中国电力, 2020, 53(6): 41-47.
ZHANG Xiaochen, ZHU Jinda, YANG Dongmei, et al. Photovoltaic inverter fault prediction technology based on t-SNE manifold learning and fast clustering algorithm [J]. Electric Power, 2020, 53(6): 41-47.
- [10] 梁志峰, 秦放, 崔方. “6·21”日环食对光伏发电及电网运行影响分析[J]. 电力系统自动化, 2021, 45(7): 1-7.
LIANG Zhifeng, QIN Fang, CUI Fang. Impact analysis of annular solar eclipse on June 21, 2020 in China on photovoltaic power generation and power grid operation [J]. Automation of Electric Power Systems, 2021, 45(7): 1-7.
- [11] 刘洪波, 王铎皓, 石鹏, 等. 基于图神经网络的多光伏场站出力短期时空预测[J]. 电网与清洁能源, 2025, 41(1): 89-96.
LIU Hongbo, WANG Duohao, SHI Peng, et al. Short term time-space prediction of multi-photovoltaic plant output based on graph neural network [J]. Power System and Clean Energy, 2025, 41(1): 89-96.
- [12] LIU W C, MAO Z Z. Short-term photovoltaic power forecasting with feature extraction and attention mechanisms [J]. Renewable Energy, 2024, 226: 120437.
- [13] ELISSAIOS S, EVANGELOS S, EFSTATHIOS S, et al. Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent long short-term memory models [J]. Renewable Energy, 2023, 216: 118997.
- [14] 殷豪, 张铮, 丁伟锋, 等. 基于生成对抗网络和 LSTM-CSO 的少样本光伏功率短期预测[J]. 高电压技术, 2022, 48(11): 4342-4351.
YIN Hao, ZHANG Zheng, DING Weifeng, et al. Short-

- term prediction of small-sample photovoltaic power based on generative adversarial network and LSTM-CSO [J]. High Voltage Engineering, 2022, 48(11): 4342-4351.
- [15]肖白,黄钰茹,姜卓,等.数据匮乏场景下采用生成对抗网络的空间负荷预测方法[J].中国电机工程学报,2020,40(24):7990-8001,8236.
- XIAO Bai, HUANG Yuru, JIANG Zhuo, et al. The method of spatial load forecasting based on the generative adversarial network for data scarcity scenarios [J]. Proceedings of the CSEE, 2020, 40(24): 7990-8001, 8236.
- [16]顾荣直,田心如,禹梁玉,等.江苏寒潮天气过程风险预评估方法研究[J].气象学报,2024,82(2):247-256.
- GU Rongzhi, TIAN Xinru, YU Liangyu, et al. Methodology in pre-assessment of the cold surge induced risks in Jiangsu province of China [J]. Acta Meteorologica Sinica, 2024, 82(2): 247-256.
- [17]王丽婕,张青山,郝颖,等.基于气象数据外推法和显著性分析的光伏自适应功率预测模型[J].太阳能学报,2025,46(2):317-325.
- WANG Lijie, ZHANG Qingshan, HAO Ying, et al. Photovoltaic adaptive power prediction model based on meteorological data extrapolation and significance analysis [J]. Acta Energetica Solaris Sinica, 2025, 46(2): 317-325.
- [18]郑珂,王丽婕,郝颖,等.基于数据集蒸馏的光伏发电功率超短期预测[J].中国电机工程学报,2024,44(13):5196-5208.
- ZHENG Ke, WANG Lijie, HAO Ying, et al. Ultra-short-term prediction of photovoltaic power based on dataset distillation [J]. Proceedings of the CSEE, 2024, 44(13): 5196-5208.
- [19]彭曙蓉,陈慧霞,孙万通,等.基于改进LSTM的光伏发电功率预测方法研究[J].太阳能学报,2024,45(11):296-302.
- PENG Shurong, CHEN Huixia, SUN Wantong, et al. Research on photovoltaic power prediction method based on improved lstm [J]. Acta Energetica Solaris Sinica, 2024, 45(11): 296-302.
- [20]卫志农,马智刚,陈胜,等.考虑光伏随机性的交直流混合配电网鲁棒机会约束安全域模型[J].中国电机工程学报,2024,44(6):2208-2220.
- WEI Zhinong, MA Zhigang, CHEN Sheng, et al. Robust chance-constrained security region model of AC/DC hybrid distribution network considering the uncertainty of photovoltaic generation [J]. Proceedings of the CSEE, 2024, 44(6): 2208-2220.
- [21]CHEN C C, CHAI L, WANG Q L. Research on stacking ensemble method for day-ahead ultra-short-term prediction of photovoltaic power [J]. Renewable Energy, 2025, 238: 121853.
- [22]孔令国,王嘉祺,韩子娇,等.基于权重调节模型预测控制的风-光-储-氢耦合系统在线功率调控[J].电工技术学报,2023,38(15):4192-4207.
- KONG Lingguo, WANG Jiaqi, HAN Zijiao, et al. On-line power regulation of wind-photovoltaic-storage-hydrogen coupling system based on weight adjustment model predictive control [J]. Transactions of China Electrotechnical Society, 2023, 38(15): 4192-4207.
- [23]YANG M, JIANG Y, ZHANG W, et al. Short-term interval prediction strategy of photovoltaic power based on meteorological reconstruction with spatiotemporal correlation and multi-factor interval constraints [J]. Renewable Energy, 2024, 237:121834.
- [24]ZHAO Y, WANG B, WANG S, et al. Photovoltaic power generation power prediction under major extreme weather based on VMD-KELM [J]. Energy Engineering, 2024, 121(12):3711-3733.

(本文责编:刘芳)

收稿日期:2025-03-26;修回日期:2025-04-30
上网日期:2025-06-17;附录网址:www.ienergy.cn

作者简介:

孙师奇(2001),男,硕士生,从事光伏功率预测方面的研究,1335900761@qq.com;

马刚*(1984),男,副教授,博士,从事新能源发电及入网、综合能源系统等方面的研究,nnmg@njnu.edu.cn;

许文俊(2002),男,硕士生,从事综合能源系统方面的研究,1819387688@qq.com;

李豪(1999),男,硕士生,从事光伏发电功率预测方面的研究,221812050@njnu.edu.cn;

马健(2001),男,硕士生,从事电力系统安全稳定控制方面的研究。

*为通信作者。