

Joint Forecasting Technology of Short-term Load and Photovoltaic Output Based on Deep Learning and Multi-source Data

Di Ge Lekang Liu Junfeng Qu Meng Zhang

Xuchang Power Supply Company of State Grid Henan Electric Power Company, Xuchang, Henan, 461000, China

Abstract

Under the background of rapid penetration of distributed generation, enhanced power load volatility and accelerated diversification of energy consumption patterns, the energy flow dynamics in the substation area presents a high degree of randomness and spatio-temporal coupling characteristics, making short-term load and photovoltaic output prediction one of the key technologies to improve the operation toughness and regulation precision of distribution network. The traditional forecasting method is difficult to describe the nonlinear correlation characteristics of load and photovoltaic output across time scales and space sources at the same time, resulting in a significant increase in the prediction bias in the high penetration photovoltaic scenario. The joint forecasting model proposed in this paper is superior to the traditional single task model and single source data model in Mae, RMSE and MAPE indexes, especially in the rapidly changing sunny overcast sunny weather conditions, showing higher robustness, and providing effective support for distribution network regulation strategy optimization, energy storage scheduling, flexible load management and distributed energy planning.

Keywords

deep learning; multi source data; station area load forecasting; photovoltaic output prediction

基于深度学习与多源数据的台区短期负荷与光伏出力联合预测技术

戈狄 刘乐康 渠俊锋 张萌

国网河南省电力公司许昌供电公司, 中国·河南 许昌 461000

摘要

在分布式电源快速渗透、电力负荷波动性增强以及用能形态加速多元化的背景下,台区的能量流动态呈现出高度的随机性与时空耦合特征,使短期负荷与光伏出力预测成为提升配电网运行韧性与调控精细度的关键技术之一。传统预测方法难以同时刻画负荷与光伏出力在跨时间尺度、跨空间源的非线性关联特征,导致预测偏差在高渗透率光伏场景下显著增大。本文提出的联合预测模型在MAE、RMSE及MAPE指标上均优于传统单任务模型和单源数据模型,特别是在快速变化的晴—阴—晴天气条件下表现出更高的鲁棒性,为配电网调控策略优化、储能调度、柔性负荷管理及分布式能源规划提供有效支撑。

关键词

深度学习; 多源数据; 台区负荷预测; 光伏出力预测

1 引言

随着分布式光伏在居民侧、工商业屋顶及公共建筑中的持续扩张,台区已成为承载电能双向流动的关键节点,其运行特性呈现出负荷与光伏出力叠加的复杂性。一方面,居民负荷受节假日、季节温差以及新型负荷(电动汽车、热泵、电采暖等)的影响呈现多峰化和高波动趋势;另一方面,光

伏出力受辐照度、云量、气温和组件方位角等影响,具有随机性与突变性。两者的共同作用使台区潮流呈现快速变化的耦合动态,增加了电压越限、反向潮流、变压器过载与谐振风险。因此,构建基于深度学习与多源数据的台区负荷—光伏联合预测技术体系,不仅能够提高预测精度,还能够为台区监测、储能调度、需求侧响应、电价优化及配网数字孪生等业务提供核心数据支撑。

【作者简介】戈狄(1985-),男,中国江苏泗洪人,硕士,高级工程师,从事新能源发电控制、电网调控运行及发电厂并网管理研究。

2 多源数据预处理与融合特征工程构建

2.1 多源异构数据的统一处理与时间一致性构建

在台区负荷—光伏出力联合预测中，多源数据的核心挑战并非单个变量的噪声控制，而在于不同数据来源之间采样粒度、时间对齐方式及语义表达存在显著差异。在实际配电网场景中，居民侧智能电表大多执行 15 min 或 5 min 采样周期，而光伏逆变器监测系统往往具备 1 min 或更高频率的输出能力，部分智能气象站甚至提供秒级辐照度与云量变化数据。如果直接采用原始数据输入深度模型，会出现时间步长不匹配、缺失值累积放大及突发事件定位不准确等问题，使模型在编码阶段产生非物理的特征偏移。因此，数据预处理首先需要建立统一的时间基准，采用向上对齐策略将高频数据通过滑动窗口平均或变分聚合方式映射至 5 min 或 15 min 时间步，使其具备负荷序列的可比性^[1]。

对于负荷数据中的异常尖峰，传统统计方法基于 3σ 或箱线图难以区分真实用户行为与计量异常，在居民区夏季空调集中开启、工业园区设备联动启动等情形中，峰值增长具有真实语义，因此需要在时间聚类框架下识别异常值，将行为类峰值保留，将计量跳变剔除，从根源上保证预测输入反映真实物理过程。在实现时间一致性的基础上，还需对间断性事件进行标注，如停电、逆变器保护动作、回路切换等状态变量同样影响负荷与光伏生成机制，其缺失会导致模型在训练阶段误把扰动当作噪声抑制，最终在真实运行中出现严重偏差。因此，在数据预处理阶段建立统一的事件时间轴并嵌入状态标签，是联合预测工程的一项基础性工作。

2.2 融合表达与高阶特征构建的深层次策略

多源数据的联合价值并非来自简单拼接，而来自变量间交互关系在预测空间中的非线性表达，特征工程因此成为连接数据与深度学习模型的关键环节^[2]。对于负荷数据，单一历史值序列无法充分刻画用户行为的波动机制，必须构建跨尺度时间表示，将短周期峰谷、日周期用能模式和周周期行为约束在统一表达空间中；这一构造可以通过基于傅里叶时间编码的周期嵌入实现，使模型具备区分工作日与周末、昼夜温差与季节迁移的能力。对于光伏出力，组件温度、辐照度、入射角之间存在高度非线性关系，仅引入逐步回归特征会带来高维噪声，因此通过物理合理性的组合特征更具稳定性，在气象特征中，全云量、云底高度与风速的交互在光伏短期突变时呈现显著增益，尤其在云团边缘经过阵列的瞬间，风速作为扰动传播速度的代理变量，直接影响辐照衰减持续时间，因此必须构建二阶交叉项而非孤立输入。在用户画像的 Embedding 表示方面，传统做法是对用户类型、户均面积、电动汽车保有量等特征进行独热编码，这一方式会造成维度膨胀且语义割裂，无法反映特征间潜在协同关系，因此需要采用低维连续映射，通过自编码器或词向量式嵌入

方式学习住宅—商用—工业用户之间的能耗分布距离，使得模型在特征空间而非数据空间捕捉台区结构差异。

3 深度学习联合预测模型的结构设计与协同优化机制

3.1 双流神经网络结构与负荷—光伏时序特征的差异化编码

在台区短期预测中，负荷与光伏出力属于两类数据生成机制完全不同的时序变量，前者反映用户行为与设备运行状态的响应结构，而后者由光照物理过程及组件特性决定，尽管二者在电能流层面存在叠加，但其时序特征并非对称或同源。因此，本研究采用双流神经网络结构，将负荷预测子模型与光伏出力预测子模型分离训练，并在高层语义空间建立融合接口，使两个任务既能够保留自身规律，又能够通过耦合层进行信息交换，从而避免在端到端模式下出现语义混淆。负荷预测子模型首先采用卷积层对历史功率序列进行局部模式提取，以识别短时波动、突发跃升与周期性峰值，随后通过 LSTM 或 GRU 网络捕获跨日尺度的行为迁移，使模型能够理解“工作日用能—周末回落”或“夜间制冷—白天制热”的跨周期模式。与此不同，光伏预测子模型在输入端引入辐照度、组件温度、风速等环境变量，并构建时间注意力机制以捕捉快速变化段的关联性，因为云团边缘通过阵列时的辐照突降具有较强的瞬态性，与负荷的缓慢变化形成显著差异。在两条子网络完成低层特征抽取后，通过共享表示层进行隐空间匹配，该表示层采用双向注意力机制，将负荷序列的周期信息与光伏序列的瞬态扰动进行交叉映射，使得模型可以在上午 10:00—14:00 光伏快速波动、夜间空调集中开启等时段识别潜在交互关系^[3]。

3.2 多任务学习与协同损失机制的预测偏差抑制策略

尽管双流结构在特征空间建立了负荷与光伏的差异化表达，但若两个任务完全独立优化，预测结果将在数据稀疏或扰动场景下出现偏离实际运行状态的“孤立化误差”。为避免这一问题，需要构建面向联合预测的多任务优化机制，使模型在训练过程中主动发现负荷与光伏之间的行为耦合，并通过统一的损失函数进行权重调节。在本研究中，损失函数不仅包含负荷预测误差项与光伏预测误差项，还引入时间相关性约束项与耦合正则项，用于在短时间尺度上限制模型出现反常偏移。在居民主导型台区，午间负荷下降与光伏出力上升往往同步出现，如果模型在该区间产生预测反向振荡，将导致实际调度中储能充放电策略被误触发。因此构建的协同项通过衡量两条预测曲线局部梯度方向的一致性，约束模型在敏感时段避免出现“差异性裂解”。不同于常见的均值平滑或误差回传，该约束项以二阶导数为参数，在梯度更新中表现为对急剧变化区域的惩罚，从而避免模型将短时扰动误识别为长期趋势。

4 模型性能验证与典型应用情境评估

4.1 基于真实台区样本的预测性能对比与误差结构分析

为了检验联合预测模型的有效性，本研究选取某省会城郊典型台区作为实验样本，该台区包含居住类用户 4152 户、商业用户 62 户、充电桩接口 19 个以及 74 套屋顶光伏设施，监测周期覆盖夏秋季连续五个月，其中包含多轮晴转阴、高温低风以及强对流天气事件。训练数据采用滑动窗口构造，保证负荷与光伏出力序列在不同时间尺度具有充分采样密度。作为对比，本研究引入传统 ARIMA、XGBoost、

单任务 LSTM 以及 Transformer 单任务模型，用以验证深度联合预测方法在负荷—光伏耦合场景中的优势。

从预测结果看，ARIMA 模型在稳定天气条件下的平均误差可达到较低水平，但在阴晴交替或居民行为突变场景中，缺乏对高阶非线性关系的捕捉能力，使其在突变区产生明显滞后响应；XGBoost 虽在单变量拟合方面具有较强能力，但对时间依赖结构的表达有限，对光伏出力快速变化的敏感性不足。基于单任务的 LSTM 在负荷预测上表现较好，但其光伏预测误差显著受到气象数据权重分配不足和瞬态扰动识别能力缺陷影响，在云团快速遮挡事件发生时往往产生过度平滑，导致预测峰谷不匹配。

表 1 台区短期联合预测模型与对比算法的性能比较（测试区间：夏秋季 5 个月）

模型类型	RMSE(W)	MAPE(%)	峰谷捕捉准确率 (%)	阴晴切换鲁棒性评分
ARIMA 时间序列	312.7	10.83	54.2	★☆☆☆☆
XGBoost 回归	241.5	8.97	61.8	★★☆☆☆
单任务 LSTM	215.3	8.21	69.4	★★★☆☆
Transformer 单任务	202.4	7.62	72.9	★★★★☆
本文双流联合模型	165.1	6.33	84.6	★★★★★

该表 1 展示了不同模型在 RMSE、MAPE、峰谷捕捉能力及复杂气象条件下的鲁棒性表现，可以清晰看出单任务模型在光伏突变区间存在系统性偏差，而本文提出的联合模型在时间依赖捕获、动态扰动识别及负荷—光伏协同表达方面表现更优，其峰谷捕捉准确率较传统方法提升 10%—20%，表明该模型不仅在统计指标上优胜，更具工程部署价值。

4.2 多天气、负荷扰动与调控反馈情景下的应用鲁棒性评估

联合预测模型不仅要在统计意义上实现误差降低，更重要的是在配电运维场景中具备可执行价值，因此需要在多维扰动条件下评估其鲁棒性。本研究将测试划分为三类典型场景，即天气快速切换场景、用户行为突变场景与设备运行扰动场景，用于验证模型在复杂运行环境中对电流量变化的响应能力。在晴—阴—晴快速交替背景下，辐照度呈现短周期不规则波动，光伏阵列组件温度受云影遮挡易产生阶段性冷却，随后在再曝晒时导致转换效率短时提升。

普通模型往往依赖平滑策略对串扰信号进行抑制，使预测曲线呈现“缓坡”形态，无法反映真实物理动态，从而在储能与调峰策略执行中触发滞后行为。而联合模型通过双流语义对齐机制识别光伏出力的瞬态突变并快速将注意力重新分配至高频气象特征，同时利用负荷分支稳定性约束避免出力曲线出现不合理振荡，使其预测表现呈现尖锐但结构合理的变化轨迹。在设备扰动场景中，当逆变器反复进入保护限流模式或充电桩群在短时段集中上线，电网状态呈现明显反馈行为，普通模型常将其归类为噪声信号而主动平滑，

但这会导致预测偏差累积并在运行策略中放大，实际调控则必须高度敏感。联合模型通过耦合正则项保留扰动区结构，从而使预测序列更接近真实运行路径，为储能提前充电、无功补偿投切以及台变过载风险评估提供可操作依据。实测结果显示，在三类扰动场景下，联合模型在 RMSE、峰值识别率以及异常事件捕获率方面均显著优于传统模型，尤其在“高波动弱相关”条件下表现最为突出，体现出深度模型在多源数据条件下的稳健泛化能力。

5 结论

本文构建了一种面向台区场景、融合多源异构数据的短期负荷—光伏出力联合预测模型，以深度学习为核心框架，通过双流神经网络、注意力机制、多任务损失函数以及特征耦合策略实现了预测精度与鲁棒性的全面提升。实际台区实验表明，该模型能够适应快速变化的天气与负荷条件，显著降低预测误差，对配电网调控体系、储能优化配置及分布式电源规划具有实际应用价值。

参考文献

- [1] 李芬,李雨欣,王亚维,等.考虑屋顶光伏热效应的短期净负荷预测[J].电力系统自动化,2025,49(9):146-156.
- [2] 张俊涛,程春田,申建建,等.考虑风光不确定性的比例可再生能源电网短期联合优化调度方法[J].中国电机工程学报,2020,40(18):5921-5931.
- [3] 高漪,周瑜,张安龙,等.整县光伏下基于个性化联邦学习的光伏出力及负荷功率预测[J].电网技术,2023,47(11):4629-4637.