

TransKANs-Fusion: 融合时空稀疏注意力与KAN的光伏预测

周子冠¹, 朱亚萍¹, 刘柱¹, 黄超¹, 曹军威², 涂国煜², 黄燕^{2,3}

1. 北京国网电力技术股份有限公司, 北京 100176

2. 清华大学 信息科学与技术国家研究中心, 北京 100084

3. 清华大学 自动化系, 北京 100084

基金项目: 北京市科技计划项目 (Z241100004224008); 新一轮第一批国家专精特新“小巨人”企业高质量发展资金项目 (国家“小巨人”新一轮第一批【2025】合同49号)

摘要: 针对间歇性和不确定性, 提出一种新型光伏功率预测方法——TransKANs-Fusion。设计了基于KAN的新型编码器结构、局部-全局稀疏注意力机制及多核一维卷积模块, 实现了对时序数据中的时空动态变化规律的有效捕捉。该方法融合了物理正则化项与数据驱动策略的优点, 增强了对复杂气象条件下光伏输出特性的描述能力, 能在保障相近计算复杂度的情况下获得高精度及高稳定性的光伏功率预测效果。

关键词: 光伏预测; 稀疏注意力; KAN; 1D Convolution; 时序特征

中图分类号: TM615 文献标识码: A 文章编号: 1672-0792(2026)02-0000-10

DOI: 10.3969/j.issn.1672-0792.2026.02.000

TransKANs-Fusion: Photovoltaic Prediction via Integrating Spatio-temporal Sparse Attention and KAN

ZHOU Ziguan¹, ZHU Yaping¹, LIU Zhu¹, HUANG Chao¹, CAO Junwei², TU Guoyu², HUANG Yan^{2,3}

1. Beijing State Grid Electric Power Technology Co., LTD., Beijing 100176, China

2. National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China

3. Department of Automation, Tsinghua University, Beijing 100084, China

Abstract: To address the inherent intermittency and uncertainty of PV power generation, under the global transition towards clean energy, this paper proposes a novel PV forecasting framework named TransKANs-Fusion. This framework integrates a newly designed encoder based on Kolmogorov-Arnold networks (KAN), a local-global sparse attention mechanism, and a multi-kernel one-dimensional convolution module, enabling effective capture of the spatiotemporal dynamic patterns in time-series data. The method combines the strengths of physics-informed regularization and data-driven approaches, enhancing its capability to characterize PV power generation under complex meteorological conditions. The proposed TransKANs-Fusion framework achieves highly accurate and stable forecasting performance while maintaining comparable computational efficiency.

Key words: photovoltaic prediction; sparse attention; KAN; 1D convolution; temporal feature

0 引言

光伏发电固有的强间歇性与气象依赖性, 导致其功率输出呈现高度非线性、时空耦合及动态不确定性, 给高精度时序建模带来根本性挑战^[1]。为应对这一问题, 数据驱动方法被广泛探索。循环神经网络^[2,3]能够捕捉时间序列依赖, 在短期预测中表现良好, 但难以建模长程关联且对极端天气适应性有限。尽管有研究通过引入变分模态分解或改进优化算法提升门控循环单元性能^[4], 或结合奇异谱分析增强其特征提取能力^[5], 但“黑箱”特性与固定门控结构仍限制了其对复杂气象突变的响应能力。用卷积神经网络虽可提取局部时序模式, 但感受野有限故而无法对全局动态演化有效表征^[6]。为兼顾局部与全局特征, 部分学者尝试构建神经网络混合架构^[7,8], 但往往因引入额外参数与计算开销, 使模型难以部署于边缘侧实时系统。

近年来, Transformer 凭借自注意力机制在光伏预测中展现出优越性能^[9,10]。然而, 其计算复杂度随序列长度呈平方增长, 在高分辨率(如 5 分钟级)多步预测场景下难以满足实时性需求。尽管稀疏注意力及高效变体被用于降低开销^[11,12], 但其在保留关键气象突变(如云团快速移动)敏感性方面仍显不足。例如, 基于概率 GRU-Transformer 的模型虽能刻画不确定性^[13], 但其推理延迟较高, 难以支撑超短期调度决策。与此同时, 传统多层感知机受限于固定激活函数(如 ReLU、Sigmoid), 其非线性表达能力有限。新兴的 KAN (Kolmogorov-Arnold networks) 通过可学习的样条基函数替代固定激活, 在提升函数逼近能力的同时增强了模型可解释性^[14]。KAN 将参数置于边而非节点, 理论上可更高效逼近高维非线性映射, 这为“气象-辐照度-发电功率”这一复杂物理过程建模提供了新思路, 但其在时序建模尤其是多源

异构数据融合中的应用方面的研究尚未深入。

此外, 多源信息的有效整合是提升预测鲁棒性的关键。实际光伏系统输出不仅受太阳辐照、温度、湿度等气象因素影响, 还与组件老化、逆变器效率、空间布局等设备状态密切相关。现有方法或依赖单一数据源^[15,16], 或仅简单拼接多源特征^[17], 未能充分挖掘其互补性。虽有研究融合多源数值天气预报并通过 Stacking 集成处理提升精度^[18], 或采用多目标优化构建区间预测以量化不确定性^[19], 但这些方法通常忽略设备健康状态的动态演化, 也未在模型内部建立物理一致性约束。部分研究尝试将物理模型与数据驱动方法结合, 通过残差修正或正则化提升预测鲁棒性^[20,21], 然而多数研究仅在损失函数中加入经验性惩罚项, 缺乏对能量守恒、光电转换效率边界等基本规律的显式建模。

综上, 当前光伏预测方法在计算效率、非线性表达能力、多源异构数据深度融合与物理可解释性 4 个维度上仍存在明显短板, 尤其在云层剧烈变化、设备部分故障等复杂工况下, 其预测稳定性显著下降。针对上述问题, 本文提出一种新型光伏功率预测方法——TransKANs-Fusion, 主要创新点: 1) 与传统 Transformer 模型相比可以有效降低计算复杂度, 能够在保持稳定精度的同时实现实时运行; 2) 设计了一个基于双 KAN 结构, 即含 Kernel-Adaptive Network 及 Kolmogorov-Arnold Network 的 KAN_MLP 处理层, 通过引入动态可学习的径向基函数作为激活函数, 增强本文预测方法的非线性表达能力和物理可解释性。

1 TransKANs-Fusion 网络

融合时空稀疏注意力与 KAN 结构的高效光伏预测方法的结构如图 1 所示, 本文提出的 TransKANs-Fusion 神经网络框架如图 2 所示。

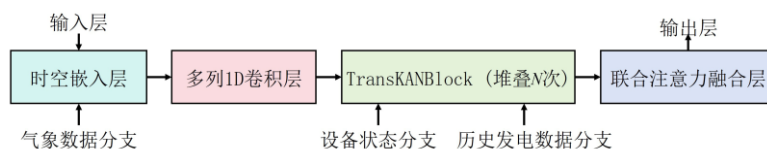


图 1 TransKANs-Fusion 方法的结构

Fig. 1 Structure of the TransKANs-Fusion method

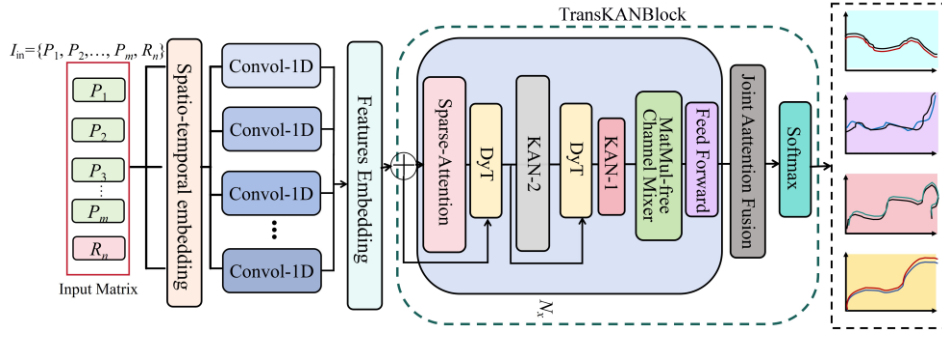


图 2 TransKANs-Fusion 神经网络框架

Fig. 2 Framework of the TransKANs-Fusion neural network

新型光伏功率预测方法的主要步骤如下:

步骤 1: 数据预处理。给定任意输入数据集 $I_{in} = \{P_1, P_2, P_3, \dots, P_m, R_n\}$ 。其中, $P_i = \{d_{i1}, d_{i2}, \dots, d_{im} | 1 \leq i \leq m, m \in N^+\}$ 是不同传感器的光伏发电数据; R_n 是所有影响因素个数 n 的集合, 支持动态增添, 其主要包括气象数据和用户行为模式。在数据预处理阶段, 依时间序列规整策略将历史发电数据与气象数据的时间序列进行对齐, 并设计特征工程、提取辐照度的累计变化率, 同时构建设备健康指数。

步骤 2: 基于步骤 1) 的预处理结果, 设计时空嵌入层 Spati-Temporal Embedding, 将输入层传输的数据和所有气象数据分支按照时间序列和空间序列特征进行时空编码。

步骤 3: 定义一个名为 TransKANBlock 的神经网络模块并进行该类的初始化。在该模块中设计一个名为 SparseAttention 的局部-全局注意力层, 并增加注意力窗口参数 w_size , 用来捕捉时空嵌入层中数据时空特征的关键信息; 同时设计一个 KAN-MLP 的网络层, 用来实现高效的特征变换和增强。引入 DyT 层, 以降低计算资源的需求并加快训练速度; 同时通过物理正则化项注入气象条件约束和无乘积融合全连接层 MatMul-free Channel Mixer, 在确保预测输出更符合物理规律的同时, 通过代替传统神经网络中的矩阵乘法运算, 降低计算成本和内存占用。

步骤 4: 设计联合注意力融合层 Joint Attention Fusion Layer。该层的作用是完成气象交叉注意力计算、门控权重生成和注意力结果融合, 对来自不

同来源并经过 KAN-MLP 处理的关键特征 (时空数据、气象数据和设备状态数据) 进行有效的整合, 充分利用各类信息的互补性, 实现对数据中复杂模式和关系的进一步捕捉, 从而提高其表达能力和泛化能力。

步骤 5: 输出层。该层的作用是完成特征整合、线性变换、激活函数设置、物理正则化以及最后光伏功率预测结果的生成。

2 数据预处理

采用时间序列规整策略识别历史发电数据与气象数据间的最优对齐路径, 并同步完成缺失值插补与异常值处理。先构建两序列的二维代价矩阵, 矩阵元素表示对应时间点的差异。然后从起始点出发逐点计算累积代价矩阵, 以确定从起点至各节点的最小路径代价。从终点回溯获得最优对齐路径, 对齐后数据中的缺失值采用插值法填补, 异常值则通过偏离均值检测方法识别并进行剔除。用式(1)(2)计算辐照度的累计变化率:

$$V = \frac{I_t - I_{t-1}}{I_{t-1}} \quad (1)$$

$$\Delta V = \sum_{i=1}^n V_i \quad (2)$$

式中: V 为辐照度变化率; I_t 和 I_{t-1} 分别为当前时间点和前一时间点的辐照度; n 为时间窗口, 设为 15 min; ΔV 为反映辐照度随时间变化趋势的辐照度累计变化率, 用于辅助分析其对发电量的影响。

构建设备健康指数 (H_I) 时, 选择辐照度、发

电量、温度、风速、风向、湿度、设备运行时间、维护记录以及故障历史记录作为关键特征，并对这些特征进行特征归一化，得到统一尺度下的特征集合 N 。健康指数由式(3)计算。

$$H_I = \sum_{i=1}^n w_i N_i \quad (3)$$

式中： w_i 为特征 i 的权重，可以通过专家经验或机器学习确定。

3 时空嵌入层及多尺度 1D 卷积

在数据预处理环节后设计了一个时空嵌入层（Spatio-temporal embedding layer），用于将输入数据的时间序列和空间序列特征进行编码。

时间序列编码。通过设置不同核数的多尺度一维卷积操作来捕捉局部特征，并动态设定分配不同时间步的权重。基于输入特征的数量，设计多个并行的一维卷积层。每个卷积层使用不同大小（如 3、5、7 等）的滤波器，以捕捉不同尺度上的局部模式。图 3 提供了该多尺度 1D 卷积的处理流程。图中多列卷积设计有助于从多个角度理解输入数据的时间依赖性。经过卷积操作，各卷积层输出的特征图会被拼接在一起，形成一个更丰富的特征表示。这一过程不仅保留了原始数据的时间顺序信息，还增强了对复杂时间依赖关系的理解。

空间序列编码。依据输入数据的时间分辨率

以及样本采样的波动规律，在线性拟合的规律范围内提取空间关系的输入特征作为空间数据编码。根据历史发电数据的空间分布情况，利用线性拟合方法提取出具有代表性的空间特征。这些特征能够反映不同地理位置之间的相互作用及其对发电量的影响。将提取出的空间特征转换为适合神经网络处理的形式。这通常涉及归一化处理：确保所有特征处于相同的尺度范围，从而避免因数值差异过大而导致的训练不稳定问题。

在完成时间和空间序列编码之后，需要将两者进行融合，形成统一的时空特征编码结构。本文将时间编码层和空间编码层的输出进行拼接，生成包含时空信息的综合特征表示。对拼接后的特征进行进一步的整合处理，再将其输入局部-全局注意力层（Sparse-Attention）增强特征间的交互作用，最终得到可用于下层设计的时空特征表示。为了有效提取光伏输出序列中的时间依赖性和空间相关性特征，采用了多尺度的一维卷积设计。这样不仅能够捕捉输入序列中的局部模式，还能增强对全局特性的理解能力。通过设置不同大小的卷积核，可以实现从多个尺度上分析输入数据，从而更好地捕捉复杂的时空特征。例如，小尺度卷积核擅长捕捉短期波动，而大尺度卷积核则更适合识别长期趋势。各个卷积层的输出特征图被拼接在一起，形成一个多通道的特征表示。这种方式优点是可同时考虑多种尺度下的信息，提高特征提取的全面性和准确性。

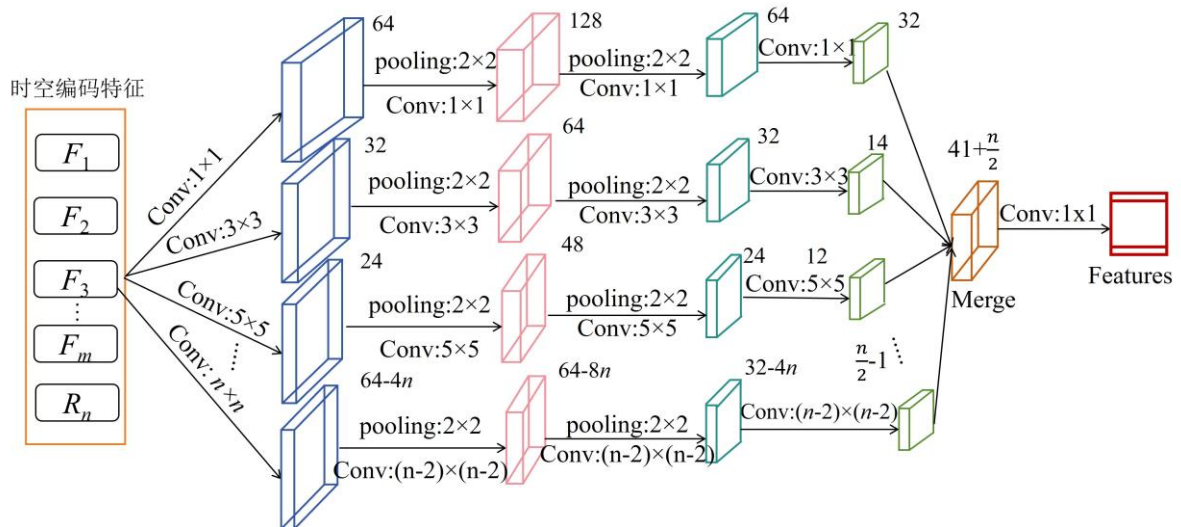


图 3 多尺度 1D 卷积层架构

Fig. 3 Architecture of the multi-scale 1D convolution operation

4 TransKANBlock 模块

多尺度卷积层的输出通过特征编码后进入到基于 Transfomer 编码器改进的 TransKANBlock 模块。该模块涉及局部-全局稀疏注意力层, KAN_MLP 处理层, 动态门控 Tanh (dynamically gated tanh, DyT) 层^[22], MatMul-free Channel Mixer^[23], Feed Forward 层, 联合注意力融合层 Joint Attention Fusion Layer 以及 SoftMax。在初始阶段, 构建一个名为 TransKANBlock 的自定义类实例。该实例接受 2 个参数, 即维度 d_model 和注意力头数 n_heads 。此模块内部集成了一个 SparseAttention 层, 用于实现局部-全局注意力机制, 并设置注意力窗口 w_size 大小为 $24h$, 以便在特定时间窗口内进行注意力计算, 从而捕捉数据中的局部和全局相关性。

然后, 在 TransKANBlock 模块中引入了一个名为 KAN_MLP 的处理层。在领域内, 本文首次尝试同时融合双 KAN-1 架构模式和 KAN-2 网络的核心特性, 以及多层感知机的优势。KAN_MLP 由一系列 KANLayer 子层堆叠而成。每个 KANLayer 首先通过执行线性变换将输入维度从 d_model 扩展至 $d_model \times 4$ 。采用动态可学习径向基函数作为激活函数以增强非线性和物理可解释性, 并通过另一线性变换将维度还原回 d_model 。这种多层次结构不仅能够有效转换和增强特征表示, 还能高效逼近复杂函数, 同时保持较高的物理可解释性。

此外, 本文方法创建了一个 PhysicsConstraint 模块, 旨在通过添加物理正则化项来确保预测输出符合已知物理规律。其中, 正则化系数变量 $coef$ 设定为 0.1。

KAN_MLP 层的输出与 PhysicsConstraint 模块的结果相加, 同时注入气象条件约束, 形成最终的该层输出。

为了提升 TransKANBlock 编码器的效率, 设计了动态门控 Tanh 的 DyT 层作为传统编码器中归一化层替代方案。DyT 是一种元素级操作, 其定义为:

$$DyT(x) = \sigma(x) \tanh(\alpha x) \quad (4)$$

式中: σ 为 Sigmoid 函数, x 为输入特征。

DyT 的设计灵感来源于观察到 Transformer 中的层归一化往往会产生类似 Tanh 的 S 形输入输出映射。通过使用 DyT, 可以在保持或甚至超越标准化性能的同时, 提高 Transformer 的效率。传统的全连接层被 KAN 结构所取代。与多层感知机相比, KAN 将可学习参数置于网络连接的边缘而非节点上。这种设计有利于复杂数据关系的捕捉, 拥有更强的学习能力和更高的参数率。对于第 L 层的 KAN, 其操作计算可以表示为:

$$K_{an}(x) = (\Phi_L \circ \Phi_{L-1} \circ \dots \circ \Phi_1)(x) \quad (5)$$

为了进一步降低计算复杂度, 本文还采用了无矩阵乘法 (matmul-free channel mixe)。BitNet 通过利用三值权重 $\{-1, 0, +1\}$, 将矩阵乘法转化为高效的三元累加。在标准密集层中, 输入向量 x 与权重矩阵 W 之间的核心操作是矩阵乘法 $y = W(x)$ 。为了提高计算效率, BitNet 引入了 BitLinear 模块, 用三元权值 $\{-1, 0, +1\}$ 中的元素代替传统的权值, 将矩阵乘法转化为高效的三元累加计算。因此与三元权矩阵 W_b 的矩阵乘法可以表示为:

$$y_b = BitLinear(x, W_b) \quad (6)$$

式中: y_b 为输出; $BitLinear$ 为三元矩阵乘法。

由于 W_b 的值有限, 乘法运算可以简化为:

$$y_b = \text{sign}(W_b) \square W_b |x \quad (7)$$

因此, 三元矩阵乘法最终转化为加减运算:

$$y_b = \sum_{j=1}^n \text{sign}(W_{ij}) |W_{ij}| x_j = \sum_{j=1}^n \text{sign}(W_{ij}) x_j \quad (8)$$

式中: $W_{ij} \in \{-1, 0, 1\}$ 。

在标准 Dense 层中, 输入向量 $u \in R^{1 \times D}$ 与权值矩阵 $W \in R^{D \times M}$ 之间的核心运算可以用矩阵乘法表示:

$$v = uA = \sum_{k=1}^D u_k A_{pk}, \quad p=1, 2, \dots, m \quad (9)$$

为了提高计算效率, 在 BitNet 引入了 BitLinear 模块, 用三元权值 $C \in \{-1, 0, +1\}$ 的元素代替传统的权值, 将矩阵乘法转化为高效的三元累加。三元权矩阵 C 的矩阵乘法为:

$$V = u \square C = \sum_{k=1}^D u_k C_{pk}, \quad C_{pk} \in \{-1, 0, +1\} \quad (10)$$

式中: $p=1,2,\dots,m$; $V^{\sim}$ 为权重取值后的输出; $\odot$ 为三元矩阵乘法。

由于 $C_{pk}^{\sim}$ 的值有限, 因此该乘法运算可简化为如下表达式子:

$$u_k C_{pk}^{\sim} \begin{cases} u_k, C_{pk}^{\sim}=1 \\ 0, W_{ij}^{\sim}=0 \\ -u_k, C_{pk}^{\sim}=-1 \end{cases} \quad (11)$$

网络层中的三元矩阵乘法可以转化为加减运算:

$$V^{\sim} = \sum_{k=1}^D u_k C_{pk}^{\sim} = \sum C_{pk}^{\sim} = 1u_k - \sum C_{pk}^{\sim} = -1u_k \quad (12)$$

联合注意力融合层主要负责气象交叉注意力的计算。利用气象数据作为 Key, 历史发电数据作为 Query, 生成注意力权重矩阵, 并结合局部-全局稀疏注意力结果, 通过 Kernel-Adaptive Network 结构生成包含三种注意力权重的向量, 并对各注意力结果进行加权求和, 综合考虑多种信息源以优化光伏发电预测。然后在输出层通过全连接层完成特征整合和非线性变换, 确保特征维度完全匹配输出需求。使用线性激活函数并应用物理正则化调整, 保证光伏的预测结果符合物理规律。这一模块各层共同作用, 可以显著提升网络对复杂环境的适应能力和预测精度。

5 案例实验

选取 2023—2025 年某光伏电站的历史监测数据为研究对象, 数据集的时间分辨率为 5 min。

考虑到不同季节的复杂气象情况, 分别对不同年度相同月份进行时空数据建模。

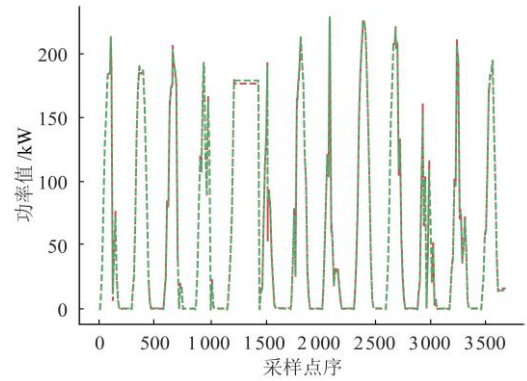
将本文方法与 Transformer encoder、GRU (gated recurrent unit)、LSTM (long short-term memory)、CNN (convolutional neural network) 及消融结构 CNN-Transformer encoder、KAN、KAN-Transformer、CNN-KAN 等具有竞争性的神经网络模型进行对比。

采用回归评价指标全面评估各类方法的预测精度, 包括平均绝对百分比误差 (MAPE)、均方根误差 (RMSE) 和平均绝对误差 (MAE)。其中, MAPE、RMSE 和 MAE 均用于量化预测值与实际

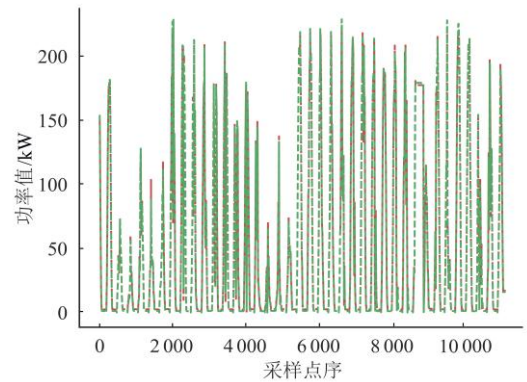
值之间的偏差, 其数值越小表示预测精度越高, 即预测结果与实际发电量之间的误差越小。此外, 本文还将使用决定系数 (R^2) 来评估网络的拟合优度, 其值越接近 1, 表明该神经网络对数据的解释能力越强, 预测结果与实际观测值之间的相关性越高。

5.1 定性评估

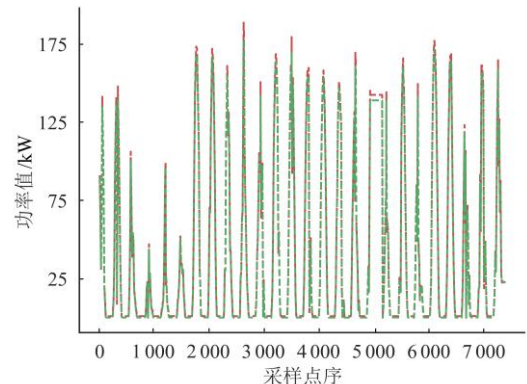
TransKANs-Fusion 支持多输入、多输出的多步预测, 即不同传感器的光伏功率、气象特征和用户行为模式等数据可以同时输入该网络进行不同类型的光伏预测。图 4 展示了本文方法的多步预测结果, 涉及超短时 (15 min)、短时 (1 h)、中长时 (2 d)、长时 (9 d) 的光伏功率预测。



(a) 15 min 预测结果



(b) 1 h 预测结果



(c) 2 d 预测结果

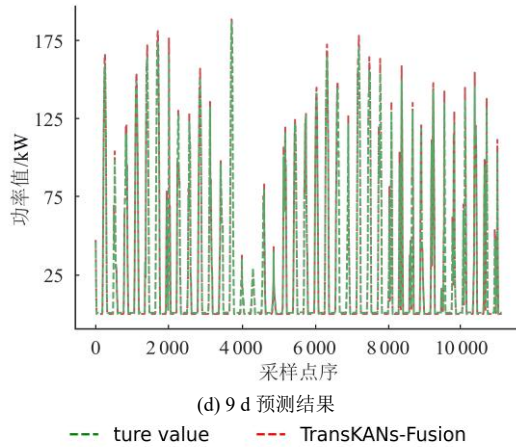


图4 TransKANs-Fusion 的整体预测表现

Fig. 4 Overall prediction performance of TransKANs-Fusion

由图4知, TransKANs-Fusion 方法在不同预测任务中展现出优异的性能, 在不同时间尺度均能较好地拟合实际光伏出力变化趋势, 预测值与真实值间的偏差较小, 体现出较强的时序建模能力和稳定性。在超短时预测中, TransKANs-Fusion 能够快速响应光照条件的突变, 准确捕捉输出功率的高频波动, 说明其适用于实时控制和微电网调度场景; 在短时预测任务中, 预测曲线与实际值保持高度一致, 具备良好的动态跟踪能力, 满足电力系统短期负荷平衡的需求; 对于中长期预测, 尽管图中曲线部分峰值点存在小幅偏差, 但整体趋势拟合良好, 表明该神经网络具有一定的泛化能力和多尺度适应性; 在长时预测任务中, 虽然受天气预报误差等因素影响, 模型预测精度有所下降, 但预测结果仍能反映出光伏的基本周期性和变化规律, 可为电网调度和能源规划提供有效参考。因此, 定性实验分析结果表明, 本文的 TransKANs-Fusion 方法在多种预测场景下均表现出较强鲁棒性与实用性。

图5展示了本文 TransKANs-Fusion 方法与所有对比方法在不同时序预测任务中的表现。

由图5知, TransKANs-Fusion 在短期预测中, 能够精确捕捉光伏输出的快速变化趋势, 其不同同步数的预测曲线与真实值曲线高度吻合, 尤其是在波动较大的时间段内表现尤为突出。相较于其他网络方法, 该方法在峰值和低谷时段均能较好地保持预测稳定性, 避免了过大的偏差, 表明该方

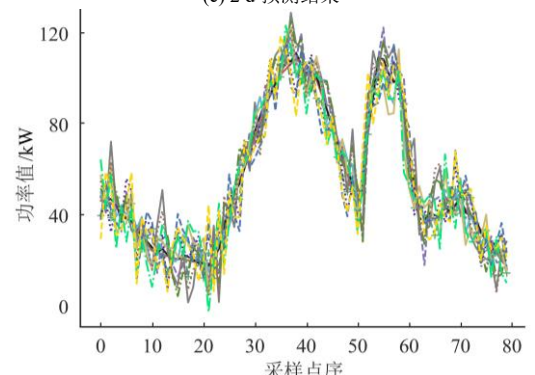
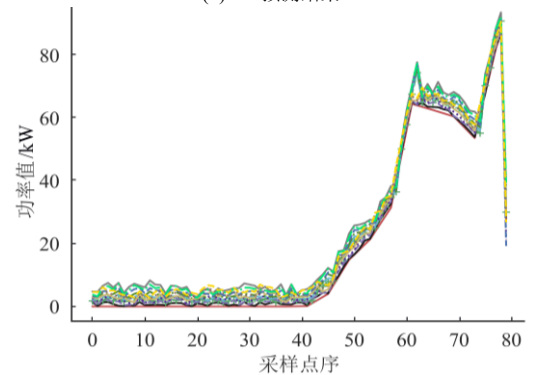
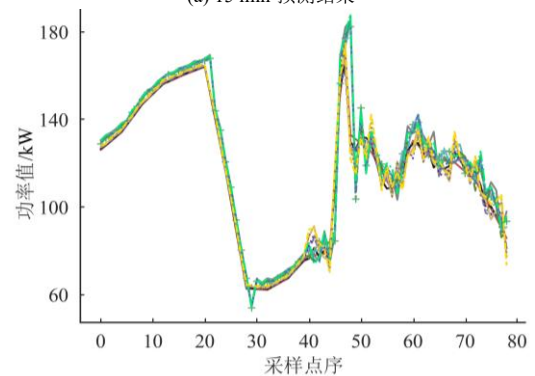
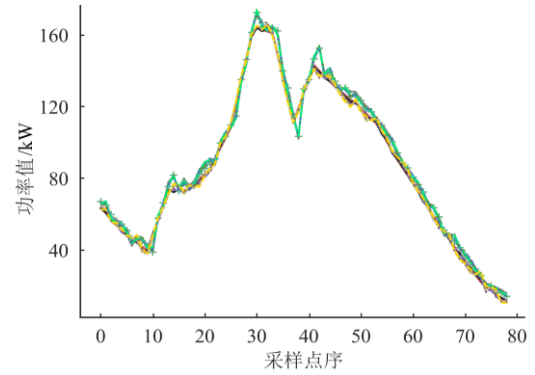


图5 TransKANs-Fusion 与对比方法的光伏功率预测表现

Fig. 5 PV prediction performance of TransKANs-Fusion

and comparative methods

<http://dlkxygc.ncepu.edu.cn>

法在处理高频数据时有较强的适应性和鲁棒性。

在中长期预测任务中，该方法能够有效捕捉到光伏输出的周期性变化规律，并在长时间序列预测中保持较高的精度，同样展现出了强大的预测能力。与 KAN-2、Transformer 等神经网络相比，TransKANs-Fusion 在预测曲线的整体趋势和局部细节上均有更佳的表现，特别是在复杂天气条件下的预测对比效果更为显著，这说明该方法在处理长时序数据时具有较好的泛化和抗干扰能力。

因此，通过融合 Transformer 和 KAN 网络的优势，TransKANs-Fusion 不仅在短期预测中表现出色，能够在高频数据中捕捉到快速变化的趋势；而且在中长期预测中也展现出了强大的泛化能力和稳定性，能够有效应对复杂多变的光伏输出特性。

5.2 定量评估

表 1 总结了各光伏发电预测方法在 4 个评价指标上的平均统计值。表中统计了所有方法在不同步数下的光伏预测任务中的表现。TransKANs-Fusion 在多个关键评价指标上不仅展示出稳定的预测性能，而且相较于其他对比模型实现了显著

的改进。

在 RMSE 指标方面，本文提出的方法获得了最低值 25.344，相比 GRU、CNN、KAN-2 和 Transformer 等神经网络模型，分别降低了约 38.3%、38.6%、33.7%和 28.5%，显示其在减少预测误差方面的相对优势。

对于 MAE 指标，TransKANs-Fusion 达到了 6.155，比上述各模型分别减少了约 67.6%、68.1%、57.2%和 56.5%，进一步证明了其降低预测偏差的能力。

在 MAPE 指标方面，TransKANs-Fusion 方法获得了 0.041 8 的最低值，相比于其他对比模型，相对误差最大降低了约 56.6%。

在决定系数 R^2 指标上，TransKANs-Fusion 平均可以达到 0.928，比所有对比方法提高了最大约 30.2%。这说明本文方法在解释光伏功率变化规律及拟合效果方面表现尤为出色。

从定量分析的角度，TransKANs-Fusion 不仅在各项评价指标上优于其他先进模型，而且在减少预测误差、提高预测精度方面实现了明显的提升。

表 1 光伏功率预测方法比较

Tab. 1 Comparison of different PV prediction methods

预测方法	评价指标的平均值			
	RMSE	MAE	MAPE	R^2
GRU	41.063	19.003	0.096 3	0.713
CNN	41.260	19.286	0.091 3	0.759
LSTM	37.524	16.740	0.072 4	0.779
KAN-2	37.559	16.862	0.079 5	0.782
CNN-KAN	37.113	15.117	0.070 1	0.834
BiLSTM	36.987	15.209	0.076 3	0.815
BiGRU	35.049	15.228	0.074 9	0.809
Transformer	33.217	12.098	0.053 9	0.871
KAN-Transformer	32.901	12.004	0.051 9	0.890
CNN-Transformer	36.958	12.825	0.063 4	0.853
LSTM-Transformer	32.159	10.473	0.057 8	0.896
GRU-Transformer	29.083	8.693	0.049 9	0.900
TransKANs-Fusion	25.344	6.155	0.041 8	0.928

5 结论

针对光伏发电功率预测中的复杂时空特征捕

<http://dlkxygc.ncepu.edu.cn>

捉与处理问题，提出了基于时空稀疏注意力机制和 KAN 结构的新型光伏功率预测方法——TransKANs-Fusion。该方法通过设计局部-全局稀疏注意力机制、双 KAN 结构以及多核一维卷积模

块, 实现了对时序数据中动态变化规律的有效捕捉, 同时增强了模型对复杂气象条件下光伏输出特性的描述能力。为弥补传统光伏预测方法在数据异构性、计算效率、物理可解释性和动态环境适应性方面的不足, 本文方法实现了自动从大量历史数据中学习并提取特征, 对多源异构数据(如气象数据、设备状态等)的有效整合, 提高了预测结果的时间连续性和一致性, 满足电力系统调度、电网稳定控制及电力市场交易的需求。实验结果显示, 在多个关键评价指标上, TransKANs-Fusion 相较于其他对比模型实现了显著改进, 其在减少预测误差、提高预测精度方面具有优越性能。

未来的研究工作可进一步探索以下几个方面: 1) 精细化天气状况的聚类划分, 针对各种复杂天气类型分别建模, 以进一步提升不同天气条件下的预测精度。2) 探索更加高效的数据处理与特征提取技术, 优化模型架构, 降低计算成本的同时保持或提升预测准确性。3) 随着智能电网的发展, 将更多维度的信息(如用户行为模式、电力市场价格波动等)纳入考量, 构建更全面、精确的光伏发电功率预测模型, 为实现清洁能源的高效利用提供强有力的技术支持。通过不断的研究和实践, 我们期望为推动全球向更清洁、可持续能源体系的转型做出更多的贡献。

参考文献:

- [1] Fu X Q, Zhang C Y, Xu Y, et al. Statistical machine learning for power flow analysis considering the influence of weather factors on photovoltaic power generation[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 36(3): 5348-5362.
- [2] Wang H Y, Xie G Q, Yang X D, et al. Photovoltaic power prediction based on improved long and short-term memory neural network model[C]//2025 International Conference on Electrical Automation and Artificial Intelligence (ICEAAI). Guangzhou, China: IEEE, 2025: 1031-1035.
- [3] Huang Y, Cao J W, Cai S X. A spatial-temporal neural network for photovoltaic power prediction[C]//2023 International Conference on Power and Energy Systems (ICPES). Chengdu, China: IEEE, 2023: 448-452.
- [4] 郭鹏, 田敏, 刘祥. 基于 VMD-IAVOA-GRU 的超短期光伏功率预测模型[J]. 计算机仿真, 2025, 42(3): 111-116.
- [5] 王玲芝, 李晨阳, 刘婧, 李程. 基于 GRO-SSA-LSTM 的短期光伏发电功率预测[J]. 太阳能学报, 2025, 46(2): 401-409.
- [6] 陈殿昊, 臧海祥, 蒋雨楠, 等. 基于多层次云图特征与宽度学习的超短期光伏功率预测[J]. 电力系统自动化, 2024, 48(22): 131-139.
- [7] Wang X G, Liu W J, Ni J H, Zhang M. Sample-wise graph-based multivariate short-term PV power forecasting[J]. IEEE Transactions on Sustainable Energy, 2025, 16(4): 3003-3014.
- [8] 王腾飞, 杨力, 孙龙, 等. 基于 C-Transformer 的光伏负荷预测方法[J]. 微电子学与计算机, 2025, 42(6): 67-74.
- [9] Yang J, He H, Zhao X M, et al. Day-ahead PV power forecasting model based on fine-grained temporal attention and cloud-coverage spatial attention[J]. IEEE Transactions on Sustainable Energy, 2024, 15(2): 1062-1073.
- [10] 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.
- [11] Song K J, Kim M, Kim H S. Graph-based large scale probabilistic PV power forecasting insensitive to space-time missing data[J]. IEEE Transactions on Sustainable Energy, 2025, 16(1): 160-173.
- [12] YANG Y, LIU Y, ZHANG, Y, et al. DEST-GNN: A double-explored spatio-temporal graph neural network for multi-site intra-hour PV power forecasting[J]. Applied Energy, 2025, 378: 124744.
- [13] Mao W Q, Zhao H Y, Huang X, et al. A short-term power prediction method for photovoltaic power generation based on GRU-Transformer model[C]//2024 International Conference on Energy, Electrical and Power Engineering (CEEPE). Chongqing, China: IEEE, 2024: 1365-1370.
- [14] Yan J, Gao Y H, Cao W P, et al. A study on the application of Kolmogorov-Arnold networks for renewable energy forecasting[C]//2025 International Conference on Machine Learning and Cybernetics (ICMLC). Guangzhou, China: IEEE, 2025: 229-234.
- [15] Qu J Q, Sun Q, Qian Z, et al. A transferable framework of PV power forecasting for cross-regional distributed PV systems using domain adversarial temporal network[J].

- IEEE Transactions on Industrial Informatics, 2025, 21(10): 8121-8132.
- [16] 韩家鹏, 丁蕾蕾, 韩崇. 基于奇异谱分析的 Transformer 神经网络光伏预测算法[J]. 软件导刊, 2025, 24(3): 8-15.
- [17] Wang S C, Huang Y. Spatio-temporal photovoltaic prediction via a convolutional based hybrid network[J]. Computers and Electrical Engineering, 2025, 123(Part A): 110021.
- [18] Behera S, Behera S. Application of stacked LSTM networks in forecasting day ahead PV power generation from onsite data[C]//2025 International Conference for Emerging Technology (INCET). Belgaum, India: IEEE, 2025: 1-5.
- [19] Ming M J, Wang R, Zha Y B, et al. Multi-objective optimization of hybrid renewable energy system with load forecasting[C]//2017 IEEE International Conference on Energy Internet (ICEI). Beijing, China: IEEE, 2017: 113-118.
- [20] YANG J X, WU T, WANG K, et al. A hybrid vmd-based arima-lstm model for day-ahead pv prediction and uncertainty analysis[C]//2022 4th International Conference on Smart Power & Internet Energy Systems (SPIES). Beijing, China: IEEE, 2022: 2009-2014.
- [21] 干逸飞, 吕品, 郑树泉. 面向光伏功率预测的残差深度学习模型[J]. 计算机应用与软件, 2024, 41(11): 101-107.
- [22] Zhu J C, Chen X, He K, et al. Transformers without normalization[J/OL]. arXiv preprint arXiv:2506.12345, 2025 [2026-02-13]. <https://arxiv.org/abs/2506.12345>.
- [23] WANG H Y, MA S M, DONG L, et al. Bitnet: scaling 1-bit transformers for large language models [EB/OL]. (2025-07-15)[2023-10-17]. <https://arxiv.org/abs/2310.11453>.
- 收稿日期: 2025-09-15
- 作者简介:
- 周子冠(1981—), 男, 正高级工程师, 研究方向为新能源电力系统、信息技术;
- 朱亚萍(1990—), 女, 工程师, 研究方向为新能源电力系统、信息技术;
- 刘柱(1983—), 男, 正高级工程师, 研究方向为新能源电力系统、信息技术;
- 黄超(1976—), 男, 工程师, 研究方向为新能源电力系统、信息技术;
- 曹军威(1973—), 男, 研究员, 研究方向为新能源系统、自动控制、人工智能在电力系统中的应用;
- 涂国煜(1980—), 女, 正高级工程师, 研究方向为自动控制、人工智能在电力系统中的应用;
- 黄燕(1989—), 女, 助理研究员, 研究方向为自动控制、计算机视觉、人工智能在电力系统中的应用。
- 通信作者: 朱亚萍