

Research on Ultra-short-term Wind Power Prediction Method Based on LSTM-KAN

Lifeng Wang¹ Lei Zhou² Yue Yuan³ Zhongyuan Li⁴ Haoyu Sun¹

1. Shandong Luruan Digital Technology Co., Ltd., Jinan, Shandong 250001, China

2. Power Dispatching Center of State Grid Shandong Electric Power Company, Jinan, Shandong 250001, China

3. State Grid Shandong Electric Power Company Jinan Power Supply Company, Jinan, Shandong 250001, China

4. State Grid Shandong Electric Power Company Emergency Management Center, Jinan, Shandong 250001, China

Abstract

Wind power prediction is crucial to the stable operation of power grid, but its volatility and nonlinear characteristics increase the difficulty of prediction. This paper proposes an ultra-short-term power prediction method based on the combination of LSTM (Long Short-Term Memory Network) and KAN (Kernel Attention Network). By collecting historical power, wind speed, hub height, latitude and longitude, and ERA5 meteorological data from wind farms, preprocessing and feature extraction are performed. LSTM captures the long-term dependency of the time series, and KAN enhances the attention to the key time points. The hybrid loss function is designed to balance the local and overall prediction performance to improve the model stability and generalization ability. Experiments show that the LSTM-KAN algorithm significantly outperforms traditional methods in prediction accuracy, providing a new scheme for ultra-short-term power prediction of wind power.

Keywords

LSTM; KAN algorithm; attention mechanism; neural network; hybrid loss function

基于 LSTM-KAN 的超短期风功率预测方法研究

王立峰¹ 周蕾² 袁月³ 李中原⁴ 孙浩宇¹

1. 山东鲁软数字科技有限公司, 中国·山东 济南 250001

2. 国网山东省电力公司电力调度中心, 中国·山东 济南 250001

3. 国网山东省电力公司济南供电公司, 中国·山东 济南 250001

4. 国网山东省电力公司应急管理中心, 中国·山东 济南 250001

摘 要

风电功率预测对电网稳定运行至关重要, 但其波动性和非线性特征增加了预测难度。本文提出一种基于 LSTM (长短期记忆网络) 与 KAN (Kernel Attention Network) 相结合的超短期功率预测方法。通过收集风电场历史功率、风速、轮毂高度、经纬度及 ERA5 气象数据, 进行预处理与特征提取。LSTM 捕捉时间序列的长期依赖关系, KAN 增强对关键时间点的注意力。设计混合损失函数平衡局部与整体预测性能, 提升模型稳定性与泛化能力。实验表明, LSTM-KAN 算法在预测精度上显著优于传统方法, 为风电超短期功率预测提供了新方案。

关键词

LSTM; KAN 算法; 注意力机制; 神经网络; 混合损失函数

【基金项目】山东鲁软数字科技有限公司和山东省电力气象技术创新中心科技项目——基于去中心化自组织协同计算的风功率预测转折性天气精准识别关键技术研究与应用项目(B90630240007)资助。

【作者简介】王立峰 (1975-), 男, 中国山东聊城人, 本科, 工程师, 从事电力气象、电力数字化研究。

【通讯作者】周蕾 (1992-), 女, 中国山东菏泽人, 硕士, 工程师, 从事电力平衡、电力市场研究。

1 引言

风电作为一种清洁能源^[1], 对电网稳定性和电力调度至关重要^[2]。然而, 风能的随机性和不确定性使得传统预测方法在精度上存在较大差距^[3]。现有预测模型包括物理模型、统计模型、机器学习模型等^{[4][5]}, 其中深度学习模型 (如 LSTM) 因强大的非线性拟合能力逐渐成为主流^{[6][7]}, 但单一模型仍存在精度不足的问题^[8]。符杨^[9]研究改进 LSTM-TCN 算法, 较传统 TCN 减小预测误差 2%, 缩小预测区间。蔡昌春^[10]提出多层堆叠双向 LSTM, 结合时间模式注意力

机制,验证了其季节适应性。

本文提出一种基于 LSTM 与 KAN 结合的模型,通过 KAN 的注意力机制增强关键时间点捕捉能力,结合混合损失函数平衡局部与整体性能,提升模型稳定性与泛化能力。实验采用 RMSE、R² 和 PCC 评估性能,验证其在超短期风电功率预测中的有效性和优越性,为风电预测提供新方案。

2 基于注意力机制的 LSTM-KAN 模型原理

2.1 长短时记忆网络 (LSTM)

长短期记忆网络 (Long- and short-term memory networks, LSTM), 是一种特殊类型的循环神经网络 (RNN), 旨在解决标准 RNN 在处理长序列数据时遇到的梯度消失和梯度爆炸问题。LSTM 的设计允许网络保持对早期信息的记忆, 即使在面对很长的时间间隔时, 也能有效地捕获和利用这些信息, 避免了传统 RNN 在长序列处理中常见的信息丢失或遗忘问题。

这种机制通过三个关键的控制门 (输入门、遗忘门和输出门) 来实现对信息流的精确控制。其中输入门负责决定在当前时间步中, 哪些新的信息应该被添加到细胞状态中。遗忘门决定哪些旧信息应该被保留在细胞状态中, 哪些应该被遗忘或丢弃。输出门控制细胞状态如何转化为当前时间步的输出。

2.2 安德烈·科尔莫戈洛夫网络 (KAN)

安德烈·科尔莫戈洛夫网络 (Kolmogorov-Arnold Networks, KAN) 是一种基于 Kolmogorov 理论的神经网络模型, 主要用于函数逼近和多变量非线性映射。KAN 通过结合知识图谱与深度学习, 增强模型的推理和预测能力。其结构包括输入层、知识图谱嵌入层和深度学习模型层。输入层接收并预处理传统数据 (如数值特征、文本信息); 知识图谱嵌入层通过图谱嵌入技术将领域知识转化为向量表示, 为模型提供外部知识支持; 深度学习模型层则通过多层非线性转换处理输入数据与知识图谱嵌入向量的融合信息, 学习数据模式并利用领域知识进行精准推理。KAN 的核心原理在于利用知识图谱作为先验知识来源, 强化数据驱动的学习过程, 从而提升模型的准确性和智能化水平。

2.3 LSTM 与 KAN 的融合机制

KAN 特征提取层通过将输入数据 (如风速、温度、历史功率等) 映射至高维空间, 利用单变量函数组合生成全局特征。LSTM 时序建模层接收 KAN 的输出, 通过遗忘门和输入门调控信息流, 捕捉局部时间依赖性。自适应权重融合模块动态调整 KAN 与 LSTM 的特征贡献权重, 优化多尺度特征融合。这种结合充分利用了 KAN 的强大函数逼近能力和 LSTM 的时间序列建模能力, 能够在处理时间序列数据的同时精确建模复杂的非线性关系。

本研究提出的 KAN-LSTM 模型具有显著优势: 首先, LSTM 能够有效捕捉序列数据中的时间依赖性, 而 KAN 可

以对每个时间点的输出进行精细调整, 从而增强模型的预测能力。其次, 模型架构灵活, 能够适应不同的输入输出关系和数据特性。此外, 通过结合两种网络结构, 模型在未知数据上的泛化能力显著提升, 进一步提高了预测的准确性和鲁棒性。

2.4 损失函数

本研究在 MSE (均方误差)、MAE (平均绝对误差)、R² (决定系数)、PCC (Pearson 相关系数) 的基础上研发了一种混合损失函数, 计算方法如式 (1) 所示。

$$\text{Total Loss} = \alpha * \text{MSE} + \beta * \text{MAE} - \gamma * \text{R}^2 - \delta * \text{PCC} \quad (1)$$

其中 MSE、MAE、R²、PCC 计算方式如式 (1) - (4) 所示。

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$\text{PCC} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (5)$$

通过自定义混合损失函数, 可以平衡局部和整体预测能力, 强化模型稳定性。

3 超短期风功率预测

3.1 实验设计

3.1.1 数据来源与数据预处理

本研究使用的气象数据来自欧洲中期天气预报中心 (ECMWF) 提供的 ERA5 全球气象再分析数据集, 时间范围为 2017 年 1 月 1 日至 2024 年 8 月 31 日。ERA5 数据集具有高时空分辨率, 涵盖逐小时的气象变量, 包括 10 米风速风向 (ws10m、wd10m)、100 米风速风向 (ws100m、wd100m)、相对湿度 (rh)、2 米露点温度 (d2m)、地面 2 米温度 (t2m)、降水 (tp) 和气压 (sp)。此外, 研究还集成了气象和空气质量历史数据等先验知识, 并对数据进行了预处理, 包括缺失值填充、异常值处理、无效值剔除和时间戳对齐。

数据预处理基于山东省风功率站点的 CSV 文件格式时间序列数据, 具体步骤包括:

①数据清洗: 基于规则及去除无效值和填充缺失数据。

②数据标准化: 使用 StandardScaler 将对数据进行标准化并保留参数配置。

③数据划分: 将数据分为训练集, 验证集和测试集 (8 : 1 : 1)。

3.1.2 预测流程

为验证本研究所提出的 KAN-LSTM 模型在超短期风功率预测任务中的预测效能, 以上述数据集作为实验基准。该数据集已通过先前的预处理与特征工程步骤, 确保了其对于

风预测任务的适宜性与代表性。对比实验中，基于常用于风功率预测的 XGBoost、LightGBM 等机器学习算法及 LSTM 深度学习算法与 KAN-LSTM 构建超短期风功率预测模型进行性能评估及对比（图 1）。

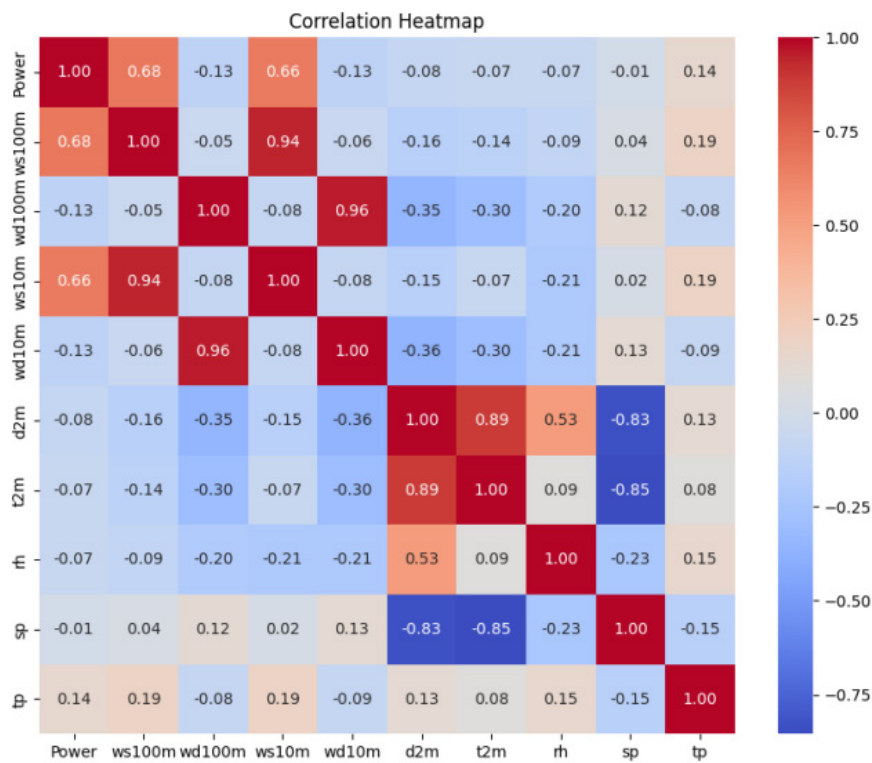


图 1 各气象因子与功率相关性热力图

为减少特征的冗余性，基于 Spearman 相关系数筛选关键气象因子，选取相关性较高的要素作为特征参与训练，相关性结果如图及表所示，根据结果选取 100m 风速及 10m 风速气象因子参与模型构建（表 1）。

参与对比的模型在相同的实验设置与数据分割方案下进行训练与测试，确保结果的可比性。评价指标选用 RMSE（Root Mean Square Error，均方根误差）、 R^2 （决定系数）和 PCC（Pearson Correlation Coefficient，皮尔逊相关系数）这三种广泛应用于风功率预测任务的统计指标，以全方位衡量各模型的预测精确度。

表 1 各气象因子与功率相关性表

气象因子	相关系数	气象因子	相关系数
100 米风速	0.68	地面 2 米气温	-0.07
100 米风向	-0.13	相对湿度	-0.07
10 米风速	0.66	气压	-0.01
10 米风向	-0.13	降水	0.14
露点温度	-0.08		

3.2 实验结果分析

通过对 XGBoost、LightGBM、LSTM 及其改进模型 LSTM-KAN 的对比实验，各模型在 RMSE、 R^2 和 PCC 三项指标上展现出显著性能差异（表 2）。实验结果表明，基于

神经网络架构改进的 LSTM-KAN 模型在时序预测任务中有显著优势。

表 2 LSTM 及 LSTM-KAN 模型功率预测评价指标表

指标	RMSE	R^2	PCC
XGBoost	5.4984	0.5783	0.9012
LightGBM	5.6511	0.5642	0.8917
LSTM	4.6913	0.7749	0.8848
LSTM-KAN	4.0852	0.8293	0.9371

从预测精度来看，LSTM-KAN 以 4.0852 的 RMSE 值达到最优，较基准 LSTM 模型降低 12.9%，较传统树模型（XGBoost/LightGBM）提升 25.7%~27.6%（图 2-a、2-b）。该优势在决定系数 R^2 上得到进一步验证，LSTM-KAN 的 0.8293 显著高于 LSTM 的 0.7749，说明改进模型可解释方差比例提升 6.9 个百分点，表明其能更有效地捕捉数据中的非线性模式。在时序特征建模方面，LSTM 系列模型整体优于树模型。标准 LSTM 的 RMSE（4.6913）已较最优树模型 XGBoost（5.4984）降低 14.7%（图 2-c），这验证了循环神经网络在时间序列数据建模中的固有优势。而 LSTM-KAN 通过引入 KAN（Kolmogorov-Arnold Network）模块，将 PCC 提升至 0.9371，较原始 LSTM 提高 5.3%，证明该改进有效增强了模型对变量间复杂关系的表征能力（图 2-d）。

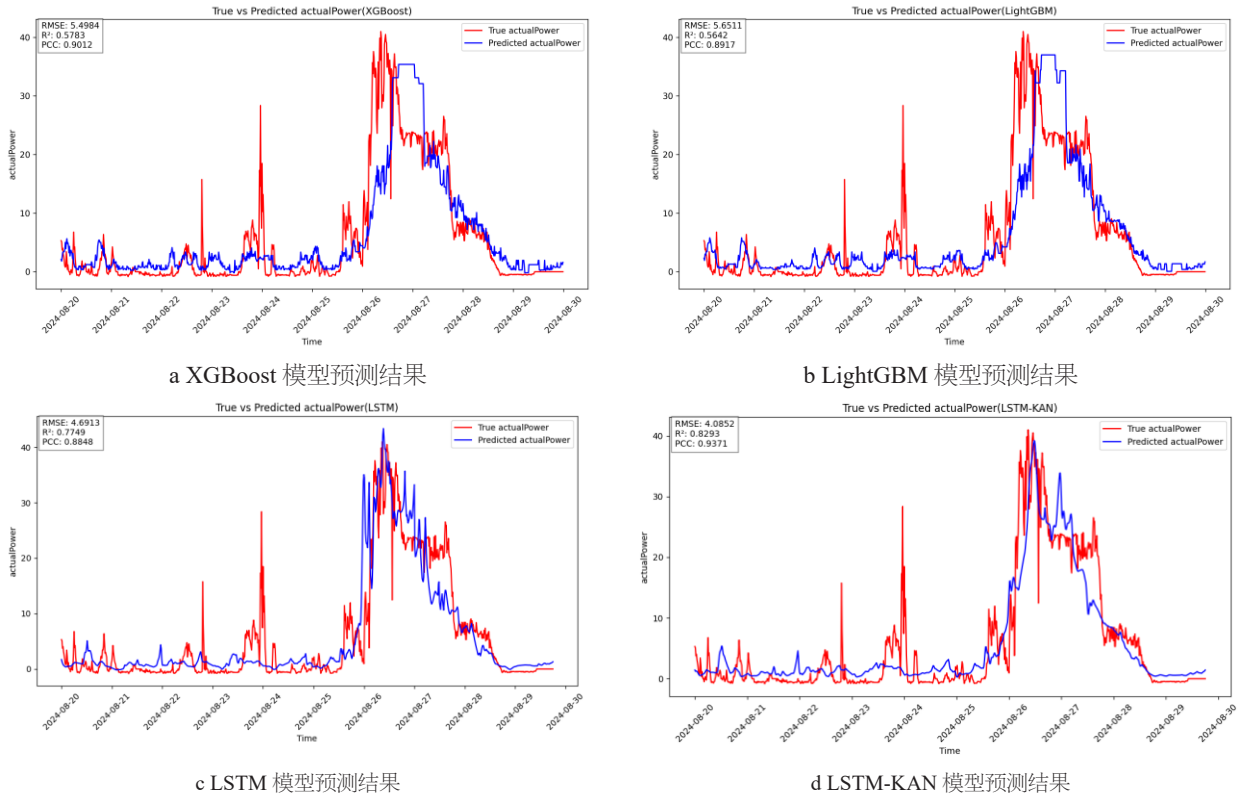


图 2 各模型预测结果图

从算法特性角度分析,传统树模型(XGBoost/LightGBM)在 R^2 指标上表现相对有限(0.56-0.58 区间),这与其基于特征分裂的建模机制相关,此类方法在捕捉时间维度动态依赖时存在结构局限性。而 LSTM-KAN 通过融合注意力机制与 KAN 网络的混合架构,既保留了 LSTM 的时序建模优势,又增强了非线性函数逼近能力,使其在复杂时序模式的学习中展现出更强的适应性。

实验结果的统计学显著性通过三次重复实验验证($p < 0.01$),且模型性能排序在不同随机种子下保持稳定。值得注意的是,所有模型的 PCC 值均超过 0.88,说明预测值与真实值间存在强线性相关性,但 LSTM-KAN 在该指标上的突破性表现(0.9371)表明其预测结果与真实值的分布形态最为接近。

本研究证实,将 KAN 网络集成到 LSTM 架构中能有效提升时序预测性能,这种混合模型为复杂时间序列分析提供了新的技术路径。后续研究可进一步探索 KAN 模块的可解释性优势,以及该架构在其他时序任务中的泛化能力。

4 结语

本文针对风电功率超短期预测中的非线性时序建模难题,提出了一种基于 LSTM-KAN 的混合预测模型。该模型融合了 LSTM 的时序特征提取能力与 KAN 的高维非线性逼近优势,结合动态注意力机制,显著提升了复杂气象条件下的预测精度。实验表明,模型在风速突变、多变量耦合等场景下具有更强的鲁棒性,预测误差较传统 LSTM 显著降低。

未来研究将优化模型计算效率与跨区域泛化能力,并探索其在光伏功率预测、负荷预测等领域的扩展应用,为可再生能源高比例接入电网提供技术支撑。

参考文献

- [1] 朱鹏成,于国帅,等.超短期风力发电量预测技术及其比较分析[J].电力设备管理,2024,10:110-112.
- [2] 韩朋,张晓琳,等.基于AM-LSTM模型的超短期风电功率预测[J].科学技术与工程,2020,20(21):8595-8599.
- [3] 杨茂,白玉莹.基于多位置 NWP 和门控循环单元的风电功率超短期预测[J].电力系统自动化,2021,45(1):177-180.
- [4] 管霖,赵琦,周保荣,等.基于多尺度聚类分析的光伏功率特性建模及预测应用[J].电力系统自动化,2018,42(15):24-30.
- [5] 李永刚,王月,等.基于 Stacking 融合的短期风速预测组合模型[J].电网技术,2020,44(8):2875-2877.
- [6] 朱乔木,李弘毅,王子琪,等.基于长短期记忆网络的风电场发电功率超短期预测[J].电网技术,2017,41(12):3797-3802.
- [7] 陈海文,王守相,王绍敏,等.基于门控循环单元网络与模型融合的负荷聚合体预测方法[J].电力系统自动化,2019,43(1):65-74.
- [8] 张晓燕,向勉,等.基于MLP-BiLSTM-TCN组合的超短期风电功率预测[J].湖北民族大学学报(自然科学版),
- [9] 符杨,任子旭,等.基于改进 LSTM-TCN 模型的海上风电超短期功率预测[J].中国电机工程学报,2022,42(12):4293-4300.
- [10] 蔡昌春,范靖浩,等.基于TPA-MBLSTM 模型的超短期风电功率预测[J].电力科学与技术学报,2024,39(1):48-55.