

DOI: 10.19783/j.cnki.pspc.250170

基于 KAN 和误差时间序列自相关性特征的 多时间尺度光伏功率预测

李嵩^{1,2}, 夏成军^{1,2}, 刘译夫^{1,2}, 汪愉康^{1,2}, 刘育成^{1,2}

(1 华南理工大学电力学院, 广东 广州 510640; 2. 广东省新能源电力系统智能运行与控制企业重点实验室, 广东 广州 510663)

摘要: 针对日前预测模型在非线性特征提取能力上的不足, 以及日内预测时功率输出波动特征复杂导致模型学习困难的问题, 提出了基于科尔莫戈洛夫-阿诺尔德网络(Kolmogorov-Arnold networks, KAN)和误差时间序列自相关性特征的多时间尺度光伏功率预测方法。首先, 在日前预测阶段, 设计了以 KAN 为基本单元的预测模型, 通过残差连接改进的深层 KAN 提取数据空间特征, 并结合多头注意力机制提取数据时序特征, 显著提升了模型对各种气候特征的提取能力。然后, 在日内预测阶段, 基于日前预测结果, 结合误差时间序列的自相关性特征进行间接预测, 显著降低了预测序列的波动幅度, 降低了模型的学习难度。最后, 基于某光伏功率预测大赛提供的数据进行实验, 结果表明, 所设计的前日预测模型相较于长短时记忆(long short-term memory, LSTM)网络、Transformer 等模型, 预测结果的均方误差(mean square error, MSE)至少降低 3.8%。所提出的日内间接预测方法相比于直接预测方法, 预测结果的 MSE 降低 38.3%。

关键词: 光伏功率预测; 多时间尺度; KAN; 误差时间序列的自相关性特征

Multi-time scale PV power forecasting based on KAN and error time series autocorrelation features

LI Song^{1,2}, XIA Chengjun^{1,2}, LIU Yifu^{1,2}, WANG Yukang^{1,2}, LIU Yucheng^{1,2}

(1. School of Electric Power, South China University of Technology, Guangzhou 510640, China; 2. Key Laboratory for Guangdong Provincial New Energy Power System Intelligent Operation and Control Enterprise, Guangzhou 510663, China)

Abstract: To address the limited nonlinear feature extraction capability of day-ahead forecasting models and the difficulty of model learning due to complex power output fluctuations in intraday forecasting, a multi-time scale photovoltaic (PV) power forecasting method based on Kolmogorov-Arnold networks (KAN) and the autocorrelation features of error time series is proposed. First, in the day-ahead forecasting stage, a forecasting model using KAN as the basic building block is designed. A deep KAN architecture enhanced with residual connections is used to extract spatial features, while a multi-head attention mechanism is employed to extract temporal features, significantly improving the model's ability to capture diverse climatic characteristics. Then, in the intraday forecasting stage, based on the day-ahead forecasting results, indirect prediction is performed by incorporating the autocorrelation features of the error time series. This approach significantly reduces the fluctuation range of the predicted sequence and lowers the learning difficulty of the model. Finally, experiments conducted using data provided by a PV power forecasting competition demonstrate that the proposed day-ahead model reduces the mean square error (MSE) by at least 3.8% compared with long-short-term memory (LSTM) and Transformer models. Compared with direct forecasting, the MSE of the forecasting results is reduced by 38.3%.

This work is supported by the National Key Research and Development Program of China (No. 2022YFB2403500).

Key words: PV power forecasting; multi-time scale; KAN; autocorrelation features of error time series

基金项目: 国家重点研发计划项目资助 (2022YFB2403500)

0 引言

加快光伏发展是推动能源低碳转型、如期实现“双碳”目标的重要举措之一^[1-3]。近年来,大规模光伏发电并网导致电源侧出力的不确定性上升,给电力系统的运行与调度带来挑战^[4-5]。对光伏出力进行可靠的预测,有利于弃光量的减少、太阳能资源的灵活利用^[6-7]。

深度学习方法现在已成为光伏预测领域众多学者的研究对象^[8-9]。目前常用的深度学习模型有:长短时记忆神经^[10-12](long short-term memory, LSTM)网络、卷积神经网络^[13-14](convolutional neural network, CNN)、Transformer^[15-17]及它们的组合模型等。然而,上述所提模型都是将不可学习的激活函数放置于神经元上,对模型的非线性拟合能力提升有限,从而限制了模型对非线性特征的提取能力。为提高神经网络模型的非线性拟合能力,文献[18]提出了科尔莫戈洛夫-阿诺尔德网络(Kolmogorov-Arnold networks, KAN),与上述网络模型不同, KAN 将可学习的激活函数放置在权重上,极大地提高了网络模型对复杂函数的拟合能力。现已有较多学者将 KAN 应用于预测领域,例如文[19]提出了一种基于 KAN 的净负荷概率预测方法,通过实验对比了包括 FCN、LSTM 和 CNN 在内的多种模型,突出了 KAN 在概率分布建模上的优越性能。文献[20]提出了一种将重构的二次分解-集成(reconstruction-based secondary decomposition-ensemble, RSDE)框架与 KAN 结合的电价预测方法, KAN 处于模型的顶层,用于将底层模型提取的时序特征映射到最终输出。实际光伏输出功率和数值天气预报之间存在非常复杂的时空映射关系,而原始的 KAN 并没有对数据的时序特征进行捕捉。文献[19]的预测方法没有充分利用输入特征的时序信息,文献[20]虽然利用底层模型提取了数据的时序信息,但底层模型并非 KAN。若能够将 KAN 的强大拟合能力应用到时序信息的提取,可进一步提高模型的预测性能。

光伏功率的预测精度受预测时间尺度的影响较大,尤其是在极端天气状况下,日前阶段几乎无法预测输出功率的急剧变化,为降低预测误差,有必要在日内时刻进行短时间尺度的滚动预测。文献[21]通过对比预测日的气象数据和历史气象数据挑选相似日集合,并结合与输出功率相关性较高的实时预测气象数据构成输入特征向量,输入至超短期预测模型中得到预测结果。文献[22]将实时采集的辐照度、风速、温度及功率信息作为特征向量进行超短期滚动预测。上述文献在日内预测时采用的是一种直接

预测功率的方法,模型需要捕捉功率序列在时序上的波动特征。通常来说,数据的波动范围越大,其规律越复杂,模型学习难度越大。实际功率受各种因素影响往往在较大范围内波动,模型直接对功率的波动规律进行学习,难以达到理想的预测效果。

为进一步提高光伏功率的预测精度,本文在日前和日内两个时间尺度上对上述文献中存在的问题进行改进。在日前预测阶段,本文以 KAN 为基本单元构建预测模型,结合残差连接和多头注意力机制提高模型对各种非线性天气特征的提取能力。在日内预测阶段,本文提出了一种基于误差时间序列自相关性特征的间接预测方法,首先利用实时信息预测日前预测误差,再将预测的误差结果与日前预测结果叠加,得到功率预测值。相比于直接预测方法,该方法能够有效降低预测序列的波动范围,提高模型的预测精度。

1 日前预测

1.1 科尔莫戈洛夫-阿诺德网络

KAN 基于 Kolmogorov-Arnold 表示定理,其核心思想为:有界域上的多元连续函数可以表示为有限单变量连续函数的两层嵌套叠加,如式(1)所示。

$$f(\mathbf{x}) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^n \phi_{qp}(x_p) \right) \quad (1)$$

式中: f 为多元连续函数; $\mathbf{x} = [x_1, \dots, x_p, \dots, x_n]$ 为多元函数的输入, n 为输出维度; Φ_q 为第 q 个外层单变量连续函数; ϕ_{qp} 为第 q 个外层的第 p 个内层单变量连续函数。

以此理论为基础,文献[18]将 ϕ_{qp} 、 Φ_q 作为可学习的激活函数 $\eta(x)$, 提出了 KAN。

$$\eta(x) = \omega \left[x / (1 + e^{-x}) + \sum_b c_b B_b(x) \right] \quad (2)$$

式中: ω 为控制函数整体大小的可学习参数; c_b 为第 b 个 B 样条曲线 $B_b(x)$ 的可学习权重参数。

为了更一般化网络,并增强网络的非线性映射能力,文献[18]利用式(3)所示的二维单变量函数矩阵定义了 KANlayer。

$$\Psi = \begin{bmatrix} \eta_{11} & \eta_{12} & \cdots & \eta_{1J} \\ \eta_{21} & \eta_{22} & \cdots & \eta_{2J} \\ \vdots & \vdots & \eta_{hj} & \vdots \\ \eta_{H1} & \eta_{H2} & \cdots & \eta_{HJ} \end{bmatrix} \quad (3)$$

式中: η_{hj} 为二维单变量函数矩阵 Ψ 的第 h 行、第 j 列元素; 下标“ H ”为 KANlayer 的输出维度,下标

“ J ”为 KANlayer 的输入维度。后文用 $\Psi(\cdot)$ 代表 KANlayer 的函数计算式, 用 (J, H) 表示 KANlayer 的维度。

基于上述理论, 式(1)即为两层 KAN 的表达式, 以二维输入变量, 一维输出变量为例, 式(1)对应的 KAN 结构如图 1 所示。在图 1 中, 输入层的 x_1 、 x_2 对应式(1)中的 x_p , 激活函数对应式(1)中的 ϕ_{qp} 、 Φ_q ; y 对应式(1)中的 $f(x)$ 。由多个 KANlayer 进行堆叠来构建深度的 KAN 可以拟合更加复杂的函数^[18]。图 1 所示的 KAN 仅由两个 KANlayer 堆叠而成(维度分别为(2,5)、(5,1)), 属于浅层 KAN, 其维度表示为(2,5,1), 增加 KANlayer 的数量即可构成深度 KAN。

1.2 基于 KAN 的日前预测模型

本文搭建的日前预测模型以 KAN 为基本单元,

将数值天气预报中各类天气特征作为输入, 输出光伏功率预测值, 包括空间特征提取模块、时序特征提取模块、时空特征融合及预测结果输出模块。预测时间间隔为 15 min, 一日共 96 个时间点。整个模型结构如图 2 所示。

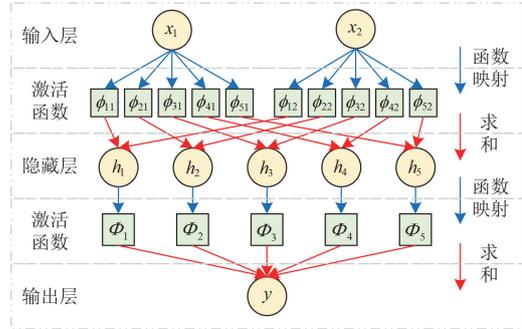


图 1 KAN 结构示意图

Fig. 1 Schematic diagram of KAN structure

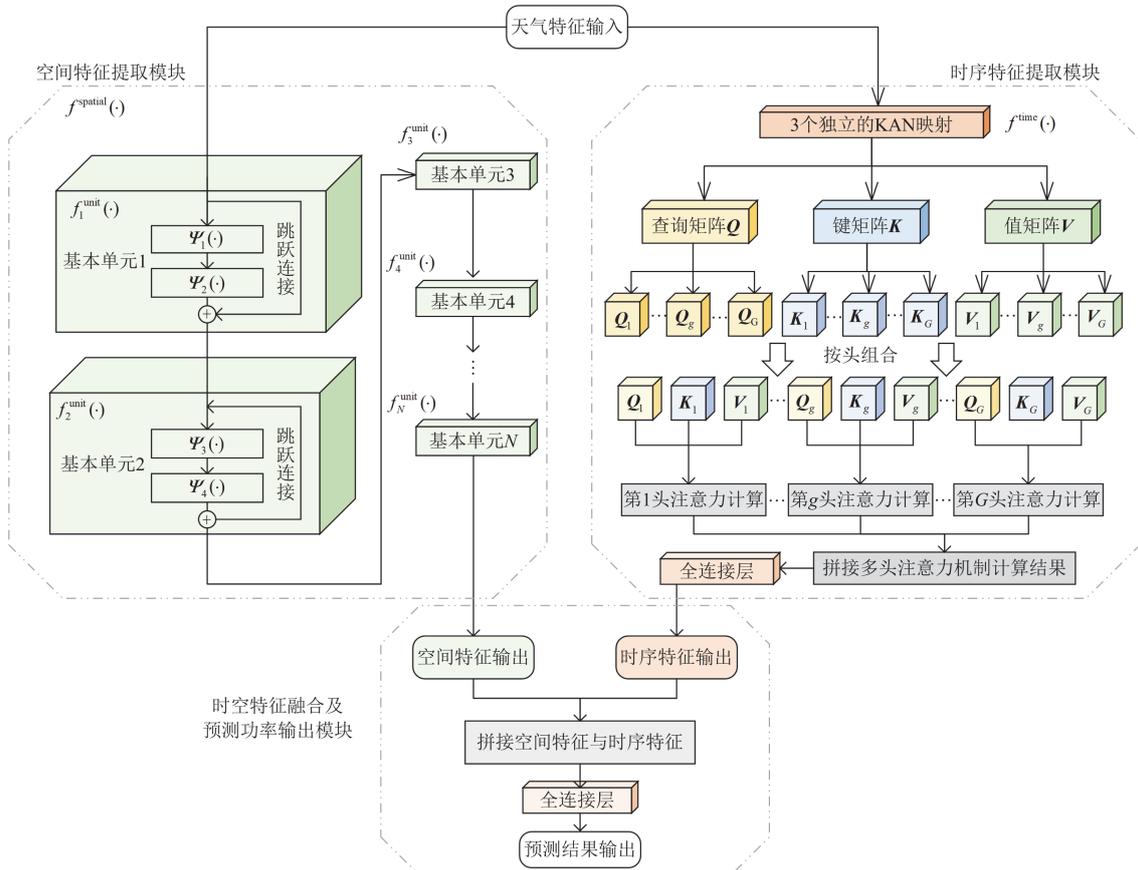


图 2 日前预测模型结构示意图

Fig. 2 Schematic diagram of the structure of the day-ahead prediction model

1.2.1 空间特征提取模块

光伏功率输出受多种气象因素的综合影响, 这些因素在空间维度上存在复杂的交互关系。因此, 在建模过程中需充分考虑气象变量在空间上的相互

作用特点。本文采用深度 KAN 作为空间特征提取模块, 用于刻画同一时间点各气象变量间的关联机制。然而, 原始深度 KAN 在网络层数加深时易出现梯度消失现象, 导致模型训练困难、收敛速度缓

慢。针对此问题, 本文借鉴文献[23]提出的残差结构, 通过引入跳跃连接对深度 KAN 进行改进, 并将其作为模型的空间特征提取模块。其结构示意图如图 2 中的“空间特征提取模块”, 对应的计算如式(4)所示。

$$\begin{cases} \mathbf{y}^{\text{spatial}} = f^{\text{spatial}}(\mathbf{X}^{\text{W}}) = f_{[U]}^{\text{unit}}(\mathbf{X}^{\text{W}}) \\ f_u^{\text{unit}}(\mathbf{X}^{\text{W}}) = \Psi_{2u}(\Psi_{2u-1}(\mathbf{X}^{\text{W}})) + \mathbf{X}^{\text{W}}, u < U \\ f_U^{\text{unit}}(\mathbf{X}^{\text{W}}) = \Psi_{2U}(\Psi_{2U-1}(\mathbf{X}^{\text{W}})) \end{cases} \quad (4)$$

式中: $\mathbf{y}^{\text{spatial}} = [y_1^{\text{spatial}}, \dots, y_t^{\text{spatial}}, \dots, y_{96}^{\text{spatial}}]^T$ 为空间特征提取模块的输出, 其中 y_t^{spatial} 为第 t 个时间点各输入变量的空间特征提取结果; $f^{\text{spatial}}(\cdot)$ 为空间特征提取模块的计算函数; \mathbf{X}^{W} 为维度为 $(96, N^{\text{W}})$ 的输入变量矩阵, 96 为一日时间点数, N^{W} 为天气特征变量维数; U 为基本单元的总数; $f_{[U]}^{\text{unit}}(\cdot)$ 为 $f_1^{\text{unit}}(\cdot)$ 到 $f_U^{\text{unit}}(\cdot)$ 的连续嵌套, 即 $f_{[u+1]}^{\text{unit}}(\cdot) = f_{u+1}^{\text{unit}}(f_{[u]}^{\text{unit}}(\cdot))$;

$f_u^{\text{unit}}(\cdot)$ 为第 u 个基本单元计算函数; $\Psi_{2u}(\cdot)$ 、 $\Psi_{2u-1}(\cdot)$ 分别为第 $2u$ 、 $2u-1$ 个 KANlayer, 维度分别为 $(2N^{\text{W}}+1, N^{\text{W}}+1)$ 、 $(N^{\text{W}}, 2N^{\text{W}}+1)$, 但 $\Psi_{2U}(\cdot)$ 的维度为 $(2N^{\text{W}}+1, 1)$ 。

此处需要注意的是, y_t^{spatial} 仅与 \mathbf{X}^{W} 的第 t 行 \mathbf{X}_t^{W} 有关, 即 $y_t^{\text{spatial}} = f^{\text{spatial}}(\mathbf{X}_t^{\text{W}})$, 当输入变量 \mathbf{X}^{W} 为矩阵时, $\Psi(\cdot)$ 会对 \mathbf{X}^{W} 的每一行向量分别进行运算, 且计算结果行数不变。之所以如此设计, 是因为空间特征提取模块专注于提取同一时间点各变量间的相互作用关系。

1.2.2 时序特征提取模块

从时序的角度来看, 天气因素对光伏功率输出的影响不仅是瞬时的, 而且还受到时间积累的影响。在建模时, 同样需要考虑各个天气变量在时序上隐含的内在特征, 时序特征提取模块用于从不同时间点的输入变量中提取出时序特征, 该模块在每个时间点的特征输出都必须综合考虑所有时间点的天气特征变量。而原始的 KAN 并没有针对数据的时序特征进行捕捉。针对此问题, 本文将多头注意力机制与 KAN 结合, 搭建时序特征提取模块, 其结构如图 2 中“时序特征提取模块”, 具体计算步骤如下所述。

1) 利用 3 个独立的 KAN 将输入特征分别映射到查询、键和值向量空间。得到查询矩阵 \mathbf{Q} 、键矩阵 \mathbf{K} 、值矩阵 \mathbf{V} 。

$$\begin{cases} \mathbf{Q} = \Psi_{\text{Q}}(\mathbf{X}^{\text{W}}) \\ \mathbf{K} = \Psi_{\text{K}}(\mathbf{X}^{\text{W}}) \\ \mathbf{V} = \Psi_{\text{V}}(\mathbf{X}^{\text{W}}) \end{cases} \quad (5)$$

式中: \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 的维度为 $(96, G)$, G 为多头注意力机制头数; $\Psi_{\text{Q}}(\cdot)$ 、 $\Psi_{\text{K}}(\cdot)$ 、 $\Psi_{\text{V}}(\cdot)$ 为 3 个独立的 KAN, 维度皆为 $(N^{\text{W}}, 2N^{\text{W}}+1, G)$, 其维度设置参考了式(1)。

2) 将 \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 按列分块为 G 头, 每个头独立进行注意力计算。

$$\mathbf{A}_g = \text{softmax}\left(\frac{\mathbf{Q}_g \mathbf{K}_g^T}{\sqrt{96}}\right) \mathbf{V}_g \quad (6)$$

式中: \mathbf{A}_g 为第 g 头注意力机制计算结果; \mathbf{Q}_g 、 \mathbf{K}_g 、 \mathbf{V}_g 分别为第 g 头查询矩阵、键矩阵、值矩阵。

3) 将多头注意力机制计算结果拼接成一个矩阵 $\mathbf{A} = [\mathbf{A}_1, \dots, \mathbf{A}_g, \dots, \mathbf{A}_G]$, 并通过全连接层输出时序特征提取结果 \mathbf{y}^{time} 。

$$\mathbf{y}^{\text{time}} = \mathbf{A} \boldsymbol{\omega}^{\text{time}} + \mathbf{b}^{\text{time}} \quad (7)$$

式中: $\mathbf{y}^{\text{time}} = [y_1^{\text{time}}, \dots, y_t^{\text{time}}, \dots, y_{96}^{\text{time}}]^T$, y_t^{time} 为第 t 个时间点的时序特征提取结果; $\boldsymbol{\omega}^{\text{time}}$ 、 \mathbf{b}^{time} 分别为时序特征提取模块全连接层的权重矩阵和偏置行向量, 皆为可学习参数。整个时序特征提取模块的映射函数用 $f^{\text{time}}(\cdot)$ 表示, 即 $\mathbf{y}^{\text{time}} = f^{\text{time}}(\mathbf{X}^{\text{W}})$ 。

此处需要注意的是, 虽然 y_t^{time} 是第 t 个时间点的时序特征提取结果, 但 y_t^{time} 与所有时间点的输入变量都有关(根据式(5)和式(6)多头注意力机制的计算原理可知), 即 $y_t^{\text{time}} \neq f^{\text{time}}(\mathbf{X}_t^{\text{W}})$ 。

1.2.3 时空特征融合及预测功率输出模块

时空特征融合模块用于将时序特征与空间特征进行融合, 并映射到光伏功率。为避免进一步加深整个模型的深度, 本文所提模型采用空间特征与时序特征并行提取的网络结构。提取完时空特征后, 将提取结果拼接, 并通过全连接层输出光伏出力预测结果, 该过程如式(8)所示。对应图 2 中的“时空特征融合及预测功率输出模块”。

$$\hat{\mathbf{p}} = \mathbf{F} \boldsymbol{\omega} + \mathbf{b} \quad (8)$$

式中: $\hat{\mathbf{p}} = [\hat{p}_1, \dots, \hat{p}_t, \dots, \hat{p}_{96}]^T$ 为整个日前预测模型输出的功率预测结果; \mathbf{F} 为时空特征提取结果拼接后的行向量, $\mathbf{F} = [y_1^{\text{spatial}}, \dots, y_{96}^{\text{spatial}}, y_1^{\text{time}}, \dots, y_{96}^{\text{time}}]$; $\boldsymbol{\omega}$ 、 \mathbf{b} 分别为时空特征融合及预测功率输出模块全连接层的权重矩阵和偏置行向量, 皆为可学习参数。

1.3 预测流程

日前预测的整个流程可分为模型训练阶段和实际应用阶段。在训练阶段中, 模型将历史天气特征作为输入, 并以实际功率为目标输出进行参数优化调整; 在实际应用阶段, 模型以相同的天气特征作为输

入, 输出预测结果。整个日前预测流程如图 3 所示。

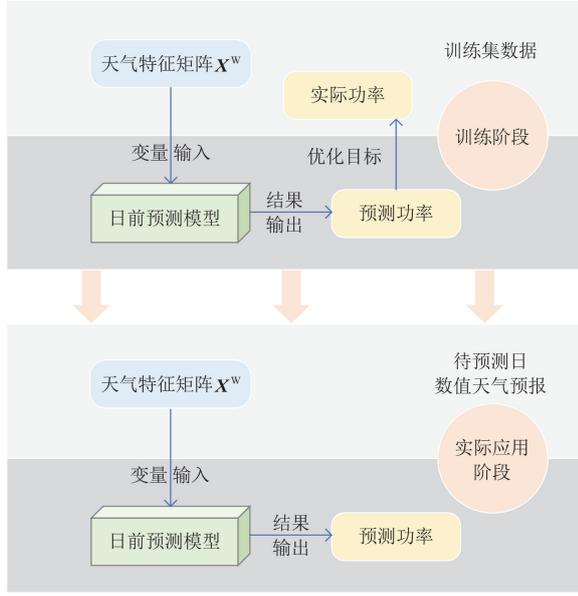


图 3 日前预测流程

Fig. 3 Day-ahead forecasting process

2 日内实时预测

日内预测通常需要提供未来 15 min~4 h 的预测结果, 而 15 min~4 h 的预测曲线可以通过 15 min 的单步外推得到, 因此, 本文主要对提前 15 min 的单步预测进行研究。

2.1 预测思路

本文将实际功率视为两部分之和: 1) 日前预测结果, 这部分在实时预测时属于已知部分; 2) 日前预测误差, 这部分通过实时信息进行滚动预测, 如图 4 所示, 其中的误差是指日前误差。

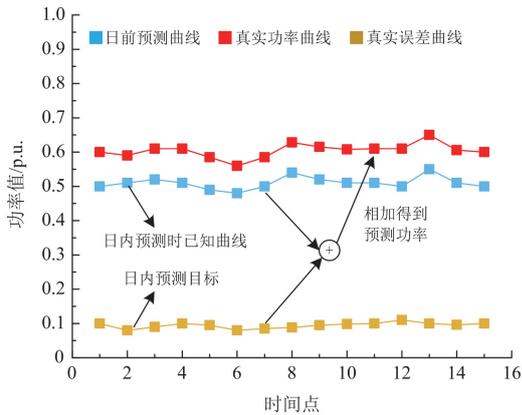


图 4 日内预测思路

Fig. 4 Intraday forecasting ideas

在日前预测中, 模型已充分利用数值天气预报等信息, 其预测误差通常在小范围内波动。通过预

测日前误差以间接预测实际功率的方式, 可有效降低预测难度。而日前误差主要源于天气预报的误差, 具有随时间累积的特性, 相邻时间点的误差值之间存在较强的自相关性, 本文利用这一自相关性规律, 来实时预测日前误差值。

2.2 预测模型的构成及其预测步骤

日内实时预测模型由 K_m 个子模型和一个权重判别器组成, K_m 为训练集划分的类别总数, 每个子模型利用实时信息对未来时刻的日前误差值进行预测。相较于日前预测, 日内预测对模型的推理时间要求更高, 由于 KAN 的推理时间较长, 日内预测不再使用 KAN 作为预测模型。本文选择 LSTM^[10] 作为日内实时预测的子模型, 处理误差时间序列的时序特征, 选择多层感知机(multilayer perceptron, MLP)^[24] 作为权重判别器, 输出每个子模型的选择权重。MLP 和 LSTM 皆采用数值矩阵进行计算, 两者推理时间相近, 而 KAN 采用函数矩阵进行运算, 推理时间较长。MLP 与 KAN 的模型维度可以进行统一比较, 两者的推理时间对比如表 1 所示。

表 1 MLP 与 KAN 推理时间对比

Table 1 Comparison of MLP and KAN inference time

模型名称(模型维度)	推理时间/s	推理样本数
MLP(4,256,512,4)	0.003	1
KAN(4,256,512,4)	7.353	1

日内实时预测采用滚动预测的形式, 假设当前时间点为 t , 对下一时间点 $t+1$ 的功率预测具体步骤如下。

1) 将第 $t-3$ 到第 t 个时间点的日前误差输入至所有子模型和权重判别器中。

2) 各子模型分别输出第 $t+1$ 个时间点的日前误差预测值, 设 $\Delta \hat{p}_{t+1,k}$ 为第 k 个子模型输出的预测值。权重判别器输出对 K_m 个子模型的选择权重向量 $\omega_{t+1}^s = [\omega_{t+1,1}^s, \dots, \omega_{t+1,k}^s, \dots, \omega_{t+1,K_m}^s]$ 。按照式(9)将 K_m 个子模型的日前误差预测值加权求和得到第 $t+1$ 时间点的最优日前误差预测值 $\Delta \hat{p}_{t+1}$ 。

$$\Delta \hat{p}_{t+1} = \sum_{k=1}^{K_m} \Delta \hat{p}_{t+1,k} \omega_{t+1,k}^s \quad (9)$$

3) 最后将 $\Delta \hat{p}_{t+1}$ 与第 $t+1$ 个时间点的日前预测值相加, 得到第 $t+1$ 个时间点的功率预测值。

2.3 模型训练方法

功率误差时间序列的变化规律在不同的环境状况下具有不同的表现。为提高各子模型学习的针对性, 本文以误差时间序列自相关性特征为数据集划分的标准, 来有效区分不同环境状况下的数据, 并采用两阶段训练的方式对模型进行训练。具体步

骤如下。

1) 训练集创建: 将日前训练集中的天气特征输入至已训练完成的日前预测模型, 得到对应的日前预测结果; 再将实际功率减去日前预测结果得到日前误差时间序列。每个误差时间序列样本以日为单位, 包含 96 个时间点。

2) 训练集划分: 利用式(10)提取误差时间序列样本的自相关性特征, 得到自相关性特征向量: $\mathbf{R}=[r_1, \dots, r_z, \dots, r_z]$, 将 \mathbf{R} 作为 K-means 算法^[25-26]的输入, 得到 K_m 个类别的样本数据, K_m 由轮廓系数^[25]和肘部法则^[26]的计算结果确定。

$$r_z = \frac{\sum_{t=1}^{96-z} (\Delta p_t - \Delta \bar{p})(\Delta p_{t+z} - \Delta \bar{p})}{\sum_{t=1}^{96} (\Delta p_t - \Delta \bar{p})^2} \quad (10)$$

式中: r_z 为误差时间序列的 z 阶自相关性系数, z 为滞后阶数, $z=1, 2, \dots, Z$, Z 为最大滞后阶数; Δp_t 为第 t 个时间点的误差值; $\Delta \bar{p}$ 为整个误差时间序列的平均值。

3) 两阶段训练: 首先用 K_m 个类别的样本数据分别训练 K_m 个子模型, 以真实误差值为子模型的目标输出优化其参数; 然后固定子模型的参数不变, 使用所有类别的样本对权重判别器进行训练, 以真实误差值作为式(9)中加权求和结果的目标, 调整权重判别器的参数。

2.4 预测评估指标

为了评估日内预测效果并对日前和日内的预测精度进行比较, 设计了 S^{MSE} 、 S^{RMSE} 、 S^{MAE} 3 个评价指标, 分别如式(11)一式(13)所示。

$$S^{\text{MSE}} = \frac{\sum_{t=1}^{96} (\Delta \hat{p}_t - \Delta p_t)^2}{\sum_{t=1}^{96} (\Delta p_t)^2} \quad (11)$$

$$S^{\text{RMSE}} = \frac{\sqrt{\sum_{t=1}^{96} (\Delta \hat{p}_t - \Delta p_t)^2}}{\sqrt{\sum_{t=1}^{96} (\Delta p_t)^2}} \quad (12)$$

$$S^{\text{MAE}} = \frac{\sum_{t=1}^{96} |\Delta \hat{p}_t - \Delta p_t|}{\sum_{t=1}^{96} |\Delta p_t|} \quad (13)$$

式中: S^{MSE} 、 S^{RMSE} 、 S^{MAE} 为日内误差与日前误差的比值, 三者分别对应均方误差、均方根误差、平均绝对误差。在日前误差为定值的情况下, 三者值越小, 说明日内预测精度越高; $\Delta \hat{p}_t$ 为第 t 个时间点

的日前误差预测值。

2.5 预测技术路线总结

本文预测方法的主要技术体现在预测方式和模型的训练方法上, 总体技术路线如图 5 所示, 其中的误差均指日前预测误差。

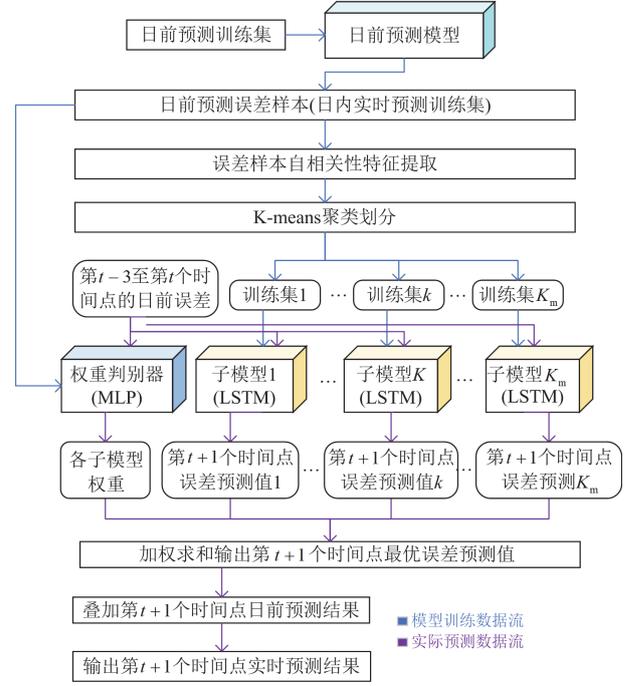


图 5 日内实时预测技术路线

Fig. 5 Intraday real-time prediction technical routes

3 算例分析

算例数据来源于某光伏功率预测大赛, 包含 2016 年—2018 年间某光伏电站 66 720 条数据, 每条数据包含一个时间点的辐照度、风速、风向、温度、压强、湿度 6 个天气特征预报值、实际光伏输出功率数值, 相邻数据间的时间间隔为 15 min。首先对原始数据进行清洗, 随后按照式(14)对数据进行 min_max 归一化处理。并将 66 720 条数据以日为单位划分为 695 个样本, 其中 600 个样本为训练集, 95 个样本为测试集。由于该数据集为脱敏数据, 故后续实验结果皆为基于 min_max 归一化后的数据值。

$$s'_i = \frac{s_i - \min(s)}{\max(s) - \min(s)} \quad (14)$$

式中: s'_i 为 s_i 归一化后数据值; s_i 为原始数据集 s 中的第 i 个数据。

3.1 日前预测

3.1.1 深度 KAN 改进效果分析

本节通过实验分析 1.2.1 节对原始 KAN 的改进效果。将辐照度、风速、风向、温度、压强、湿度

6个天气特征作为改进前、后深度KAN的输入,以实际功率为两者的目标输出,对两者的参数进行调整,记录两者在训练过程中的损失值,如图6所示。

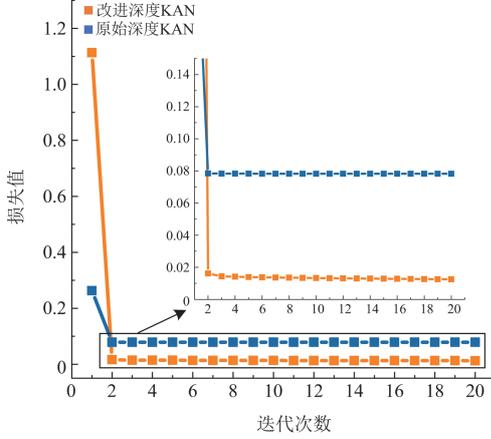


图6 改进前后深度KAN训练过程对比

Fig. 6 Comparison of the deep KAN training process before and after the improvement

从图6中可以观察到,改进后的深度KAN与原始深度KAN在完成第二轮迭代后就收敛到了各自的最优值,但改进后的深度KAN最终损失值明显低于原始深度KAN,说明本文对原始深度KAN的改进效果显著,有效提高了其拟合精度。

3.1.2 非线性特征提取能力分析

非线性拟合能力主要是强调模型的非线性拟合精度,而非线性特征提取能力则重点关注模型能否充分利用非线性输入特征来提高拟合精度。

本节对引言中所提模型和本文模型进行非线性特征提取能力对比。具体方法是:逐步将更强的非线性特征作为模型的输入,同时观察各模型在测试集上预测精度的变化,推测出模型对新增非线性特征的提取能力。在此之前,需要依据Spearman相关系数和Pearson相关系数来筛选有效输入特征并评估特征的非线性程度,计算表达式分别如式(15)、式(16)所示,各天气变量与实际功率的相关性系数计算结果如表2所示。

$$\rho = 1 - \frac{6 \sum_{l=1}^L (r_{x_l} - r_{y_l})^2}{L(L^2 - 1)} \quad (15)$$

式中: ρ 为变量 X 、 Y 的Spearman相关系数; r_{x_l} 、 r_{y_l} 分别为样本 l 在 X 、 Y 中的秩次; L 为样本总数。

$$r^P = \frac{\sum_{l=1}^L (x_l - \bar{x})(y_l - \bar{y})}{\sqrt{\sum_{l=1}^L (x_l - \bar{x})^2 \sum_{l=1}^L (y_l - \bar{y})^2}} \quad (16)$$

式中: r^P 为变量 X 、 Y 的Pearson相关系数; x_l 、 y_l 为 X 、 Y 的第 l 个样本; \bar{x} 、 \bar{y} 为 X 、 Y 的样本均值。

表2 天气变量与实际功率的相关性系数计算结果

Table 2 Calculation results of the correlation coefficient between weather variables and actual power

	辐照度	湿度	风速	温度	压强	风向
Pearson	0.895	0.409	0.371	0.649	0.682	0.571
Spearman	0.970	0.796	0.826	0.856	0.865	0.860

各天气变量与实际功率之间的Spearman相关系数都大于0.79,说明各天气特征都属于有效输入特征;Pearson相关系数反映的是变量间的线性关系强度,基于Pearson相关系数的差异将特征按非线性程度划分为4类:线性特征(辐照度)、普通特征(温度、压强)、非线性特征(风向)、强非线性特征(风速、湿度),据此构建4种不同的输入方案:

方案一:线性特征输入;

方案二:线性+普通特征输入;

方案三:线性+普通+非线性特征输入;

方案四:线性+普通+非线性+强非线性特征输入。

以测试集所有样本的平均预测精度为判别标准比较各模型对特征提取的效果。每个样本预测精度的评价指标采用均方误差(e^{MSE})、均方根误差(e^{RMSE})、平均绝对误差(e^{MAE})、决策系数(R^2)。上述指标的计算公式分别如式(17)~式(20)所示。

$$e^{\text{MSE}} = \frac{1}{96} \sum_{t=1}^{96} (p_t - \hat{p}_t)^2 \quad (17)$$

$$e^{\text{RMSE}} = \sqrt{\frac{1}{96} \sum_{t=1}^{96} (p_t - \hat{p}_t)^2} \quad (18)$$

$$e^{\text{MAE}} = \frac{1}{96} \sum_{t=1}^{96} |p_t - \hat{p}_t| \quad (19)$$

$$R^2 = 1 - \frac{\sum_{t=1}^{96} (p_t - \hat{p}_t)^2}{\sum_{t=1}^{96} (p_t - \bar{p})^2} \quad (20)$$

式中: p_t 为第 t 个时间点的真实功率值; \hat{p}_t 为第 t 个时间点的预测功率值; \bar{p} 为96个时间点的真实功率平均值。

利用训练好的模型在测试集上进行预测,各个方案实验结果如表3所示,表中的加粗数字表示该方案所有模型中的最优指标。其中,变化率是指同一模型相邻方案(当前方案相比上一方案)之间指标值的变化百分数。由于CNN-LSTM模型、CNN-

BiLSTM 模型和本文模型都是针对多输入特征而设计的, 而输入特征方案一中只有一个输入特征, 故方案一未对 CNN-LSTM 模型、CNN-BiLSTM 模型

和本文模型进行实验。其中, 本文模型的空间特征提取模块层数 N 和注意力头数 G 均设为 6。

表 3 各模型非线性特征提取能力对比实验结果

Table 3 Comparison experimental results of the nonlinear feature extraction ability of various models

方案类别	模型名称	e^{MSE}		e^{MAE}		e^{RMSE}		R^2	
		值	变化率%	值	变化率%	值	变化率%	值	变化率%
方案一	LSTM	0.0142	—	0.0666	—	0.1153	—	0.8467	—
	BiLSTM	0.0140	—	0.0657	—	0.1146	—	0.8489	—
	KAN	0.0154	—	0.0690	—	0.1204	—	0.8337	—
	Transformer	0.0147	—	0.0682	—	0.1170	—	0.8416	—
方案二	LSTM	0.0117	-17.27	0.0564	-15.30	0.0989	-14.21	0.8732	3.13
	BiLSTM	0.0116	-17.21	0.0560	-14.76	0.0990	-13.61	0.8749	3.06
	KAN	0.0121	-21.67	0.0565	-18.03	0.1000	-16.91	0.8697	4.32
	Transformer	0.0114	-21.88	0.0549	-19.52	0.0980	-16.25	0.8763	4.12
	CNN-LSTM	0.0128	—	0.0619	—	0.1074	—	0.8618	—
	CNN-BiLSTM	0.0128	—	0.0618	—	0.1078	—	0.8613	—
	本文模型	0.0116	—	0.0548	—	0.0978	—	0.8751	—
方案三	LSTM	0.0118	0.93	0.0566	0.39	0.0997	0.80	0.8720	-0.13
	BiLSTM	0.0119	3.18	0.0573	2.37	0.1011	2.15	0.8709	-0.45
	KAN	0.0123	2.44	0.0578	2.18	0.1026	2.55	0.8665	-0.37
	Transformer	0.0118	3.45	0.0557	1.37	0.0993	1.27	0.8720	-0.49
	CNN-LSTM	0.0130	2.01	0.0626	1.18	0.1088	1.33	0.8590	-0.32
	CNN-BiLSTM	0.0130	0.94	0.0622	0.64	0.1082	0.36	0.8600	-0.15
	本文模型	0.0111	-4.21	0.0532	-2.85	0.0948	-3.06	0.8803	0.60
方案四	LSTM	0.0106	-10.57	0.0529	-6.68	0.0941	-5.56	0.8855	1.55
	BiLSTM	0.0109	-8.72	0.0527	-8.10	0.0949	-6.16	0.8822	1.29
	KAN	0.0109	-11.84	0.0547	-5.28	0.0973	-5.15	0.8823	1.82
	Transformer	0.0108	-8.85	0.0531	-4.62	0.0955	-3.76	0.8833	1.30
	CNN-LSTM	0.0126	-3.66	0.0612	-2.27	0.1067	-1.91	0.8641	0.60
	CNN-BiLSTM	0.0126	-2.90	0.0613	-1.43	0.1069	-1.18	0.8640	0.47
	本文模型	0.0102	-8.03	0.0507	-4.87	0.0904	-4.58	0.8899	1.09

对表 3 实验结果进行分析。

1) 在方案三和方案四中, 本文模型的 e^{RMSE} 、 e^{MSE} 、 e^{MAE} 、 R^2 皆为最优, 在方案二中, 本文模型的 MAE 和 RMSE 最优, 而在 e^{MSE} 和 R^2 指标上, 本文模型与最优的 Transformer 模型仅有细微差异。说明当输入变量相同时, 本文模型能够更好地提取数据的特征, 从而达到更高的预测精度。

2) 表格中红色数据代表该模型在新增输入特征后, 其预测指标值反而发生恶化的情况。方案三相比于方案二, 加入了“风向”特征作为模型的输入, 基于前文的分析, 风向特征是属于非线性的有效输入特征。可以看到, 除本文模型外其他模型的所有指标都发生了恶化, 本文模型在各个指标上都有较大程度的优化, 说明了其他模型在非线性的特征提取能力上存在不足。

3) 从方案三变化至方案四时, 新增了强非线性特征风速与湿度。虽然本文模型的指标优化程度并非最优, 但指标值始终优于其他模型, 表明本模型在实际应用中具有更强的非线性特征提取能力和预测性能。

4) 在各个方案中对比 KAN 和本文模型, 可以观察到, 本文模型的预测误差更小, 这主要是因为 KAN 并未对输入数据的时序特征进行捕捉, 而本文模型将注意力机制与 KAN 结合, 提取了数据的时序特征, 达到了更好的预测效果。

将全部有效特征作为各个模型的输入, 在测试集中随机抽取 4 个样本日进行可视化展示, 结果如图 7 所示。可以观察到, 在天气状况正常或稍有波动的情况下(天气类别 1、2、3), 本文模型相比于现有模型都有明显的优势, 而在天气状况波动较大或

非常极端的情况下(天气类别 4), 本文模型与现有模型差距不大, 且预测精度不够理想。事实上, 极端天气状况下的预测精度主要受天气预报准确度的影

响, 难以通过优化预测模型的方式来有效降低日前误差。为降低极端天气带来的影响, 需要在日内时刻进行误差预测来修正。

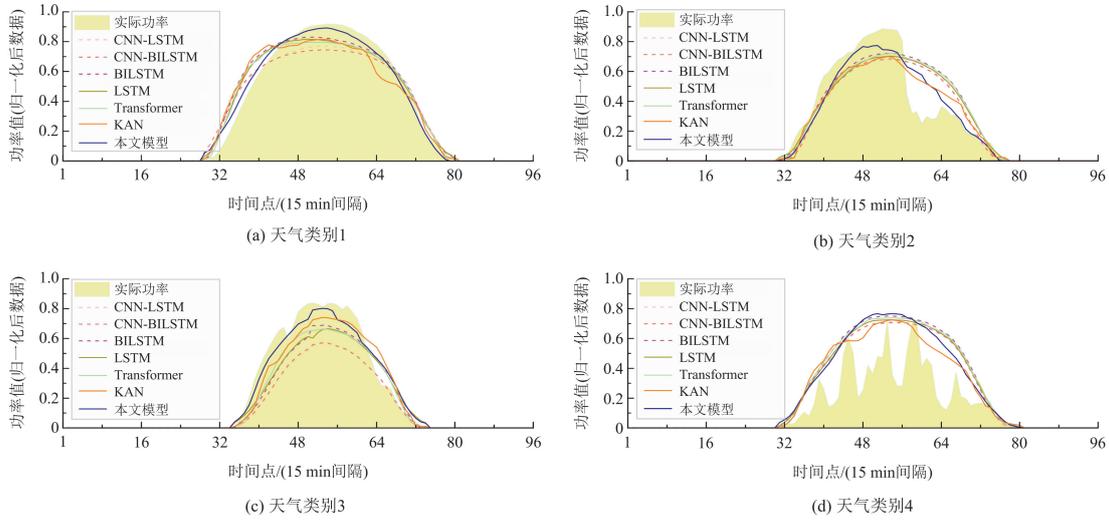


图 7 各模型预测结果可视化展示

Fig. 7 Visual display of prediction results of various models

3.2 日内实时预测

3.2.1 训练集创建

按 2.3 节中步骤 1) 的要求创建日内实时预测的训练集, 其中日前预测模型选择本文 1.2 节设计模型, 天气特征的输入方案选择方案四。

3.2.2 多方法对比分析

方法一: 本文方法, 详见第 2 节。式(10)中最大滞后阶数设置为 5, 综合肘部法则和轮廓系数的计算结果(两者计算结果如图 8 所示), 选取 K-means 算法聚类数目为 4, 聚类结果如表 4 所示。

误差作为输入, 第 $t+1$ 个时间点的日前误差作为输出。最后叠加日前误差预测结果和日前预测结果得到预测功率。

方法三: 利用所有数据训练一个 LSTM 模型。日内实时预测时, 将第 $t-3$ 到第 t 个时间点的功率作为输入, 预测第 $t+1$ 个时间点的功率。

以 2.4 节设计的 3 个指标来评估测试集的预测效果, 比较各方法优劣, 结果如表 5 所示。

表 5 各方法日内实时预测评价结果

Table 5 Evaluation results of real-time prediction of various methods within a day

方法	评价指标			平均值
	S^{RMSE}	S^{MAE}	S^{MSE}	
一	0.484	0.417	0.276	0.392
二	0.506	0.460	0.289	0.419
三	0.629	0.611	0.447	0.563

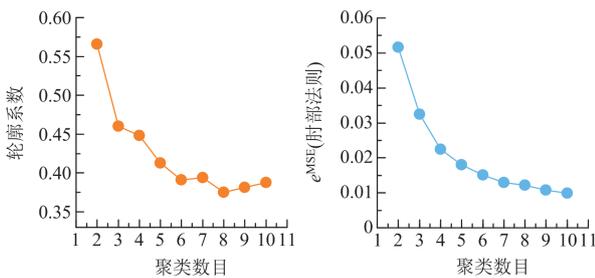


图 8 肘部法则和轮廓系数计算结果

Fig. 8 Calculation results of elbow rule and contour coefficient

表 4 训练样本聚类结果

Table 4 Clustering results of training samples

聚类类别	1	2	3	4
样本数量	155	164	192	89

方法二: 利用所有数据训练一个 LSTM 模型。日内实时预测时, 将第 $t-3$ 到第 t 个时间点的日前

对实验结果进行分析。

1) 对比方法二和方法三可知, 通过预测误差来间接预测实际功率的方法比直接预测实际功率的方法预测精度更高, 因为功率序列的波动相比于误差序列的波动更大, 对于模型来说, 学习实际功率序列的波动规律更加复杂, 所以预测效果不佳。

2) 方法一和方法二都是间接预测方法, 不同之处在于, 方法一(本文方法)对数据集进行了划分, 并训练了权重判别器来实时选择预测模型。从实验结果来看, 本文方法进一步提高了预测精度。

3.2.3 本文方法深入分析

为说明数据划分的必要性以及权重判别器在模型选择中的重要作用,使用方法一模型中的子模型 1—子模型 4(后文简称子模型 1—子模型 4)、方法一模型(其预测结果为权重判别器对各子模型预测结果的加权值)和方法二模型对测试样本进行逐个预测。对于每一个测试集样本,预测结果 S^{MSE} 、 S^{RMSE} 、 S^{MAE} 平均值最小的模型即为该测试样本的最优模型,并将该模型的最优样本数加 1。所有测试样本预测完成后,计算每个模型的最优比率,即该模型最优样本数 B 与测试集样本总数 S 之比,如式(21)所示。

$$R = \frac{B}{S} \times 100\% \quad (21)$$

为了区分数据划分和权重判别器的作用,表 6

表 6 各模型最优比率计算结果

Table 6 Optimal ratio calculation results of various models

模型类别	子模型 1	子模型 2	子模型 3	子模型 4	方法二模型	方法一模型
最优比率	38	16	29	6	11	—
$R/\%$	17	16	12	6	8	41

图 9 展示了日内、日前预测结果对比。其中,日前预测采用图 2 的模型,输入特征方案采用方案四。显然,在日前预测较为准确的情况下,本文方法能够进一步提高预测精度;在日前预测不够理想的情况下,本文方法能较大程度修正日前预测误差,实时减小极端天气带来的影响。进一步细致观察可以发现,在天气类别 2 的第 60—75 个时间点、天气类别 3 的第 50—70 个时间点以及天气类别 4 的第

的最优比率计算结果分为两类。

1) 第一类为子模型 1—子模型 4 和方法二模型之间的比较。从结果中可以看到,利用部分样本进行训练的子模型 1,最优比率远高于利用所有样本训练的方法二模型,说明了误差时间序列的特征变化多样,单个模型学习到所有样本的特征较为困难,有必要对样本进行划分训练。在整个测试集中,子模型 1—子模型 4 的最优比率皆不为 0,说明在不同的情况下,宜使用不同的子模型进行预测,因此,需要权重判别器来对各个子模型进行选择。

2) 第二类为方法一模型及子模型 1—子模型 4、方法二模型之间的比较。此时,方法二模型的最优比率进一步降低,方法一模型的预测结果最优比率最高。说明本文设计的权重判别器能够自适应误差的特征变化,根据已知信息对各个子模型进行有效判别。

40—70 时间点,输出功率波动较大,预测曲线出现了滞后于实际功率曲线的现象。事实上,当天气状态稳定时,功率变化规律性强,模型能够准确捕捉功率变化趋势;而在天气剧烈波动时,功率变化呈现随机性,模型在训练过程中发现功率变化几乎无规律可循,趋向于采用前一刻的实际值作为预测值来降低平均预测误差,这是模型在训练过程中学习到的最优预测策略。

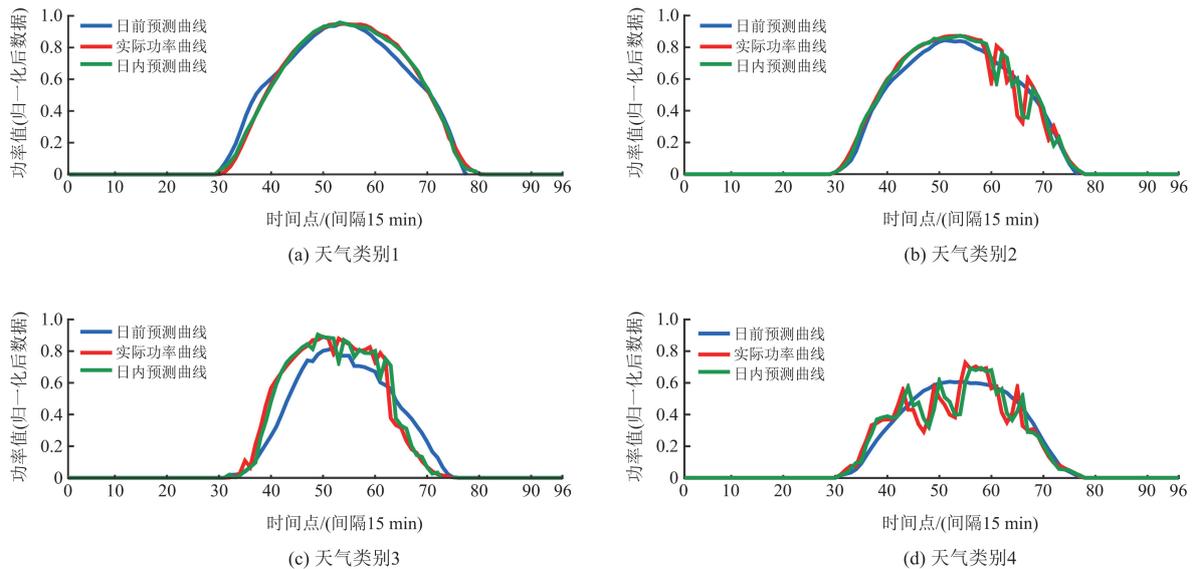


图 9 各天气状况下日内、日前预测结果对比

Fig. 9 Comparison of intraday and day-ahead forecast results under various weather conditions

4 结论

本文对光伏功率进行了日前和日内提前 15 min 两个时间尺度的预测, 得出以下结论。

1) 针对原始 KAN 在时序特征提取能力上的不足, 本文引入多头注意力机制加以改进, 与原始 KAN 相比, 本文模型预测结果的 e^{MSE} 、 e^{RMSE} 、 e^{MAE} 分别降低 6.4%、7.3%、7.1%。与 Transformer、LSTM 等模型相比, 本文所提日前预测模型的非线性特征提取能力更强, 预测结果的 MSE 较 LSTM 降低了 3.8%。

2) 本文提出的日内间接预测方法根据误差时间序列的自相关性特征进行样本聚类, 并引入权重判别器动态分配子模型权重, 有效提升了预测精度。相比于直接预测方法, 本文方法预测结果的 e^{MSE} 、 e^{RMSE} 、 e^{MAE} 分别降低了 38.3%、23.1%、31.8%。

本文从模型结构入手提升了预测精度, 选取辐照度、风速、风向、温度、压强、湿度 6 项气象特征作为输入示例。考虑到光伏功率还受云层、日照角度、地理位置及季节性等多因素影响, 本文尚未全面纳入, 后续研究将融合更多数据源, 提升预测的全面性与准确性。

参考文献

- [1] 许洪华, 邵桂萍, 鄂春良, 等. 我国未来能源系统及能源转型现实路径研究[J]. 发电技术, 2023, 44(4): 484-491.
XU Honghua, SHAO Guiping, E Chunliang, et al. Research on China's future energy system and the realistic path of energy transformation[J]. Power Generation Technology, 2023, 44(4): 484-491.
- [2] 张海涛, 李文娟, 李雪峰, 等. 基于变分模态分解和时间注意力机制 TCN 网络的光伏发电功率预测[J]. 电测与仪表, 2024, 61(12): 156-163.
ZHANG Haitao, LI Wenjuan, LI Xuefeng, et al. Photovoltaic power forecasting based on TPA-TCN model and variational modal decomposition[J]. Electrical Measurement & Instrumentation, 2024, 61(12): 155-163.
- [3] 付小标, 侯嘉琪, 李宝聚, 等. 一种二模态天气分型方法及其在光伏功率概率预测的应用[J]. 发电技术, 2024, 45(2): 299-311.
FU Xiaobiao, HOU Jiaqi, LI Baoju, et al. A two-modal weather classification method and its application in photovoltaic power probability prediction[J]. Power Generation Technology, 2024, 45(2): 299-311.
- [4] 陈阳, 谢丽蓉, 马兰, 等. 考虑光伏预测误差兼顾平抑波动的双层储能运行策略[J]. 电力工程技术, 2023, 42(1): 70-79.
CHEN Yang, XIE Lirong, MA Lan, et al. Double-layer energy storage operation strategy that takes into account PV output error and smooths out fluctuations[J]. Electric Power Engineering Technology, 2023, 42(1): 70-79.
- [5] 杨康, 李蓝青, 李艺丰, 等. 一种新型分布式光伏出力区间预测方法[J]. 发电技术, 2024, 45(4): 684-695.
YANG Kang, LI Lanqing, LI Yifeng, et al. A novel distributed photovoltaic output interval prediction method[J]. Power Generation Technology, 2024, 45(4): 684-695.
- [6] 杜文娟, 郝向坤, 陈珏. 光伏场经柔直并网振荡稳定性分析与抑制方法研究[J]. 电力工程技术, 2024, 43(3): 2-11, 51.
DU Wenjuan, HAO Xiangkun, CHEN Jue. Oscillation stability analysis and mitigation method of photovoltaic field connected to the grid via VSC-HVDC[J]. Electric Power Engineering Technology, 2024, 43(3): 2-11, 51.
- [7] 韩晓, 王涛, 韦晓广, 等. 考虑阵列间时空相关性的超短期光伏出力预测[J]. 电力系统保护与控制, 2024, 52(14): 82-94.
HAN Xiao, WANG Tao, WEI Xiaoguang, et al. Ultrashort-term photovoltaic output forecasting considering spatiotemporal correlation between arrays[J]. Power System Protection and Control, 2024, 52(14): 82-94.
- [8] LIU C, LI M, YU Y, et al. A review of multitemporal and multispatial scales photovoltaic forecasting methods[J]. IEEE Access, 2022, 10: 35073-35093.
- [9] 陈瑶, 陈晓宁. 基于自适应 Kmeans 和 LSTM 的短期光伏发电预测[J]. 电测与仪表, 2023, 60(7): 94-99.
CHEN Yao, CHEN Xiaoning. Prediction of short-term photovoltaic power generation based on adaptive Kmeans and LSTM[J]. Electrical Measurement & Instrumentation, 2023, 60(7): 94-99.
- [10] 王海军, 居蓉蓉, 董颖华. 基于时空关联特征与 B-LSTM 模型的分布式光伏功率区间预测[J]. 中国电力, 2024, 57(7): 74-80.
WANG Haijun, JU Rongrong, DONG Yinghua. Distributed photovoltaic power interval prediction based on spatio-temporal correlation feature and B-LSTM model[J]. Electric Power, 2024, 57(7): 74-80.
- [11] ZHANG H, SUN H, KANG L, et al. Prediction of health level of multiform lithium sulfur batteries based on incremental capacity analysis and an improved LSTM[J]. Protection and Control of Modern Power Systems, 2024, 9(2): 21-31.
- [12] 史加荣, 殷诏. 基于 GRU-BLS 的超短期光伏发电功率预测[J]. 智慧电力, 2023, 51(9): 38-45.
SHI Jiarong, YIN Zhao. Prediction of ultra short term photovoltaic power generation based on GRU-BLS[J]. Smart Power, 2023, 51(9): 38-45.
- [13] 欧阳福莲, 王俊, 周杭霞. 基于改进迁移学习和多尺

- 度 CNN-BiLSTM-Attention 的短期电力负荷预测方法[J]. 电力系统保护与控制, 2023, 51(2): 132-140.
- OUYANG Fulian, WANG Jun, ZHOU Hangxia. Short-term power load forecasting method based on improved hierarchical transfer learning and multi-scale CNN-BiLSTM-Attention[J]. Power System Protection and Control, 2023, 51(2): 132-140
- [14] 杨建, 常学军, 姚帅, 等. 基于 WT-CNN-BiLSTM 模型的日前光伏功率预测[J]. 南方电网技术, 2024, 18(8): 61-69.
- YANG Jian, CHANG Xuejun, YAO Shuai, et al. Day-ahead photovoltaic power forecasting based on WT-CNN-BiLSTM model[J]. Southern Power System Technology, 2024, 18(8): 61-69.
- [15] 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.
- YANG Pengwei, ZHAO Liping, CHEN Junfa, et al. Distributed photovoltaic ultra-short-term power forecasting method based on temporal analog matching approach and Transformer network modeling[J]. Electric Power, 2024, 57(12): 60-70.
- [16] 王瑞, 靳鑫鑫, 逯静. SVM-D-PE-BP-Transformer 短期光伏功率预测[J]. 电网与清洁能源, 2024, 40(8): 141-150.
- WANG Rui, JIN Xinxin, LU Jing. Short-term photovoltaic power forecasting based on SVM-D-PE-BP-Transformer[J]. Power System and Clean Energy, 2024, 40(8): 141-150.
- [17] 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, 59.
- HUANG Li, GAN Hengyu, LIU Xingju, et al. Ultra-short-term photovoltaic power generation prediction based on Transformer encoder[J]. Smart Power, 2024, 52(5): 16-22, 59.
- [18] LIU Z, WANG Y, VAIDYA S, et al. Kan: Kolmogorov-arnold networks[J]. arXiv preprint arXiv: 2404.19756, 2024.
- [19] 刘栋, 郭国栋, 辛蜀骏, 等. 基于 KAN 的可解释性净负荷概率预测方法[J]. 电力系统自动化, 2025, 49(15): 123-132.
- LIU Dong, GUO Guodong, XIN Shujun, et al. Probabilistic prediction method for interpretable net load based on Kolmogorov-Arnold network[J]. Automation of Electric Power Systems, 2025, 49(15): 123-132.
- [20] 成润坤, 王辉, 刘达, 等. 融合 RSDE 框架与 KAN 算法的现货电价异质模型集成预测[J]. 中国电机工程学报, 2024, 44(24): 9645-9658.
- CHENG Runkun, WANG Hui, LIU Da, et al. Ensemble prediction of spot electricity prices using heterogeneous models by integrating the RSDE framework and KAN algorithm[J]. Proceedings of the CSEE, 2024, 44(24): 9645-9658.
- [21] 李练兵, 高国强, 陶鹏, 等. 基于 SFLA 和 MSISSA-ANFIS 的超短期光伏功率动态预测方法[J]. 太阳能学报, 2024, 45(10): 326-335.
- LI Lianbing, GAO Guoqiang, TAO Peng, et al. Ultra-short-term PV power dynamic prediction method based on SFLA and MSISSA-ANFIS [J]. Acta Energetica Sinica, 2024, 45(10): 326-335.
- [22] 黄泽, 毕贵红, 谢旭, 等. 基于 MBI-PBI-ResNet 的超短期光伏功率预测[J]. 电力系统保护与控制, 2024, 52(2): 165-176.
- HUANG Ze, BI Guihong, XIE Xu, et al. Ultra-short-term PV power prediction based on MBI-PBI-ResNet[J]. Power System Protection and Control, 2024, 52(2): 165-176.
- [23] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [24] 余凯峰, 吐松江·卡日, 张紫薇, 等. 基于级联 MCNN-MMLP 双残差网络的短期负荷预测[J]. 电力系统保护与控制, 2025, 53(2): 151-162.
- YU Kaifeng, TUSONGJIANG-Kari, ZHANG Ziwei, et al. Short-term load forecasting based on a cascade MCNN-MMLP double residual network[J]. Power System Protection and Control, 2025, 53(2): 151-162.
- [25] 张晓珂, 张辉, 戴小然, 等. 基于层次聚类和 BiLSTM 的光伏短期功率预测模型[J]. 智慧电力, 2024, 52(9): 41-48.
- ZHANG Xiaoke, ZHANG Hui, DAI Xiaoran, et al. Photovoltaic short-term power forecasting model based on hierarchical clustering & BiLSTM[J]. Smart Power, 2024, 52(9): 41-48.
- [26] 彭勃, 李耀东, 龚贤夫. 基于自编码的改进 K-means 光伏能源数据清洗方法[J]. 计算机科学, 2024, 51(增刊 1): 725-729.
- PENG Bo, LI Yaodong, GONG Xianfu. Improved K-means photovoltaic energy data cleaning method based on autoencoder[J]. Computer Science, 2024, 51(S1): 725-729.

收稿日期: 2025-02-24; 修回日期: 2025-07-01

作者简介:

李嵩(2001—), 男, 硕士研究生, 研究方向为新能源出力预测及电力系统优化调度; E-mail: 452785748@qq.com

夏成军(1974—), 男, 通信作者, 博士, 副教授, 研究方向为电力系统分析运行与控制。E-mail: cjxia@scut.edu.cn

(编辑 张颖)