

融合物理约束与多尺度特征的 Mamba-Transformer 超短期风电功率预测

余勇祥, 匡相宇, 韩 建, 高 波, 曾 晗, 李泽文

(华东交通大学电气与自动化工程学院, 南昌 330013)

摘 要: 提出一种融合物理约束与多尺度特征的 Mamba-Transformer 模型, 用于超短期风电功率预测。该模型通过自适应噪声完备经验模态分解风电功率数据, 提取不同频率下的非线性与非平稳特征, 并将分解后的模态函数与风电场气象数据输入 Mamba 模型, 挖掘局部时变特性。再利用 Transformer 模型的自注意力机制捕捉多尺度特征间的长距离依赖关系。最后, 结合一维 Jensen 尾流模型构建物理约束, 通过自适应加权机制整合物理模型与数据驱动模型预测结果。实验结果表明, 该文提出的模型与 Transformer 相比, 均方根误差 (RMSE) 和平均绝对误差 (MAE) 分别降低 41.29% 和 50.26%。该模型在提高风电功率预测精度的同时, 展现出较高的泛化能力。

关键词: 风电功率预测; 物理约束; 自适应噪声完备经验模态分解; Mamba; Transformer; 一维 Jensen 尾流模型

中图分类号: TM614

文献标志码: A

0 引 言

在应对能源危机和环境保护的背景下, 中国正加速发展风能等清洁能源, 风能作为清洁能源, 已成为中国能源结构中的重要一环^[1]。然而, 风能的间歇性和波动性对风电场的运行优化和电网的安全稳定性提出较高要求。

风电功率预测作为风电场高效运行的重要技术, 通常根据时间尺度分为超短期、短期、中期和长期预测。其中, 超短期风电功率预测, 通常在 0~4 h 内, 对电力系统的频率控制、备用容量调整和风力机优化运行至关重要, 因此本文聚焦于超短期预测。提高预测精度可减少能量损失, 提高能效和稳定性。

近年来, 根据应用方法的不同, 较为流行的超短期风电预测模型可分为统计模型、机器学习模型和组合模型。统计模型中, 刘名德等^[2]构建了一种融合经验模态分解 (empirical mode decomposition, EMD)、长短期记忆网络 (long short-term memory, LSTM) 和自回归积分滑动平均模型 (autoregressive integrated moving average, ARIMA) 的 EMD-LSTM-ARIMA 模型来提高风速预测精度, 从而提高风电预测精度; 袁晓辉等^[3]利用自回归分数阶积分滑动平均模型探索风电时间序列数据中隐藏的线性成分, 利用最小二乘支持向量机提取数据中的非线性成分, 实现风电功率预测。机器学习模型中, 向玲等^[4]将 LSTM 与 Vision Transformer 相结合, 通过 LSTM 提取

特征时空特征, 再结合多头自注意力机制建立特征与输出之间的联系, 从而提升风电功率预测精度; 李扬等^[5]提出联邦深度强化学习方法, 以分散的方式共享预测模型, 有效保护数据隐私。组合模型方面, 刘永前团队^[6]提出“动态匹配与在线建模”策略, 基于特征相似性动态选择训练样本, 并结合数值天气预报风速, 用在线建模应对概念漂移问题; 王颖等^[7]提出一种改进的灰狼优化算法增强全局搜索能力, 并用于优化 Informer 模型超参数。

尽管上述数据驱动的预测模型在风电功率预测中取得显著效果, 但其预测过程通常依赖历史数据特征, 可解释性仍有待提升。而现有的物理-数据融合方法虽提升了可解释性, 但仍存在一定的局限。例如, 郭永等^[8]通过在 LSTM 模型的损失函数中引入物理损失函数作为正则化项, 但该方法在面对复杂风电场数据时, 存在泛化能力不足的问题。LSTM 引入了门控机制来弥补传统循环神经网络 (recurrent neural network, RNN) 的缺陷, 能对短期依赖进行有效捕捉。但在处理长序列复杂的风电场数据时, 仍易出现梯度消失问题。因此, 如何在充分表征短期波动特征的同时, 有效建模长期时序依赖, 并引入物理约束提升预测结果的合理性, 仍是超短期风电功率预测中需要解决的关键问题。

为解决上述问题, 本文提出一种融合物理约束与多尺度特征的 Mamba-Transformer 模型 (以下简称 MT 模型)。该模型利用自适应噪声的完备经验模态分解 (complete ensemble

收稿日期: 2025-01-14

基金项目: 国家自然科学基金 (52367015; 52277148); 中国博士后科学基金 (2024M750897); 江西省自然科学基金 (20232BAB214061)

通信作者: 余勇祥 (1991—), 男, 博士、副教授, 主要从事电力系统储能技术方面的研究。xyyu366@ecjtu.edu.cn

empirical mode decomposition with adaptive noise, CEEMDAN) 风电功率数据进行分解, 提取多个本征模函数 (intrinsic mode function, IMF), 将分解后的 IMF 和风电场气象历史数据组成多尺度特征输入 MT 模型。在风电功率预测场景中, Mamba 通过选择记忆和遗忘机制, 专注捕捉数据的短期趋势和局部波动; 而 Transformer 则利用自注意力机制捕捉数据的长时间依赖和全局时序依赖。同时, 引入一维 Jensen 尾流模型计算物理预测值, 并通过自适应加权机制, 将数据驱动预测值与物理预测值之间的均方误差 (mean squared error, MSE) 损失函数作为物理约束项, 从而实现动态调整物理约束的引导作用, 提高模型遵循风力机尾流动力学规律的合理性。

1 数据分解方法

CEEMDAN 算法是在 EMD 的基础上提出的。该算法通过添加自适应白噪声并经过多次迭代, 解决了模态混叠问题, 同时也避免了引入新噪声^[9]。

CEEMDAN 分解的具体步骤如下:

1) 假设 $s(t)$ 表示原始信号; $n_i(t)$ 为第 i 次向原信号中添加的标准正态白噪声; ε_0 为信噪比。在第 i 次实验中, 向原始信号加入白噪声, 构造含噪信号 $s_i(t)$, 如式(1)所示。

$$s_i(t) = s(t) + \varepsilon_0 n_i(t), \quad i = 1, 2, \dots, k \quad (1)$$

式中: k ——实验的总次数。

2) 对每个信号 $s_i(t) (i = 1, 2, \dots, k)$ 进行 EMD 分解, 得到第 i 个信号的第一个 IMF, 记为 $F_1^i(t)$ 。然后, 计算所有实验中第一个 IMF 的平均值, 得到 CEEMDAN 的第一个 IMF 分量, 记为 $\overline{F_1(t)}$, 如式(2)所示。

$$\overline{F_1(t)} = \frac{1}{k} \sum_{i=1}^k F_1^i(t) \quad (2)$$

根据第一个 IMF 分量, 计算第一阶段的残余分量 $r_1(t)$, 如式(3)所示。

$$r_1(t) = s(t) - \overline{F_1(t)} \quad (3)$$

3) 对每次迭代得到的残余分量进行类似的操作。首先, 对第 j 个 IMF 分量进行 EMD 分解。为此, 先给该 IMF 分量添加噪声, 构造新的含噪信号, 然后对新信号进行 EMD 分解, 经过 k 次迭代得到第 $j+1$ 个 IMF 分量 $\overline{F_{j+1}(t)}$, 如式(4)所示。

$$\overline{F_{j+1}(t)} = \frac{1}{k} \sum_{i=1}^k H_j \{r_j(t) + \varepsilon_j H_j[n_i(t)]\} \quad (4)$$

式中: H_j ——EMD 处理的第 j 个模态分量; $r_j(t)$ ——第 j 个 IMF 分量分解后的残差信号; ε_j ——添加到第 j 个 IMF 中的噪声信号。

4) 重复上述步骤, 直到残余分量不能再继续分解为止, 最终得到 m 个 IMF 分量, 则原始信号 $s(t)$ 的 CEEMDAN 分解结果如式(5)所示。

$$s(t) = \sum_{n=1}^m \overline{F_n(t)} + r_m(t) \quad (5)$$

式中: m ——分解得到的 IMF 数量; $\overline{F_n(t)}$ ——CEEMDAN 的第 n 个 IMF 分量; $r_m(t)$ ——残余分量。

从 CEEMDAN 分解的流程可知, CEEMDAN 分解在计算第二个 IMF 分量及之后的 IMF 分量时, 是将白噪声添加到一个残余分量中, 而不是直接添加到原始信号中, 因此分解重构后的数据信号噪声较少^[10], 从而增强了多尺度特征的细节表示。

2 融合物理约束和多尺度特征的 MT 模型的构建方法

2.1 Mamba 模型原理

2.1.1 状态空间模型

状态空间模型 (state space models, SSM) 可通过使用一阶微分方程来表示系统内部状态的演变, 以及另一组方程来描述潜在状态与输出序列之间的关系, 输入序列 $x(t) \in \mathbf{R}^D$ 可通过潜在状态 $h(t) \in \mathbf{R}^N$ 映射到输出序列 $y(t) \in \mathbf{R}^N$ 如式(6)所示。

$$\begin{aligned} h(t)' &= Ah(t) + Bx(t) \\ y(t) &= Ch(t) \end{aligned} \quad (6)$$

式中: $A \in \mathbf{R}^{N \times N}$ 、 $B \in \mathbf{R}^{N \times 1}$ 和 $C \in \mathbf{R}^{1 \times N}$ ——可学习的矩阵; $h(t)$ ——隐状态序列; $h(t)'$ ——隐状态序列的导数; $x(t)$ ——输入序列; $y(t)$ ——输出序列。

2.1.2 结构化状态空间模型

结构化状态空间 (structured state space sequence, S4) 模型源于 SSM, 可通过其内部状态的存储和更新机制有效捕获长时间序列的依赖关系。S4 利用了 HiPPO (high-order polynomial projection operator) 方法使得状态更新更具动态适应性^[11], 通过对连续系统的参数使用零阶保持 (zero-order hold, ZOH) 方法离散化处理, 从而改进长距离依赖关系的建模^[12-13], 以实现递归和卷积两种计算模式的结合, 通过步长 Δ 离散化的 S4 表达式如式(7)所示。

$$\begin{cases} h_t = \bar{A}h_{t-1} + \bar{B}x_t \\ y_t = Ch_t \end{cases} \quad (7)$$

式中: $\bar{A} = \exp(\Delta A)$ 和 $\bar{B} = (\Delta A)^{-1} [\exp(\Delta A) - I] \Delta B$ ——离散化后的矩阵; I ——单位矩阵; h_t ——离散时间步 t 的隐状态序列; h_{t-1} ——离散时间步 $t-1$ 的隐状态序列; x_t ——离散时间步 t 的输入序列; y_t ——离散时间步 t 的输出序列。

2.1.3 Mamba 模型

Mamba 模型建立在 SSM 和 S4 的基础上, 通过引入选择性状态空间模型 (selective SSM) 和硬件感知的并行算法来增强模型在长序列建模中的适应性和效率^[14], Mamba 模型结构示意图如图 1 所示。

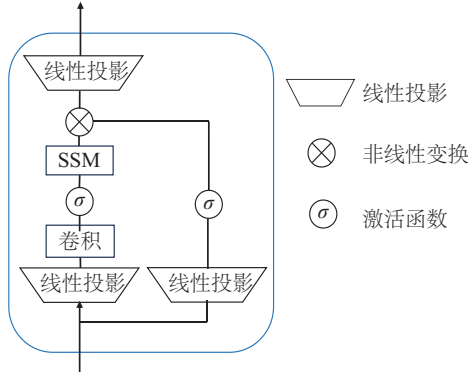


图1 Mamba模型结构示意图

Fig. 1 Schematic of Mamba model structure

Mamba 模型的接收形状为 (b, l, d) 的输入序列张量 $\mathbf{x}$, 其中, b 表示批次大小, l 表示序列长度, d 表示每个时间步的特征维度。通过基于 S4 的离散化处理, Mamba 扩展了输入特征维度, 生成特征映射后, 利用卷积操作和 SiLU (Sigmoid Linear Unit) 激活函数增强特征的非线性能力。在特征映射处理完成后, 模型引入输入依赖的动态参数生成机制, 以实现选择性状态更新。因此, 模型通过离散化状态矩阵和选择性参数递归更新隐状态 h_t , 并通过输出映射矩阵计算当前时间步的输出 y_t , 表达式为:

$$\begin{cases} h_t = \bar{A}h_{t-1} + s_B(\mathbf{x}_t)x_t \\ y_t = s_C(\mathbf{x}_t)h_t \end{cases} \quad (8)$$

式中: $s_B(\mathbf{x}_t)$ 和 $s_C(\mathbf{x}_t)$ ——通过输入依赖的线性投影生成的矩阵。

Mamba 模型通过选择性记忆和遗忘机制, 更侧重局部信息学习, 并通过卷积操作, 易捕捉短期依赖, 从而有效捕捉风电场数据中的局部动态信息。但由于缺乏长期记忆机制, Mamba 难以有效捕捉长期依赖。

2.2 Transformer 模型

Transformer 是基于自注意力机制 (self-attention), 且可在不使用传统的递归和卷积操作的情况下处理序列数据的模型, 这使得它具有良好的并行化能力。在风电功率预测场景中, Transformer 能有效捕捉跨越多个时间步的依赖, 从而捕捉风电场数据长期依赖关系。Transformer 模型的主要组成部分包括编码器 (Encoder) 和解码器 (Decoder), 每个部分由多个相同的层堆叠而成^[15]。Transformer 结构如图 2 所示。

Transformer 每个层包含以下部分:

1) 自注意力机制: 计算时序数据中不同时间点之间的关系, 帮助模型理解一个时间步的值如何受到其他时间步的影响。表达式如式 (9) 所示。

$$A(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{a}}\right)V \quad (9)$$

式中: $A(\cdot)$ ——注意力函数; $\text{Softmax}(\cdot)$ ——归一化函数; a ——

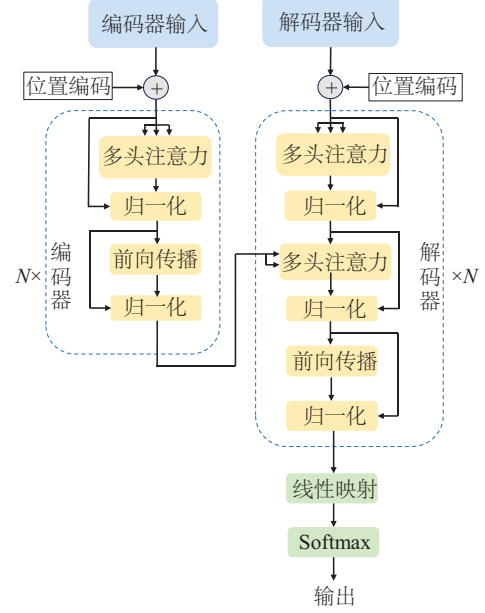


图2 Transformer模型结构示意图

Fig. 2 Schematic of Transformer model structure

输入序列的维度; Q 、 K 和 V ——查询 (query)、键 (key) 和值 (value) 矩阵。

2) 多头注意力机制: 通过多个并行的注意力头计算不同子空间的自注意力, 并将结果拼接合并。

3) 时间编码: 为序列中的每个元素添加时间位置信息, 使模型能理解和利用数据中的时间顺序和依赖关系, 从而更好地处理时序数据。

4) 前向传播: 由全连接层和 ReLU 激活函数构成, 用于进一步处理自注意力机制的输出。

2.3 基于风电场尾流效应的物理预测模型

一维 Jensen 尾流模型 (下文简称 Jensen 模型) 于 1983 年由丹麦国家 Risø 实验室提出, 用于描述多台风力机间的相互作用和尾流影响^[16-17]。与二维和三维 Jensen 模型相比, Jensen 模型计算复杂度低, 能快速计算尾流风速和功率预测, 更适合超短期风电功率预测的应用场景。Jensen 模型简化示意图如图 3 所示。

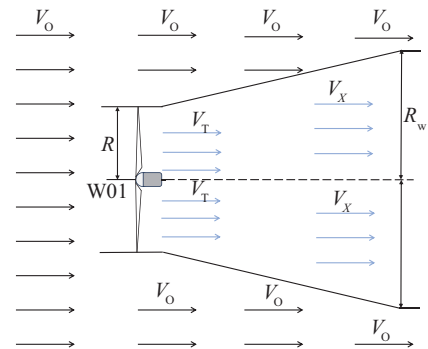


图3 Jensen模型示意图

Fig. 3 Schematic of Jensen's model

图 3 中, W_{01} 为上游风力机; R 为上游风力机的风轮半径, 也可称扫风半径, 工程计算中等于风轮直径 D 的一半; R_w 表示风力机尾流半径; V_0 表示自由风速, 即风进入风力机前的风速; V_T 表示风刚吹过风力机叶片的速度; V_x 表示尾流风速。

在 Jensen 模型中, 尾流的半径随下游距离呈线性扩展。假设尾流扩展角是恒定的, 尾流半径 R_w 在上下游风力机距离 X 处的表达式为:

$$R_w(X) = R + KX \quad (10)$$

式中: R ——扫风半径, m; K ——尾流扩展系数, 通常取值在 0.075~0.1 之间, 本文取 0.075; X ——上下游风力机之间的实际距离, m。

计算得到尾流半径后, 在风力机之间的距离 X 处的尾流风速 V_x 为:

$$V_x = V_0 \left[1 - \frac{2\alpha}{(1 + KX/R)^2} \right] \quad (11)$$

式中: V_0 ——自由风速, m/s; α ——诱导因子, 表示风力机对迎面风速造成减速的量化, 表达式为 $\alpha = 1 - \sqrt{1 - C_T}$; C_T ——推力系数, 反映风力机对迎面风施加的阻力。

将诱导因子 α 代入尾流风速 V_x :

$$V_x = V_0 \left[1 - 2 \left(1 - \sqrt{1 - C_T} \right) \left(1 + \frac{KX}{R} \right)^{-2} \right] \quad (12)$$

计算出尾流风速 V_x 后, 将 V_x 代入工程常用得风电功率计算公式, 即可得到风力机的输出功率。工程中常用的功率计算公式为:

$$P = \frac{1}{2} \rho S C u^3 \quad (13)$$

式中: ρ ——风场空气密度, kg/m³; C ——风力机的功率转换效率; S ——风力机旋翼扫风面积, m²; u ——风力机的尾流风速, m/s。

基于 Jensen 尾流模型的物理功率预测值计算步骤可归纳为: 1) 将上下游两个风力机之间的实际距离 X 和扫风半径 R 代入式(10)计算尾流半径 R_w ; 2) 结合诱导因子 α , 将其代入式(12)中, 计算得到尾流风速 V_x ; 3) 将尾流风速 V_x 代入式(13)计算得到物理功率预测值。

2.4 建立融合物理约束与多尺度特征的 MT 模型

2.4.1 构建基于 CEEMDAN 的多尺度 MT 模型

为提取风电功率数据中的多时间尺度特征, 构建基于 CEEMDAN 的多尺度 MT 模型。采用 CEEMDAN 算法对风电功率数据分解, 生成具有不同频率特性的 IMF 分量。低频 IMF 分量反映长期趋势, 有助于识别风电功率的周期性趋势。高频 IMF 分量反映短期波动, 帮助捕捉瞬时气象变化。将分解出的各 IMF 分量与风电气象数据组成多尺度特征。模型结构如图 4 所示。

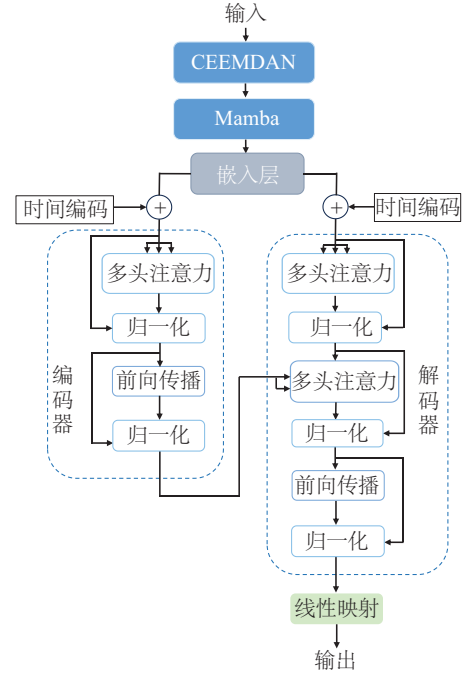


图4 基于CEEMDAN的多尺度MT模型结构

Fig. 4 CEEMDAN-based multiscale MT model structure

MT模型通过结合Mamba和Transformer的优势, 充分发挥了它们在捕捉多尺度特征的互补性。多尺度特征输入至Mamba模型后, Mamba模型通过选择性状态空间机制与卷积操作, 实现动态筛选与超短期预测任务相关的关键信息, 从而精确捕捉多尺度特征中的短期趋势和瞬时变化, 生成经过优化的潜在特征表示。

但Mamba在全局趋势的捕捉能力有限, 而全局趋势是多尺度特征的重要组成部分。此时, Transformer则通过自注意力机制捕捉多尺度特征中的长期依赖和全局趋势, 可有效弥补Mamba在全局趋势捕捉的不足。

将Mamba优化处理后的潜在表示通过独立嵌入层输送至Transformer, 使得Transformer不仅能接收到包含局部波动的潜在表示, 且可利用自注意力机制捕捉全局趋势。因此MT模型在准确捕捉多尺度特征的整体趋势的同时, 又能聚焦局部动态, 从而在风电功率预测中获得更高的精度和稳定性。

2.4.2 物理约束

为结合物理规律与数据驱动模型的灵活性, 本文提出一种融合物理约束和多尺度特征的MT模型, 该模型通过动态加权机制将物理约束嵌入数据驱动模型的损失函数中。融合物理约束与多尺度特征的MT模型运行流程如图5所示。

首先, 上游风力机的数据用于构造Jensen模型, 将上游风力机的尾流扩展系数 K 、风力机间距 X 、推力系数 C_T 、上游风力机的自由风速 V_0 与上游风力机扫风半径 R 等参数代入式(11)~式(13)计算下游风力机的物理功率预测值 $\hat{p}_i$ 。将下游风力机数据输入基于CEEMDAN的多尺度MT模型中, 输出

数据驱动模型功率预测值 $\hat{y}_i$ 。

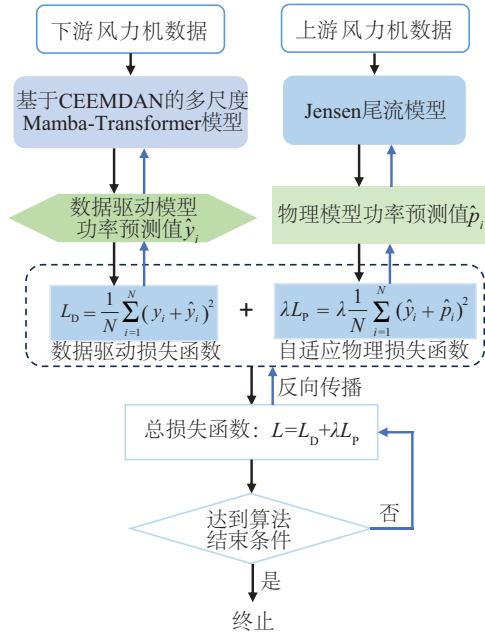


图5 模型的运行流程

Fig. 5 Flowchart of model

其次,基于物理规律与数据驱动特征的互补性,构造物理损失函数。通过计算数据驱动预测值 $\hat{y}_i$ 与物理预测值 $\hat{p}_i$ 之间的均方误差(mean squared error, MSE)作为物理损失函数,并与数据驱动的 MSE 损失函数结合。数据驱动模型损失函数 L_D 与物理损失函数 L_P 表达式为:

$$L_D = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (14)$$

$$L_P = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - \hat{p}_i)^2 \quad (15)$$

式中: N ——样本数量; $\hat{y}_i$ ——第 i 个样本预测值; y_i ——第 i 个样本真实值。

最后,通过自适应加权机制,将 L_D 和 L_P 结合。融合物理约束与多尺度特征的 MT 模型的总损失函数为:

$$L = L_D + \lambda L_P \quad (16)$$

式中: L ——模型的总损失函数; λ ——自适应权重,能依据模型当前拟合的误差动态调整物理功率预测值 $\hat{p}_i$ 与数据驱动模型预测值 $\hat{y}_i$ 之间的关系,以更精确地反映物理规律对模型的约束作用。

3 算例与结果分析

3.1 实验描述

3.1.1 实验数据集选取

本文的实验数据集来自中国新疆某风电场,具体选取该风电场中编号为 W01 和 W02 的两台风力机作为研究对象。W02 的上游风力机为 W01。两台风力机的数据收集时间为 2019 年 1 月 1 日—12 月 31 日,采样时间的间隔为 15 min。

本文实验将以 W01 和 W02 的 2019 年数据集为基础,聚焦于风电功率超短期预测。将 W01 和 W02 的历史数据按 7:2:1 的比例划分为训练集、验证集与测试集。以 W02 作本文的预测对象,用 W02 的测试数据集对模型预测结果的不同评价指标进行比较和分析,风力机 W01 的参数见表 1。

表1 风力机 W01 的参数

Table 1 Parameters of wind turbine W01

风轮直径 D/m	推力系数 C_T	W01 与 W02 的距离 X/D
80	0.80	4.0

3.1.2 数据分解

使用 CEEMDAN 算法对风电功率数据进行分解。将 W02 的原始非线性、非平稳功率分解为一系列具有不同频率特征的 IMF 和残差分量 r 。

通过对分解结果的分析,IMF₁~IMF₄ 为高频分量,主要反映风电功率数据中的短期波动信息。而 IMF₅~IMF₁₀ 为低频分量,揭示了风电功率的长期变化趋势,反映了风力机运行中的平稳特性和宏观波动规律,结果如图 6 所示。

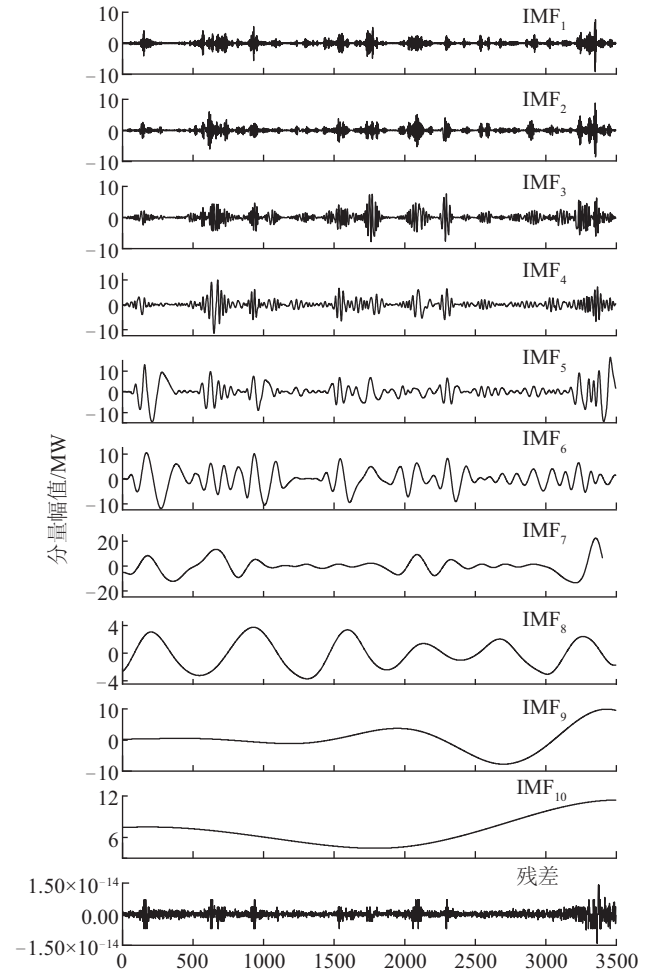


图6 分解结果

Fig. 6 Decomposition results

3.1.3 评估指标

为评估模型的预测精度与稳定性,本文选择以下 3 种评估指标,均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)和拟合优度(R-Square, R^2),计算公式分别为:

$$E_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{\theta=1}^n (y_{\theta} - \hat{y}_{\theta})^2} \quad (17)$$

$$E_{\text{MAE}} = \frac{1}{n} \sum_{\theta=1}^n |y_{\theta} - \hat{y}_{\theta}| \quad (18)$$

$$R^2 = 1 - \frac{\sum_{\theta=1}^n (y_{\theta} - \hat{y}_{\theta})^2}{\sum_{\theta=1}^n (y_{\theta} - \bar{y})^2} \quad (19)$$

式中: n ——样本数量; $\hat{y}_{\theta}$ ——第 θ 个样本预测值; y_{θ} ——第 θ 个样本真实值; $\bar{y}$ ——所有样本预测值的均值。

3.2 预测步骤

融合物理约束与多尺度特征的 MT 模型进行超短期风电功率预测,具体步骤如下:

- 1)数据集清洗与处理。检查数据集完整性,处理数据集中存在的异常值和缺失值。
- 2)数据集划分。本文的实验以 W02 作为预测对象,将 W01 和 W02 的 2019 年数据集以 7:2:1 的比例划分为训练集、验证集与测试集。
- 3)数据分解。加载划分完的数据集,用 CEEMDAN 算法对 W02 功率进行分解,得到不同频的 IMF。
- 4)归一化。将 3)中 IMF 按比例与 W02 的训练集构成多尺度特征,并将其进行归一化处理,统一量纲。
- 5)模型训练。将 W01 的训练集输入到 Jensen 模型,用于计算物理功率预测值。同时,将 W02 的训练集输入至 MT 模型,生成数据驱动功率预测值。
- 6)构造物理约束。训练完后输出两个结果:一个是基于

W01 数据的物理功率预测值,另一个基于 W02 数据的数据驱动模型预测值。将二者之间的 MSE 损失函数作为物理约束项,与数据驱动 MSE 损失函数作自适应加权。

7)反向传播。通过损失函数反向传播。同时,在每个训练周期后,将 W02 的验证集输入模型,计算验证误差以监控模型性能。

8)测试模型。在 W02 的测试集上将预测结果反归一化,再通过计算评估指标和可视化,对模型进行评估。

3.3 对比实验结果与分析

3.3.1 全年数据对比实验

为验证基于 CEEMDAN 的多尺度 MT 模型的预测精度,使用 LSTM(I)、TCN(II)、Transformer(III)、CNN-Transformer(IV)、Attention-GRU(V)与基于 CEEMDAN 的多尺度 MT 模型(VI)共 6 组预测模型作对照,通过对比不同模型的评估指标来判断。本实验采用 3.1.1 节所述数据集进行模型训练、验证与测试,6 组模型的评估指标见表 2,功率预测结果曲线如图 7 所示。

表 2 全年数据对比结果

Table 2 Comparative results for annual data

模型	R^2	$E_{\text{RMSE}}/\text{MW}$	E_{MAE}/MW
I	0.949	3.400	1.230
II	0.987	1.055	0.689
III	0.985	1.143	0.969
IV	0.989	0.933	0.613
V	0.988	0.966	0.705
VI	0.993	0.671	0.482

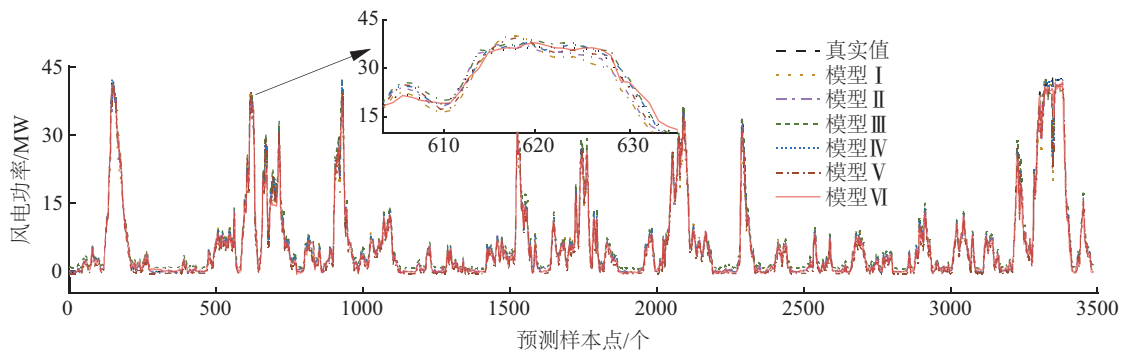


图 7 6 组模型的功率预测结果曲线

Fig. 7 Power prediction result curves for 6 models

根据表 2,本文提出的模型 VI 在所有评估指标中表现最优, R^2 为 0.993, E_{RMSE} 为 0.671 MW, E_{MAE} 为 0.482 MW,表明其在特征捕捉、预测误差最小和稳定性方面均优异。模型 I 的

表现最差, E_{RMSE} 和 E_{MAE} 分别为 3.400 和 1.230 MW。与模型 IV 相比,模型 VI 在 E_{RMSE} 和 E_{MAE} 上分别降低 0.262 和 0.131 MW。同时,从图 7 中可看出,模型 VI 在预测风电功率数据时能更

有效捕捉到数据中的波动和非线性特征,表现出比其他模型更强的波动捕捉能力。

3.3.2 季节变化对比实验

为验证基于 CEEMDAN 的多尺度 MT 模型的泛化能力与对季节性变化的适应性,将 3.1.1 小节所述数据按春、夏、秋、冬 4 个季节进行划分,并采用与 3.3.1 节相同的 6 组模型进行对比分析。以 W02 风力机作为预测对象,评估模型在不同季节条件下的预测精度与稳定性。6 组模型评估指标见表 3。

根据表 3 可知,模型 VI 在春、夏、秋、冬四季的评估指标中均表现优异,指标均为最佳。以冬季为例,模型 VI 的 R^2 达到 0.994,显著高于模型 I 的 0.969 与模型 II 的 0.981。同时,模型 VI 的 E_{RMSE} 和 E_{MAE} 分别为 0.704 和 0.518 MW,均处于最低,优于其他 5 组模型的 E_{RMSE} 和 E_{MAE} ,说明模型 VI 在寒冷季节条件下依然能有效捕捉风电功率的变化趋势,展现出良好的预测能力和泛化能力。而在春季、夏季和秋季中也表现出同样的结果。根据预测实验结果表明,模型 VI 在所有季节中均表现出较高的预测精度和泛化能力,远优于其他模型。

3.4 消融实验结果与分析

3.4.1 数据驱动模型消融实验

为验证基于 CEEMDAN 的多尺度 MT 模型各模块对整体性能的贡献,采用消融实验方法逐步去除模型的不同组件得到 5 种变体模型^[18]。实验设计的逻辑严谨性在于每次仅改变单一模块,确保结果反映特定组件的影响,将 5 个变体模型和本文提出的模型进行对比,分别为 MT(A)、Transformer (B)、Mamba (C)、CEEMDAN-Transformer (D)、CEEMDAN-Mamba (E) 与基于 CEEMDAN 的多尺度 MT 模型 (F)。实验数据采用 3.1.1 节所述数据集,以上 6 组模型的评估指标见表 4,预测值和实际值对比如图 8 所示。

根据表 4,去除 Mamba 模块后,模型 D 的 E_{MAE} 增至 0.625 MW, E_{RMSE} 增至 1.105 MW,表明 Mamba 模块在局部波动的捕捉上起到重要作用。去除 Transformer 模块后,模型 E 的 E_{MAE} 增至 0.771 MW, E_{RMSE} 增至 1.316 MW,影响最大。Transformer 模块能捕捉风电功率的全局依赖关系和长期趋

表 3 季节变化对比结果

Table 3 Comparative results of seasonal variations

季节	模型	R^2	E_{RMSE}/MW	E_{MAE}/MW
春季	I	0.958	2.335	1.556
	II	0.981	1.560	1.028
	III	0.984	1.478	0.939
	IV	0.988	1.191	0.769
	V	0.987	1.061	0.676
	VI	0.991	0.670	0.643
夏季	I	0.967	1.732	1.337
	II	0.984	0.923	0.801
	III	0.981	0.809	0.965
	IV	0.988	0.721	0.550
	V	0.986	0.783	0.670
	VI	0.992	0.521	0.454
秋季	I	0.971	1.912	1.262
	II	0.981	1.638	1.179
	III	0.983	1.546	1.378
	IV	0.985	1.453	0.946
	V	0.986	1.454	0.917
	VI	0.992	0.913	0.596
冬季	I	0.969	2.528	1.875
	II	0.981	1.921	1.571
	III	0.982	1.816	1.266
	IV	0.987	1.594	1.343
	V	0.986	1.635	1.488
	VI	0.994	0.704	0.518

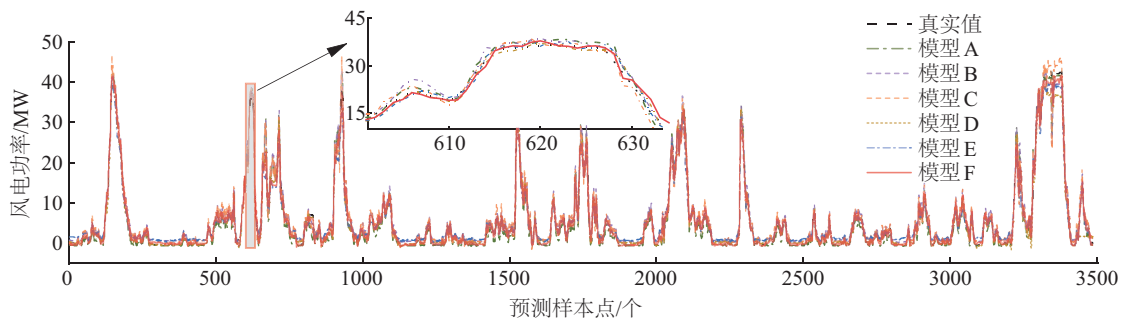


图 8 消融实验的功率预测结果曲线

Fig. 8 Power prediction result curves for ablation experiments

表4 数据驱动模型的消融结果

Table 4 Results of ablation of data-driven model

模型	R^2	$E_{\text{RMSE}}/\text{MW}$	E_{MAE}/MW
A	0.990	1.063	0.633
B	0.985	1.143	0.969
C	0.972	1.182	0.887
D	0.981	1.105	0.625
E	0.978	1.316	0.771
F	0.993	0.671	0.482

势,去除后,模型在短期波动性预测和整体精度上明显下降。去除 CEEMDAN 模块后,模型 A 的 E_{MAE} 升至 0.633 MW, E_{RMSE} 增至 1.063 MW,说明 CEEMDAN 得到不同频率的特征,实现了短期波动与长期趋势的分离。模型 B 和模型 C,由于是独立的模型,其表现较差。根据实验表明,本文提出的模型 F 在风电功率预测中具有较强的准确性、稳定性和误差控制能力。

由此可见,本文提出的模型在所有评估指标中表现最佳,去除 Mamba,Transformer 和 CEEMDAN 模块后,模型的预测精度和稳定性显著下降,表明各模块在局部波动的捕捉、全局依赖关系捕捉和多尺度特征对风电功率预测均起到关键作用。

3.4.2 物理约束模块消融实验

为验证引入物理约束对风电功率预测准确性和稳定性的提升,本实验采用 3.1.1 节所述数据集,对比未引入物理约束的 CEEMDAN 多尺度 MT 模型(模型 1)和融合物理约束的 MT 模型(模型 2)之间的评估指标。

为进一步分析物理约束模型对风电功率预测精度的影响,除运用 3.1.3 节的评估指标外,本节还将采用误差百分比(percentage of error, PE)进行对比分析。将误差百分比分为 3 个区间:误差在 5% 以内的样本比例(p_{5-})、误差在 5%~20% 之间的样本比例(p_{5-20})和误差超过 20% 的样本比例(p_{20+})^[19]。其中, p_{5-} 越大越好,表示预测结果接近实际值; p_{5-20} 越小越好,

表示大部分预测误差较小; p_{20+} 越小越好,表示大误差较少,模型预测结果更为准确。实验数据采用 3.1.1 节所述数据集,计算公式如式(20)~式(22)所示。

$$p_{5-} = \frac{k}{j} \times 100\% \quad (20)$$

$$p_{5-20} = \frac{l}{j} \times 100\% \quad (21)$$

$$p_{20+} = \frac{m}{j} \times 100\% \quad (22)$$

式中: k ——误差百分比小于或等于 5% 的样本数; l ——误差百分比在 5%~20% 的样本数; m ——误差百分比大于 20% 的样本数; j ——样本总数。

评估指标见表 5,两种模型的误差百分比如图 9 所示,预测值和实际值对比如图 10 所示。

表5 引入物理约束的消融结果

Table 5 Results of ablation introducing physical constraints

模型	R^2	$E_{\text{RMSE}}/\text{MW}$	E_{MAE}/MW
1	0.993	0.671	0.482
2	0.997	0.517	0.331

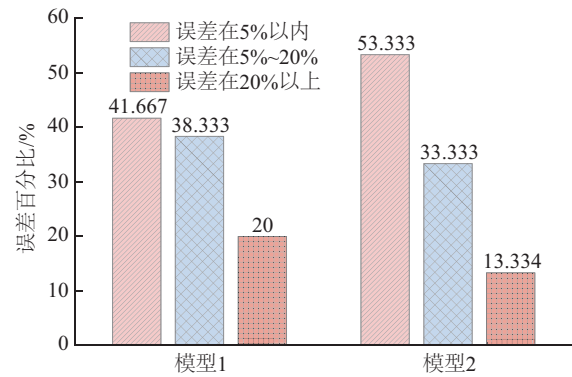


图9 模型1与模型2的误差百分比

Fig. 9 Percentage of error between model 1 and model 2

根据表 5 的结果,模型 2 在所有指标上均优于模型 1。在 R^2 上,模型 2 达到 0.997,相较于模型 1 的 0.993 进一步提

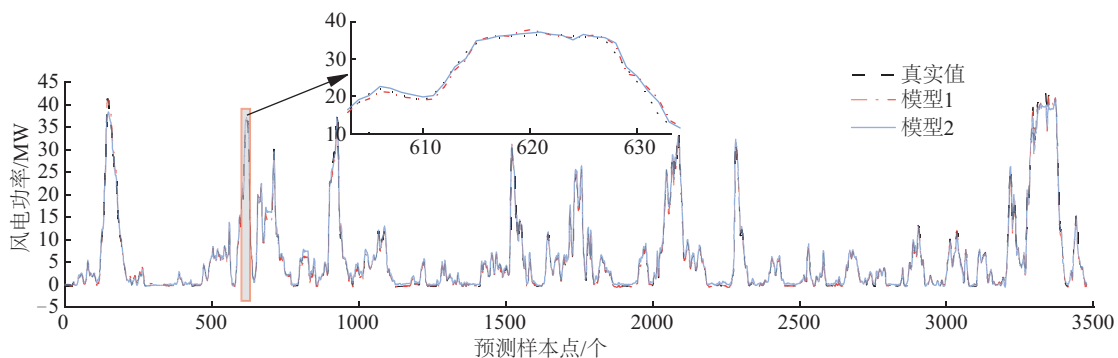


图10 模型1和模型2的功率预测结果曲线

Fig. 10 Power prediction result curves for model 1 and model 2

升。 E_{RMSE} 从 0.671 MW 降至 0.517 MW,降低 0.154 MW; E_{MAE} 从 0.482 MW 降至 0.331 MW,降低 0.151 MW。实验结果表明,引入物理约束后,不仅提高预测精度,且表现出更优的泛化能力。

由图 9 可知,模型 2 在 p_{5-} 的值为 53.333%,高于模型 1 的 41.667%,说明模型 2 的误差均小于模型 1。同时,模型 2 在 p_{5-20} 和 p_{20+} 上的值分别为 33.333% 和 13.334%,低于模型 1 的 38.333% 和 20.000%。由图 10 可知,模型 2 通过物理约束使得预测结果更符合风力机组的物理规律,提升了预测精度和增加了预测模型的可解释性,表明引入物理约束可有效提高模型的准确性和稳定性。

4 结 论

为解决当前风电功率预测模型存在可解释性较差与现有物理-数据融合模型难以同时捕捉短期波动和长期依赖的问题,本文引入基于 Jensen 尾流模型的物理约束,提出一种融合物理约束与多尺度特征的 MT 超短期风电功率预测。主要结论如下:

1)通过引入基于 Jensen 尾流模型的物理约束,模型能将数据驱动与尾流动力学规律相结合,使预测结果更具物理合理性,从而提升了模型的可解释性。

2) CEEMDAN 算法通过对风电功率数据进行分解,得到短期波动与长期趋势特征,并结合风电场气象数据的外部影响信息,进一步丰富了多时间尺度特征的表达能力。通过将分解后的模态函数与风电场气象数据共同输入 Mamba 模型,充分捕捉多尺度局部动态特征,并利用 Transformer 模型捕捉全局信息和长距离时序依赖。模型实现了多时间尺度特征与气象信息的有效融合,能有效提高模型的预测效果与稳定性。

3)实验结果表明,本文提出的预测模型在各项评估指标上优于不同预测模型。与未引入物理约束的模型相比,加入物理约束后, R^2 由 0.993 提升至 0.997, p_{5-} 提升 11.667 个百分点, E_{RMSE} 降低 0.154 MW。这些结果表明物理约束在提升模型预测精度和增强预测稳定性方面具有积极作用。

本文提出的融合物理约束与多尺度特征的 MT 模型,通过实验验证,证明了物理约束和多尺度特征能提升模型对风电功率变化规律的适应能力。根据实验结果,本文提出的模型在各项评估指标上的表现优于其他预测模型。未来工作可进一步考虑多个风力机尾流效应的影响,从而进一步提升模型的可解释性、预测精度和应用价值。

[参考文献]

- [1] 舒印彪,赵勇,赵良,等.“双碳”目标下我国能源电力低碳转型路径[J].中国电机工程学报,2023,43(5):1663-1671.
- SHU Y B, ZHAO Y, ZHAO L, et al. Study on low carbon

energy transition path toward carbon peak and carbon neutrality [J]. Proceedings of the CSEE, 2023, 43 (5) : 1663-1671.

- [2] LIU M D, DING L, BAI Y L. Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction [J]. Energy conversion and management, 2021, 233: 113917.
- [3] YUAN X H, TAN Q X, LEI X H, et al. Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine[J]. Energy, 2017, 129: 122-137.
- [4] XIANG L, FU X, YAO Q T, et al. A novel model for ultra-short term wind power prediction based on Vision Transformer[J]. Energy, 2024, 294: 130854.
- [5] LI Y, WANG R N, LI Y Z, et al. Wind power forecasting considering data privacy protection: a federated deep reinforcement learning approach [J]. Applied energy, 2023, 329: 120291.
- [6] LI Y H, WANG H, YAN J, et al. Ultra-short-term wind power forecasting based on the strategy of “dynamic matching and online modeling” [J]. IEEE transactions on sustainable energy, 2025, 16(1): 107-123.
- [7] 王颀,刘兴杰,梁英,等.一种基于 MGWO-Informer 的超短期风电功率预测方法[J].太阳能学报,2024,45(11):477-485.
- WANG J, LIU X J, LIANG Y, et al. An ultra-short-term wind power prediction method based on MGWO-Informer [J]. Acta energiae solaris sinica, 2024, 45 (11) : 477-485.
- [8] 郭永,王冰,陈玉全,等.融合精细化气象因素与物理约束的深度神经网络模型在短期风电功率预测中的应用[J].电网技术,2024,48(4):1455-1465, I0022, I0023, I0024.
- WU Y, WANG B, CHEN Y Q, et al. Application of deep learning model integrating refined meteorological factors and physical constraints in short-term wind power prediction [J]. Power system technology, 2024, 48 (4) : 1455-1465, I0022, I0023, I0024.
- [9] 刘凡,李捍东,覃涛.基于 CEEMDAN-AsyHyperBand-MultiTCN 的短期风电功率预测[J].太阳能学报,2024,45(1):151-158.
- LIU F, LI H D, QIN T. Short-term wind power prediction based on CEEMDAN-AsyHyperBand-MultiTCN [J]. Acta energiae solaris sinica, 2024, 45(1): 151-158.
- [10] 郭喜峰,王凯泽,单丹,等.多角度基于 CEEMDAN-CNN-BiLSTM 模型的锂离子电池 RUL 预测[J].太阳能学报,2024,45(7):181-189.

- GUO X F, WANG K Z, SHAN D, et al. RUL prediction for lithium ion batteries based on CEEMDAN-CNN-BiLSTM model from multiple perspectives [J]. *Acta energiae solaris sinica*, 2024, 45(7): 181-189.
- [11] GU A, DAO T, ERMON S, et al. Hippo: recurrent memory with optimal polynomial projections[J]. *Advances in neural information processing systems*, 2020, 33: 1474-1487.
- [12] GU A, GOEL K, RÉ C. Efficiently modeling long sequences with structured state spaces [C]//International Conference on Learning Representations (ICLR). 2022.
- [13] GU A, JOHNSON I, GOEL K, et al. Combining recurrent, convolutional, and continuous-time models with linear state-space layers[J]. *Advances in neural information processing systems*, 2021, 34: 572-585.
- [14] GU A, DAO T. Mamba: linear-time sequence modeling with selective state spaces [PP/OL]. V2. arXiv (2024-05-31). <https://doi.org/10.48550/arXiv.2312.00752>.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, California, USA, 2017: 6000-6010.
- [16] 杨伟峰, 文云峰, 李立, 等. 考虑疲劳载荷的风电场分散式频率响应策略[J]. *电力自动化设备*, 2022, 42(4): 55-62.
- YANG W F, WEN Y F, LI L, et al. Decentralized frequency response strategy for wind farm considering fatigue load [J]. *Electric power automation equipment*, 2022, 42(4): 55-62.
- [17] 胡阳, 张冲, 房方, 等. 基于主动尾流控制的风电机群协同优化调度[J]. *动力工程学报*, 2024, 44(4): 566-574.
- HU Y, ZHANG C, FANG F, et al. Cooperative and optimal scheduling of wind turbine groups based on active wake control [J]. *Journal of Chinese Society of Power Engineering*, 2024, 44(4): 566-574.
- [18] 李练兵, 高国强, 吴伟强, 等. 考虑特征重组与改进 Transformer 的风电功率短期日前预测方法[J]. *电网技术*, 2024, 48(4): 1466-1476, I0025, I0027-I0029.
- LI L B, GAO G Q, WU W Q, et al. Short-term day-ahead wind power prediction considering feature recombination and improved transformer [J]. *Power system technology*, 2024, 48(4): 1466-1476, I0025, I0027-I0029.
- [19] GUAN S J, WANG Y S, LIU L M, et al. Ultra-short-term wind power prediction method based on FTI-VACA-XGB model [J]. *Expert systems with applications*, 2024, 235: 121185.

MAMBA-TRANSFORMER ULTRA-SHORT-TERM WIND POWER PREDICTION BY INTEGRATING PHYSICAL CONSTRAINTS AND MULTISCALE FEATURES

Yu Yongxiang, Kuang Xiangyu, Han Jian, Gao Bo, Zeng Han, Li Zewen

(School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: We propose a Mamba-Transformer model integrating physical constraints and multi-scale features for ultra-short-term wind power forecasting. The proposed model employs complete ensemble empirical mode decomposition with adaptive noise to capture nonlinear and nonstationary patterns in wind power data across multiple frequency components. The decomposed modal functions are then fed into the Mamba model alongside meteorological data from the wind farm to uncover local time-varying properties. The self-attention mechanism of the Transformer model is then employed to capture long-range dependencies among multi-scale features. Finally, a one-dimensional Jensen wake model is integrated to establish physical constraints, integrating the prediction results of physical models and data-driven models through an adaptive weighting mechanism. Experimental results demonstrate that compared to the Transformer model, the proposed model reduces root mean square error (RMSE) and mean absolute error (MAE) by 41.29% and 50.26%, respectively. This model enhances wind power forecasting accuracy while exhibiting strong generalization capabilities.

Keywords: wind power forecasting; physical constraint; CEEMDAN; Mamba; Transformer; one-dimensional Jensen wake model