

基于信息重组和TCN-LSTM-MHSA的超短期风电功率预测

陈磊^{1,2}, 黄凯阳^{1,2}, 张怡^{1,2}, 蔡坤哲^{1,2}, 陈禹^{1,2}, 张志瑞^{1,2}

(1. 华北理工大学电气工程学院, 河北唐山 063210; 2. 河北省风光氢储安全监测与智能运行技术创新中心, 河北唐山 063210)

摘要: 为提高风电场的运行效率, 保障电力系统的稳定运行, 提出一种基于信息重组和TCN-LSTM-MHSA的超短期风电功率预测方法。采用变分模态分解将功率数据拆分, 根据样本熵的评估结果将子序列重组为信息类型不同的高熵序列、中熵序列以及低熵序列; 通过时间卷积网络(temporal convolutional network, TCN)探索序列中的特征表示, 利用长短期记忆网络(long short-term memory, LSTM)进一步处理数据特征, 依靠多头自注意力机制(multi-head self-attention, MHSA)并行学习不同的注意力表示, 以此构造风电功率预测模型。使用输出容量和风机构成不同的两个数据集作为基准验证模型, 结果表明所提模型具有更好的预测能力。

关键词: 风电; 功率预测; 时间卷积网络; 长短期记忆网络; 多头自注意力机制

Ultra-Short-Term Wind Power Forecasting Based on Information Recombination and TCN-LSTM-MHSA

CHEN Lei^{1,2}, HUANG Kaiyang^{1,2}, ZHANG Yi^{1,2}, CAI Kunzhe^{1,2}, CHEN Yu^{1,2}, ZHANG Zhirui^{1,2}

(1. Department of Electrical Engineering, North China University of Science and Technology, Tangshan, Hebei 063210, China; 2. Hebei Wind-Solar-Hydrogen-Storage System Security Monitoring and Intelligent Operation Center of Technology Innovation, Tangshan, Hebei 063210, China)

Abstract: To improve the operational efficiency of wind farms and ensure the stable operation of the power system, an ultra-short-term wind power forecasting method based on information recombination and TCN-LSTM-MHSA is proposed. Using variational mode decomposition to split the power data, the subsequences are reorganized into high entropy sequences, medium entropy sequences, and low entropy sequences with different information types based on the evaluation results of sample entropy; Extracting feature representations of data through temporal convolutional network (TCN), further processing data features using the long short-term memory network (LSTM), and parallel learning of different attention representations using multi-head self-attention mechanism (MHSA) to construct a wind power prediction model. Using two datasets with different output capacity and fan composition as benchmarks to validate the model, the results show that the model in this paper has better prediction ability.

Key words: wind power; power forecasting; temporal convolutional network; long short-term memory; multi-head self-attention

0 引言

风电相较于传统能源拥有显著的环保、可持续

优势, 根据国家能源局数据, 截至2023年11月底, 风电装机容量达410 GW, 同比增长17.6%^[1]。在风电快速发展的同时, 其波动的特点导致并网过程对整体的稳定性造成较大的扰动^[2-4]。精准的预测风电发电功率能够更有效地规划和管理电力系统, 降低因风电波动性而引起的电网扰动^[5-6]。

预测方法可以分为物理模型、统计模型、机器学习模型3大类^[7-8], 物理模型和统计模型虽然方

基金项目: 国家重点研发计划资助项目(2021YFE0190900); 教育部产学研合作协同育人项目(230802495182120)。

Foundation item: Supported by the National Key Research and Development Program of China (2021YFE0190900); Ministry of Education Collaborative Industry and Education Partnership for Holistic Development Program (230802495182120).

法成熟,但是面对风电功率预测中复杂的环境变化时应对能力不足,导致预测效果不理想^[9-10]。相比于前两种模型,机器学习模型主要通过算法从数据中学习规律和模式来进行预测,能够更灵活地适应非线性关系、处理多变数据。此外,面对风电功率预测的复杂环境情况,算法可以相互组合,形成集成模型或深度学习结构,提高整体性能。文献[11]提出一种 IWOA-SA-Elman 模型,通过 IWOA-SA (improved whale optimization algorithm and simulated annealing) 算法对 Elman 网络权重进行优化,提高了风电功率预测的精度,但该方法缺少对功率数据中非线性和非平稳性信号的处理,面对复杂数据时应对能力不足,同时较少的训练数据也使得模型存在过拟合风险,泛化能力不足。文献[12]采用 VMD (variable mode decomposition) 分解数据,随后通过 CNN-GRU (convolutional neural network and gated recurrent unit) 模型提取数据时空特征,对风电功率的波动性拥有良好的处理能力,然而该模型未考虑信息重组在数据处理中的必要性,复杂的风电数据会分解出过多的子序列,对其信息调控手段的缺失使模型计算负担大幅增加。文献[13]提出一种 CGAN-CNN-LSTM (conditional generative adversarial network-convolutional neural network-long short term memory network) 的模型,有着良好的预测与泛化能力,但 CNN 模型在处理时间序列数据时灵活性不足,特征关系捕捉不够理想。文献[14]采用 BiLSTM-AM (bidirectional long short term memory and attention mechanism) 模型进行预测,以注意力机制调控信息权重,并使用智能算法修正参数设置,但注意力机制随着序列长度的增长参数量线性增加,在处理长序列数据时效率低下且无法充分表征序列中各种关系。

根据上述分析,本文提出一种基于信息重组和 TCN-LSTM-MHSA 的超短期风电功率预测方法。通过 VMD 将功率序列数据分解,从多尺度捕捉数据中的非线性和非平稳性在不同频率上的变化特征,并以样本熵评估结果为依据进行信息重组,得到包含不同信息特征的熵值序列,显著降低了子序列维度与计算负担;融合时间卷积网络更适用于处理时间序列数据且与长短期记忆网络结合后更为灵活强大的特点,以及 MHSA 机制将

信息分为多块并行学习能够有效提高计算效率降低模型参数量的特点,构造 TCN-LSTM-MHSA 模型对各序列进一步处理,有效提升了对于风电功率的超短期预测精度。通过在实例数据上对模型预测效果的对比分析,验证了本文方法的优越性和可靠性。

1 数据分解与信息重组

1.1 变分模态分解

VMD 可以将数据分解为多个子序列,每个序列表示原序列的一种模式,对复杂数据具有较强的适应性,能够更准确地捕捉信号的时频特性^[15-17]。计算式如式(1)所示。

$$\begin{cases} \min_{\{y_v\}, \{\omega_v\}} \left\{ \sum_v \left\| \frac{d}{dt} \left[\left(\delta(t) + \frac{j}{\pi t} \right) \times y_v(t) \right] e^{-j\omega_v t} \right\|_2^2 \right\} \\ \text{s.t. } \sum_v y_v = y \end{cases} \quad (1)$$

式中: $\{y_1, y_2, \dots, y_v\}$ 为原始序列的模态分量; $\{\omega_1, \omega_2, \dots, \omega_v\}$ 为各分量中心频率; v 为分解数量; $\delta(t)$ 为狄拉克函数; y 为原始数据序列; t 为时刻。

1.2 样本熵

面对风电复杂的时间序列数据,仅对其进行单纯分解会产生较多的分量,各分量间不可避免的存在部分程度上的关联。在此基础上直接建模计算不仅使得计算量倍增,且潜在着过拟合风险。

样本熵可以有效地衡量数据特性,在各个领域中发挥着重要作用。通过计算和分析数据的样本熵能够更方便地探索其中存在的规律以及潜藏信息。这种量化数据复杂性与信息量的方法可以有效地指导对各子序列的分析,提高处理过程的效率及准确性。为此,本文使用样本熵对各子序列进行评估,根据评估情况调整分量结构,从而完成对数据信息的重组。计算进程如下所示^[18-19]。

原始数据 $\mathbf{X} = [x(p), p = 1, 2, \dots, P]$ 为 Q 维序列,根据采集顺序构建向量序列 $\mathbf{X}_q(i)$, $\mathbf{X}_q(i)$ 中各元素表示如下。

$$\mathbf{X}_q(i) = \{x(i), x(i+1), \dots, x(i+q-1)\} \quad (2)$$

式中顺序变量 $i \in [1, P-q+1]$, q 为维度数, P 为原始数据数量。

元素 $\mathbf{X}_q(i)$ 与 $\mathbf{X}_q(j)$ 间距 $dis[\mathbf{X}_q(i), \mathbf{X}_q(j)]$ 如式(3)所示。

$$\text{dis}[X_q(i), X_q(j)] = \max_{k=0, \dots, q-1} |x(i+k) - x(j+k)| \quad (3)$$

式中 $j(1 \leq j \leq P-q, j \neq i)$ 为原始数据数量。

对于 $\text{dis}[X_q(i), X_q(j)]$, 以 $B^{(q)}(r)$ 的形式记录 j 位于 $[1, P-q]$ 且 $j \neq i$ 的数量并求取 i 位于 $[1, P-q]$ 时的平均值, 即为两个序列匹配 q 个点的概率。如式(4)所示。

$$B^{(q)}(r) = \frac{1}{P-q} \sum_{i=1}^{P-q} B_i^q(r) \quad (4)$$

式中: $B_i^q(r)$ 为 j 位于 $[1, p-q]$ 且 $j \neq i$ 的数量; r 为相似容差。

达到 $q+1$ 维时, 以 $A^{(q)}(r)$ 的形式记录 j 位于 $[1, P-q]$ 且 $j \neq i$ 时 $\text{dis}[X_q(i), X_q(j)] \leq r$ 的数量并求取 i 位于 $[1, P-q]$ 时的平均值, 即为两个序列匹配 $q+1$ 个点的概率。

$$A^{(q)}(r) = \frac{1}{P-q} \sum_{i=1}^{P-q} A_i^q(r) \quad (5)$$

式中 $A_i^q(r)$ 为 j 位于 $[1, p-q]$ 且 $j \neq i$ 时 $\text{dis}[x_q(i), x_q(j)] \leq r$ 的数量。

当 P 有限时样本熵 $SE(q, r, P)$ 计算如下。

$$SE(q, r, P) = -\ln \left[\frac{A^q(r)}{B^q(r)} \right] \quad (6)$$

2 预测模型构建

2.1 时间卷积网络

TCN 融合了 CNN 思想并针对时间序列数据特点改进而来, 具有更好的并行能力和更为灵活的感受野, 提高了模型的训练和推理能力^[20-22]。根据 dilation 参数的增长, 卷积核在时间维度上跳越间隔增大, 感受视野随之增加, 从更长的距离上捕获依赖关系。因果卷积的使用保证了模型的因果性, 使得模型在预测 y_t (目标值) 时仅依赖于 t (预测时间点) 前的信息, 不会泄露未来信息。此外, 残差块在卷积层间加入跳跃连接, 可以跨越多个卷积层, 能够有效地传播梯度并加速网络的训练。以 kernel size=2, dilations=[1, 2, 4] 为例, TCN 结构如图 1 所示。计算式如式(7)所示。

$$(F_d * X)(x_t) = \sum_{k=1}^K f_k x_{t-(K-k)d} \quad (7)$$

式中: F_d 为卷积核; f_k 为卷积核第 k 个权重; $*$ 表示膨胀卷积运算; d 为 dilation 参数; $X =$

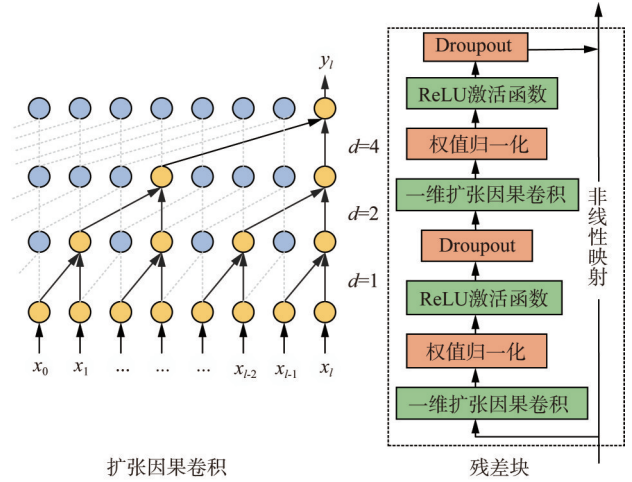


图 1 TCN 结构

Fig. 1 Structure of TCN

$(x_1, x_2, \dots, x_T)$ 为时间序列数据; t 为当前时刻; K 为 kernel size 参数; k 为 K 的索引。

2.2 长短期记忆网络

LSTM 解决了传统网络结构的不足, 在较长的时间序列上可以更好地对数据进行处理。引入的记忆单元、门控机制使其对数据序列适应性增高, 可以更好的去对数据信息进行控制^[23]。

图 2 为 LSTM 内部结构, 可以观察到模型首先通过计算 w_t 作为遗忘门对前一时刻状态信息 c_{t-1} 进行有选择的丢弃。随后, 输入门门控信号 r_t 对 x_t 中重要部分进行记忆并更新记忆 c_t 。完成以上操作后 l_t 控制输出门将信息输出。具体计算式见表 1。

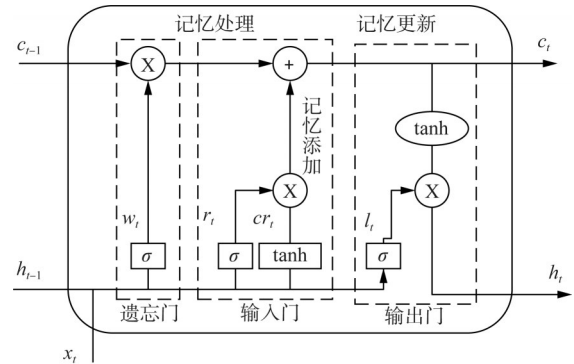


图 2 LSTM 网络结构

Fig. 2 Network structure of LSTM

x_t, h_t, h_{t-1} 分别为 t 时刻的信息输入、隐藏状态输出、以及 t 时刻的隐藏状态输入; w_t, r_t, l_{t-1} 分别对应不同的门控结构; cr_t 为补充信息; c_t 为记忆信息; W_w, W_r, W_c, W_l 分别为遗忘门、输入

表 1 LSTM 各部分计算式

Tab. 1 Calculation formulas for each part of LSTM

结构	表达式
遗忘门	$w_t = \sigma(W_w \cdot [h_{t-1}, x_t] + b_w)$
输入门	$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$
临时记忆	$cr_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c)$
记忆更新	$c_t = w_t \odot c_{t-1} + r_t \odot cr_t$
输出门	$l_t = \sigma(W_l \cdot [h_{t-1}, x_t] + b_l)$
隐藏状态	$h_t = l_t \odot \tanh(c_t)$

门、临时记忆、输出门的权重矩阵； b_w 、 b_r 、 b_c 、 b_l 分别为遗忘门、输入门、临时记忆、输出门的偏置； σ 、 $\tanh$ 为激活函数，分别将数据统一到 $[0, 1]$ 、 $[-1, 1]$ 范围内。

2.3 多头自注意力机制

MHSA 在自注意力中引入多头机制，各头可以在不同的子空间并行学习存在差异的注意力表示并进行融合，能够更好地捕获数据信息，且对于序列数据中各部分的关注更为灵活^[24-26]。

MHSA 流程如图 3 所示，对于输入信息 X 根据头的数量 h 进行切分，每一部分分别乘以其对应的权重矩阵 $[W_h^Q, W_h^K, W_h^V]$ 得到计算矩阵 $[Q_h, K_h, V_h]$ 。其中， W_h^Q 、 W_h^K 、 W_h^V 、 Q_h 、 K_h 、 V_h 分别表示第 h 个 Query、Key、Value 权重矩阵与计算矩阵。每一组计算矩阵对应产生一个输出 Y_h 。将输出进行串联，此时的结果为各头输出拼接而来，维度远大于输入信息 X 的维度，为统一维度，拼接结果乘以权重矩阵 W 得到最终输出 Y 。

2.4 TCN-LSTM-MHSA 模型构造

本文通过 TCN 在时间序列上优秀的处理能力

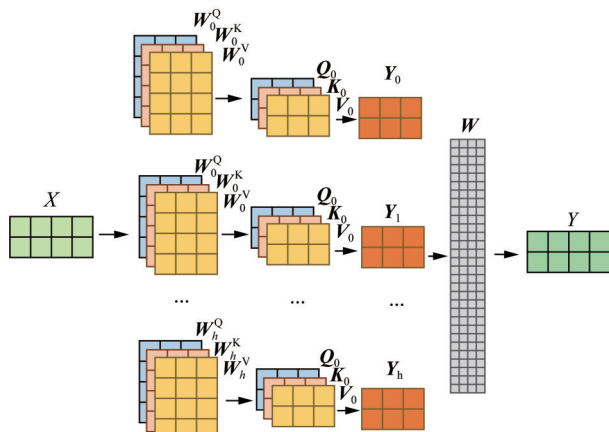


图 3 MHSA 流程

Fig. 3 Process of MHSA

对输入的风电功率数据和环境特征信息进行卷积操作，为后续模型提供丰富的特征表示。LSTM 通过门控机制对 TCN 提取的特征序列进一步处理，学习和记忆重要数据信息，更有效地理解时间序列数据的动态特征。MHSA 作为最后一部分模型，从全局范围内动态捕捉不同位置的相关性和重要信息，使用其自身特点以多头并行的机制同时学习不同的注意力表示，提升了模型的表征能力和泛化能力。结合 TCN、LSTM、MHSA 的优点构造了 TCN-LSTM-MHSA 模型，实现了高精度的风电功率预测，模型结构如图 4 所示，具体构造流程如下。

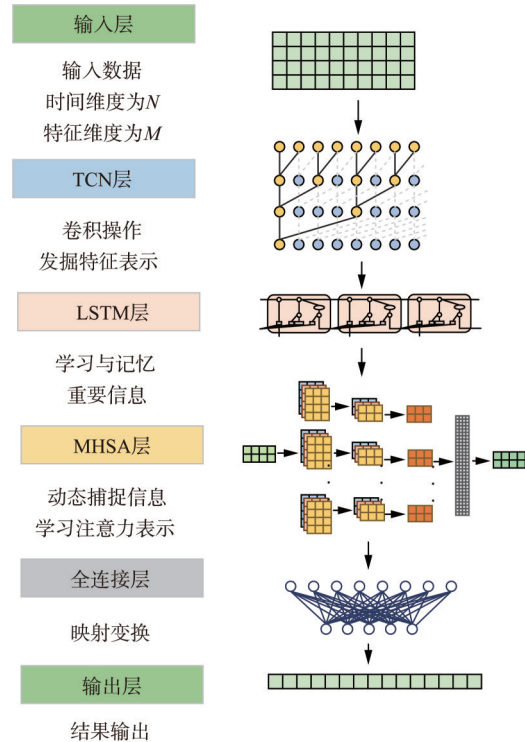


图 4 TCN-LSTM-MHSA 模型结构

Fig. 4 Structure of TCN-LSTM-MHSA model

1) 输入层：将时间维度为 N 、特征维度为 M 的数据信息构造成为 $N \times M$ 的信息矩阵输入模型，每个时间维度 N_i 对应 M 个数据信息。

2) TCN 层：对输入的时间序列矩阵信息进行卷积处理，使用包含非线性映射和两层卷积单元的残差块一个，卷积核数量 14，空间丢失因子 0.05。

3) LSTM 层：接收 TCN 的数据信息，并使用单隐层，64 神经元的 LSTM 结构对每个时间步上的输入进行处理，根据门控机制学习和记忆重要信息。

4) MHSA 层: 增强全局信息捕获能力, 以 4 头、44 键的结构并行关注不同子空间下的注意力表示, 灵活捕捉不同类型信息。

5) 全连接层与输出层: 对张量进行变换, 将模型学习到的高级特征从隐藏表示空间映射到输出空间。

3 模型评价与预测流程

3.1 模型评价

为客观地对模型预测性能进行评估, 本文使用平均绝对误差(ξ_{MAE})、均方根误差(ξ_{RMSE})、决定系数(ξ_{R^2})作为纵向比较模型的评价指标。考虑到 ξ_{MAE} 、 ξ_{RMSE} 以数值反映模型效果, 在横向对比时面对不同容量数据集指标解释能力不足。根据相关标准, 选用引入装机容量 C 的平均绝对误差(ξ_{NMAE})与均方根误差(ξ_{NRMSE})作为横向比较模型的评定指标。

$$\xi_{MAE} = \frac{\sum_{g=1}^n |y'_g - y_g|}{n} \quad (8)$$

$$\xi_{RMSE} = \sqrt{\frac{\sum_{g=1}^n (y'_g - y_g)^2}{n}} \quad (9)$$

$$\xi_{R^2} = 1 - \frac{\sum_{g=1}^n (y_g - y'_g)^2}{\sum_{g=1}^n (y_g - \bar{y}_g)^2} \quad (10)$$

$$\xi_{NMAE} = \frac{1}{n} \sum_{g=1}^n \left| \frac{y'_g - y_g}{C} \right| \quad (11)$$

$$\xi_{NRMSE} = \sqrt{\frac{1}{n} \sum_{g=1}^n \left(\frac{y'_g - y_g}{C} \right)^2} \quad (12)$$

式中: n 为数据数量; y_g 为原样本; y'_g 为 g 时刻的预测结果; $\bar{y}_g$ 为 y_g 的平均值。

3.2 模型构造

依照前文所述内容, 构造基于信息重组和 TCN-LSTM-MHSA 的风电功率预测模型, 模型预测流程如图 5 所示。

具体步骤如下。

1) 筛选并剔除数据中因设备故障或信息交互等问题造成的异常数据, 使用样条插值对缺失的数据信息进行填补, 增加数据信息的可靠性。

2) 完善数据后通过 VMD 将功率序列拆分为多个子序列, 使用 SE 对每个子序列进行评估, 以评估结果为依据对数据信息进行重组, 进而得到高熵

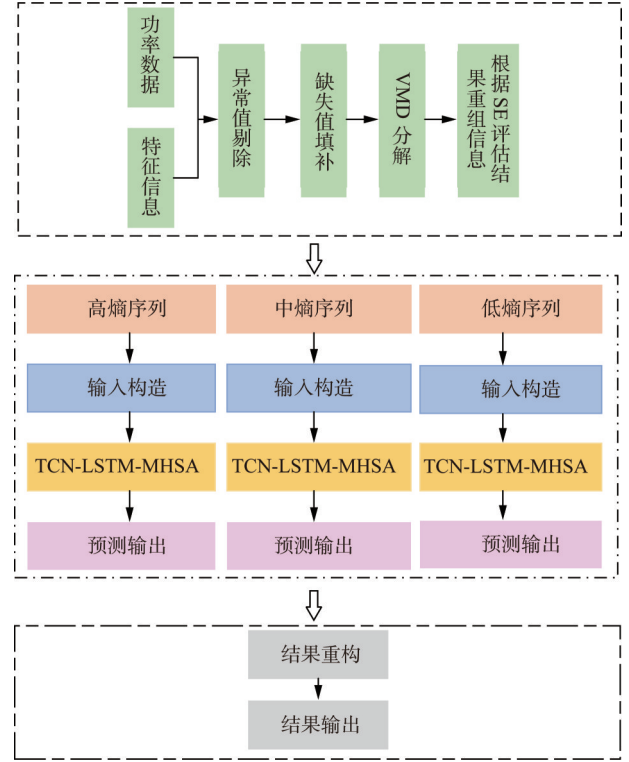


图 5 模型预测流程

Fig. 5 Prediction process of the model

序列、中熵序列、低熵序列。

3) 各序列分别与特征信息结合, 构造 $N \times M$ 维信息矩阵, 并对每一组信息矩阵以 8:1:1 的比例进行划分。

4) 建立 TCN-LSTM-MHSA 模型, 使用划分的数据集分别进行模型的训练、验证、测试。

5) 对各组模型输出值进行组合调整并输出预测值。

4 实例分析

4.1 实验数据及参数设置

本文实例分析与验证部分采用文献[27]作者于 Figshare 网站上所提供的国家电网可再生能源发电量预测竞赛数据, 提前对未来 15 min 风电功率的输出值进行预测, 实验数据详细内容见表 2。

数据采集时间为 2019 年 1 月 1 日至 2020 年 12 月 31 日, 其中风速、风向两部分分别包含了 10 m、30 m、50 m 高度以及轮毂高度相关数据, 不同高度风速、风向数据与气温、气压等气象特征共同组成 11×7016 的特征信息矩阵。特征信息矩阵结合历史功率信息构成预测模型的输入信息矩阵, 对未来的

表 2 实验数据具体信息

Tab. 2 Specific information on experimental data

数据集名称	A	B
数据来源	Figshare 网站	
数据组成	风速、风向、气温、气压、湿度、功率	
数据量/个	70 176	
时间粒度/min	15	
输出容量/MW	200	66
风机型号	GW3000/110	FD98A-1500、FD166A-2000
风机数量/台	67	20、18

风电功率进行预测，其中 80 %用于训练，其余数据以 1：1 的比例分别用于验证与测试。模型相应参数设置如表 3 所示。

表 3 各层级相应参数设置

Tab. 3 Parameter settings for each level

所属层级	名称	参数值
TCN 层	卷积核	14
	空间丢失因子	0.05
	残差块	1
LSTM 层	隐含层神经元	64
MHSA 层	注意力头	4
	通道数	44

4.2 数据处理

以数据集 A 为例，风电功率数据信息如图 6 所示。根据其变化情况可以看出，风电的功率存在周期性循环特征，且随着时间维度的推进数据信息存在一定程度的演进特点。同时由于风速的不稳定以及不确定性数据中包含了部分随机信息。

风电功率数据根据时间排序构成 N 维序列 S 。

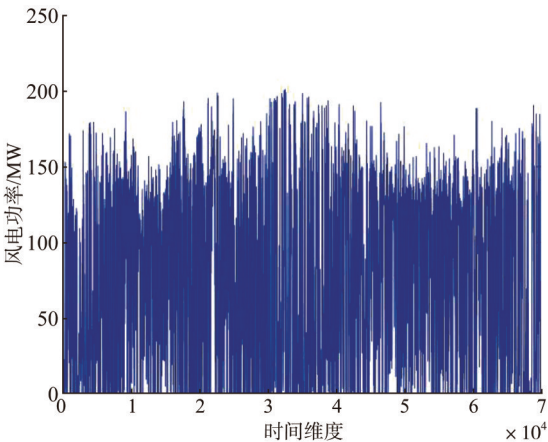


图 6 风电功率数据信息

Fig. 6 Data information on wind powers

$$S = [S_1, S_2, \dots, S_N] \tag{13}$$

为了发掘数据的非线性和非平稳特征，更为准确地分析其变化情况，使用 VMD 对序列 S 进行拆分。以增幅是否超过 50 % 为依据，选择模态数 $K=9$ ，原始数据与分解数据如图 7 所示。

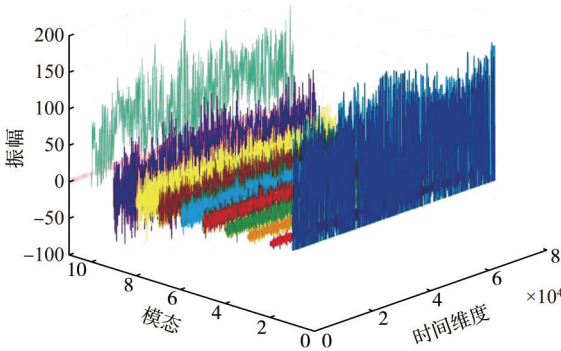


图 7 VMD 分量构成

Fig. 7 Composition of VMD components

根据 SE 计算结果(表 4)对序列进行评估，熵值小于 0.5 时认为序列混乱程度较低，信息相对集中。0.5 至 0.8 之间认为序列规律性与随机性并存，信息相对分散但仍存在结构性特征。熵值大于 0.8 时，序列混乱程度较高，信息分散，结构性特征难以识别。因此，以熵值 0.5、0.8 作为划分界限，将各序列划归不同熵值类型的集合，更为方便、准确的进行信息重组。由数据信息不稳定、构成复杂的高熵集合 $\{Res\}$ 得到高熵序列 SF_1 ；中度波动且具有一定程度规律性的中熵集合 $\{IMF4\sim IMF9\}$ 重组为中熵序列 SF_2 ；较为稳定规律的低熵集合 $\{IMF1\sim IMF3\}$ 重组为低熵序列 SF_3 。

$$SF_1 = [sf_{11}, sf_{12}, \dots, sf_{1N}] \tag{14}$$

$$SF_2 = [sf_{21}, sf_{22}, \dots, sf_{2N}] \tag{15}$$

$$SF_3 = [sf_{31}, sf_{32}, \dots, sf_{3N}] \tag{16}$$

式中： sf 根据角标的不同分别表示重构后不同序列的元素组成。

表 4 各子序列样本熵

Tab. 4 Sample entropy of each subsequence

IMF	熵值	IMF	熵值
IMF1	0.292 3	IMF6	0.664 2
IMF2	0.294 6	IMF7	0.674 3
IMF3	0.438 8	IMF8	0.671 3
IMF4	0.548 6	IMF9	0.622 3
IMF5	0.621 0	Res	0.958 2

4.3 信息重组效果分析

为分析信息重组的效果, 在数据集 A 上保证其余输入相同的情况下对 TCN-LSTM-MHSA 模型采用信息重组与不采用信息重组的两种选择进行对比分析, 图 8 展示了不同选择对模型各方面的影响。

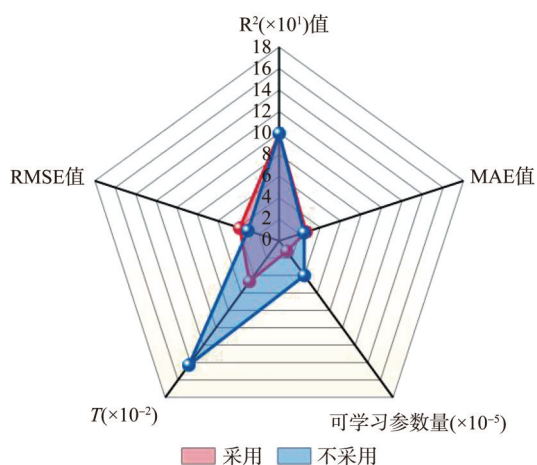


图 8 不同选择对模型的影响

Fig. 8 Impact of different choices on the model

图 8 中可以观察到, 采用信息重组对决定系数 ξ_{R^2} 与平均绝对误差 ξ_{MAE} 影响并不明显, 仅对均方根误差 ξ_{RMSE} 有所提升。但从训练时间 T 与模型可学习参数量方面来看, 采用信息重组能够有效地降低可学习参数量并缩短模型训练时间, 相比于不采用的情况效果提升在 2.5 倍以上。因此, 虽然不使用信息重组对 ξ_{RMSE} 有所降低, 但较小的优势下模型付出了训练时间成倍增加、可学习参数量大幅上升的代价。这一结果表明信息重组对模型是必要的, 其对模型训练速度的提升以及计算负担的降低有显著效果。

4.4 预测模型纵向改进对比

为确保本文模型的有效性, 检验模型改进效果, 采用输出容量 200 MW 的数据集 A 进行预测性能验证。图 9 为不同模型在数据集 A 上的预测结果对比, 表 5 为不同模型在此数据集上的性能表现评价。

通过对图 9 与表 5 的分析, 各模型均有相对优秀的预测水平, 其中本文模型的精度最高, 决定系数 ξ_{R^2} 达 0.995 5, 体现模型对复杂的风能数据具有较好的解释能力。从平均绝对误差方面能够清晰地观察到模型性能的提升, ξ_{MAE} 由 5.926 4 MW 显著

降到 2.613 4 MW, 反映了本文模型对降低整体误差的显著效果。均方根误差也呈现相同的变化趋势, 本文模型的 ξ_{RMSE} 仅为 3.847 4 MW, 表明考虑大误差项时模型策略能够在降低整体误差方面具有良好效果。

图 9 可以观察到部分预测值出现负数, 原因是设计过程考虑了实际运行可能会出现各种情况, 其中包括风力资源暂时不足的问题。在这一情况下, 风速低于切入风速, 为保障风电场自身电力需求, 需要从电网获取支持以维持运行, 导致功率方向反向, 出现负功率。这种情况并非异常, 而是在特定运行状态下的表现。

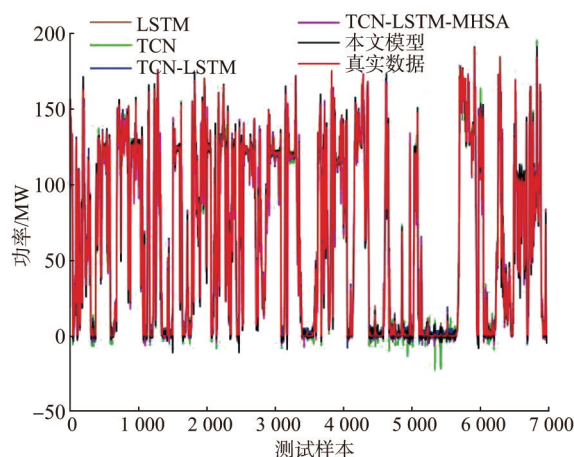


图 9 数据集 A 上各模型预测结果对比

Fig. 9 Comparison of prediction results of each model on data set A

表 5 数据集 A 上各模型性能表现评价

Tab. 5 Performance evaluation of models on data set A

模型	ξ_{MAE}/MW	ξ_{RMSE}/MW	ξ_{R^2}
TCN	5.926 4	8.887 9	0.976 0
LSTM	5.582 1	9.058 8	0.975 0
TCN-LSTM	4.429 9	7.395 7	0.983 4
TCN-LSTM-MHSA	4.086 1	7.203 2	0.984 2
本文模型	2.613 4	3.847 4	0.995 5

4.5 预测模型横向对比

为确保本文模型的有效性, 以数据集 A 为基准对功率的输出进行预测, 并横向比较各模型效果。对比模型参考文献[12-14]所提结构, 同时考虑到对比的公平性, 在参考结构的基础上进行了部分修改得到 VMD-CNN-GRU-SA、VMD-CNN-LSTM-AM(SE)模型, 使之与本文模型具有相似的结构与

复杂度,能够更好地捕捉序列中的依赖关系。以此减少评估过程的偏差,提高性能差异的可解释性与模型之间的可比性。图 10 为横向对比模型在数据集 A 上的预测结果,表 6 为各对比模型在该数据集上的表现评价。

分析图 10 与表 6 可以发现,各横向对比模型决定系数 ξ_{R^2} 均达到 0.99 以上,本文模型可达 0.9955,拥有杰出的拟合效果,能够较好地评释数据的变异性。平均绝对误差 ξ_{NMAE} 降幅分别为 22.02%、27.22%,均方根误差 ξ_{NRMSE} 降幅分别为 16.16%、29.15%,表明本文模型性能更为优异,能够很好地捕捉数据中的复杂变化关系。

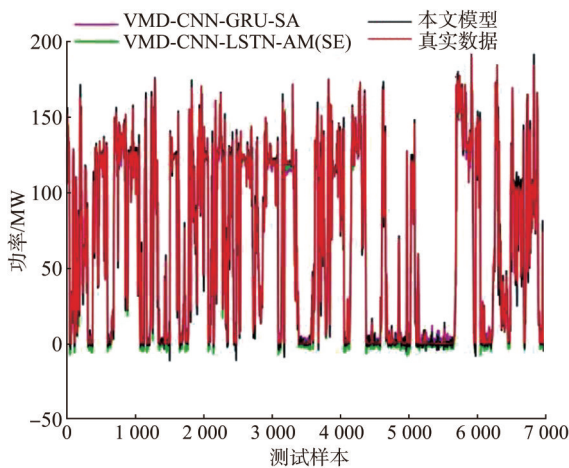


图 10 数据集 A 上各横向对比模型预测结果

Fig. 10 Prediction results of horizontal comparison models on data set A

表 6 数据集 A 上各对比模型性能表现评价

Tab. 6 Performance evaluation of various comparison models on data set A

模型	ξ_{NMAE}	ξ_{NRMSE}	ξ_{R^2}
VMD-CNN-GRU-SA	0.016 8	0.022 9	0.993 7
VMD-CNN-LSTM-AM(SE)	0.018 0	0.027 1	0.991 1
本文模型	0.013 1	0.019 2	0.995 5

考虑到模型的鲁棒性与泛化能力直接影响着其在实际应用中的可靠性,采用风机构成、输出容量均不同的数据集 B 进一步验证,各模型表现见图 11 与表 7。

对比分析图 10—11、表 6—7,在不同的数据基准下本文模型的决定系数 ξ_{R^2} 均在 0.995 以上,这一结果显著优于对比模型。在平均绝对误差 ξ_{NMAE} 与均方根误差 ξ_{NRMSE} 方面,本文模型相较对比模型

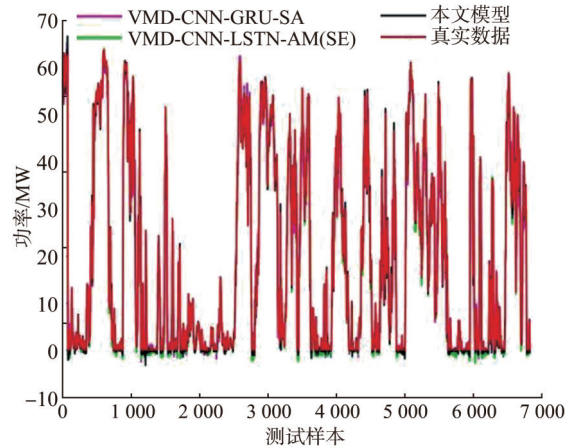


图 11 数据集 B 上各横向对比模型预测结果

Fig. 11 Prediction results of horizontal comparison models on data set B

表 7 数据集 B 上各对比模型性能表现评价

Tab. 7 Performance evaluation of various comparison models on data set B

模型	ξ_{NMAE}	ξ_{NRMSE}	ξ_{R^2}
VMD-CNN-GRU-SA	0.015 7	0.022 7	0.994 5
VMD-CNN-LSTM-AM(SE)	0.017 3	0.025 2	0.993 2
本文模型	0.011 5	0.016 8	0.997 0

最高降幅分别达 33.53%、33.33%,性能突出。以上结果表明,本文模型面对不同数据时,性能表现稳定,鲁棒性强,且能够捕捉到数据中的普遍规律,适应不同的数据特征,具有良好的泛化能力。

5 结论

本文提出一种基于信息重组和 TCN-LSTM-MHSA 的超短期风电功率预测方法,通过 VMD 分解数据并根据样本熵评估结果将信息重组为不同序列;随后结合 TCN、LSTM、MHSA 的优势构造 TCN-LSTM-MHSA 模型对各序列进行预测。经过在实际数据上的验证分析,结论如下。

1) 采用 VMD 分解数据并结合样本熵重组子序列信息能够更好地提取数据特征,同时降低子序列维度,减少计算负担。

2) 构造的 TCN-LSTM-MHSA 模型能够深入地探查数据间的关系,显著提升模型预测的准确性。

3) 在不同的数据集上对本文模型进行了纵向改进对比以及横向对比,结果表明:在不同的输出容量以及风机构成下,本文模型的预测性能优于对比模型。

参考文献

- [1] 国家能源局. 国家能源局发布 1-11 月份全国电力工业统计数据 [EB/OL]. (2023-12-20)[2024-01-10]. https://www.nea.gov.cn/2023-12/20/c_1310756286.htm.
- [2] XIAO H, PEI W, DENG W, et al. Enhancing risk control ability of distribution network for improved renewable energy integration through flexible DC interconnection [J]. *Applied Energy*, 2021(284): 116387. 1-116387. 16.
- [3] 吴慧军, 郭超雨, 苏承国, 等. 基于 EEMD-GRU-MC 的短期风功率组合预测方法 [J]. *南方电网技术*, 2023, 17(2): 66-73.
WU Huijun, GUO Chaoyu, SU Chengguo, et al. Combined prediction method for short-term wind power based on EEMD-GRU-MC [J]. *Southern Power System Technology*, 2023, 17(2): 66-73.
- [4] DING Y, CHEN Z, ZHANG H, et al. A short-term wind power prediction model based on CEEMD and WOA-KELM [J]. *Renewable Energy*, 2022(189): 188-198.
- [5] LIU H, CHEN C, LV X, et al. Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods [J]. *Energy Conversion and Management*, 2019(195): 328-345.
- [6] LI H. Short-term wind power prediction via spatial temporal analysis and deep residual networks [J]. *Frontiers in Energy Research*, 2022(10): 920407. 1-920407. 9.
- [7] 刘新宇, 蒲欣雨, 李继方, 等. 基于贝叶斯优化的 VMD-GRU 短期风电功率预测 [J]. *电力系统保护与控制*, 2023, 51(21): 158-165.
LIU Xinyu, PU Xinyu, LI Jifang, et al. Short-term wind power prediction of VMD-GRU based on Bayesian optimization [J]. *Power System Protection and Control*, 2023, 51(21): 158-165.
- [8] 刘栋, 魏霞, 王维庆, 等. 基于 VMD-WPE 和 SSA-ELM 的短期风电功率预测研究 [J]. *太阳能学报*, 2022, 43(12): 360-367.
LIU Dong, WEI Xia, WANG Weiqing, et al. Short term wind power forecasting based on VMD-WPE and SSA-ELM [J]. *Acta Energetica Solaris Sinica*, 2022, 43(12): 360-367.
- [9] 卢继平, 曾燕婷, 喻华, 等. 基于改进 AWNN 的风电功率超短期多步预测 [J]. *太阳能学报*, 2021, 42(1): 166-173.
LU Jiping, ZENG Yanting, YU Hua, et al. Ultra-short-term wind power multi-step forecasting based on improved AWNN [J]. *Acta Energetica Solaris Sinica*, 2021, 42(1): 166-173.
- [10] 王红君, 谢煜轩, 赵辉, 等. 改进黑猩猩算法和 LSSVR-BiLSTM 双尺度模型的短期风功率预测 [J]. *重庆理工大学学报(自然科学)*, 2023, 37(9): 243-252.
WANG Hongjun, XIE Yuxuan, ZHAO Hui, et al. Short-term wind power prediction based on improved chimp algorithm and LSSVR-BiLSTM dual scale model [J]. *Journal of Chongqing University of Technology (Natural Science)*, 2023, 37(9): 243-252.
- [11] 刘吉成, 朱玺瑞, 于晶. 基于 IWOA-SA-Elman 神经网络的短期风电功率预测 [J]. *太阳能学报*, 2024, 45(1): 143-150.
LIU Jicheng, ZHU Xirui, YU Jing. Short-term wind power prediction based on IWOA-SA-Elman neural network [J]. *Acta Energetica Solaris Sinica*, 2024, 45(1): 143-150.
- [12] ZHAO Z, YUN S, JIA L, et al. Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features [J]. *Engineering Applications of Artificial Intelligence*, 2023(121): 105982. 1-105982. 14.
- [13] ZHANG J, ZHAO Z, YAN J, et al. Ultra-short-term wind power forecasting based on CGAN-CNN-LSTM model supported by lidar [J]. *Sensors*, 2023, 23(9): 4369. 1-4369. 20.
- [14] 章志晃, 徐启峰, 林穿. 基于 SSA-BiLSTM-AM 的短期风电功率预测 [J]. *国外电子测量技术*, 2023, 42(3): 46-51.
ZHANG Zhihuang, XU Qifeng, LIN Chuan. Short-term wind power prediction based on SSA-BiLSTM-AM [J]. *Foreign Electronic Measurement Technology*, 2023, 42(3): 46-51.
- [15] 张未, 余成波, 王士彬, 等. 基于 VMD-LSTM-LightGBM 的多特征短期电力负荷预测 [J]. *南方电网技术*, 2023, 17(2): 74-81.
ZHANG Wei, YU Chengbo, WANG Shibin, et al. Multi-featured short-term power load forecasting based on VMD-LSTM-LightGBM. [J]. *Southern Power System Technology*, 2023, 17(2): 74-81.
- [16] 李英顺, 阚宏达, 郭占男, 等. 基于数据预处理和 VMD-LSTM-GPR 的锂离子电池剩余寿命预测 [J]. *电工技术学报*, 2024, 39(10): 3244-3258.
LI Yingshun, KAN Hongda, GUO Zhannan, et al. Prediction of remaining useful life of lithiumion battery based on data preprocessing and VMD-LSTM-GPR [J]. *Transactions of China Electrotechnical Society*, 2024, 39(10): 3244-3258.
- [17] JIAO X, ZHANG D, SONG D, et al. Wind speed prediction based on VMD-BLS and error compensation [J]. *Journal of Marine Science and Engineering*, 2023, 11(5): 1082. 1-1082. 22.
- [18] ZHUANG D, LIU H, ZHENG H, et al. The IBA-ISMO method for rolling bearing fault diagnosis based on VMD-sample entropy [J]. *Sensors*, 2023, 23(2): 991. 1-991. 19.
- [19] WIDODO A, SHIM M C, CAESARENDRA W, et al. Intelligent prognostics for battery health monitoring based on sample entropy [J]. *Expert Systems with Applications*, 2011, 38(9): 11763-11769.
- [20] BAI S, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. [J]. *arXiv preprint arXiv:1803.01271*, 2018.
- [21] 刘辉, 凌宁青, 罗志强, 等. 基于 TCN-LSTM 和气象相似日集的电网短期负荷预测方法 [J]. *智慧电力*, 2022, 50(8): 30-37.
LIU Hui, LING Ningqing, LUO Zhiqiang, et al. Power grid short term load forecasting method based on TCN-LSTM and meteorological similar day sets [J]. *Smart Power*, 2022, 50(8): 30-37.
- [22] 宋技峰, 彭小圣, 杨子民, 等. 基于偏差补偿 TCN-LSTM 和梯级迁移策略的短期风电功率预测 [J]. *南方电网技术*, 2023, 17(12): 71-79.
SONG Jifeng, PENG Xiaosheng, YANG Zimin, et al. Short-term wind power prediction based on deviation compensation TCN-LSTM and step transfer strategy [J]. *Southern Power System Technology*, 2023, 17(12): 71-79.
- [23] 程先龙, 保佑智, 何度江, 等. 基于 EMD-ELM-LSTM 的短期风电功率预测 [J]. *昆明理工大学学报(自然科学版)*, 2023, 48(6): 78-87.

(下转第 110 页 Continued on Page 110)