

基于 DCT-Channel Attention-TCN-Transformer 混合模型的短期风电功率预测方法

江善和, 徐小艳, 陈文胜

(安庆师范大学 电子工程与智能制造学院, 安徽 安庆 246011)

摘要: 针对风电功率预测数据的特征复杂性和高波动性难题,文章提出了一种基于离散余弦变换-通道注意力-时间卷积网络-Transformer(DCT-Channel Attention-TCN-Transformer)的混合模型。该方法引入基于 DCT 的通道注意力机制,旨在捕捉输入信号的频率特征与多尺度信息,增强模型对不同通道特征的选择性;利用 TCN 提取时间依赖性特征,降低预测的不确定性;结合 Transformer 模型的自注意力机制,实现对长时间序列数据的全局依赖建模,提升预测精度。基于 Bangalore 某风电场数据预测结果表明:该模型的决定系数(R^2)达到 0.998;相较于 RNN,LSTM,GRU 和 TCN-Transformer,所提方法的平均绝对误差(MAE)分别降低 85.53%,75.69%,78.60%和 71.10%,均方根误差(RMSE)分别降低 80.61%,73.65%,74.26%和 72.20%。研究表明,文章所提混合模型预测方法具有优异的预测性能与稳健性。

关键词: 离散余弦变换;通道注意力;时间卷积网络;Transformer;风电功率预测

中图分类号: TK81; TM614 **文献标志码:** A **文章编号:** 1671-5292(2026)08-1069-10

0 引言

随着全球对可再生能源需求的日益增长,风力发电已成为现代能源体系的重要组成部分^[1]。然而,受风资源间歇性和不确定性的影响,风电功率预测面临极大挑战^[2],尤其是短期预测(通常指未来数小时至一日)。短期预测的精度直接关系到电网调度、能量管理及系统的稳定运行。实现高精度的短期风电功率预测,不仅能降低风电波动引发的电网不稳定风险,还能优化发电计划并降低备用容量需求,显著提高风电并网的安全性和可靠性。

现有风电功率预测模型及优化方法虽已取得阶段性成果,但仍存在预测精度低、模型鲁棒性差等问题。自回归移动平均(ARMA)^[3]、差分自回归移动平均(ARIMA)^[4]等传统统计预测方法难以有效捕捉风电数据的非线性特征,使预测性能受到极大限制。为此,诸多学者引入支持向量机(SVM)^[5]、随机森林(RF)^[6]、极限学习机(ELM)^[7]等机器学习方法,提高了复杂数据的处理能力,但在面对风电数据的长时间依赖性和多尺度特征时仍显不足。

随着深度学习算法的提出和应用,长短期记

忆网络 LSTM^[8]、门控循环单元(GRU)^[9]、循环神经网络(RNN)^[10]、卷积神经网络(CNN)^[11]等模型在时间序列预测领域得到了广泛应用,取得了较好的应用效果。深度学习模型在处理时序数据中优势突出,但在长时间依赖、频率信息提取和通道间相互依赖性的处理上仍存在着局限性,导致整体预测精度难以进一步提升。由此可见,单一预测模型难以全面捕捉风电数据中的复杂非线性与多尺度特征。

为了提高预测模型的准确性和鲁棒性,研究者提出了采用模型组合的方式^[12,13],融合不同模型优势以弥补单一模型的局限性,以此改善风电功率预测结果的精度和稳定性。文献[14]构建了融合变分模态分解(VMD)、模糊熵(FE)、卷积神经网络(CNN)和双向长短期记忆神经网络(BiLSTM)的组合预测模型,发挥多类算法优势,但该模型在处理数据的时间依赖性时仍存在局限性。徐钊提出了 TCN-Wpsformer 混合模型^[1],先借助位置编码提取数据特征并将其与位置编码拼接,再引入 TCN 完成模型搭建。该模型通过位置编码增强了特征提取,但其对长时间序列数据的处理能力仍受限于位置编码的设计,无法充分捕捉全

收稿日期: 2025-03-16。

基金项目: 国家自然科学基金项目(51607004);安徽省自然科学基金项目(2008085MF197)。

作者简介: 江善和(1975-),男,博士,教授,研究方向为新能源电力系统优化、智能计算与优化算法等。E-mail:jshxllw@163.com

局依赖性。

针对上述问题,本文提出了一种基于离散余弦变换-通道注意力-时间卷积网络-Transformer (DCT-Channel Attention-TCN-Transformer) 的混合模型用于短期风电功率预测。首先,引入 DCT 增强的通道注意力机制,利用 DCT 对数据进行频域转换,精准选择和响应风电数据中各通道有效特征,充分捕捉数据频率信息,增强模型对关键特征的敏感性,提升整体预测性能;其次,采用 TCN 处理序列数据,提取时间依赖性特征,结合残差连接应对梯度消失问题,提高模型的稳定性;最后,将 Transformer 自注意力机制与 TCN 融合,强化长时间序列数据的全局依赖建模,进一步提升模型预测精度和泛化能力。所提混合模型可高效提取频域和时域特征,大幅提高复杂时序数据的处理能力,有效增强风电功率预测模型的预测准确性和鲁棒性。

1 离散余弦变换-通道注意力

1.1 离散余弦变换(DCT)

DCT 可实现时域信号向频域信号的转换^[15],将信号分解为若干余弦基函数的加权和。该方法广泛应用于数据压缩、图像处理和信号分析领域,在捕捉信号频率成分方面优势明显。时序数据转换至频域后,可更清晰地揭示信号中的周期性特征和低频分量。

对于长度为 N 的一维序列 x_n ,其 DCT 的变换公式为

$$x_k = \sum_{n=0}^{N-1} x_n \cdot \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad (1)$$

式中: x_k 为频域中的第 k 个 DCT 系数,表示信号在对应频率下的幅度。

对时序数据应用 DCT,能够将原始信号分解为多组不同频率成分的线性组合,直观开展频域能量分布特性分析。一维序列 x_n 的 DCT 逆变换表达式为

$$x_n = \frac{1}{N} \left\{ X_0 + 2 \sum_{k=1}^{N-1} x_k \cdot \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \right\} \quad (2)$$

式中: x_n 为时域信号的第 n 个样本。

1.2 通道注意力机制(Channel Attention)

Channel Attention 由 SENet 提出^[16],核心为 Squeeze-and-Excitation (SE) 模块,能够反映全局信息、捕捉通道间的依赖关系,增强特征表示能力,其结构如图 1 所示。

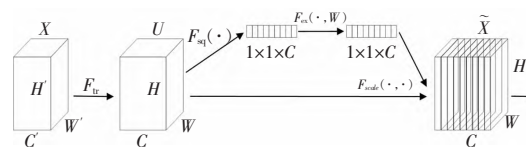


图 1 SENet 模块

Fig.1 SENet module

SE 模块由 Squeeze 和 Excitation 两部分组成。Squeeze 模块通过全局平均池化(Global Average Pooling)提取全局空间信息;Excitation 模块利用全连接层和非线性激活函数(如 ReLU 和 Sigmoid)捕捉通道间的相互关系并生成注意力向量;最终将输入特征各通道通过与注意力向量对应的元素相乘进行权重调整。

给定输入特征 $X \in \mathbf{R}^{H' \times W \times C}$,经卷积运算完成 F_u 提取后,可得到输入特征图 $U \in \mathbf{R}^{H \times W \times C}$ 的各通道 $u_c \in \mathbf{R}^{H \times W}$,其中, H', H, W 为特征图空间维度, C 为通道量, $c \in (1, 2, \dots, C)$ 。SE 模块具体操作流程如下。

①Squeeze 阶段:通过全局平均池化压缩空间维度,生成通道描述符 $z_c \in \mathbf{R}^C$ 。

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (3)$$

②Excitation 阶段:利用全连接层和激活函数生成通道注意力权重 $s \in \mathbf{R}^C$ 。

$$s = F_{ex}(z, W) = \sigma(W_2 \cdot \delta(W_1 \cdot z)) \quad (4)$$

式中: σ 为 sigmoid 函数; W_1, W_2 为权重矩阵; δ 为 ReLU 非线性激活函数。

③Reweight 阶段:将学习权重作用于原始特征图,完成通道级特征增强。

$$\tilde{X} = F_{scale}(s, U) = [s_1 u_1, s_2 u_2, \dots, s_C u_C] \quad (5)$$

该机制通过学习各通道的重要性权重,动态调整模型对不同通道特征的关注意度,增强关键信息的提取能力。对于多通道时序数据,通道注意力机制可充分捕捉通道间的互补信息,强化适配预测任务的有效特征。

2 关键时序建模方法

2.1 时间卷积网络(TCN)

TCN 是基于 CNN 构建的处理序列数据专用建模架构^[17]。该网络通过改进卷积神经网络结构实现时序建模,核心包含因果卷积、膨胀卷积与残差连接 3 大模块。图 2 为 kernel size=3, dilations=[1, 2, 4] 的 TCN 结构。网络采用一维因果卷积确保

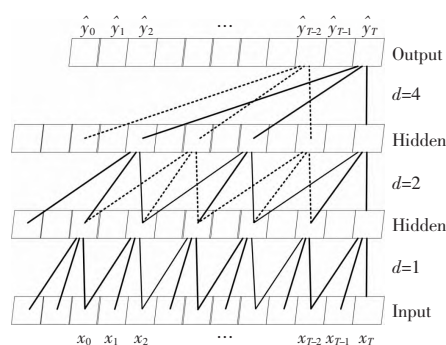


图2 TCN 结构图

Fig.2 TCN architecture diagram

输出仅依赖历史输入,通过堆叠指数增长的膨胀卷积层,使感受野呈指数扩展。该结构仅需 $O(\log N)$ 层即可覆盖长度为 N 的序列,大幅提升长程依赖建模效率。

为缓解深层网络的梯度消失问题,各卷积模块均引入残差连接,通过跳跃连接保持梯度通路。残差模块包含一个引出系列变换 F 的分支,其输出被添加到输入 x ,具体运算式为

$$O = \text{Activation}(x + F(x)) \quad (6)$$

2.2 Transformer 模型

Transformer 模型由 Vaswani A 于 2017 年提出,主要用于机器翻译、文本生成等序列到序列的任务学习^[18],其核心创新在于多头自注意力机制,编码器-解码器模型结构如图 3 所示。

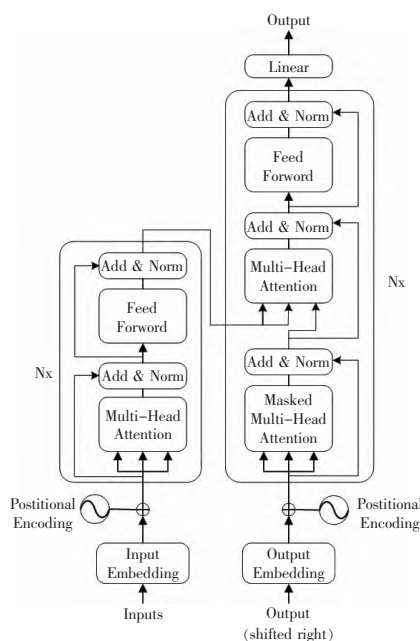


图3 编码器-解码器结构图

Fig.3 Encoder-decoder architecture diagram

给定输入序列 X , 经线性变换生成查询矩阵

($Q=XW^Q$)、键矩阵($K=XW^K$)和值矩阵($V=XW^V$),其中, W^Q, W^K, W^V 为用于线性变换的权重矩阵。自注意力计算通过缩放点积形式实现,其表达式为

$$\text{Attention}(Q, K, V) = \text{Softmax}(QK^T / \sqrt{d_k})V \quad (7)$$

式中: d_k 为键向量维度,作为缩放因子用于防止点积值过大,以稳定梯度。

多头机制将注意力运算扩展到 h 个独立子空间:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_h)W^0 \quad (8)$$

式中: W^0 为线性变换的权重矩阵。

各注意力头 head_i 可通过式(9)计算:

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (9)$$

各注意力头通过参数矩阵独立完成计算,最终拼接后经线性变换输出。编码器堆叠多个包含多头注意力和前馈网络(FFN)的模块,各子层采用残差连接与层归一化操作。解码器额外引入掩码机制,限制当前位置仅能关注历史信息,以此保持模型自回归运算特性。

3 DCT-Channel Attention-TCN-Transformer 混合模型

3.1 离散余弦变换-通道注意力

原始多通道时序数据兼具复杂的非线性和非平稳性,且掺杂大量的噪声干扰。为此,本文采用一种融合离散余弦变换与通道注意力机制的处理方法,优化风电功率预测时间序列数据解析效果。首先对多通道输入数据应用 DCT 频域变换,深度挖掘时序波动中潜藏的频率特征,强化模型对数据周期变化与趋势模式的感知能力;其次通过通道注意力机制对频域特征进行加权处理,动态调整特征通道的表达强度,实现对重要频域特征的强化与无效信息的抑制,从而增强对关键特征的关。离散余弦变换-通道注意力(DCT-Channel Attention)融合机制结构如图 4 所示。

输入序列 $X \in \mathbf{R}^{C \times T}$ 经离散余弦变换和通道注意力机制融合处理的具体步骤如下:

步骤 1, 将输入特征图沿着通道维度分成 n 个子组 $V=(v_0, v_1, \dots, v_{n-1})$, 设定 $n=C$;

步骤 2, 对各子组进行离散余弦变换,依据式(1)将时域信号转换为频域信号,有 $F^n = \text{DCT} \cdot (V)$, $F^n \in \mathbf{R}^T$ 为频域表示;

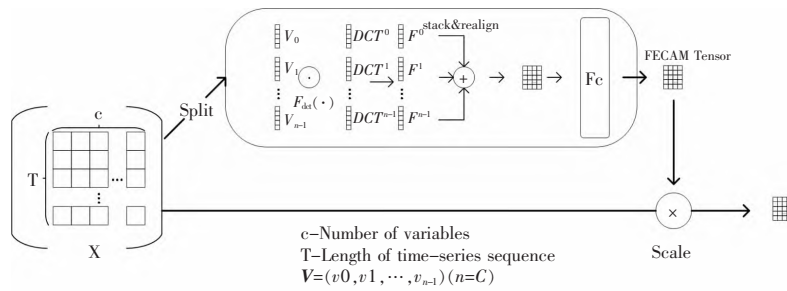


图4 DCT-Channel Attention 结构

Fig.4 DCT-Channel attention architecture

步骤 3,将所有通道 DCT 变换结果组合成频域特征矩阵 $F \in \mathbf{R}^{C \times T}$;

步骤 4,将频域特征矩阵 F 输入通道注意力机制,提取各通道的重要性权重。该机制通过压缩和激励操作生成注意力向量 $\alpha=[\alpha_1, \alpha_2, \dots, \alpha_c] \in \mathbf{R}^c$,每个通道对应一个权重,整个过程记为 $\alpha=\text{SE}(F)$;

步骤 5,将注意力权重应用于频域特征,对每个通道频域 F_i 进行加权处理,得到加权频域特征 $F_i=\alpha_i F_i$,进一步组合为 $F \in \mathbf{R}^{C \times T}$;

步骤 6,依据式(2)对加权后的频域特征进行

逆离散余弦变换(IDCT),有 $x_i(t)=\text{IDCT}(F)$, $x_i \in \mathbf{R}^T$ 为第 i 个通道的反变换结果,将频域信号还原为时域表示;

步骤 7,将所有通道的反变换结果 $\{x_1, x_2, \dots, x_c\}$ 拼接为最终加权并还原的时域特征矩阵 $X \in \mathbf{R}^{C \times T}$,作为后续模型的输入。

3.2 混合模型

本文所提 DCT-Channel Attention-TCN-Transformer 混合模型,可通过多层次特征处理与建模提升预测性能,模型整体结构如图 5 所示。

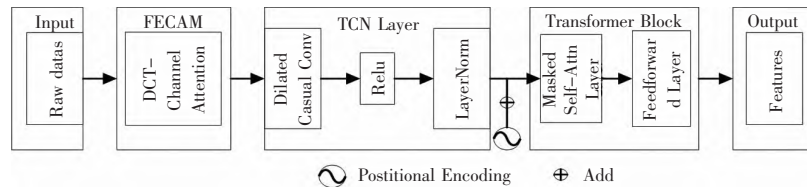


图5 DCT-Channel Attention-TCN-Transformer 结构

Fig.5 DCT-channel attention-TCN-transformer architecture

首先,引入通道注意力机制,对经过 DCT 的频域特征进行加权处理,使模型能更有效地捕捉数据中的关键频率成分。加权后的特征被输入到 TCN-Transformer 模块,进行深层时序建模与后续预测。

通过因果卷积和扩展卷积,TCN 能够捕捉长时间序列的依赖关系,处理长时依赖的时序数据,对风电功率的逐时或逐日预测效果显著。TCN 的设计确保了时间顺序的正确性,同时支持可变长度序列输入,在复杂的风电功率预测任务中具备良好的实用性。

作为 TCN 的补充,Transformer 通过自注意力机制增强全局依赖建模能力,其中多头注意力机制可以同时关注不同位置的特征,充分捕捉风电功率数据中的全局模式与长距离依赖,实现全局视角下的预测,进一步提高混合模型的整体预测

精度。

3.3 时间复杂度分析

在风电功率预测等对时间敏感性要求较高的场景中,模型的计算时间复杂度与推理效率直接影响其实用价值。因此,除了提高预测精度,提升模型的实时响应能力同样至关重要。本文所提 DCT-Channel Attention-TCN-Transformer 混合模型在结构设计中充分考虑了效率与性能的平衡。

从离散余弦变换与通道注意力机制方面看,对于输入张量 $X \in \mathbf{R}^{C \times T}$,各通道进行一次一维 DCT,其时间复杂度为 $O(T \log T)$,总复杂度为 $O(CT \log T)$,实际可采用快速余弦变换算法在 GPU 支持下高效完成。SE 通道注意力机制主要包括全局平均池化 $[O(T)]$ 、两层全连接层及激活函数运算,其整体复杂度为 $O(C)$,在整个混合模型中占比很小,不会对模型实时性构成显著负担。

TCN 模块的计算复杂度主要来自因果膨胀卷积操作。设输入序列长度为 T 、特征通道数为 C 、卷积核大小为 k 、网络层数为 L ,则单层膨胀卷积的复杂度为 $O(CkT)$,四层残差结构的总复杂度为 $O(LCkT)$ 。由于采用了膨胀率按指数增长策略,有效扩大了感受野而不显著增加计算量。同时,因果卷积允许并行计算,相较于 LSTM 或 GRU 等循环神经网络具有更高的并行性与推理效率。

Transformer 模块的计算复杂度主要集中于多头自注意力机制和前馈神经网络。对长度为 T 、模型维度为 d 的序列,单层多头注意力机制的计算复杂度为 $O(T^2d)$,前馈网络复杂度为 $O(Td^2)$ 。尽管自注意力的时间复杂度仍呈二次增长,但本文采用的是三层浅层结构(3 层 Encoder-Decoder,4 头注意力, $d=32$),且模型输入仅采用 30 h 短时序历史数据,实际推理时间可控,实验结果表明该混合模型整体运行效率良好。

4 实验设计及结果

4.1 实验设计

为验证所提 DCT-Channel Attention-TCN-Transformer 混合模型在短期风电功率预测中的有效性和优越性,本文选取 RNN^[10],LSTM^[8],GRU^[9]和 TCN-Transformer^[19]4 种主流短期风电功率预测模型进行性能对比。模型验证数据采用 Bangalore 某风电场 2007 年 1 月-2021 年 7 月的实测风电数据,采集间隔 1h,除风电功率外,同步收录气温、露点温度、相对湿度、风速、风向、气压等气象数据,以风电功率为预测目标变量,其余为与风电生成相关的特征变量。

实验流程分为数据预处理、模型训练与模型测试三个阶段。其中时间特征编码采用 time features 函数完成,频率按小时级别设定。实验基于 Python3.6 实现。在数据预处理中,对原始数据进行清洗和归一化处理,以确保输入数据的质量和模型的收敛性。数据集按比例划分为训练集(2007 年 1 月-2014 年 3 月)和测试集(2014 年 4 月-2021 年 7 月)。采用滑动窗口构建时序样本,输入序列长度设为 30,学习率为 0.000 1,丢弃率为 0.1。

4.2 关键参数敏感性分析

混合模型的预测性能依赖关键参数的设定,

本文仅对 5 个关键参数开展敏感性分析,其余参数采用经验取值并保持不变。DCT 频率分量保留数反映了信号在不同频率上的特征,故该参数的设置至关重要;TCN 的卷积核尺寸与网络层数对模型特征学习能力、计算效率及特征提取效果影响显著;Transformer 模型的关键参数主要是注意力头数和嵌入维度。因此,本文主要对混合模型中这 5 个关键参数进行寻优,选取 Bangalore 风电场实测数据进行测试,以平均绝对误差(MAE)作为考核指标。

考虑到实验过程中无法枚举参数取值范围的所有设置,故选取若干组典型参数组合完成参数敏感性分析。混合模型在不同参数组合下的 MAE 预测结果如表 1 所示。

表 1 不同参数组合下的混合模型预测性能结果
Table 1 Performance results of different parameter combinations for hybrid model

DCT 频率 分量保留数	TCN		Transformer (注意力头数/嵌入维度)		
	卷积核 尺度	网络 层数	2/64	4/128	8/256
8	3	4	15.423 71	15.417 32	15.408 95
	5	4	15.672 84	15.668 12	15.653 29
	3	6	15.532 67	15.526 81	15.518 43
16	3	4	13.102 57	12.115 93	12.132 85
	5	4	12.214 93	12.208 65	12.197 31
	3	6	12.184 72	12.076 38	12.162 54
24	3	4	12.183 45	12.176 92	12.165 83
	5	4	12.231 67	12.225 14	12.214 02
	3	6	12.202 58	12.194 31	12.181 49
32	3	4	12.218 73	12.212 45	12.201 16
	5	4	12.267 84	12.261 27	12.249 83
	3	6	12.238 14	12.229 81	12.217 95

由表 1 可知:TCN 采用 4 级残差块结构,每块包含 2 层因果膨胀卷积、通道数为 32、感受野为 63;Transformer 采用 3 层编解码架构,模型维度为 32、隐藏层维度为 128。在此基础上,模型最优参数组合为 DCT 频率分量保留数为 16,TCN 卷积核尺寸为 3、网络层数为 4,Transformer 注意力头数为 4、嵌入维度为 128。在该组合参数下,模型 MAE 最优值为 12.115 93,即为本文所提混合模型的最优参数设置。

4.3 结果分析

4.3.1 预测性能分析

图 6 为本文所提混合模型在 2014 年 4 月-

2021 年 7 月的风电功率预测曲线。

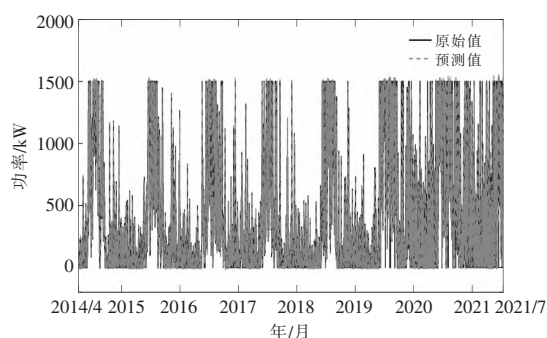


图 6 本文混合模型的预测结果

Fig.6 Prediction results of the proposed model

由图 6 可知,模型通过引入 DCT 增强频率信息提取能力,结合通道注意力机制提升不同特征通道的选择性,使预测曲线与真实值高度拟合,表现出较低的误差和更强的鲁棒性。图 7 为对应预测误差散点分布。

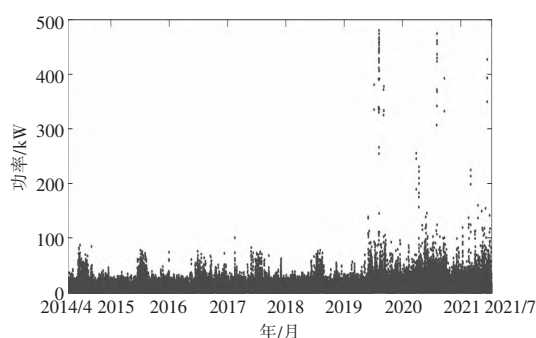


图 7 本文混合模型的预测误差散点图

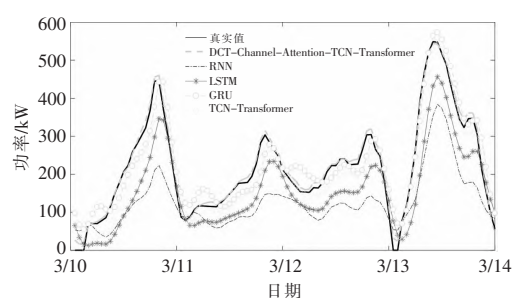
Fig.7 Prediction errors of the proposed model

由图 7 可以看出,该模型的误差散点集中在 0~50 内,表明所提方法能够有效捕捉风电数据中的多尺度特征和频率信息,显著提高短期风电功率预测精度。

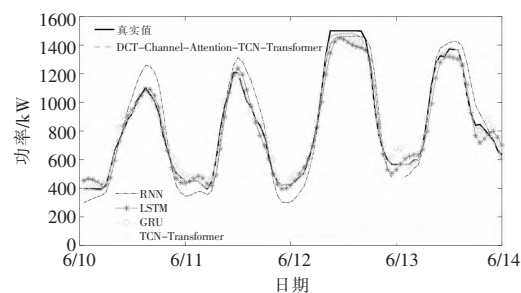
4.3.2 性能比较分析

为进一步验证各模型在风电功率高波动期间及复杂局部信息提取方面的泛化预测能力,本文选取同一年度不同季节实测样本进行对比实验,包括 2016 年春季(3 月 10~14 日)、夏季(6 月 10~14 日)、秋季(9 月 10~14 日)和冬季(12 月 10~14 日)的时序数据,共计 384 组样本。所有对比模型均沿用前文统一参数配置。图 8 为 5 种模型的预测性能对比结果,从中可以直观观测出不同模型的预测性能差异。

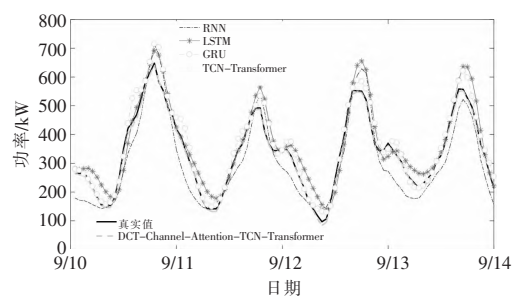
由图 8 可知,本文所提混合模型可精准捕捉数据的长期趋势与短期波动。在春季实测数据中,



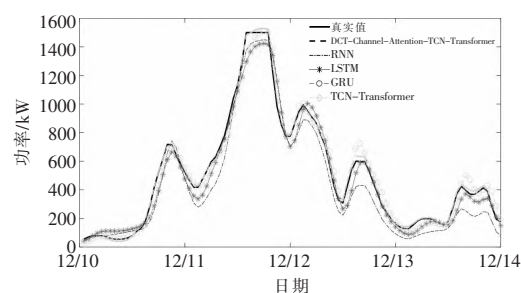
(a)春季时段不同模型预测结果对比



(b)夏季时段不同模型预测结果对比



(c)秋季时段不同模型预测结果对比



(d)冬季时段不同模型预测结果对比

图 8 不同季节时段不同模型预测结果对比

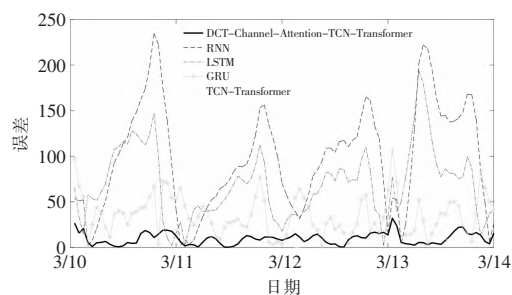
Fig.8 Comparison of prediction results of different seasons for different models

针对第 35~48 时间点功率骤升、第 80~96 时间点功率骤降两类突变工况,模型预测曲线与实际工况高度契合;而 RNN,LSTM 在峰值处存在明显滞后或过平滑现象。冬季实测数据中存在大量极端值和复杂波动,本文模型通过 DCT 增强局部特征提取和通道注意力机制动态加权关键信息,有效抑制了高频波动下的预测过冲;夏季多峰形态时

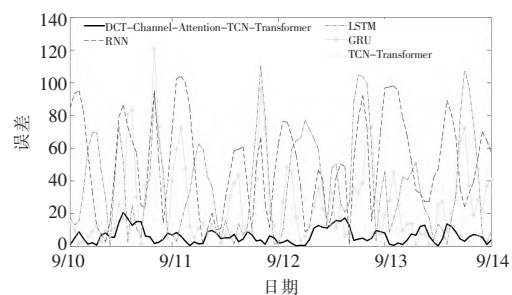
序数据中,本文模型预测误差波动范围明显低于其他模型,复杂局部信息提取性能优越;相比之下,GRU 与 TCN-Transformer 模型因缺乏多尺度特征融合,对第 50~60 时间点次峰等的预测偏差较大。

总体而言,DCT-Channel Attention-TCN-Transformer 模型在不同季节的预测曲线与实际观测值高度一致,能够准确刻画风电功率的时序波动,对功率快速上升、下降场景具备极强的动态响应能力。RNN 模型在建模长期依赖关系时存在梯度消失问题,导致预测曲线波动范围偏小,难以反映真实趋势;LSTM 模型引入门控机制,一定程度上缓解了该问题,平稳阶段预测效果尚可,但剧烈波动区域仍存在滞后。GRU 作为 LSTM 的简化版本,表现介于 RNN 和 LSTM 之间,在部分区段能够有效捕捉功率趋势,但整体拟合度略低于 LSTM。TCN-Transformer 模型凭借时间卷积模块提升了对局部特征的感知能力,预测结果相对平滑,但在剧烈变化区段灵敏度不及本文所提模型。

为直观对比各模型的预测误差,图 9 给出了春季和秋季时段的预测误差曲线。



(a) 春季时段预测误差结果



(b) 秋季时段预测误差结果

图 9 不同模型预测误差结果对比

Fig.9 Comparison results of prediction errors of different models

由图 9 可知,DCT-Channel Attention-TCN-Transformer 模型误差幅度最小,整体分布最集中,

预测值稳定性强,可有效抑制极端误差。该模型通过引入 DCT 频域特征提取与通道注意力机制,增强了对复杂风电时序特征的表达能力,在不同季节中均保持低误差。相比之下,RNN 模型误差波动大、稳定性差,对非线性和突变特征建模能力不足;LSTM 模型的误差分布相对紧凑,但在突变区段仍出现误差峰值,对局部动态变化响应不够及时;GRU 模型在多数时间点误差较小,但极端波动时仍存在短时预测偏差;TCN-Transformer 模型多数时间点的误差控制较为合理,但受限于其通道间特征融合能力,整体误差仍高于本文所提模型。

为定量比较各模型的预测性能,本文选取决定系数(R^2)、平均绝对误差(MAE)、均方根误差(RMSE)3 项评价指标开展综合性能评估,各模型指标计算结果如表 2 所示。

表 2 5 种不同预测方法的评价指标结果

Table 2 Evaluation index results of five prediction methods

预测方法	R^2	MAE	RMSE
RNN ^[10]	0.950 51	83.771 31	107.032 33
LSTM ^[8]	0.973 21	49.841 75	78.745 77
GRU ^[9]	0.971 92	56.638 31	80.618 99
TCN-Transformer ^[19]	0.975 91	41.937 78	74.674 55
DCT-Channel Attention-TCN-Transformer	0.998 14	12.115 93	20.743 82

由表 2 可知, R^2 随模型复杂度提升而逐步增大,本文所提模型的 R^2 数值最高,即本文所提模型与实测数据拟合效果最优。误差指标对比结果显示,相较于 RNN,LSTM,GRU 和 TCN-Transformer 模型,本文所提模型的 MAE 分别降低 85.53%,75.69%,78.60%和 71.10%,RMSE 分别降低 80.61%,73.65%,74.26%和 72.20%,表明该模型在捕捉数据趋势和细节方面具有更高的预测精度和稳定性。

RNN 的 MAE 和 RMSE 指标在所有模型中最高,表明其在处理风电功率波动方面的局限性。LSTM 与 GRU 两种基于递归神经网络的模型在处理短期依赖关系方面表现出一定的优势,但在处理长时间依赖和复杂非线性特征时,预测精度有所降低。TCN-Transformer 模型结合了 TCN 和 Transformer 的长短期特征提取能力,其性能表现优于 LSTM 和 GRU 模型,但仍低于本文所提模型。

4.3.3 运行效率分析

为验证各模型预测计算效率,本文在相同实验环境下,统计 5 种模型的训练时长、总体推理耗时、内存占用及模型参数量等,结果如表 3 所示,训练与推理时间均采用平均值 $\pm$ 标准差表示。

表 3 5 种不同模型的运行效率对比结果
Table 3 Comparative results of time efficiency of five different models

预测方法	训练时间 s/epoch	推理时间 ms/sample	内存占 用/MB	参数量 M
RNN ^[10]	4.2 $\pm$ 0.3	52.6 $\pm$ 2.1	680	0.78
LSTM ^[8]	6.8 $\pm$ 0.5	63.4 $\pm$ 3.2	890	1.25
GRU ^[9]	5.9 $\pm$ 0.4	57.1 $\pm$ 2.8	820	1.12
TCN-Transformer ^[19]	11.5 $\pm$ 0.7	14.9 $\pm$ 1.3	1 450	2.83
DCT-Channel Attention- TCN-Transformer	13.2 $\pm$ 0.9	17.6 $\pm$ 1.5	1 680	3.15

由表 3 可知,本文所提模型推理耗时 17.6 ms,较 RNN,LSTM,GRU 大幅缩短,提速 3~4 倍,仅较 TCN-Transformer 高出 18.1%,表明模型轻量化设计效果良好。该模型的单轮迭代训练耗时为 13.2 s,较其他模型有所增加,主要是 DCT 预处理和多模块协同训练占用部分时间,整体仍在可接受范围。模型内存占用与参数量随模型复杂度上升而增加,本文借助通道注意力机制实现特征压缩,避免了显存占用激增。总体来说,本文所提模型在实现 R^2 为 0.998 预测精度的同时,兼顾了实时响应能力,完全满足短期风电功率预测的时效性要求。

图 10 给出了本文所提模型的训练损失和测试损失随训练轮数的变化。

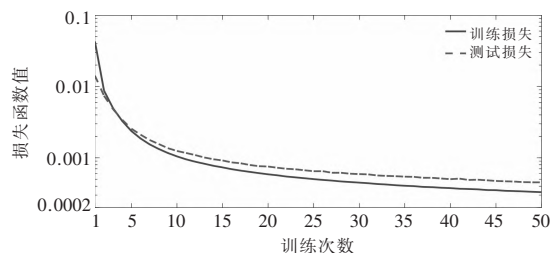


图 10 本文混合模型的训练与测试损失曲线
Fig.10 Training and testing loss curves of the proposed model

由图 10 可知,训练损失随训练轮次的增加逐步下降,表明模型在训练集上逐渐收敛。测试损失在初始阶段明显下降,表明模型在早期能够较好地适应测试数据。训练后期,训练损失趋于平稳,表明模型对训练样本拟合效果达到最优;

测试损失同样趋于稳定,表明模型具有良好的泛化性能。训练过程中,模型能够有效提取风电功率数据中的多尺度特征,并利用通道注意力机制增强对不同频率信息的捕捉能力。损失收敛趋势充分验证了 DCT-Channel Attention 机制与 TCN-Transformer 架构的有效结合。通过在通道维度引入 DCT 增强特征选择性,模型可精准识别风电功率数据中的频率信息,进一步提升对风电功率变化趋势的预测精度。同时,时间卷积网络擅长捕捉时间序列中的短期依赖关系,Transformer 结构则有效加强了对长时间依赖关系的建模能力。

5 结论

针对短期风电功率预测中的非线性特征捕捉难、功率波动剧烈及长时间依赖建模困难等问题,本文提出了一种基于 DCT-Channel Attention-TCN-Transformer 的混合模型。该方法通过对输入特征的增强、时间序列特征的捕捉及自注意力机制和多头注意力机制的应用,实现了风电功率的高精度预测。

基于实际风电场运行数据的仿真验证表明,本文所提模型可在保障高效运算效率的前提下,大幅提升短期风电功率预测精度。该模型不仅显著提高了 R^2 值,还大幅降低了 MAE 和 RMSE,在捕捉风电功率数据中的频率特征和通道间相互依赖性方面,综合性能优于其他对比模型。实验结果证实,混合离散余弦变换和通道注意力机制的融合能够有效提升预测精度,本文所提模型在风电功率预测中具有更高的准确性和鲁棒性。

参考文献:

- [1] 徐钊,谢开贵,王宇,等.基于 TCN-Wpsformer 混合模型的超短期风电功率预测[J].电力自动化设备,2024,44(8):54-61.
- [2] 唐新姿,顾能伟,黄轩晴,等.风电功率短期预测技术研究进展[J].机械工程学报,2022,58(12):213-236.
- [3] 孙荣富,张涛,和青,等.风电功率预测关键技术及应用综述[J].高电压技术,2021,47(4):1129-1143.
- [4] Yunus K, Thiringer T, Chen P. ARIMA -based frequency -decomposed modeling of wind speed time series[J].IEEE Transactions on Power Systems,2015,31(4):2546-2556.
- [5] 岳晓宇,彭显刚,林俐.鲸鱼优化支持向量机的短期风

- 电功率预测[J]. 电力系统及其自动化学报,2020,32(2):146-150.
- [6] Ye L,Li Y,Pei M,et al.A novel integrated method for short-term wind power forecasting based on fluctuation clustering and history matching [J].Applied Energy, 2022,327:120131.
- [7] 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.
- [8] 王伟,刘宏伟,陈永杰,等.基于 LSTM 循环神经网络的风力发电预测[J].可再生能源,2020,38(9):1187-1191.
- [9] 薛阳,王琳,王舒,等.一种结合 CNN 和 GRU 网络的超短期风电预测模型 [J]. 可再生能源,2019,37(3):456-462.
- [10] Niu Z,Yu Z,Tang W,et al.Wind power forecasting using attention-based gated recurrent unit network[J]. Energy, 2020,196:117081.
- [11] 贾睿,杨国华,郑豪丰,等.基于自适应权重的 CNN-LSTM&GRU 组合风电功率预测方法 [J]. 中国电力, 2022,55(5):47-56,110.
- [12] 韩自奋,景乾明,张彦凯,等.风电预测方法与新趋势综述[J].电力系统保护与控制,2019,47(24):178-187.
- [13] 田书欣,韩雪.基于正交小波变换的 LSTM-ARIMA 海上风速组合预测模型[J].智慧电力,2023,51(7):39-43,50.
- [14] 姜建国,杨效岩,毕洪波.基于 VMD-FE-CNN-BiLSTM 的短期光伏发电功率预测[J].太阳能学报,2024,45(7):462-473.
- [15] Jiang M,Zeng P,Wang K,et al.FECAM:Frequency enhanced channel attention mechanism for time series forecasting[J].Advanced Engineering Informatics,2023,58:102158.
- [16] Wang J,Luan Z,Yu Z,et al.Superpixel segmentation with squeeze-and-excitation networks [J].Signal Image and Video Processing,2022,16(5):1161-1168.
- [17] Tang P,Du P,Xia J,et al.Channel attention -based temporal convolutional network for satellite image time series classification [J].IEEE Geoscience and Remote Sensing Letters,2022,19:1-5.
- [18] Vaswani A,Shazeer N,Parmer N,et al.Attention is all you need [A].Proceedings of the 31st International Conference on Neural Information Processing Systems [C].Red Hook:Curran Associates,2017:6000-6010.
- [19] 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,2023,45(1):1-9.

Short-term wind power forecasting method based on DCT-channel attention-TCN-transformer hybrid model

Jiang Shanhe, Xu Xiaoyan, Chen Wensheng

(School of Electronic Engineering and Intelligent Manufacturing, Anqing Normal University, Anqing 246011, China)

Abstract: To address the challenges of complex and high volatility characteristics of wind power prediction data, a hybrid model prediction method based on Discrete Cosine Transform-Channel Attention-Temporal Convolutional Network-Transformer (DCT-Channel Attention-TCN-Transformer) is proposed in this study. Firstly, the frequency features and multi-scale information from the input signals by introducing a channel attention mechanism derived from the Discrete Cosine Transform (DCT) have been captured effectively, and the selectivity of the model to different channel features is enhanced. Then, the time-dependent feature of the data has been extracted through a Time Convolutional Network (TCN) to further reduce the uncertainty of the prediction. Finally, the global modeling for long time series data can be realized to further improve the prediction accuracy of wind power by applying the Transformer model's self-attention mechanism. The wind power data prediction results from a wind farm in Bangalore show that the coefficient of determination of the proposed approach is 0.998. Compared to RNN, LSTM, GRU, and TCN-Transformer, the average absolute error (MAE) index of the proposed method was reduced by 85.53%, 75.69%, 78.60%, and 71.10%, respectively. Meanwhile, the root mean square error (RMSE) index was also decreased by 80.61%, 73.65%, 74.26%, and 72.20%, respectively. The hybrid model proposed in this paper has excellent prediction performance and strong robustness.

Keywords: Discrete Cosine Transform (DCT); channel attention; time convolutional network; transformer; wind power forecasting