

# 基于VMD-Transformer-LSTM-XGBoost的短期风电机组出力混合预测模型

马虎林<sup>1</sup>, 李文清<sup>2</sup>, 刘志月<sup>2</sup>, 马子旭<sup>1</sup>, 朱新彧<sup>1</sup>, 王健<sup>3</sup>, 施雅蓉<sup>1</sup>, 赵学靖<sup>2\*</sup>

<sup>1</sup>中广核甘肃民勤第二风力发电有限公司, 甘肃 武威

<sup>2</sup>兰州大学数学与统计学院, 甘肃 兰州

<sup>3</sup>甘肃轩岳生态科技有限公司, 甘肃 兰州

收稿日期: 2025年12月21日; 录用日期: 2026年1月12日; 发布日期: 2026年1月23日

## 摘要

风电功率时间序列具有明显的非平稳性和多尺度波动特征, 使高精度短期预测面临较大挑战。针对传统模型难以同时刻画趋势、周期及高频扰动等不同时间尺度结构的问题, 本文提出一种融合自适应变分模态分解(VMD)、模糊熵复杂度分析、Transformer-LSTM深度特征提取与XGBoost回归的两阶段短期风电功率预测方法。首先, 以理论功率序列为分解对象, 通过贝叶斯优化在训练集上自适应确定VMD的模态数与惩罚参数, 并采用严格的零数据泄露策略。随后, 利用模糊熵度量各IMF的复杂度特征, 将其重构为低频趋势、中频周期与高频扰动三类协同模态(Co-IMFs), 以增强输入特征的物理可解释性与稳定性。在特征提取阶段, 构建融合Transformer全局依赖建模能力与LSTM局部时序记忆能力的DeepBlock网络, 并通过贝叶斯优化确定其最优结构与训练参数, 最终由XGBoost完成非线性回归预测。基于甘肃瓜州某风电场2023~2025年15分钟分辨率数据的实验结果表明, 所提出方法在MAE、RMSE与R<sup>2</sup>等指标上均优于多种基准模型及消融模型, 验证了该两阶段多尺度混合框架在复杂风电功率预测任务中的有效性。

## 关键词

风电功率预测, 变分模态分解, Transformer, LSTM, XGBoost, 多尺度分析

## A Short-Term Wind Power Forecasting Method Based on VMD-Transformer-LSTM-XGBoost

Hulin Ma<sup>1</sup>, Wenqing Li<sup>2</sup>, Zhiyue Liu<sup>2</sup>, Zixu Ma<sup>1</sup>, Xinyu Zhu<sup>1</sup>, Jian Wang<sup>3</sup>, Yarong Shi<sup>1</sup>, Xuejing Zhao<sup>2\*</sup>

<sup>1</sup>CGN Gansu Minqin No. 2 Wind Power Generation Co., Ltd., Wuwei Gansu

<sup>2</sup>School of Mathematics and Statistics, Lanzhou University, Lanzhou Gansu

\*通讯作者。

文章引用: 马虎林, 李文清, 刘志月, 马子旭, 朱新彧, 王健, 施雅蓉, 赵学靖. 基于 VMD-Transformer-LSTM-XGBoost 的短期风电机组出力混合预测模型[J]. 统计学与应用, 2026, 15(1): 265-282. DOI: 10.12677/sa.2026.151025

## Abstract

Wind power time series exhibit pronounced non-stationarity and multiscale fluctuations, posing significant challenges for high-accuracy short-term forecasting. To address the difficulty of conventional models in simultaneously capturing trend, periodic, and high-frequency components, this study proposes a two-stage short-term wind power forecasting framework integrating adaptive Variational Mode Decomposition (VMD), fuzzy-entropy-based complexity analysis, Transformer-LSTM deep feature extraction, and XGBoost regression. Theoretical power is first selected as the decomposition target, and the VMD mode number and penalty parameter are adaptively determined on the training set via Bayesian optimization under a strict zero-data-leakage strategy. The intrinsic mode functions (IMFs) are then reconstructed into three collaborative components—low-frequency trend, mid-frequency periodicity, and high-frequency disturbance—based on fuzzy entropy, enhancing the interpretability and stability of the input features. A DeepBlock network combining Transformer-based global dependency modeling and LSTM-based local temporal learning is employed for feature extraction, followed by XGBoost to perform nonlinear regression. Experiments conducted on 15-minute resolution data from a wind farm in Guazhou, Gansu Province (2023~2025) demonstrate that the proposed method consistently outperforms multiple benchmark and ablation models in terms of MAE, RMSE, and  $R^2$ , validating the effectiveness of the proposed multiscale two-stage forecasting framework.

## Keywords

Wind Power Forecasting, VMD, Transformer, LSTM, XGBoost, Multiscale Analysis

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

### 1.1. 研究背景

在全球能源结构向低碳化转型的背景下，风能因其清洁性和可再生性成为增长最快的新能源形式之一[1]。然而，受风速随机波动、气象条件变化以及机组运行状态等多重因素影响，风电功率序列呈现出显著的非平稳性和多尺度特征，既包含由湍流和局地扰动引起的高频随机波动，也包含由天气系统演化和昼夜变化驱动的周期结构，并叠加长期趋势成分[2]。这种复杂的动态特性显著增加了短期预测难度，并对电力系统的调度与运行带来挑战[3]。

为提升风电功率预测精度，研究者提出了物理模型、统计模型及数据驱动模型等多种方法[4]。物理模型依赖数值天气预报，难以在短时间尺度内提供高分辨率信息；传统统计模型如 ARIMA 对线性结构较为敏感，难以刻画风电功率中的强非线性特征；近年来，LSTM、Transformer 等深度学习模型在时间序列预测中表现出一定优势，但在直接建模原始风电序列时，仍难以充分应对多尺度结构与噪声干扰问题[5]。因此，从信号结构层面提升输入特征质量，并结合不同模型在特征提取与回归建模方面的优势，仍

是提升预测性能的重要研究方向[6] [7]。

## 1.2. 现有研究的局限性

尽管已有研究证明信号分解(如 EMD、VMD)和深度学习均能提高预测性能,但仍存在以下不足[8]:

- 1) 分解对象多以风速为主,风速到功率的非线性映射可能放大分解误差[9]。
- 2) VMD 等方法的分解参数常采用固定设定,缺乏对数据统计特性的自适应能力。
- 3) 深度学习模型多采用端到端预测结构,特征提取与回归任务耦合度较高,泛化稳定性有限。
- 4) 部分研究在分解与特征工程阶段未严格区分训练与测试数据,存在潜在的数据泄露风险[10]。

## 1.3. 本文的主要贡献

为解决上述问题,本文提出基于 VMD-Transformer-LSTM-XGBoost 的短期功率预测框架,主要贡献如下:

- 1) 以理论功率而非风速作为 VMD 的分解对象,减少多阶段非线性映射带来的误差累积。
- 2) 设计自适应 VMD 参数优化机制,使用贝叶斯优化在训练集上自动确定最优  $K$  与  $\alpha$ ,并实施严格零数据泄露策略。
- 3) 提出基于模糊熵(FE)的协同模态重构方法(Co-IMFs),将 IMF 按复杂度划分为趋势、周期和扰动三类,构建具有物理含义的多尺度输入结构。
- 4) 构建融合 Transformer 与 LSTM 的 DeepBlock 深度编码结构[11],实现多尺度时序依赖的统一建模。
- 5) 采用两阶段建模策略:深度特征提取 + XGBoost 回归器,显著提升预测稳定性与泛化性能。
- 6) 通过消融实验系统验证各模块的独立贡献,证明所提模型具有显著优势。

## 2. 模型与方法

### 2.1. VMD-Transformer-LSTM-XGBoost 模型

为提高风电功率预测的精度与鲁棒性,本文提出一种由三部分组成的预测框架,如图 1 所示。

该框架包括:VMD 分解与协同模态重构模块、Transformer-LSTM 深度特征提取与贝叶斯超参数优化模块,以及 XGBoost 回归预测模块。首先,对理论功率序列进行 VMD 分解,并基于模糊熵将 IMF 重构为低频趋势、中频周期和高频扰动三类协同模态(Co-IMFs),以获得结构清晰且稳定的多尺度输入特征。随后,将 Co-IMFs 与风速等变量输入 Transformer-LSTM 网络提取时序特征,其关键超参数通过 TPE-贝叶斯优化确定。最后,利用 XGBoost 回归器学习深度特征与功率之间的非线性映射,输出最终预测结果。

### 2.2. 自适应变分模态分解(VMD)

VMD 通过求解约束变分问题,将复杂非平稳信号分解为若干有限带宽的本征模态函数(IMFs)。本文采用标准 VMD 模型,并通过 ADMM 算法进行求解。

#### 2.2.1. VMD 对理论功率 $y(t)$ 进行分解[8]

VMD 通过求解一个约束变分问题,迭代地寻找一组模态及其对应的中心频率,其约束模型为:

$$\begin{aligned} \min_{\{u_k\}, \{\omega_k\}} & \left\{ \sum_k \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} & \sum_k u_k(t) = y(t) \end{aligned} \quad (1)$$

其中,  $\{u_k\}$  和  $\{\omega_k\}$  分别是有限带宽的模态分量(IMFs)及其中心频率;  $k$  是待提取的模态数量。引入二次惩

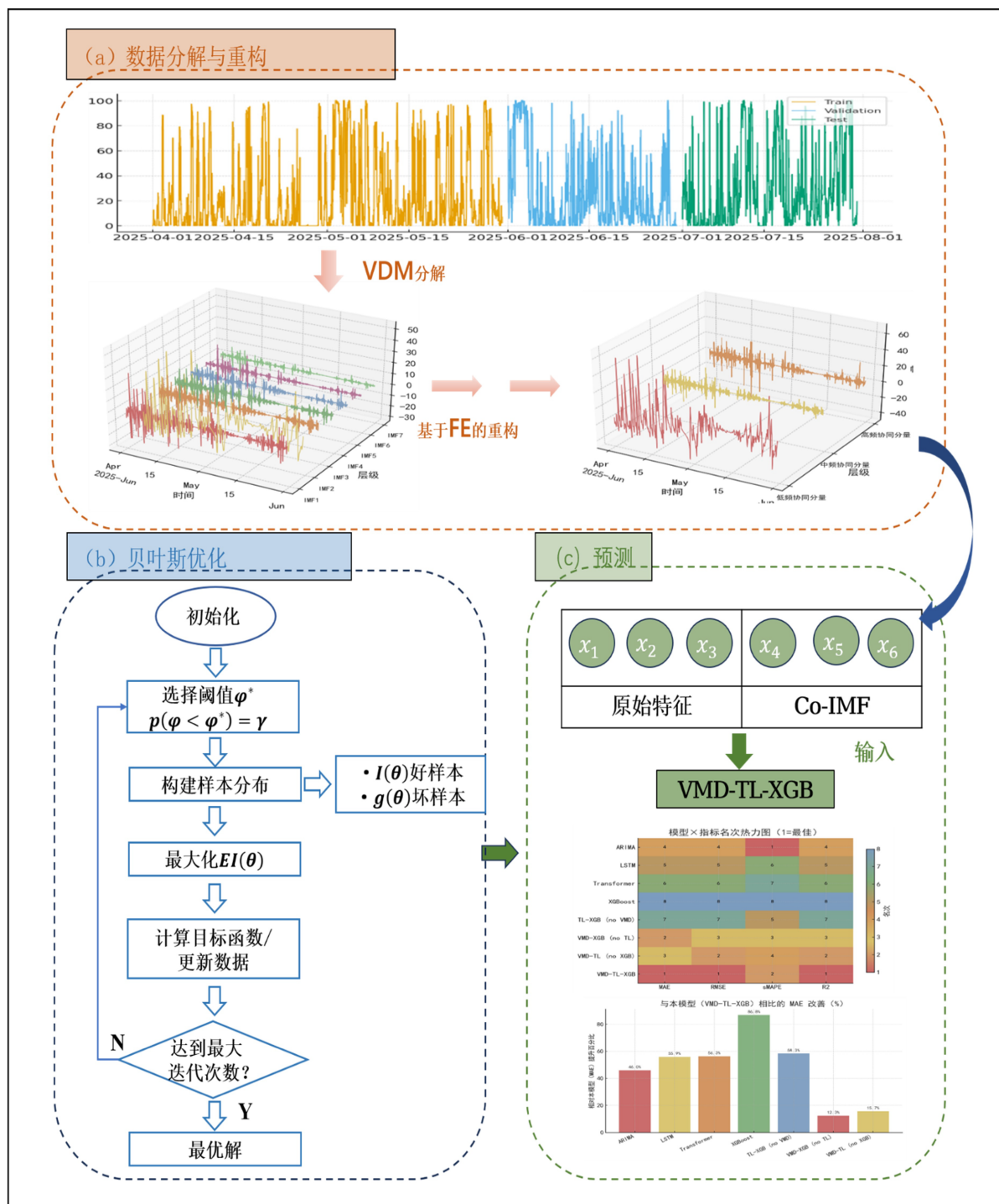


Figure 1. VMD-transformer-LSTM-XGBoost model

图 1. VMD-Transformer-LSTM-XGBoost 模型

罚因子  $\alpha$  和拉格朗日乘子  $\lambda$  将约束问题转化为增广拉格朗日形式:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_k \left\| \partial_t \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) e^{j\omega_k t} \right\|_2^2 + \left\| h(t) - \sum_k u_k(t) \right\|_2^2 + \lambda(t) \left( y(t) - \sum_k u_k(t) \right) \quad (2)$$



其中鞍点求解方程中的极小值公式(1)。使用交替方向乘法器(ADMM), 公式如下:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{y}(w) - \sum_{p \neq k} \hat{u}_p(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha(w - w_k)^2} \quad (3)$$

$$w_k^{n+1} = \frac{\int_0^\infty w |\hat{u}_k^{n+1}(w)|^2 dw}{\int_0^\infty |\hat{u}_k^{n+1}(w)|^2 dw} \quad (4)$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau \left( \hat{y}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega) \right) \quad (5)$$

其中  $n$  是迭代索引, HATS 表示傅里叶变换。迭代将继续, 直到

$$\sum_k \frac{\|\hat{u}_k^{n+1}(\omega) - \hat{u}_k^n(\omega)\|_2^2}{\|\hat{u}_k^n(\omega)\|_2^2} < \varepsilon \quad (6)$$

以  $\varepsilon > 0$  为收敛门槛。

### 2.2.2. 分解目标的选取: 出力(理论功率)

本文选择理论功率而非风速作为 VMD 的分解对象[12]。相比风速序列, 理论功率已通过功率曲线完成非线性映射, 其统计尺度与预测目标更为一致, 有助于减少多阶段建模中的误差累积。同时, 该策略实现了任务解耦: VMD 用于刻画理想工况下功率信号的多尺度结构, 而后续模型负责学习理想功率向实际出力的偏移关系, 从而隐式表征限功率及运行扰动等不确定因素, 提升模型的整体鲁棒性。

### 2.2.3. VMD 参数优化与零数据泄露策略

为避免分解参数对预测结果的主观影响, 本文采用基于 TPE 的贝叶斯优化方法: 仅在训练集上确定  $(K_{\text{train}}, \alpha_{\text{train}})$ , 并在验证集上独立优化得到  $(K_{\text{val}}, \alpha_{\text{val}})$  用于早停与超参数选择[13] [14]。

全局模态数取  $K_{\text{max}} = \max(K_{\text{train}}, K_{\text{val}})$ 。测试集分解时严格固定使用训练集参数  $(K_{\text{train}}, \alpha_{\text{train}})$ , 若模态数不足则以零填充对齐至  $K_{\text{max}}$ 。该流程确保测试集不参与任何参数优化, 从而保证评估公平可靠[15]。

## 2.3. Fuzzy Entropy (FE)复杂度度量

为量化每个 IMF 的复杂度(不规则性/粗糙度), 用作聚类与重构依据。其参数为嵌入维度  $m$ , 相似容差  $r$  (一般取原序列标准差的 0.1~0.25), 时间延迟  $\tau$  (默认 1)。

$$\text{FE}(x; m, r, N) = -\ln \left( \frac{\Phi^{m+1}(r)}{\Phi^m(r)} \right) \quad (7)$$

其中  $\Phi^m(r)$  为在嵌入维度  $m$  下、相似度阈值  $r$  内的匹配概率。对 IMF 序列  $u_k = \{u_k(t)\}_{t=1}^N$  逐一计算  $FE_k$  [16]。

## 2.4. 基于 FE 的三类协同分量重构

基于各 IMF 的复杂度特征, 本文将其重构为三类协同模态(Co-IMFs), 分别对应高频扰动、中频周期和低频趋势成分, 与频域分析结果保持一致。具体的, 对 VDM 分解得到的第  $K$  个 IMF 计算其模糊熵(参数取  $m = 2$ 、 $r = 0.2\text{std}(u_k)$ 、 $\tau = 1$ ), 并根据训练集上模糊熵分布的三分位阈值  $T_1$  与  $T_2$  进行分组: 当

$FE_k < T_1$  时归为低频趋势模态, 当  $T_1 \leq FE_k < T_2$  时归为中频周期模态, 当  $FE_k \geq T_2$  时归为高频扰动模态。最终, 将同组 IMF 相加得到三类协同模态分量  $\{c_1, c_2, c_3\}$ , 作为后续预测模型的输入特征, 如式(8)所示。

$$c_j(t) = \sum_{k \in G_j} u_k(t), \quad j = 1, 2, 3. \tag{8}$$

2.5. Transformer-LSTM-XGBoost 两阶段预测模型

本研究的核心是一个两阶段预测架构, 它将复杂的预测任务解耦为特征提取和回归预测两个子任务。  
1) 贝叶斯超参数优化  
基于树状 Parzen 估计器(TPE)的贝叶斯优化方法, 对 Transformer-LSTM 网络的关键超参数进行自动搜索, 以获得验证集上表现最优的参数组合[6], 超参数搜索空间结果如表 1 所示。

Table 1. Bayesian hyperparameter optimization search space  
表 1. 贝叶斯超参数优化搜索空间

参数类别	超参数类别	搜索范围
模型结构	Transformer 层数	[1, 2, 3, 4]
	注意力头数	[2, 4, 6, 8]
	模型维度	[16, 32, 64, 128]
	LSTM 隐藏单元数	[64, 128, 256, 512]
	LSTM 编码/解码层数	[1, 2]
	学习率	[1e-5, 1e-2]
	批量大小	[32, 64, 128, 256]
	Dropout 率	均匀分布[0.0, 0.5]

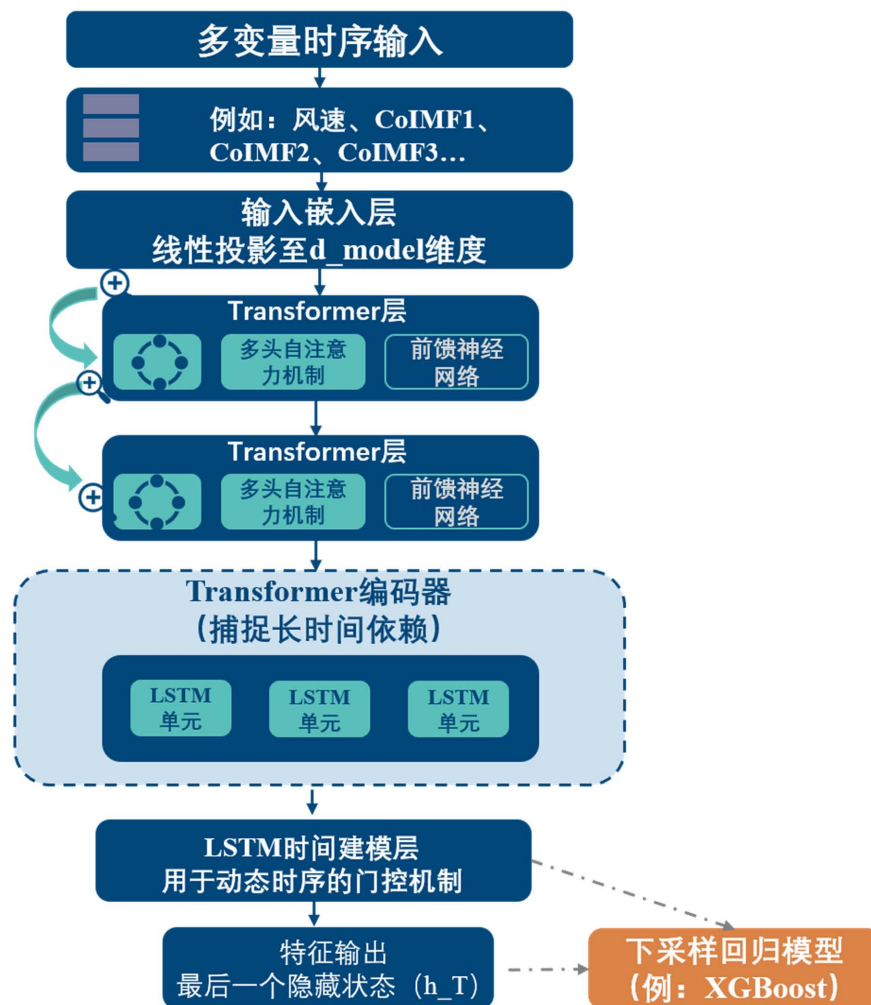
2) 第一阶段: Transformer-LSTM 深度特征提取器。  
该阶段将 VMD 分解后的多维时间序列输入编码为固定长度的深度特征向量。DeepBlock 由 Transformer 编码器与 LSTM 组成: Transformer 用于建模长程依赖, LSTM 用于进一步提炼时序动态特征[17][18]。  
其内部结构如图 2 所示:  
该模块的数学表达可简化为:

$$h_t = \text{LSTM}\left(\text{Transformer}\left(\text{InputProjection}\left(X_{t-T:t-1}\right)\right)\right)$$

其中  $X$  是输入的多维序列, 由 VMD 分解得到的 IMFs、重构信号和原始出力在时间窗口内构成。 $h_T$  是提取出的深度特征向量。

DeepBlock 模型结构及经参数调整方法得到最佳参数选择, 具体结果如下表 2 所示。  
3) 第二阶段: XGBoost 回归预测器  
将第一阶段得到的特征向量  $h_T$  作为输入, 由 XGBoost 学习其与目标出力之间的非线性映射, 并输出最终预测结果。

XGBoost 是梯度提升算法的一种高效实现, 通过迭代地训练一系列决策树, 并让每棵新树去拟合前一轮所有树预测结果的残差, 最终将所有树的预测结果加权求和得到最终输出。其优势在于: 高性能, 在众多数据科学竞赛和实际应用中被证明具有顶尖的预测精度; 鲁棒性, 内置正则化项, 能有效防止过拟合。



**Figure 2.** Internal structure of the DeepBlock network  
**图 2.** DeepBlock 深度网络内部结构

**Table 2.** Model structure and Hyperparameter setting table  
**表 2.** 模型结构及超参数设置表

模块	超参数	值
Transformer	层数	8
	注意力头数	8
	模型维度	32
	前馈维度	512
	Dropout 率	0.252
LSTM	编码器	1
	解码器层数	2
	隐藏状态维度	256
	批量大小	128
训练	学习率	1.46e-4

3. 数据描述

3.1. 数据来源与处理

本研究采用中国甘肃省瓜州大梁某风电场的运行数据，时间范围为 2023 年 8 月 1 日至 2025 年 7 月 31 日，时间分辨率为 15 分钟。

研究中使用的主要变量包括：

分解与特征输入：出力(kW)，代表在特定风速下风力发电机应输出的理论功率，作为 VMD 分解的对象和模型的核心输入特征来源；风速(m/s)，作为驱动变量，反映风资源条件与机组受力状态。

预测目标：出力(kW)，代表风力发电机的理论输出功率。

数据经清洗、异常值剔除与前向填充处理后，按时间顺序划分为训练集(2023 年 8 月 1 日至 2025 年 1 月 31 日)、验证集(2025 年 2 月 1 日至 2025 年 4 月 30 日)和测试集(2025 年 5 月 1 日至 2025 年 7 月 31 日)。训练集用于模型参数学习，验证集用于超参数优化，测试集仅用于最终性能评估，以避免数据泄露。

3.2. 描述性统计

1) 相关性分析

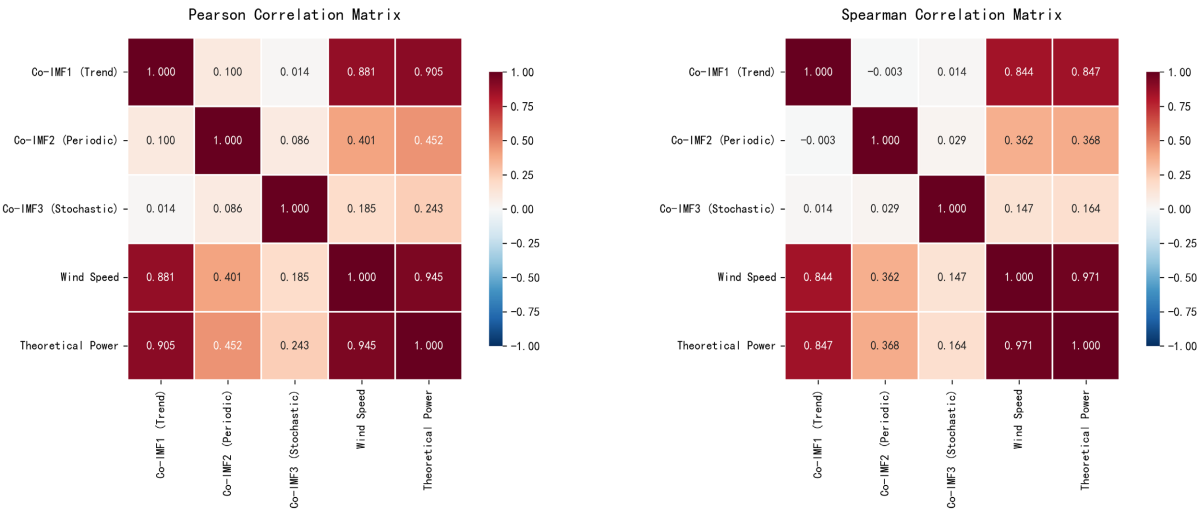


Figure 3. Correlation analysis heatmap  
图 3. 相关热力图

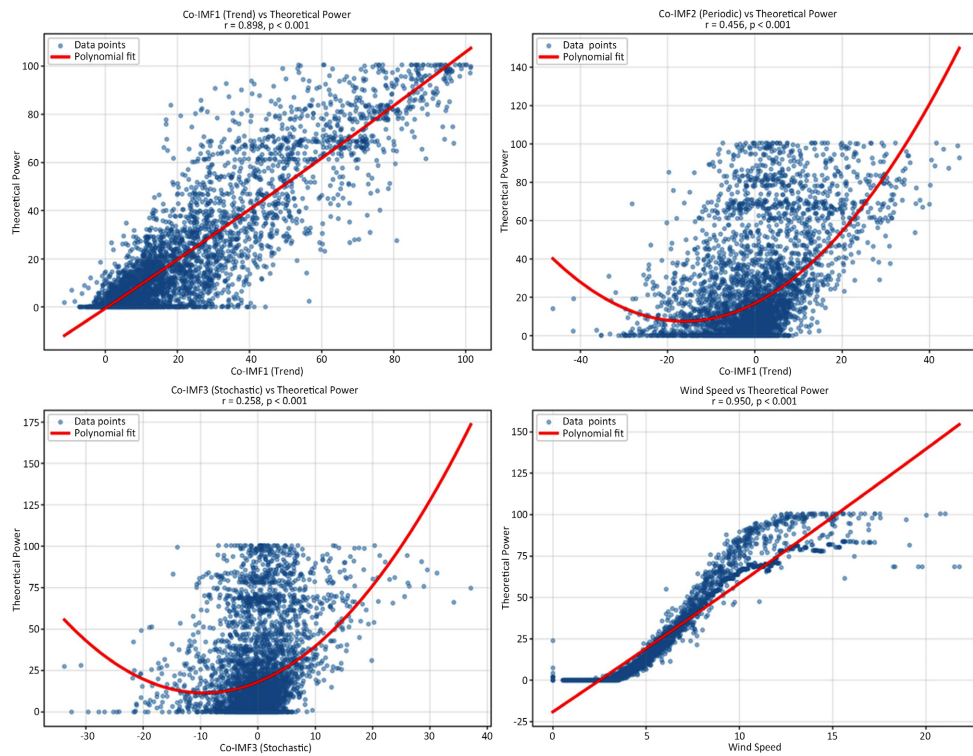
相关性热力图(图 3)显示，风速与理论功率呈显著正相关(Pearson  $r \approx 0.95$ , Spearman  $r \approx 0.97$ )；低频协同模态 Co-IMF1 与功率及风速同样保持较高相关性，表明其能够有效刻画长期变化趋势。相比之下，Co-IMF2 与 Co-IMF3 与功率的相关性较弱，反映周期与扰动成分在功率解释中的次要作用，且各变量间整体相关性较低，多重共线性风险有限。

2) 变量关联散点拟合

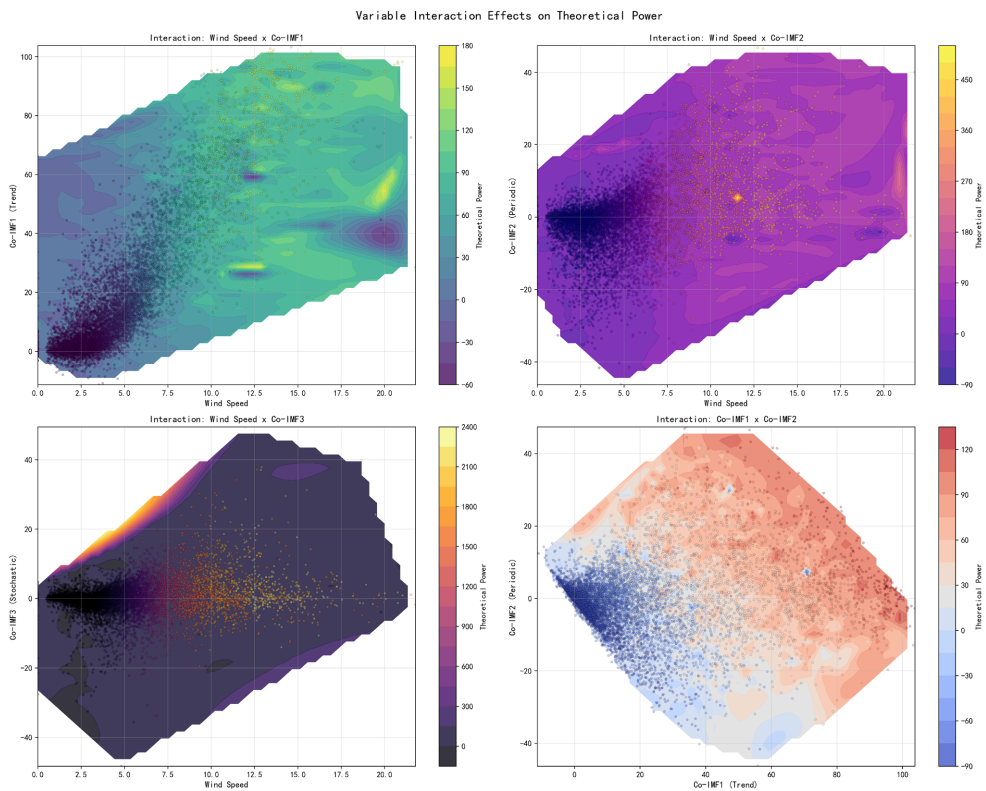
散点拟合图(图 4)表明风速 - 功率关系符合典型功率曲线，CoIMF1 与功率呈明显线性相关，CoIMF2 与 CoIMF3 的非线性特征更明显，说明它们主要反映趋势之外的结构性与随机变化。

3) 二维交互效应

二维交互图(图 5)显示：风速方向功率整体单调增大；当风速与 CoIMF1 同时处于高水平时，功率明显增强，表明趋势项对风速 - 功率关系具有调制作用；涉及 CoIMF3 的交互较弱，主要表现为局部扰动。



**Figure 4.** Scatter plot with polynomial fit for variable correlation  
**图 4.** 变量关联散点拟合图



**Figure 5.** 2D interaction effect  
**图 5.** 二维交互效应



#### 4) Co-IMF 分量 ACF 与 PACA 分析图

图 6 显示 Co-IMF1 自相关衰减最慢, 为典型低频趋势; Co-IMF2 在低滞后阶显著, 为周期项; Co-IMF3 自相关迅速衰减, 属于高频噪声型成分。三类模态的时序结构差异验证了 VMD + FE 分解的合理性。

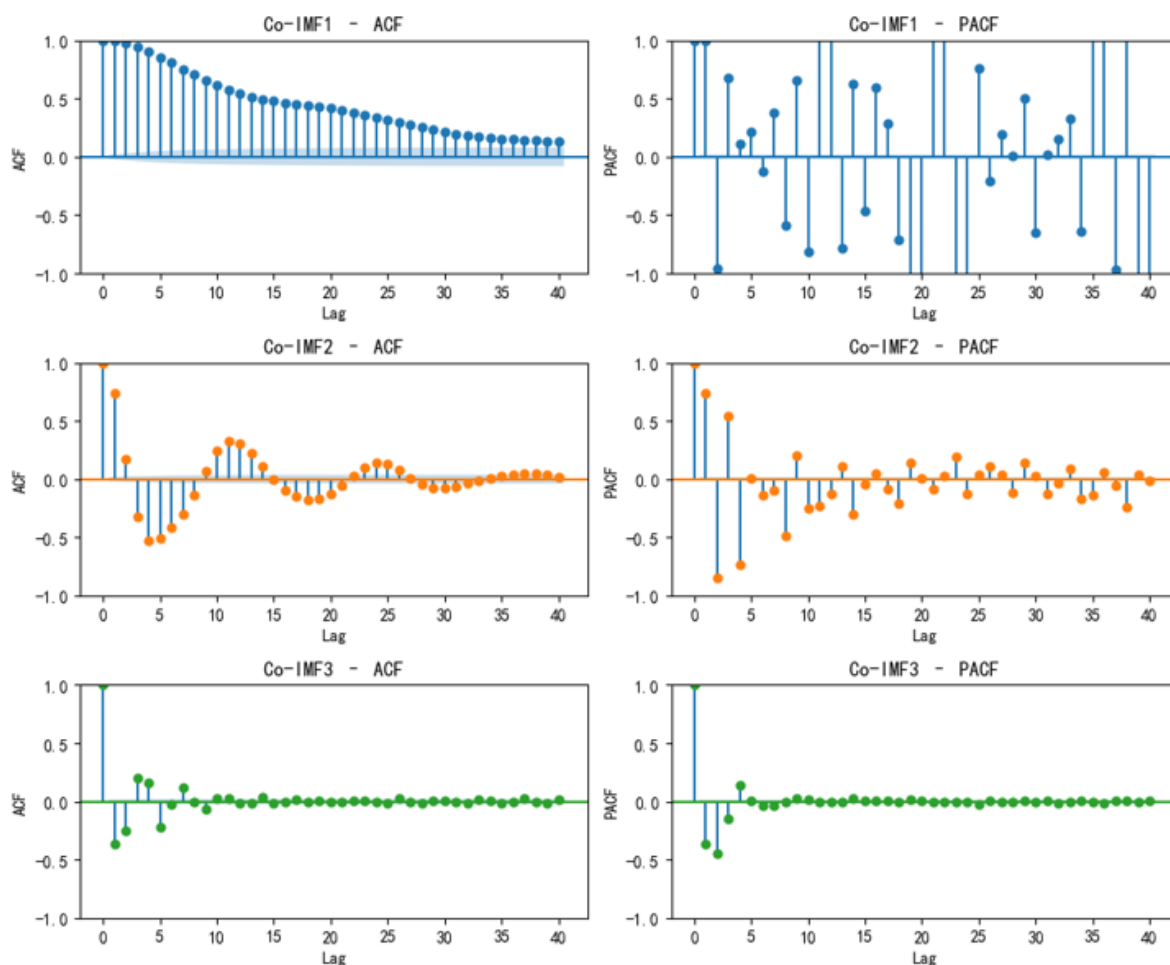


Figure 6. ACF and PACF plots for Co-IMF components  
图 6. Co-IMF 分量的自相关(ACF)与偏自相关(PACF)分析图

## 4. 实验分析

### 4.1. 实验设计与基准模型

为了评估本文提出的 VMD-Transformer-LSTM-XGBoost (V-TL-XGB)模型的性能, 选取 ARIMA、LSTM、Transformer 与 XGBoost 作为基准模型, 并设计多种消融模型以分析各模块的作用。所有模型均采用相同的数据划分与预测策略进行比较:

自回归积分滑动平均模型(ARIMA): 是一种经典的线性时间序列模型, 通过差分、自回归项和滑动平均项刻画序列的线性依赖关系, 常作为统计学基准模型用于性能对比。

长短期记忆网络(LSTM): 是一种能够建模时间序列长期依赖关系的循环神经网络, 在风电功率等时序预测任务中被广泛用作深度学习基准模型[19]。

**Transformer 模型(Transformer):** 是一种基于自注意力机制的深度学习架构,能够有效建模时间序列中的长程依赖关系,已广泛应用于时序预测任务中[20]。

**极限梯度提升(XGBoost):** 是一种基于梯度提升决策树的回归算法,具有良好的非线性拟合能力和泛化性能,在时间序列回归问题中表现稳定[21] [22]。

在与上述基准模型进行对比的基础上,为进一步分析所提出模型(V-TL-XGB)中各模块的作用,本文设计了一系列消融实验(Ablation Study),通过移除或替换关键组件以量化其对预测性能的影响,具体设置如下:

1) TL-XGB: 移除 VMD 分解,直接使用原始特征输入 Transformer-LSTM-XGBoost,以评估分解预处理的作用。

2) VMD-XGB: 移除 Transformer-LSTM,将 VMD 分解得到的分量直接输入 XGBoost,以分析深度特征提取模块的贡献。

3) VMD-TL: 移除 XGBoost,在 Transformer-LSTM 后接线性输出层,用于比较两阶段回归结构与单阶段预测结构的差异[23]。

本文提出的完整模型记为 V-TL-XGB (Adaptive VMD),该模型融合自适应 VMD 分解、Transformer-LSTM 深度特征提取以及 XGBoost 回归预测,通过与上述消融模型的性能对比,可系统评估各组成模块对整体预测性能的贡献。

## 4.2. 评价指标

为定量评估模型的预测性能,我们采用以下五个广泛使用的评价指标:

平均绝对误(MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

均方根误差(RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (10)$$

对称平均绝对百分比误差(sMAPE):

$$sMAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \quad (11)$$

决定系数(R<sup>2</sup>):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

## 4.3. 实验结果

### 4.3.1. 自适应 VMD 分解与重构

#### 1、VMD 分解

由于出力序列具有显著非平稳性,图 7 给出了风速与理论功率的时序波动。为降低序列复杂度,本文采用 VMD 将理论功率分解为若干子序列;通过贝叶斯优化确定参数  $K=8$ 。图 8 展示了训练集理论功率及其分解得到的 IMF1~IMF8。各模态呈现由高频到低频的多尺度结构: IMF1~IMF2 为高频扰动(气流扰动及测量噪声等),IMF3~IMF6 为中频周期波动(反映日内至数日尺度变化),IMF7~IMF8 为低频趋势项

(表征长期缓变背景) [24]。

2、重构

图 9 展示了测试集所有 IMF 分量的频域分布情况。不同 IMF 在归一化频率轴上呈现出明显分离的窄带特性，且顺序保持由高频到低频的自然递进关系。

根据各 IMF 的模糊熵分布(如图 10 所示)，不同模态在复杂度上呈现显著差异：IMF1~IMF3 的模糊熵最高，对应高频扰动成分；IMF4~IMF6 处于中等水平，主要反映周期性波动；IMF7~IMF8 的模糊熵最低，表现为平稳的低频趋势。

基于上述复杂度分析，将 IMF 进一步归纳为三类协同模态(Co-IMFs)：

- 1) 低频趋势项 Co-IMF1：由熵值最低的 IMF7 + IMF8 组成，主要表征整体长期趋势；
- 2) 中频周期项 Co-IMF2：由 IMF4~IMF6 构成，捕捉周期性波动与天气短期变化；
- 3) 高频扰动项 Co-IMF3：由 IMF1~IMF3 构成，对应湍流扰动与随机噪声。

三类 Co-IMFs 能够有效刻画原始功率序列的多尺度动态特征，其分解后重构结果如图 11 所示。

3、模态数 K 的适应性分析

该重构的核心是根据数据的固有复杂度分布进行分组，而非依赖于固定的  $K$  值。为验证其稳健性，我们固定 VMD 的惩罚参数  $\alpha$ ，在训练集上分别测试了模态数  $K = 6, 8, 10, 12$  的情况。表 3 展示了不同  $K$  值下，VMD-TL-XGB 模型在测试集上的性能指标。

通过上表 3 可以发现，当  $K$  从 6 变化到 12 时，模型的预测性能(1 日预测 RMSE: 5.03~5.19；完整测试集 RMSE: 40.31~43.85)保持稳定，证明 Co-IMF 重构规则能有效提取核心多尺度特征，对  $K$  值变化不敏感。虽然  $K = 12$  时取得最优精度，但  $K = 8$  在精度损失极小的情况下，训练效率显著更高，验证了通过贝叶斯优化自适应选择  $K = 8$  的合理性。该规则能普遍适用于不同分解粒度，自动生成物理意义清晰的输入特征。

4.3.2. 预测结果及可视化

为评估模型在实际应用场景下的性能，本文采用基于滑动窗口的单步滚动预测策略，对测试集进行逐时预测。模型在每一时刻仅利用此前可获得的历史特征进行预测，以模拟真实部署条件下的运行过程 [25]。

图 12、图 13 分别给出了 1 天和 3 天滚动预测结果。可以看出，所提出模型能够较好跟踪功率变化的整体趋势，并在波动区间及局部峰值附近保持合理响应，表明该两阶段混合框架具备稳定的趋势刻画能力和一定的短时波动捕捉能力。

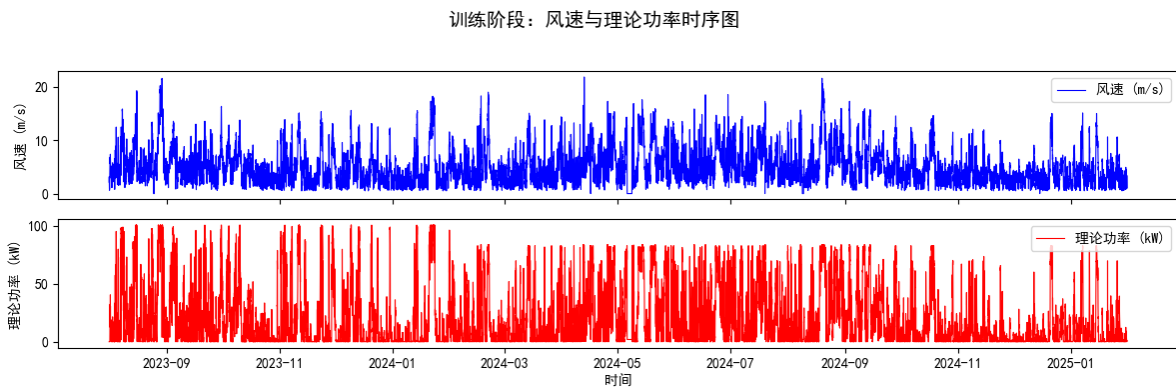
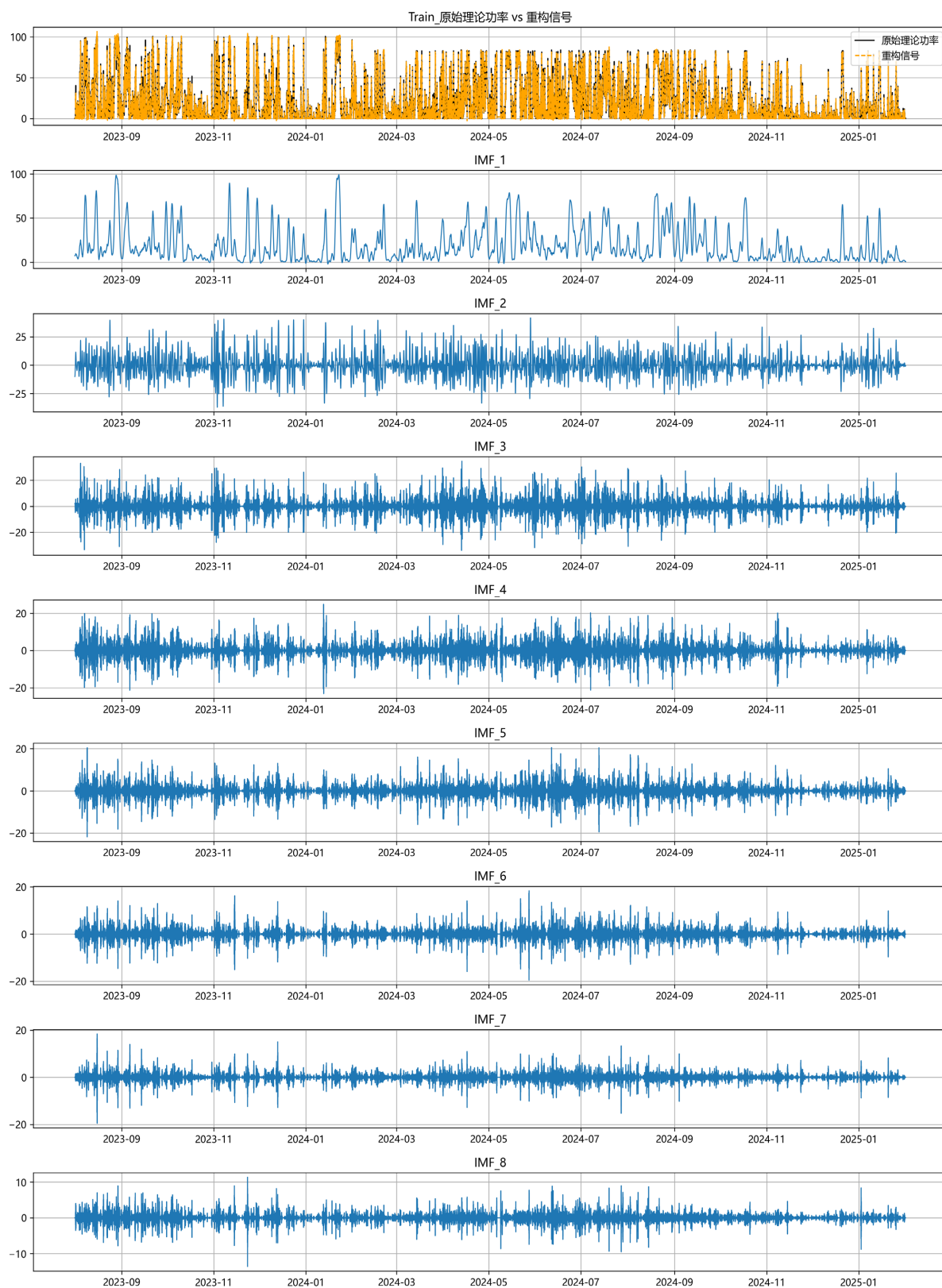


Figure 7. Time series of wind speed and theoretical power  
图 7. 风速与理论功率时序图



**Figure 8.** VMD decomposition of the training power series  
**图 8.** 训练集理论功率 VMD 分解

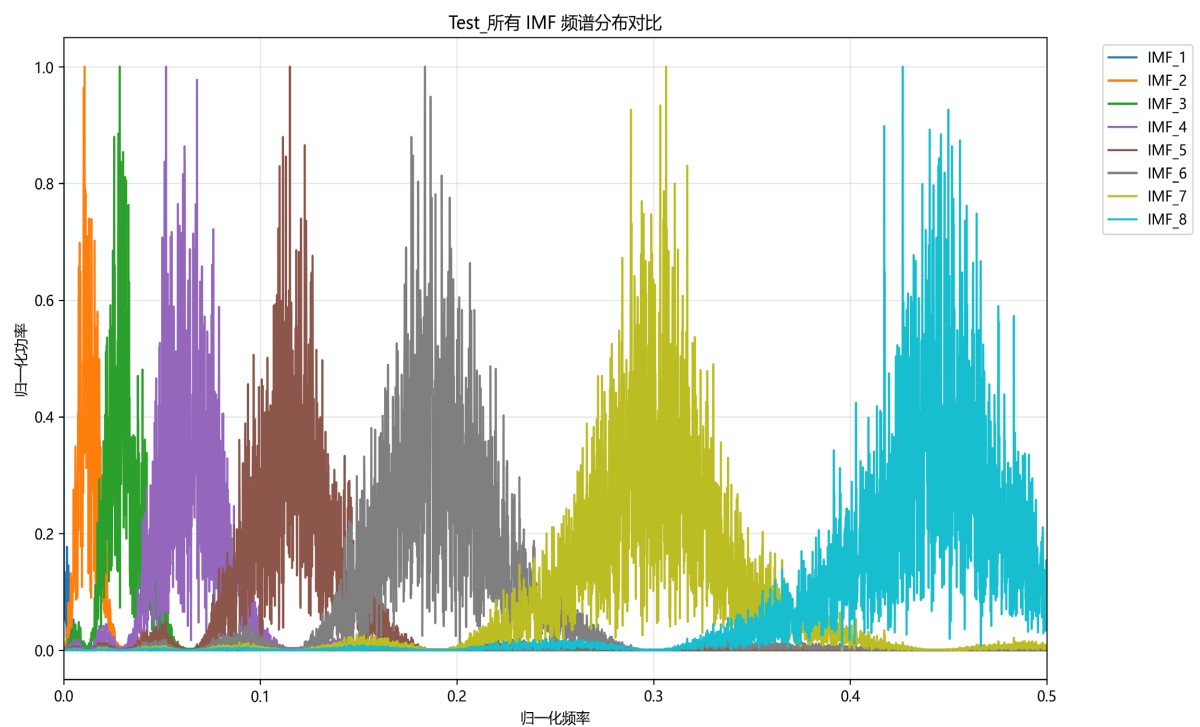


Figure 9. IMF spectral distributions (test set)

图 9. IMF 频谱分布(测试集)

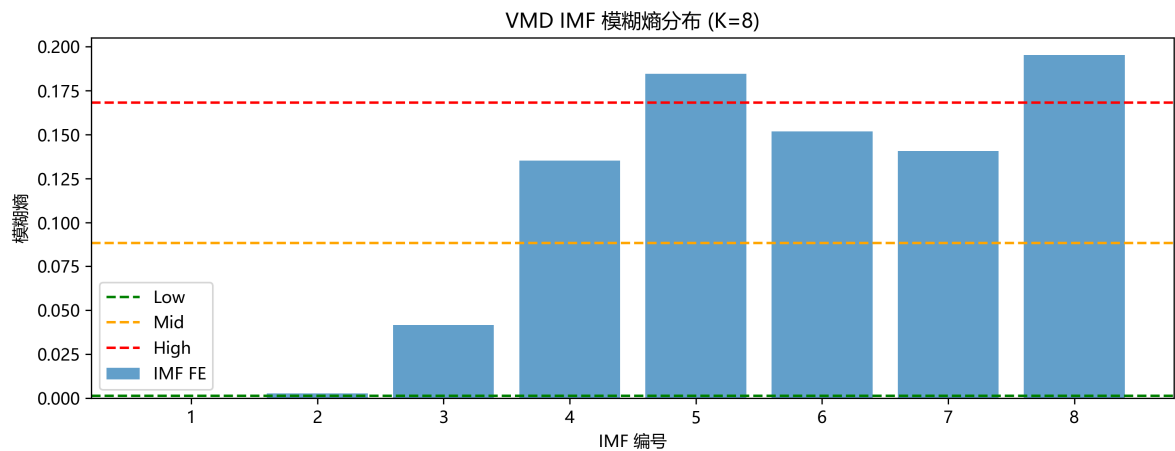


Figure 10. Fuzzy entropy distribution (test set)

图 10. 模糊熵分布(测试集)

Table 3. Comparison of model performance under different numbers of modalities  $K$

表 3. 不同模态数  $K$  下的模型性能对比

$K$	训练时间(s)	1 日预测 RMSE	3 日预测 RMSE	完整测试集 RMSE	完整测试集 MAE
6	167.88	5.19	14.13	43.85	26.05
<b>8</b>	<b>119.20</b>	<b>5.14</b>	<b>13.97</b>	<b>43.20</b>	<b>25.73</b>
10	251.53	5.19	14.56	43.00	25.69
12	247.93	5.03	13.40	40.31	24.02



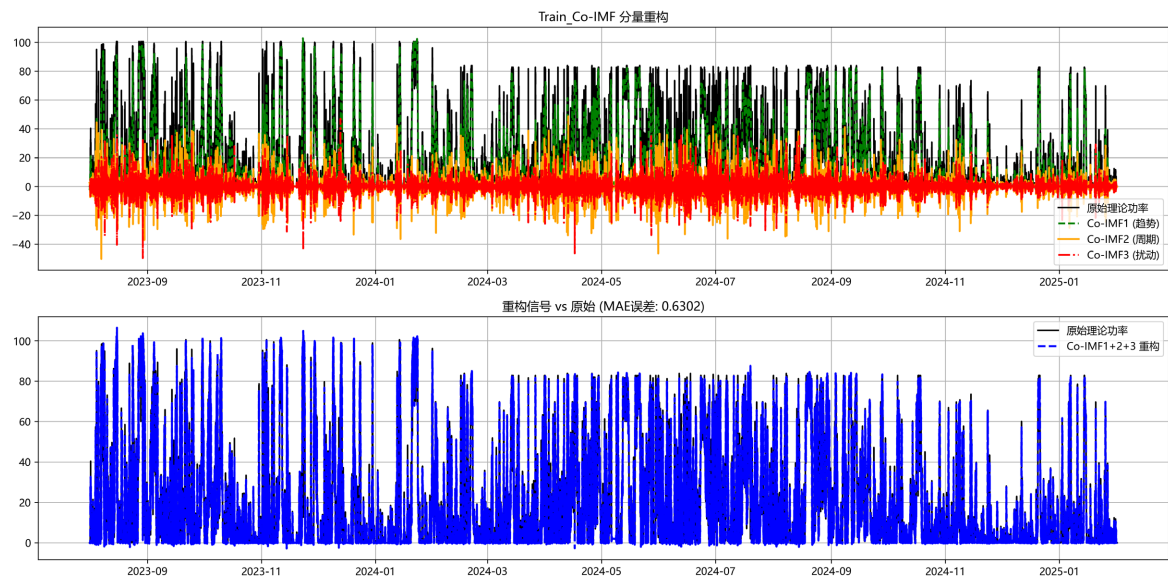


Figure 11. Co-IMF reconstructed components

图 11. 重构图 Co-IMF 图

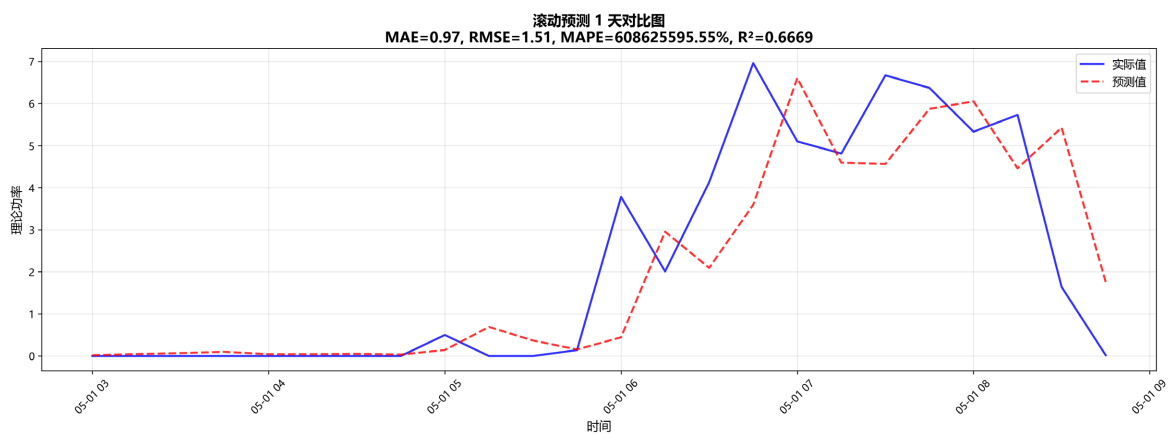


Figure 12. One-day rolling forecast

图 12. 1 天滚动预测结果

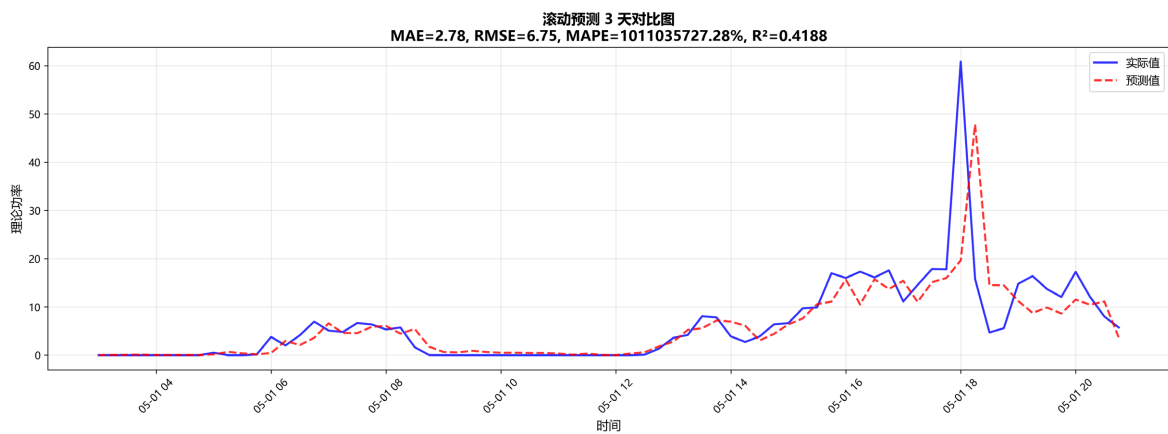


Figure 13. Three-day rolling forecast

图 13. 3 天滚动预测结果

4.3.3. 模型性能对比分析

从表 4 可见，所提出的混合模型整体性能显著优于各单一基准模型，表明信号分解、深度特征提取与回归建模相结合的策略具有明显优势。进一步的消融结果显示，引入 VMD 分解能够有效缓解功率序列的非平稳性，而基于 Transformer-LSTM 的深度特征提取相比直接使用 IMF 特征具有更强的表达能力。同时，将特征提取与回归预测解耦的两阶段建模结构在精度上优于单阶段预测模型，验证了该架构在风电功率预测任务中的合理性。

Table 4. Performance comparison of all models in the one-day forward prediction task (96 steps)

表 4. 所有模型在向前预测一天(96 步)任务中的性能对比

模型	MAE	RMSE	sMAPE	R <sup>2</sup>
ARIMA	24.2652	34.5624	105.0101	-0.1987
LSTM	6.1079	10.2749	58.4953	0.8932
Transformer	24.0422	33.0306	100.2331	-0.0934
XGBoost	6.2222	10.4213	59.6989	0.8910
TL-XGB	4.3953	6.3447	44.9099	0.8285
VMD-XGB	3.8153	5.7690	55.1911	0.9666
VMD-TL	3.4199	5.9665	49.7672	0.9643
VMD-TL-XGB	3.2356	4.8246	49.6973	0.9752

5. 模型计算复杂度分析

为评估 VMD-Transformer-LSTM-XGBoost (V-TL-XGB)模型在实际工程部署中的可行性，本节从计算复杂度的角度，定量分析模型在训练与推理阶段的时间与空间开销，并与主要基准模型进行对比。基于此分析，进一步讨论模型在不同应用场景下的适用性，并提出潜在的简化策略。

5.1. 计算开销量化

实验在配备 RTX 4090GPU (24 GB 内存)上进行，软件环境为 Python 3.10。为公平比较，所有模型的训练与推理均基于相同的输入时间窗口(历史 96 个时间步)和测试集(2025 年 5 月 1 日至 7 月 31 日)。表 5 统计了各模型完成一轮完整训练集训练的平均时间，以及对整个测试集进行单步滚动预测的平均单样本推理时间。

Table 5. Comprehensive comparison of computational time and performance of various models

表 5. 各模型计算时间与性能综合对比

模型	训练时间(秒)	单步推理时间(ms)	测试集 RMSE
ARIMA	13.646	0.67	34.5624
LSTM	102.327	5.00	10.2749
Transformer	211.629	8.50	33.0306
XGBoost	70.527	2.50	10.4213
VMD-TL-XGB	119.196	5.83	4.8246

5.2. 复杂度来源分析

VMD-TL-XGB 模型的计算时间开销主要来源于三个部分：

1、VMD 分解：其计算复杂度约为  $O(K \cdot N \log N)$ ，其中  $K$  为模态数， $N$  为序列长度。在本离线训练框架中，仅在训练集和验证集上进行一次，其时间成本可分摊，对在线推理影响较小。

2、Transformer-LSTM DeepBlock：这是训练阶段的主要耗时来源，其中自注意力机制和循环结构在长序列与高维特征下带来较高的计算开销，贝叶斯超参数搜索进一步增加了训练时间。

3、XGBoost 回归器：其复杂度取决于树的规模与数量，其训练效率高于深度网络，但集成结构使得推理速度略低于单模型。

### 5.3. 适用场景分析

#### 1、日前规划与短期预测

场景特征：该场景主要面向未来数小时至数日的发电计划制定、电力市场竞价及备用容量配置，对预测精度要求较高。

模型适配分析：本文提出的 VMD-TL-XGB 模型适用于此类场景。尽管其训练时间高于部分基准模型，但显著优于单一 Transformer 结构，且通过两阶段建模获得了明显的精度优势(RMSE = 4.82)。考虑到日前规划通常以日或周为更新周期，该模型在精度与计算成本之间具有较好的可行性。

#### 2、实时调度与超短期预测

场景特征：该场景关注未来数分钟至 1 小时内的功率波动响应，用于自动发电控制与频率调节，对推理效率和系统鲁棒性要求较高。

模型适配分析：由于 VMD-TL-XGB 的推理过程包含信号分解、深度网络前向传播及回归计算，其在极端实时性要求下可能存在一定限制。针对该场景，可考虑采用简化模型。在消融实验中，VMD-XGB 在保持较高预测精度(RMSE = 5.77)的同时显著降低了计算复杂度，在精度与效率之间取得了较好平衡。

#### 3、面向不同场景的模型简化与加速策略

为使模型更适应对效率要求苛刻的环境，未来工作可考虑以下策略：

1) 模型轻量化：使用知识蒸馏技术，将完整的 V-TL-XGB 模型的知识迁移至一个更小、更快的学生网络(如精简的 LSTM 或时序卷积网络)。

2) 架构搜索：在贝叶斯优化目标中引入推理时间或参数规模约束，自动搜索精度 - 效率权衡最优的 DeepBlock 架构。

3) 特征工程固化：设定训练阶段确定的 VMD 参数及 Co-IMF 重构规则，减少在线计算开销。

4) 模型量化与剪枝：采用量化与剪枝技术降低 Transformer-LSTM 与 XGBoost 的计算与存储成本。

## 6. 结语

针对风电功率序列非平稳性强、多尺度特征显著的问题，本文提出了一种由 VMD 分解、Transformer-LSTM 特征提取与 XGBoost 回归组成的两阶段预测框架。通过贝叶斯优化自适应确定 VMD 参数，并结合模糊熵实现协同模态重构，该方法有效提升了输入特征的多尺度表达能力。基于真实风电场数据的实验结果表明，所提模型在单步及滚动预测任务中均优于多种基准模型，能够较好刻画功率变化趋势并响应波动区间。总体而言，该框架在提升短期预测精度的同时兼顾了工程可用性，未来可结合数值天气预报与不确定性建模方法，进一步拓展其在调度与安全评估中的应用。

## 基金项目

本项目得到教育部“春晖计划”合作科研项目资助。

## 参考文献

- [1] Ackermann, T. and Söder, L. (2002) Wind Energy Technology and Current Status: A Review. *Renewable and Sustainable Energy Reviews*, **6**, 67-127.
- [2] International Energy Agency (IEA) (2023) World Energy Outlook 2023. IEA.
- [3] 陈启鑫, 吕睿可, 郭鸿业, 等. 面向需求响应的电力用户行为建模: 研究现状与应用[J]. 电力自动化设备, 2023, 43(10): 23-37.
- [4] Zhang, Z., Liu, H. and Zhu, Y. (2021) A Review on Wind Power Prediction Methods Based on Deep Learning and Big Data. *Frontiers in Energy Research*, **9**, Article ID: 682180.
- [5] 张耀聪. 基于深度学习神经网络的可再生能源“源荷”预测研究[D]: [硕士学位论文]. 兰州: 兰州理工大学, 2023.
- [6] Lim, B., Arik, S.O., Loeff, N. and Pfister, T. (2021) Temporal Fusion Transformers for Interpretable Multi-Horizon Time-Series Forecasting. *Nature Machine Intelligence*, **3**, 182-193.
- [7] 黄坤鹏. 风电功率超短期预测及可用发电功率计算[D]: [硕士学位论文]. 北京: 北京交通大学, 2022.
- [8] Liu, Z. and Zhao, X. (2026) LSTM-Transformer with Decomposition-Reconstruction for Enhanced Solar Irradiance Forecasting Incorporating Meteorological Variables. *Renewable Energy*, **258**, Article 124971. <https://doi.org/10.1016/j.renene.2025.124971>
- [9] Li, J., Wang, J. and Han, X. (2020) Short-Term Wind Power Prediction Based on Improved LSTM Model. *Energy*, **190**, Article 116400.
- [10] Zheng, K., Wang, D., Li, L., et al. (2021) A Hybrid Wind Power Forecasting Model Using CNN-LSTM and Attention Mechanism. *Applied Energy*, **301**, Article 117502.
- [11] Dragomiretskiy, K. and Zosso, D. (2014) Variational Mode Decomposition. *IEEE Transactions on Signal Processing*, **62**, 531-544. <https://doi.org/10.1109/tsp.2013.2288675>
- [12] 王宏伟, 赵鑫, 张子龙, 等. 基于 VMD 和 LSTM 的短期风电功率预测研究[J]. 太阳能学报, 2021, 42(12): 3052-3061.
- [13] Wang, J., Liu, H. and Tian, H. (2020) Wind Speed Forecasting Based on EMD and Deep Learning. *Energy Conversion and Management*, **208**, Article 112599.
- [14] Li, Z., Liu, B. and Liu, Y. (2021) Hybrid VMD-LSTM Model for Short-Term Wind Speed Forecasting. *Energy Reports*, **7**, 2563-2575.
- [15] Liu, H. and Tian, H. (2018) Hybrid Forecasting Model Based on EMD and XGBoost for Wind Speed Prediction. *Renewable Energy*, **126**, 766-781.
- [16] Zhang, Y., Wang, P. and Zhang, J. (2022) Wind Power Forecasting Based on Variational Mode Decomposition and Deep Learning Hybrid Model. *Applied Energy*, **323**, Article 119654.
- [17] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2021) Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 11106-11115. <https://doi.org/10.1609/aaai.v35i12.17325>
- [18] 宋立新, 刘志伟, 王琦, 等. 基于 Transformer 的多步风功率预测模型[J]. 电力系统自动化, 2022, 46(15): 127-136.
- [19] Wang, H., Zhang, Y., Chen, Z., et al. (2022) A Hybrid Model Combining LSTM and XGBoost for Short-Term Load Forecasting. *Energy*, **239**, Article 122293.
- [20] 李文斌, 贾旭东, 梁志鹏, 等. 基于 VMD-Transformer-XGBoost 的混合风电功率预测方法[J]. 电网技术, 2023, 47(8): 2862-2874.
- [21] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [22] Liu, H., Wang, Y. and Zhang, Z. (2023) Hybrid VMD-CNN-LSTM-XGBoost Model for Wind Power Forecasting. *Energy Conversion and Management*, **282**, Article 116909.
- [23] Bergstra, J., Yamins, D. and Cox, D.D. (2013) Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms. *Proceedings of the 12th Python in Science Conference (SciPy 2013)*, Austin, Texas, 24-29 June 2013, 13-20. <https://doi.org/10.25080/majora-8b375195-003>
- [24] Snoek, J., Larochelle, H. and Adams, R.P. (2012) Practical Bayesian Optimization of Machine Learning Algorithms. *Advances in Neural Information Processing Systems (NIPS)*, **25**, 2951-2959.
- [25] 邢伟, 周涛, 张晨. 基于贝叶斯优化的 XGBoost 风电预测模型[J]. 可再生能源, 2020, 38(9): 1325-1332.