

# 基于优化 VMD 与 TCN-ISE-Pyraformer 的短期电力负荷预测

张 翔, 李红月

(安徽理工大学 电气与信息工程学院, 安徽 淮南 232001)

**摘要:** 针对传统预测模型在捕捉多特征负荷数据时难以兼顾全局和局部特征的问题, 提出一种基于麻雀搜索算法 (SSA) 优化变分模态分解 (VMD) 与改进挤压-激励 (ISE) 模块、时序卷积网络 (TCN) 和 Pyraformer 的组合预测模型。首先, 采用 SSA 优化 VMD 参数, 将波动性较强的动态负荷序列分解为多个平稳的模态分量, 以降低原始数据的非平稳性; 随后, 将分解得到的本征模态函数输入 TCN 以提取局部特征, 并通过 ISE 模块自适应分配权重, 有效抑制噪声干扰; 最后, 将加权后的特征输入 Pyraformer 以捕获全局特征, 并输出最终的预测结果。为验证模型性能, 采用 2 个地区的真实电力负荷数据集进行仿真实验。结果表明: 在 2 个算例中, 该模型的决定系数分别达 0.994 9 与 0.984 2, 均优于对比模型。这一结果验证了所提模型在同时捕捉多特征负荷数据全局和局部特征方面的优势, 展现出更高的预测精度和稳定性。

**关键词:** 短期电力负荷预测; 时序卷积网络; Pyraformer; ISE 模块; 变分模态分解; 麻雀搜索算法; 金字塔注意力模型

**中图分类号:** TM 715 **文献标志码:** A **doi:** 10.12415/j.issn.1671-7872.25006



## Short-term Power Load Forecasting Based on Optimized VMD and TCN-ISE-Pyraformer

ZHANG Xuan, LI Hongyue

(School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China)

**Abstract:** Aiming at the problem that traditional prediction models struggle to simultaneously capture global and local features of multifeature load data, a hybrid prediction model based on sparrow search algorithm (SSA)-optimized variational mode decomposition (VMD) and improved squeeze and excitation (ISE) module, and temporal convolutional network (TCN) and Pyraformer was proposed. First, the SSA was employed to optimize VMD parameters, decomposing the highly oscillatory dynamic load sequence into multiple stationary modal components, thereby reducing the non-stationarity of the original data. Then, the obtained intrinsic mode functions was input into the TCN model to capture the local features of the data, while the ISE module adaptively assigned appropriate weights to the extracted features, thereby reducing the impact of redundant information on the prediction results. Finally, the weighted data was fed into the Pyraformer model to capture the global features and generated the final prediction results. To validate the model's performance, real-world power load datasets from two regions were used for simulation experiments. The results show that in both cases, the proposed model achieves the coefficients of

收稿日期: 2025-01-14

基金项目: 安徽省高校自然科学基金重点项目 (KJ2021A0471)

作者简介: 张翔 (2001—), 男, 安徽宿州人, 硕士生, 主要研究方向为电力系统负荷预测。

通信作者: 李红月 (1978—), 女, 安徽宿州人, 副教授, 主要研究方向为电力系统智能监测与控制。

引文格式: 张翔, 李红月. 基于优化 VMD 与 TCN-ISE-Pyraformer 的短期电力负荷预测 [J]. 安徽工业大学学报(自然科学版), 2025, 42(3):263-275.

determination is 0.9949 and 0.9842, respectively, outperforming other comparative models. This verifies the proposed model's superiority in simultaneously capture global and local features of multifeature load data, demonstrating higher prediction accuracy and stability.

**Keywords:** short-term power load forecasting; temporal convolutional network; Pyraformer; ISE module; variational mode decomposition; sparrow search algorithm; pyramidal attention model

电能作为现代社会发展的基石,深刻影响着生产与生活的方方面面,是经济增长、民生改善和社会稳定的重要保障。精准的负荷预测不仅能够帮助电力公司优化资源配置、降低运营成本,还能在高负荷时段防止电网过载,提高电力系统的可靠性。现有的负荷预测方法主要分为三大类:统计学方法、基于传统机器学习的方法和深度学习组合模型的方法。其中,经典统计学方法包括多元线性回归<sup>[1]</sup>、自回归移动平均(auto-regression and moving average model, ARMA)<sup>[2]</sup>和差分自回归移动平均(autoregressive integrated moving average model, ARIMA)<sup>[3]</sup>等。这些方法虽然拟合速度快,但对电力负荷这种非线性序列数据拟合效果较差。传统机器学习负荷预测方法主要有支持向量机(support vector regression, SVM)<sup>[4]</sup>、随机森林(random forest, RF)<sup>[5]</sup>等。这些方法虽然能够学习数据中的非线性关系,但难以有效捕捉电力负荷数据的时间依赖性和长期趋势。深度学习方法克服了传统机器学习方法的不足,具备强大的海量处理数据的能力,能够捕捉数据间的时间依赖性,同时获得更好的鲁棒性。用于电力负荷预测的深度学习方法包含反向传播(back propagation, BP)神经网络<sup>[6]</sup>、卷积神经网络(convolutional neural networks, CNN)<sup>[7]</sup>、长短时记忆(long short-term memory, LSTM)网络<sup>[8]</sup>和注意力机制<sup>[9]</sup>等。然而,深度学习方法仍存在一定的局限性,模型通常依赖于大量的数据输入,且电力负荷数据受到多种复杂因素的影响,单一模型难以有效捕捉复杂的非线性关系。因此,开展改进深度学习方法以精确预测电力负荷的研究,对于电力系统的安全稳定和经济运行具有重要意义,同时也有助于推动能源转型,促进深度学习技术的创新与应用。这一研究也是能源与人工智能交叉领域的热点方向之一。

针对单一模型难以捕捉负荷与影响因素间复杂非线性关系的问题,部分学者提出采用组合模型的方法。王耀辉等<sup>[10]</sup>提出了结合CNN与双向长短时记忆(BiLSTM)网络的预测模型,并采用改进的沙猫群优化算法对其进行优化,实验结果表明:该模型在供热负荷预测中较单一BiLSTM模型具有更高的准确性,但在处理高复杂度数据时,预测精度有所下

降。An等<sup>[11]</sup>提出了利用麻雀搜索算法(sparrow search algorithm, SSA)对CNN和SVM进行优化的混合预测模型,验证了SSA在模型参数优化中的效果,但模型的泛化性能较弱,且对数据全局特征的捕捉不足。石卓见等<sup>[12]</sup>利用改进的自适应噪声完全集合经验模态分解(improved complementary ensemble empirical mode decomposition with adaptive noise, ICEEMDAN)与变分模态分解(variational mode decomposition, VMD)进行二次模态分解,并结合Informer模型进行建模,充分考虑电价、气象等因素对负荷的影响,证明了信号分解对预测准确性提升的作用,但ICEEMDAN和VMD的组合增加了计算复杂度,且模态混叠问题仍影响预测精度。姚程文等<sup>[13]</sup>采用CNN与门控循环单元(gate recurrent unit, GRU)组合对浙江省某地区电力负荷数据进行预测,结果表明组合模型的预测精度优于单一GRU模型,但建模长序列依赖存在不足。王清等<sup>[14]</sup>利用VMD对负荷数据进行分解,并采用时序卷积网络(temporal convolutional network, TCN)模型进行预测,验证了VMD分解的优越性。但VMD的分解效果受参数(如个数 $K$ 和惩罚因子 $\alpha$ )的影响较大,仅依靠经验确定这些参数可能导致分解效果不理想<sup>[15]</sup>。

尽管组合模型相比于单一模型有所改进,但仍难以同时捕捉负荷数据的全局和局部特征,并难以同步考虑多种影响因素与电力负荷之间的非线性关系。而改进的挤压和激励(improved squeeze and excitation, ISE)模块通过增强特征选择、抑制信息冗余、提升泛化能力以及多尺度特征融合,显著提升模型性能,在处理复杂数据时表现优异。鉴于此,本文提出基于SSA优化VMD与TCN-ISE-Pyraformer组合的电力负荷预测模型,克服传统组合模型难以同时捕捉负荷数据全局和局部特征的问题,以期实现对电力负荷数据特征的有效捕捉,提升预测精度和稳定性。

## 1 组合预测模型的构建

本文提出的优化VMD与TCN-ISE-Pyraformer组合预测模型的系统框图如图1。为避免人为因素

对分解质量的影响, 采用经麻雀搜索算法 (SSA) 优化的变分模态分解 (VMD) 对负荷数据进行自适应分解, 以提取不同频率的本征模态分量 (intrinsic mode function, IMF), 确保分解结果能够准确反映电力负荷数据的复杂性。将分解后的 IMF 和特征数据输入 TCN-ISE-Pyraformer 进行处理: 时序卷积网络 (TCN) 通过因果膨胀卷积构建多时间尺度的感受野, 提取负荷数据的局部特征; 通过改进的挤压和激励 (ISE) 模块进行全局平均池化降维, 并利用全连接

层和 Sigmoid 激活函数计算特征权重, 通过乘法运算调整特征重要性, 以抑制冗余信息干扰并提高模型泛化性能, 将处理后的数据输入 Pyraformer, 利用金字塔自注意力机制分层提取长期依赖特征, 在有效降低计算复杂度的同时精准捕捉电力负荷的周期性变化和长期趋势, 解决传统模型难以兼顾全局与局部特征以及负荷与影响因素间复杂非线性关系的问题; 最终, Pyraformer 输出的特征经过全连接层进行维度映射与信号重构后生成电力负荷预测值。

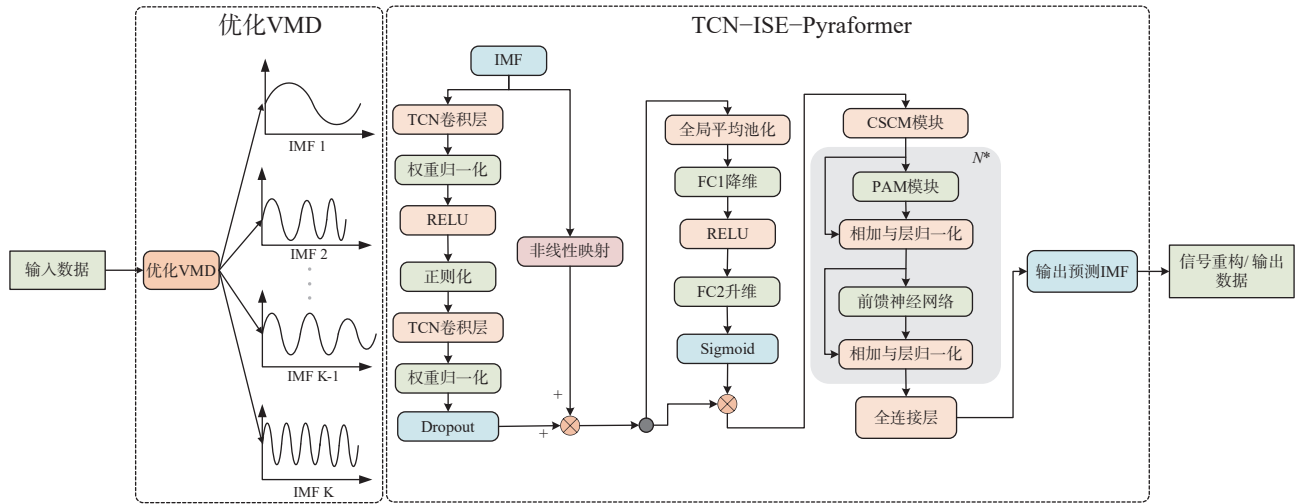


图1 优化 VMD 与 TCN-ISE-Pyraformer 模型系统框图

Fig. 1 Block diagram of the optimized VMD and TCN-ISE-Pyraformer network system

### 1.1 麻雀搜索算法优化变分模态分解的构建

变分模态分解 (VMD) 是一种自适应信号处理方法, 相比传统经验模态分解 (empirical mode decomposition, EMD), VMD 通过优化框架实现模态分量的同步分解, 能够有效克服 EMD 的模态混叠和频域混淆问题<sup>[16]</sup>。VMD 通过求解进化变分问题的最优解获得理论方案, 并在迭代过程中保持优化策略的稳定性。其中任何信号的分解都可被表述为:

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

式中:  $u_k(t)$  为第  $k$  个模态分量;  $f(t)$  为原始负荷序列;  $\delta(t)$  为单位脉冲函数;  $\partial_t$  为偏导运算符;  $\omega_k$  为第  $k$  个中心频率;  $j$  为虚数单位。

通过引入惩罚因子  $\alpha$  与拉格朗日乘法算子  $\lambda$ , 将约束问题转变为非约束型变分问题, 从而得到式 (2):

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) \cdot u_k(t) \right] e^{-j\omega_k t} \right\|^2 + \\ & \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \end{aligned} \quad (2)$$

通过引入交替方向乘法法可有效最小化式 (2), 相应更新中心频率和频域中的估计模式, 从而得到式 (3), (4):

$$\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega) / 2$$

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \hat{\lambda}(\omega) / 2}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k(\omega)| d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega} \quad (4)$$

式中:  $\hat{u}_k^{n+1}(\omega)$  为当前余量的维纳滤波;  $\omega_k^{n+1}$  为当前模态的中心频率;  $\hat{f}(\omega)$ ,  $\hat{u}_k(\omega)$ ,  $\hat{\lambda}(\omega)$  分别为  $f(t)$ ,  $u_k(t)$ ,  $\lambda(t)$  的傅里叶变换;  $n$  为迭代次数。

由上述 VMD 的过程及原理可看出: 模态分解数  $K$  和惩罚因子  $\alpha$  对 VMD 分解的精度具有显著影响。由于实际待分解信号的复杂性和非平稳性, 人为设置模态分解数  $K$  和惩罚因子  $\alpha$  较为困难, 且容易导致分解结果的随机性。麻雀搜索算法 (SSA) 是基于麻雀的社会性行为将种群划分为不同角色的个体以实现全局优化。具体而言, 麻雀群体根据觅食行为主要分为发现者、加入者和警戒者<sup>[17]</sup>。适应度值最高的个体作为发现者, 负责全局搜索以寻找最优解; 适应度值次高的个体作为加入者, 紧随发现者

进行局部搜索; 剩余个体充当警戒者, 负责警惕环境变化, 避免种群陷入局部最优。因此, 本文利用麻雀搜索算法 (SSA) 优化 VMD 的 2 个关键参数, 确保 VMD 能够更好地适配电力负荷数据的分解特性, 为后续预测模型提供高质量的模态分量。具体采用包络熵 ( $E$ ) 优化, 其主要过程如下:

- 1) 设置 SSA 参数及初始化种群。
- 2) 对于当前的  $K$  与  $\alpha$ , 使用 VMD 算法对电力负荷数据进行分解, 得到  $K$  个本征模态分量。
- 3) 根据式 (5) 计算每个本征模态分量的包络熵:

$$\begin{cases} E = -\sum_{m=1}^M p_m \lg p_m \\ p_m = a(m) / \sum_{m=1}^M a(m) \end{cases} \quad (5)$$

式中:  $p_m$  为事件发生的概率;  $a(m)$  为由希尔伯特变

换 (Hilbert transform) 得到的包络信号;  $M$  为样本数。

4) 更新麻雀个体位置, 比较各位置对应的包络熵 ( $E$ ), 不断更新最小适应度值。

5) 循环迭代步骤 2)~4), 直至确定全局最小适应度值或达到最大迭代次数。

6) 利用最优的  $K$  与  $\alpha$  对数据进行 VMD 分解。

### 1.2 时间卷积网络的构建

时间卷积网络 (TCN) 采用因果膨胀卷积作为核心结构, 通过一维卷积替代传统神经网络, 具有参数精简、计算高效、并行性强与精度优越等特点<sup>[18]</sup>。TCN 结构如图 2。图中:  $d$  为输入层的采样率 (如  $d=2$  时每两个数据点中仅选取一个作为输入);  $y$  表示每次卷积操作中连续选取的卷积核数量, 用于加权求和。此外, TCN 通过引入残差连接有效缓解网络性能随深度增加而退化的问题。

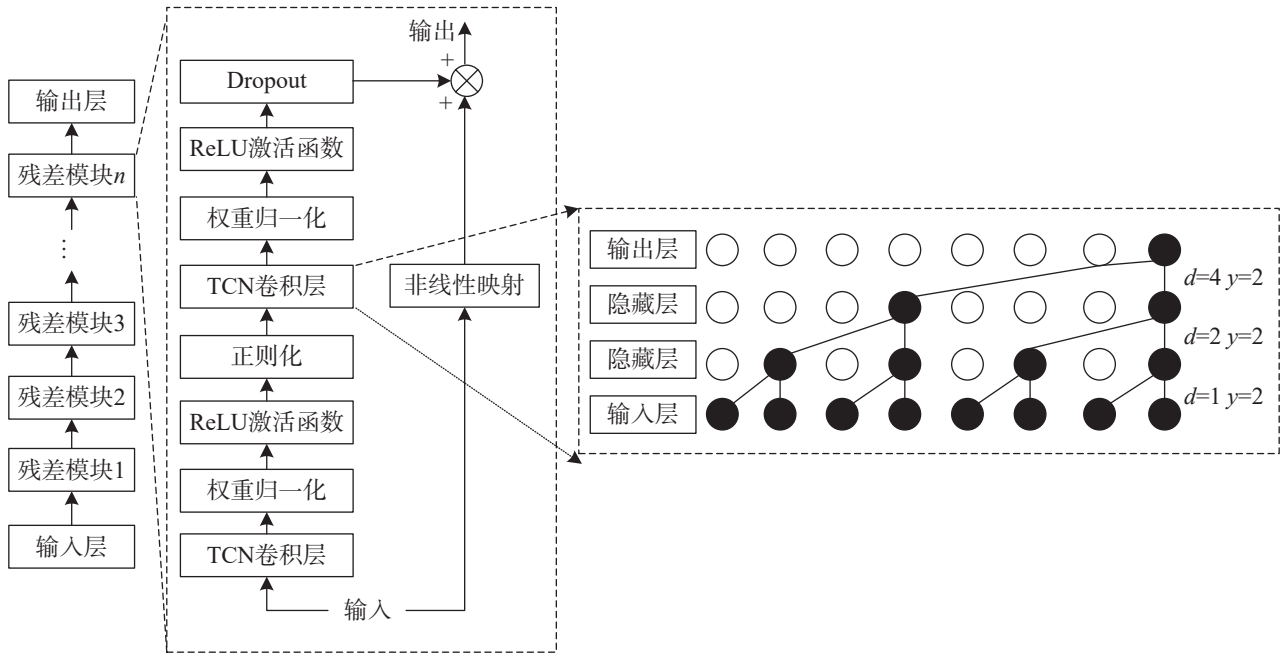


图2 TCN 网络结构

Fig. 2 TCN network structure

#### 1.2.1 因果膨胀卷积

因果卷积通过对卷积核的位置进行限制, 确保每个时间步的输出仅依赖当前及之前的时间步, 而不会受到未来时间步数据的影响。主要目的是解决传统卷积结构中普遍存在的信息泄露问题, 从而使 TCN 能够更好地应用于电力负荷预测时间序列任务。其计算表达式为:

$$z_y = \sum_{y=0}^{Y-1} w_y \cdot x_y \quad (6)$$

式中:  $w_y$  为卷积核的权重;  $Y$  为卷积核的大小;  $x_y$  为当前与过去的输入值。

而膨胀卷积的主要目的是在不增加参数量和计算成本的前提下, 扩展卷积核的感受野, 从而有效捕获输入信号的长程依赖关系, 解决与历史数据相关的信息丢失问题<sup>[19]</sup>。TCN 将膨胀卷积与因果卷积相结合, 这种结合不仅扩展了模型的视野, 使其能够捕获输入序列中的长范围依赖关系, 还避免了信息泄露问题。

#### 1.3 改进挤压-激励模块的构建

SE 模块是 Hu 等<sup>[20]</sup>提出的一种基于通道注意力的机制, 能够自适应地调整每个通道的权重, 从而显著提升网络的特征表达能力。SE 模块的结构如图 3。

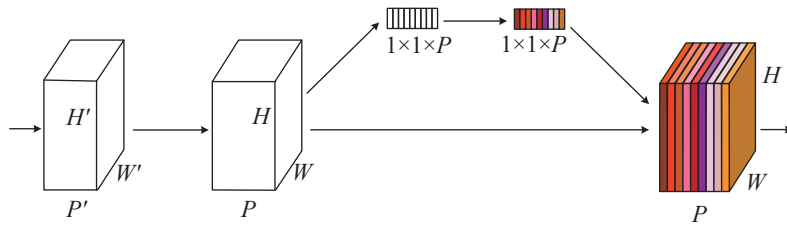


图3 SE 模块示意图

Fig. 3 Schematic diagram of the SE module

如图 3, SE 模块主要通过挤压与激励操作来实现对特征通道的动态加权。通过挤压操作对输入特征的空间维度进行全局平均池化, 将每个通道的空间信息压缩为一个全局描述值, 计算公式如下:

$$z_p = \frac{1}{HW} \sum_{h=1}^H \sum_{w=1}^W X(h, w) \quad (7)$$

式中:  $H, W, P$  分别表示特征图的高度、宽度和通道数;  $p$  表示某一通道;  $X$  表示特征;  $h, w$  分别表示某一高度和宽度。

通过激励操作对通道描述值进行非线性变换(如使用 Sigmoid 函数计算等), 以学习不同通道权重; 再将通道权重与原始输入特征逐通道相乘, 重新调整特征图的通道权重<sup>[21]</sup>。本文将原始 SE 模块进行改进, 将原有的挤压操作替换为更适合电力负荷数据的时序维度平均池化, 而非对空间维度操作。通过对每个通道的时间序列取平均值, 确保获取通道全局特征, 从而能够更准确地分配各个特征的权重。其计算公式如下:

$$z_a = \frac{1}{A} \sum_{a=1}^A X(a) \quad (8)$$

式中:  $A$  表示特征的时间步数;  $a$  表示某个时间步。

### 1.4 Pyraformer 模型的构建

Pyraformer 是针对时间序列预测任务设计的模型, 旨在解决长时间依赖关系建模和 Transformer 模型存在的计算复杂度问题。该模型通过引入金字塔注意力模块, 显著提升模型对长序列数据的处理能力, 并有效降低计算资源的需求。其流程图如图 4。

#### 1.4.1 金字塔注意力模块

为了捕捉不同范围的时间依赖关系, 文中引入金字塔注意力模块 (pyramidal attention module, PAM)<sup>[22]</sup>。该模块通过递归缩减输入序列的长度, 形成类似金字塔结构的层次化表示, 能够有效捕获长序列数据中的全局和局部特征, 同时显著降低计算复杂度。其核心思想是通过滑动窗口平均池化或其他聚合方法逐层缩减序列长度, 使得高层包含全局信息, 低层保留细节信息, 从而在不同时间尺度上提取特征, 有效降低模型的复杂性。如图 5 所示, 从内到外, 蓝色虚线表示每个节点的自注意力, 橙色虚线

表示同尺度内节点之间的信息交互, 绿色虚线表示不同尺度间节点的信息交互, 红色点画线表示任意两节点间信息交互所需的最大信息传播路径<sup>[23]</sup>。通过这种层次化设计, PAM 能够在保证特征提取能力的同时, 显著提升计算效率, 适用于处理长序列数据中的复杂依赖关系。

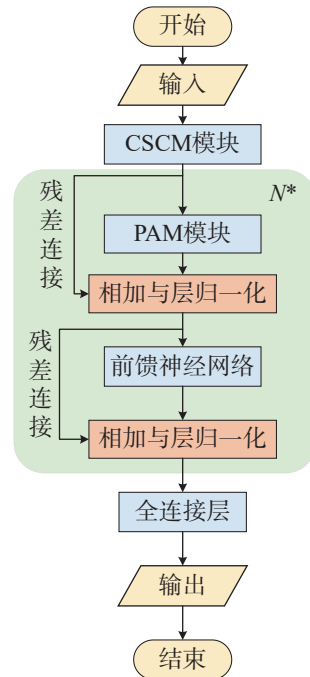


图4 Pyraformer 网络流程图

Fig. 4 Flowchart of Pyraformer network

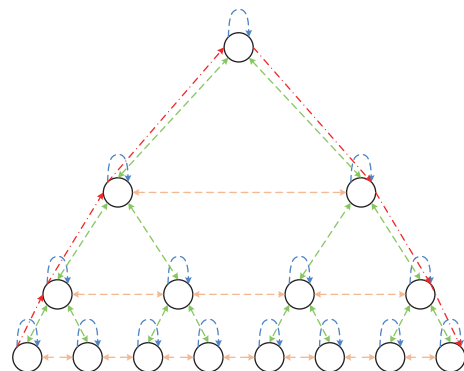


图5 金字塔注意力模块

Fig. 5 Pyramid attention module

#### 1.4.2 粗粒度构造模块

粗粒度构造模块 (coarse-scale construction

module, CSCM) 的目标是初始化金字塔图中粗尺度的节点, 促进后续 PAM 在这些节点之间进行信息交互, 其结构如图 6。其中,  $B$  为批量大小,  $D$  为向量特征维度。CSCM 使用多个卷积层, 卷积层的核大小和步幅均为  $C$ , 且按时间维度顺序嵌入序列<sup>[24]</sup>, 从而生成长度为  $L/C^{s+1}$  的序列。将这些从精细到粗糙的序列整合后输入 PAM。此外, 为了降低参数量与计算成本, 在特征传递至堆叠卷积层之前, 先通过全连接层进行降维处理, 待所有卷积操作完成后, 再通过全连接层恢复至原始维度。这一设计在减少了参数量的同时提升了计算效率, 使模型在处理长序列电力负荷数据时兼具高效性与灵活性。

综上, 基于 SSA-VMD 优化的 TCN-ISE-Pyraformer 预测流程如图 7。首先对收集的电力负荷数据进行预处理以降低干扰; 然后利用优化的 VMD 分解数据, 并通过相关性分析筛选关键变量; 最后, 将处理后的数据划分为训练集、验证集与测试集, 输入 TCN-ISE-Pyraformer 进行训练与预测。

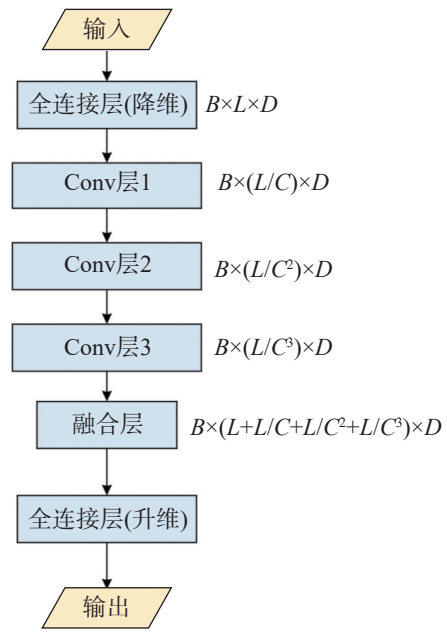


图6 粗粒度构造模块

Fig. 6 Coarse-grained construction module

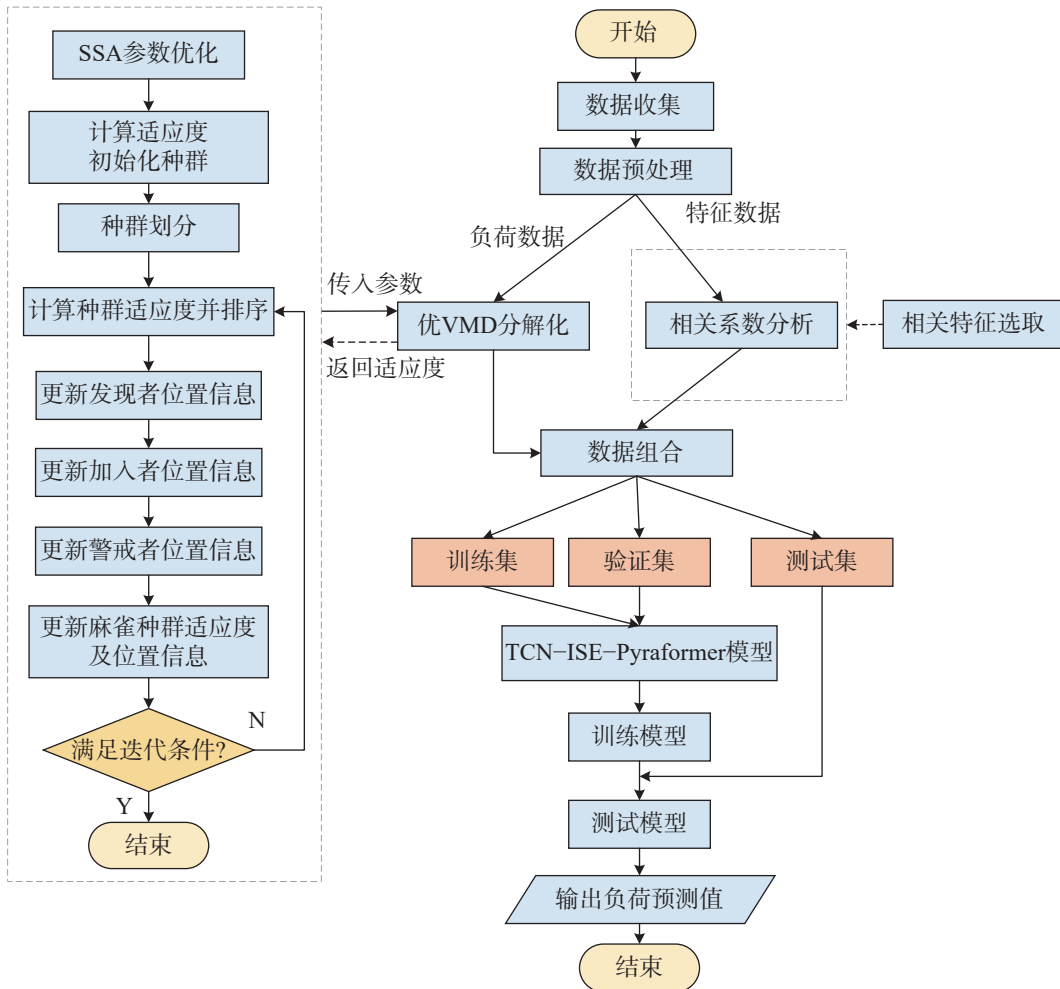


图7 基于 SSA 优化 VMD 与 TCN-ISE-Pyraformer 组合预测模型的流程图

Fig. 7 Flowchart of the combined prediction model based on SSA-optimized VMD and TCN-ISE-Pyraformer

## 2 组合预测模型的仿真验证

为验证本文所提组合预测模型的可靠性,采用澳大利亚某地区电力负荷数据集(算例1)进行仿真实验。该数据集包括2006年1月1日0时30分至2011年1月1日0时0分的电力负荷以及电价等影响电力负荷因素的数据,采样间隔为30 min,每天有48个采样点,特征参数信息如表1。

表1 特征参数信息

Tab. 1 Characteristic parameter information

特征	描述
电价	采样点时间内一度电的价格
露点温度	水蒸气与水达到平衡状态时的温度
干球温度	被测量空气的真实热力学温度
湿球温度	绝热饱和温度
湿度	大气干燥程度

选取上述数据集中2010年共17519条具有6个维度的数据进行仿真实验,将这些数据按7:2:1的比例划分为训练集、验证集与测试集。硬件平台选择 Intel Core i9-13980HX CPU 与 NVIDIA GeForce RTX 4060 GPU,软件平台采用 Python3.10 环境及 Pytorch2.2 深度学习框架。

### 2.1 数据预处理

在算例分析前,对数据集进行完整性和异常性

分析,以检测缺失值和异常值。鉴于短时期内电力负荷波动较小,采用前向填充法,对个别缺失值与异常值进行插值处理,确保数据的连续性和合理性。对预处理后的部分原始数据进行可视化分析,结果见图8。结果表明电力负荷数据具有较高的波动性和复杂的时序特征。

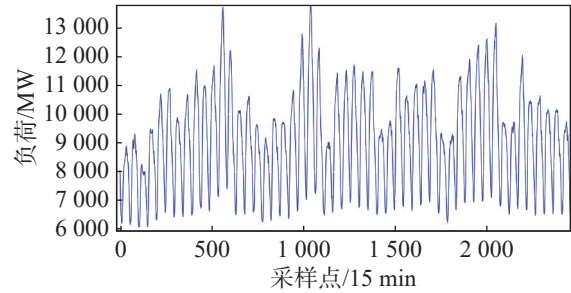


图8 部分原始负荷数据

Fig. 8 Some raw load data

原始数据包含电价、湿度等5个与电力负荷相关的特征。但是,由于电力负荷具有明显的时间特征,为准确地刻画负荷的日变化规律,进一步引入24 h周期内的时间信息作为关键特征。该特征能够反映白天用电高峰与夜间用电低谷的差异,从而提升模型对时间依赖性的理解。为了评估各特征与电力负荷的关系,采用皮尔逊相关系数( $\rho$ )进行分析并进行可视化,结果如图9。

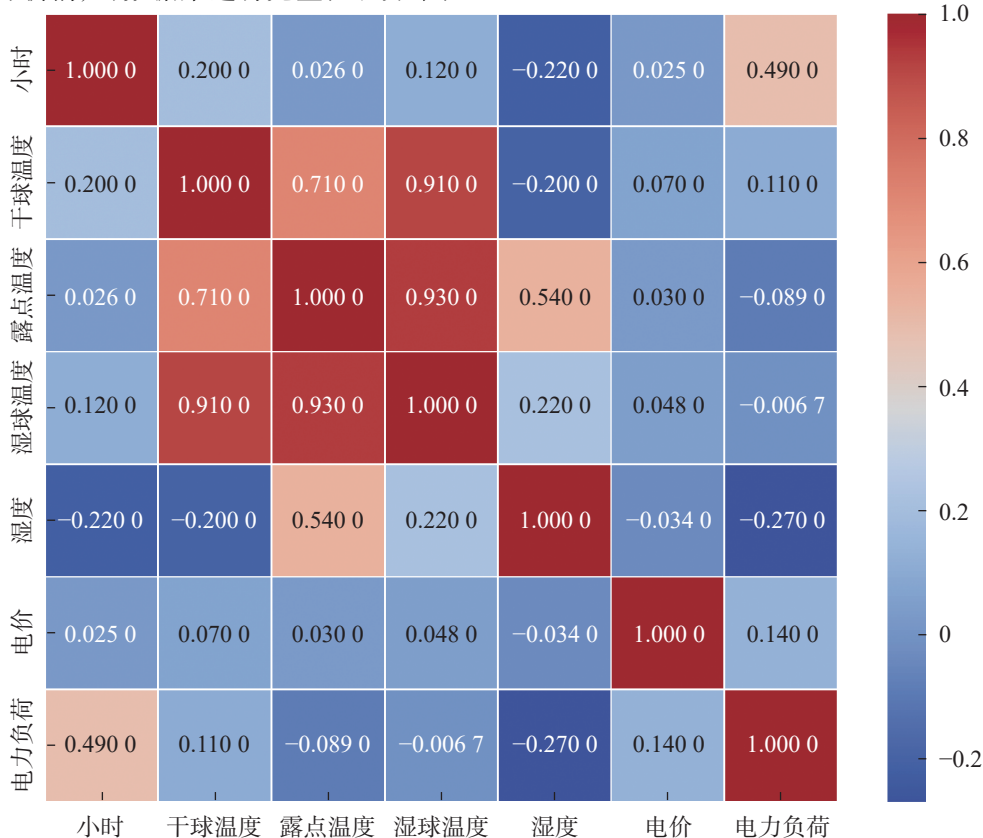


图9 相关性分析结果

Fig. 9 Results of the correlation analysis

图9中,  $\rho$ 的取值范围在  $[-1, 1]$ , 当  $\rho > 0$  时, 表明2个数据正相关; 当  $\rho < 0$  时, 表明2个数据负相关; 当  $\rho = 1$  或  $\rho = -1$  时, 表明2个数据完全相关; 如果  $\rho = 0$ , 表明2个数据不相关。

由图9可得: 电力负荷与小时特征的相关性最强, 而与湿球温度的相关性最弱。为了降低计算复杂度并加快模型训练速度, 选取小时、电价、湿度和干球温度4个与电力负荷相关性较高的特征进行后续仿真。此外, 采用滑动窗口法将这4个特征与电力负荷数据划分为输入特征和目标值, 并按照7:2:1的比例划分为训练集、验证集和测试集。为增强模型对电力负荷内部特征的提取能力, 并消除不同特征之间因单位差异带来的影响, 本文对输入数据进行归一化处理。具体的归一化表达式如下:

$$\bar{x} = \frac{x - \min x}{\max x - \min x} \quad (9)$$

式中:  $\bar{x}$ 为归一化后的数据;  $x$ 为需要归一化的数据。

## 2.2 优化VMD的结果分析

为验证SSA优化VMD在电力负荷信号分解及预测准确性提升方面的优势, 对比分析3种方案的预测性能: SSA优化VMD分解、传统EMD分解以及未分解的原始电力负荷数据。仿真参数设置:

SSA预警值为0.8, 发现者比例为0.2, 意识到危险的麻雀比例为0.1, 种群数量为50, 迭代30轮。除VMD分解数  $K$  和惩罚因子  $\alpha$  外, 其余参数保持默认值, 即步长 ( $\tau$ ) 为0、直流分量控制参数 (DC) 为0、初始化方式 (init) 为1、收敛容忍误差为  $1e-7$ 。在分解过程中,  $K$  过大会导致过度拟合, 而  $K$  过小会导致分解不充分; 同时,  $\alpha$  过大可能会导致模态混叠, 而  $\alpha$  过小会引入噪声。本文通过SSA确定上述2个参数最终的寻优范围和最优结果, 结果如表2。

表2 超参数寻优范围及最优值

Tab. 2 Hyperparameter optimization range and optimal value

待优化超参数	$K$	$\alpha$
寻优范围	4~10	1 000~5 000
最优值	4	5 000

由表2可知: 对于波动性较强的电力负荷数据, 惩罚因子 ( $\alpha$ ) 的选取至关重要, 其取值范围直接影响分解效果的稳定性和准确性。利用SSA优化获得的最优参数对电力负荷数据进行VMD分解, 结果如图10。同时, 采用传统EMD对电力负荷进行分解, 结果如图11。

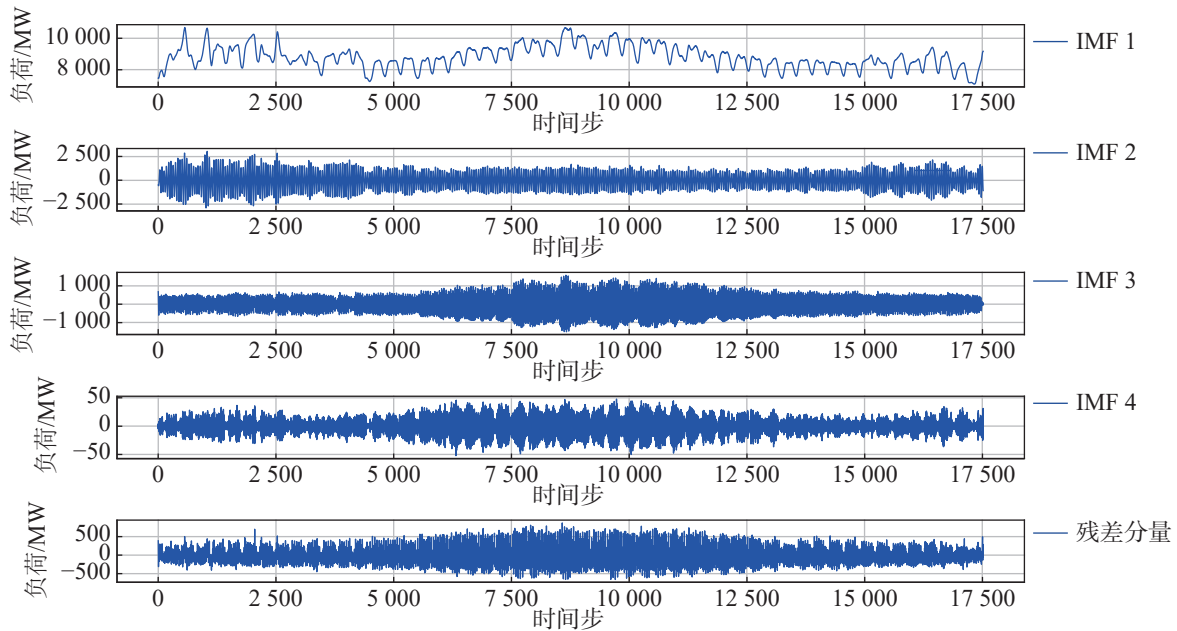
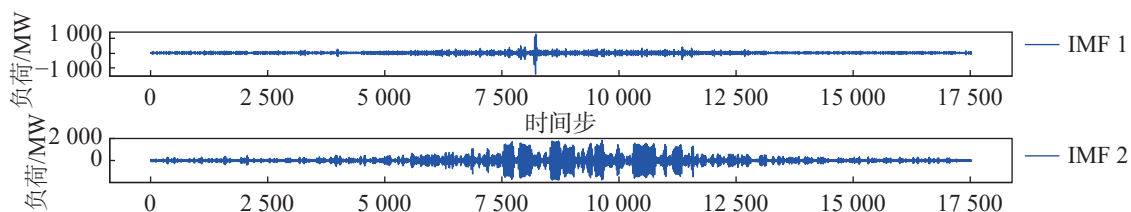


图10 优化VMD的分解结果

Fig. 10 Optimized decomposition results of VMD



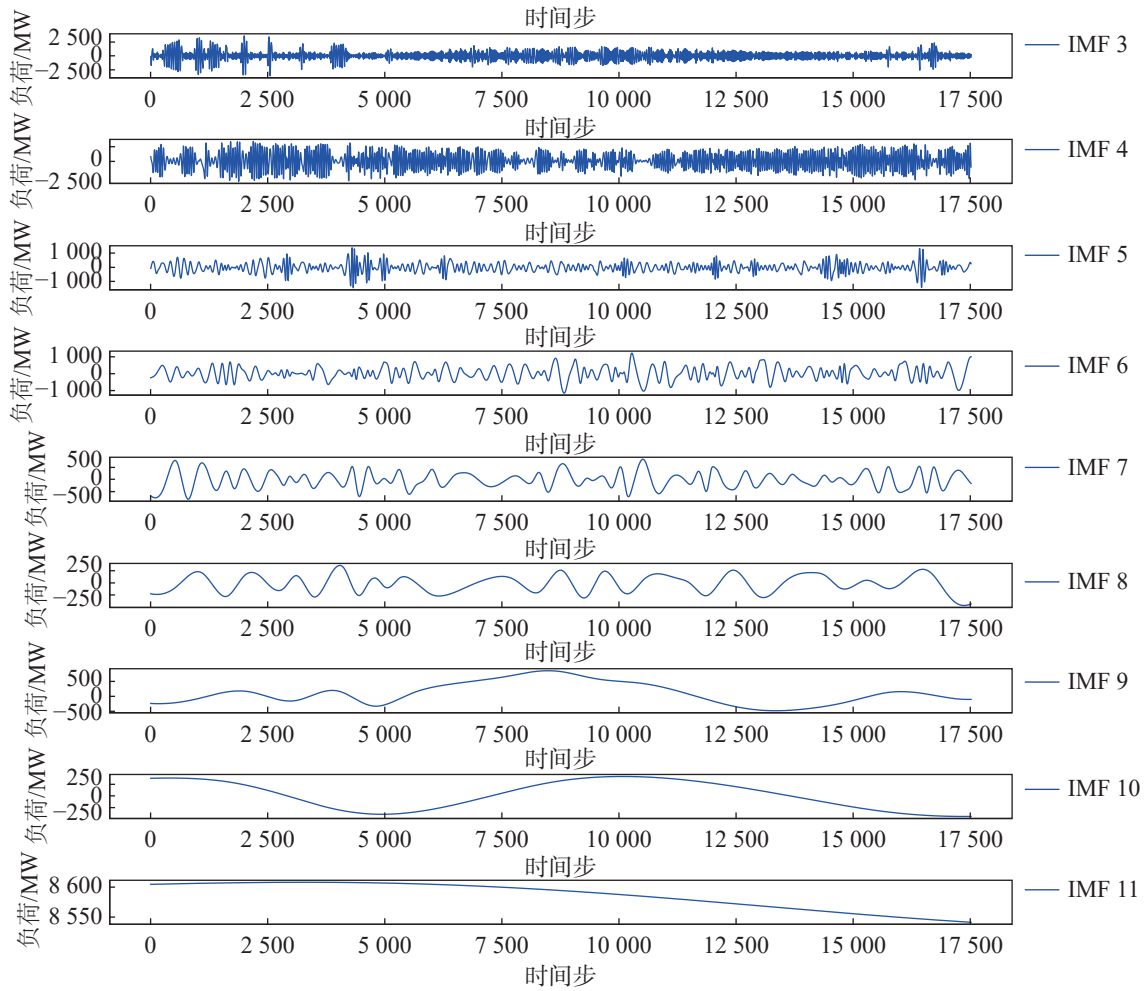


图11 EMD 分解结果

Fig. 11 Results of EMD decomposition

本文利用样本熵 (sample entropy, SE) 评估 EMD 分解所得各 IMF 的复杂度, 参考文献 [25] 的方法将复杂度相近的 IMF 进行聚合, 从而减少冗余模态并简化模型输入。IMF 的具体聚合方案及对应样本熵计算结果如表 3。

表 3 IMF 的聚合方案及对应样本熵计算结果

Tab.3 IMF aggregation scheme and corresponding SE calculation results

IMF聚合结果	参与分量	样本熵计算结果
Co-IMF1	IMF1	0.882 8
	IMF2	0.346 9
	IMF3	0.527 7
	IMF4	0.470 4
	IMF5	0.166 8
Co-IMF2	IMF6	0.074 1
	IMF7	0.035 1
	IMF8	0.013 9
Co-IMF3	IMF9	0.002 6
	IMF10	0.001 1
Co-IMF4	IMF11	0.000 2

如表 3 所示, 将 SE 计算结果位于  $[0.1\sim 1.0]$ ,  $[0.01\sim 0.10]$ ,  $[0.001\sim 0.100]$  以及小于 0.001 的 IMF 分别进行聚合。随后, 将优化后的 VMD 分解结果、聚合后的 EMD 分解结果以及未分解的负荷预处理数据分别输入 TCN-ISE-Pyraformer 模型进行预测。在模型参数设置方面, 选取 Adam 优化器, 滑动窗口大小为 48, 特征数为 5 个, 训练次数为 50 轮, 学习率为 0.001, 批量大小为 256。再使用默认参数对各模型进行仿真, 并采用单变量调整法优化关键参数, 最终确定以下结构配置: TCN 网络由 3 个残差块堆叠而成, 每个残差块卷积核大小为 3, 每层输出通道数分别为 32, 64, 128; ISE 隐藏层维度为 128; Pyraformer 网络注意力层数为 4, 注意力头数为 8。为了防止模型复杂度增加导致的过拟合, 在模型中加入 Dropout 层, 设置 Dropout 的概率为 0.1。3 种分解方法的预测结果如图 12。

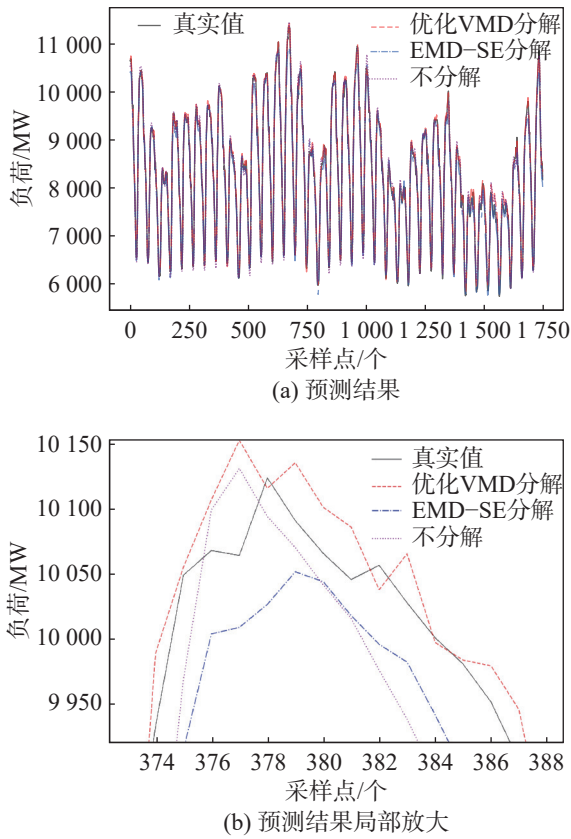


图12 3种分解方法的预测结果

Fig. 12 Prediction results of the 3 decomposition methods

为了全面评估本文所提模型的预测精度,选取平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE) 以及决定系数 ( $R^2$ ) 作为评估指标。其中, MAE 与 RMSE 的值越小而决定系数的值越大, 表明模型预测的结果越精准。3种分解方法的预测误差指标如表4。

表4 3种分解方法的预测误差指标

Tab. 4 Prediction error indicators for the 3 decomposition methods

分解方式	MAE/MW	RMSE/MW	$R^2$
优化VMD分解	70.170 5	90.978 4	0.994 9
聚合EMD分解	112.283 5	144.520 5	0.987 2
不分解	112.888 8	145.445 7	0.987 0

由表4可知: 相比于采用EMD分解以及未分解, 采用SSA优化VMD分解预测的各项误差指标均表现出显著优势。其中, MAE分别降低42.113, 42.718 MW; RMSE分别降低53.542, 54.467 MW; 而 $R^2$ 分别提高了0.79%, 0.80%。这表明采用SSA优化VMD分解在减少预测误差和提升拟合效果方面表现出色, 使模型能够更好地捕获复杂的时间序列

特性, 预测结果更接近于真实值。

### 2.3 模型性能优势的结果分析

为进一步验证本文所提模型的性能优势, 将其与VMD-LSTM, SVM, VMD-CNN-LSTM, VMD-Transformer-BiLSTM模型的仿真结果进行对比分析。为确保仿真结果的公平性, 消除数据处理方式对仿真结果的影响, 所有模型的输入数据均经过相同的预处理, 且分解方法均采用SSA优化的VMD。模型参数设置方面, SVM模型选用RBF核函数、惩罚因子为100。其他模型的超参数在相同部分保持一致: 滑动窗口大小为48、特征数为5个、训练轮次50轮、学习率0.001、批次大小256; LSTM的隐藏层维度为128、num\_layers为2层; CNN-LSTM的卷积核数量32、卷积核大小为3、padding为1、隐藏层维度为128、num\_layers为2层; Transformer-BiLSTM中Transformer的注意力头数为8、隐藏层维度为128、num\_layers为2层。对于TCN-ISE-Pyraformer模型, 参数设置与上节同。此外, 为防止过拟合, 在3个组合模型中加入Dropout层, Dropout概率均设置为0.1。5种模型的预测结果如图13。同时, 采用MAE, RMSE,  $R^2$ 对5种模型的预测效果进行定量评估, 结果如表5。

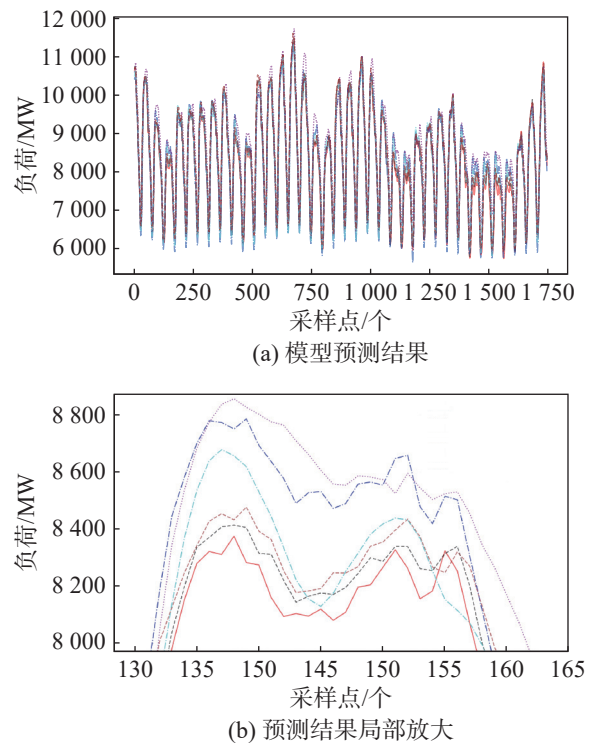


图13 5种模型对算例1的预测结果

Fig. 13 Prediction results of of five models for case 1

表 5 5 种模型对算例 1 的预测误差指标

**Tab. 5 Prediction error metrics of five models for case 1**

模型	MAE/MW	RMSE/MW	$R^2$
VMD-LSTM	201.8	252.617 0	0.960 9
SVM	416.019 1	483.706 0	0.856 7
VMD-CNN-LSTM	170.157 8	214.759 8	0.971 8
VMD-Transformer-BiLSTM	147.924 2	175.641 0	0.981 1
本文模型	70.170 5	90.978 4	0.994 9

由表 5 可知: 相较 VMD-LSTM, SVM, VMD-CNN-LSTM, VMD-Transformer-BiLSTM 模型, 本文所提模型的 MAE 分别下降 131.629 5, 345.848 6, 99.987 3, 77.753 6 MW; RMSE 分别下降 161.638 6, 392.727 2, 123.781 3, 84.662 7 MW; 而决定系数分别提升 3.54%, 16.14%, 2.39%, 1.41%。综上表明, 本文所提模型的预测效果最佳, 能够更准确地捕捉电力负荷数据的复杂特性, 具有较高的预测准确性和可靠性。

2.4 模型泛化能力的结果分析

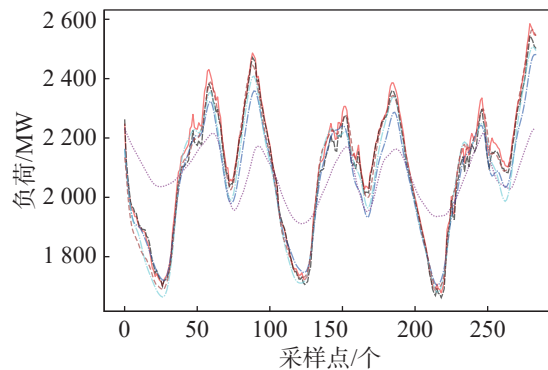
为了进一步验证本文所提模型的泛化能力, 选取印度北部某一邦的电力负荷数据 (算例 2) 进行测试。该数据集涵盖该地区 2018 年全年电力负荷数据, 同时包含体感温度等影响负荷变化的相关因素。数据采样间隔为 15 min, 每天共 96 个采样点, 选取 6 月份的数据进行仿真分析。为确保仿真分析的公正性与科学性, 除必要的模型划分和输入调整外, 所提模型及对比模型参数与前文算例保持一致。对算例 2 进行相同的特征分析后, 最终选取体感温度、湿度、实际温度、露点 4 个特征进行仿真分析。对本电力负荷数据进行 VMD 分解的寻优范围及最优结果如表 6。由表 6 可知: 与算例 1 相比, 算例 2 的惩罚因子 ( $\alpha$ ) 值有所降低。

表 6 算例 2 的超参数寻优范围及最优值

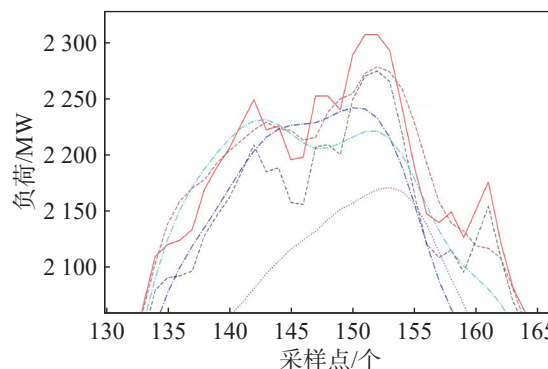
**Tab. 6 Hyperparameter optimization ranges and optimal values for case 2**

待优化超参数	$K$	$\alpha$
寻优范围	4~10	1 000~5 000
最优值	4	1 888

利用寻优得到的最优参数对电力负荷数据进行 VMD 分解处理, 并将分解后的模式分量与选取的特征参数经归一化后输入模型, 按照算例 1 的仿真流程进行测试, 最终得到的模型预测结果如图 14。同时, 采用 MAE, RMSE,  $R^2$  指标对模型的预测结果进行评估, 结果如表 7。



(a) 模型预测结果



(b) 预测结果局部放大

— 真实值; ··· 本文模型; - - VMD-LSTM; - · - SVM; ··· VMD-CNN-LSTM; ··· VMD-Transformer-BiLSTM。

图 14 5 种模型对算例 2 的预测结果

Fig. 14 Prediction results of of five models for case 2

表 7 5 种模型对算例 2 的预测误差指标

**Tab. 7 Prediction error metrics of five models for case 2**

模型	MAE/MW	RMSE/MW	$R^2$
VMD-LSTM	54.629 5	66.232 7	0.907 1
SVM	143.298 1	167.724 9	0.404 3
VMD-CNN-LSTM	43.315 4	52.025 8	0.942 7
VMD-Transformer-BiLSTM	25.981 5	32.714 9	0.977 3
本文模型	23.057 9	27.280 3	0.984 2

由表 7 可知: 相较于 VMD-LSTM, SVM, VMD-CNN-LSTM, VMD-Transformer-BiLSTM 模型, 本文所提模型的 MAE 分别下降 31.571 6, 120.240 2, 20.257 4, 2.923 51 MW; RMSE 分别下降 38.952 4, 140.444 5, 24.745 5, 5.434 5 MW; 而  $R^2$  分别提升 8.50%, 143.46%, 4.42%, 0.71%。综上所述, 本文提出的模型不仅在不同数据集上展现了高精度的预测能力, 还能够有效适应不同区域的电力负荷特性, 表现出较强的泛化能力。这进一步验证了本文所提模型在电力负荷预测任务中的实用性与可靠性。

### 3 结论

针对电力负荷数据非线性、波动性及不确定性的特征,以及传统组合预测模型难以同时捕捉多特征负荷数据全局和局部特征的问题,提出一种基于优化VMD与TCN-ISE-Pyraformer的短期电力负荷组合预测模型,通过在2个不同地区的真实电力负荷数据集上进行仿真分析,得出以下主要结论:

1) 通过引入麻雀搜索算法(SSA)对VMD的2个关键参数进行优化,使VMD能够更精确地分解电力负荷数据,有效减少原始序列中的噪声干扰,提升模型对负荷特征的捕捉能力。

2) TCN-ISE-Pyraformer组合模型能够同时捕捉多特征负荷数据的全局和局部特征。其中,改进的挤压-激励(ISE)模块通过自动学习特征的重要性,有效减少冗余信息;而Pyraformer作为一种改进的Transformer架构,有效解决传统Transformer计算复杂度与预测耗时间长的问题。

以上结论表明,本文提出的模型在捕捉电力负荷数据的全局与局部特征、提升预测精度和稳定性方面具有显著优势,能够有效应对电力负荷预测中的复杂性和不确定性。

#### 参考文献:

- [1] 方磊, 楚成博, 何映虹, 等. 基于自适应辛几何模态分解-多元线性回归-卷积长短时记忆的台区电力负荷预测[J/OL]. 现代电力, (2024-12-26). <https://doi.org/10.19725/j.cnki.1007-2322.2023.0112>.
- FANG L, CHU C B, HE Y H, et al. Power load forecasting in distribution areas based on adaptive symplectic geometric mode decomposition-multiple linear regression-convolutional long short-term memory [J/OL]. Modern Electric Power, (2024-12-26). <https://doi.org/10.19725/j.cnki.1007-2322.2023.0112>.
- [2] KIM S Y, JUNG H W, PARK J D, et al. Weekly maximum electric load forecasting for 104 weeks by seasonal ARIMA model[J]. Journal of the Korean Institute of Illuminating and Electrical Installation Engineers, 2014, 28(1):50-56.
- [3] 严启华, 罗亮, 陈振宇, 等. 基于差分自回归移动平均模型、长短期记忆网络模型及相关模型的无人机俯仰角预测[J]. 科学技术与工程, 2024, 24(5):2142-2149.
- YAN Q H, LUO L, CHEN Z Y, et al. UAV pitch angle prediction based on ARIMA and LSTM related models[J]. Science Technology and Engineering, 2024, 24(5):2142-2149.
- [4] SI J C, WANG Y Y, GUAN Y C, et al. Load forecasting based on multi-core learning support vector machine (SVM)[J]. Journal of Physics: Conference Series, 2024, 2876(1):012035.
- [5] SAYED H A, WILLIAM A, SAID A M. Smart electricity meter load prediction in Dubai using MLR, ANN, RF, and ARIMA[J]. Electronics, 2023, 12(2):389.
- [6] WANG H J, LU X Y. Research on short-term forecasting of power load based on big data BP neural network[J]. Journal of Physics: Conference Series, 2022, 2401(1):012077.
- [7] WANG C, LI X, SHI Y, et al. Load forecasting method based on CNN and extended LSTM[J]. Energy Reports, 2024, 12:2452-2461.
- [8] LI B L, SHAO Y, LIAN Y F, et al. Bayesian optimization-based LSTM for short-term heating load forecasting[J]. Energies, 2023, 16(17):6234.
- [9] 杨世海, 薛冰, 李磊, 等. 基于注意力机制的CNN-BiGRU超短期省间现货购电需求预测[J/OL]. 电力系统及其自动化学报, (2024-12-26). <https://doi.org/10.19635/j.cnki.csu-epsa.001547>.
- YANG S H, XUE B, LI L, et al. Ultra-short-term inter-provincial spot electricity purchase demand forecasting based on attention mechanism with CNN-BiGRU [J/OL]. Proceedings of the CSU-EPSA, (2024-12-26). <https://doi.org/10.19635/j.cnki.csu-epsa.001547>.
- [10] 王耀辉, 薛贵军, 赵广昊. 基于改进沙猫群算法优化CNN-BiLSTM的热负荷预测[J]. 现代电子技术, 2024, 47(14):20-29.
- WANG Y H, XUE G J, ZHAO G H. Optimised CNN-BiLSTM for heat load prediction based on improved sand cat swarm algorithm[J]. Modern Electronics Technique, 2024, 47(14):20-29.
- [11] AN W H, GAO B, LIU J H, et al. Predicting hourly heating load in residential buildings using a hybrid SSA-CNN-SVM approach[J]. Case Studies in Thermal Engineering, 2024, 59:104516.
- [12] 石卓见, 冉启武, 徐福聪. 基于聚合二次模态分解及Informer的短期负荷预测[J]. 电网技术, 2024, 48(6):2574-2583.
- SHI Z J, RAN Q W, XU F C. Short-term load forecasting based on aggregated secondary decomposition and informer[J]. Power System Technology, 2024, 48(6):2574-2583.
- [13] 姚程文, 杨苹, 刘泽健. 基于CNN-GRU混合神经网络的负荷预测方法[J]. 电网技术, 2020, 44(9):3416-3424.
- YAO C W, YANG P, LIU Z J. Load forecasting method based on CNN-GRU hybrid neural network[J]. Power System Technology, 2020, 44(9):3416-3424.
- [14] 王清, 陈祉如, 李贵民, 等. 基于VMD与TCN的台区短期负荷预测算法研究[J]. 哈尔滨理工大学学报, 2024, 29(2):121-129.

- WANG Q, CHEN Z R, LI G M, et al. Research on short-term load forecasting algorithm based on VMD and TCN[J]. Journal of Harbin University of Science and Technology, 2024, 29(2):121-129.
- [15] 刘杰, 从兰美, 夏远洋, 等. 基于 DBO-VMD 和 IWOA-BiLSTM 神经网络组合模型的短期电力负荷预测 [J]. 电力系统保护与控制, 2024, 52(8):123-133.
- LIU J, CONG L M, XIA Y Y, et al. Short-term power load prediction based on DBO-VMD and an IWOA-BiLSTM neural network combination model[J]. Power System Protection and Control, 2024, 52(8):123-133.
- [16] 胡威, 张新燕, 李振恩, 等. 基于优化的 VMD-mRMR-LSTM 模型的短期负荷预测 [J]. 电力系统保护与控制, 2022, 50(1):88-97.
- HU W, ZHANG X Y, LI Z E, et al. Short-term load forecasting based on an optimized VMD-mRMR-LSTM model[J]. Power System Protection and Control, 2022, 50(1):88-97.
- [17] 王斌斌, 孙丽江. 基于 VMD-SSA-BiLSTM 网络下的短期电力负荷预测 [J]. 上海电机学院学报, 2024, 27(5):274-279,298.
- WANG B B, SUN L J. Short-term power load forecasting based on VMD-SSA-BiLSTM network[J]. Journal of Shanghai Dianji University, 2024, 27(5):274-279,298.
- [18] 郭渊, 张雪成, 董振标, 等. 基于 TCN-BiLSTM-AM 的居民住宅短期电力负荷预测 [J]. 现代电子技术, 2024, 47(19):100-108.
- GUO Y, ZHANG X C, DONG Z B, et al. Short-term residential power load prediction based on TCN-BiLSTM-AM[J]. Modern Electronics Technique, 2024, 47(19):100-108.
- [19] 赵小强, 杨秦乐, 王涛, 等. 基于 TCNMS-BiLSTM 的短期电力负荷预测 [J/OL]. 控制工程, (2024-11-24). <https://doi.org/10.14107/j.cnki.kzgc.20240422>.
- ZHAO X Q, YANG Q L, WANG T, et al. Short-term power load forecasting based on TCNMS-BiLSTM [J/OL]. Control Engineering of China, (2024-11-24). <https://doi.org/10.14107/j.cnki.kzgc.20240422>.
- [20] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8):2011-2023.
- [21] ZHU R, TANG J, HE X S, et al. SE-MACConvLSTM: a deep learning framework for short-term traffic flow prediction combining squeeze-and-excitation network and multi-attention convolutional LSTM network[J]. PloS One, 2024, 19(12):e0312601.
- [22] LIU X, WANG Y, LONG G, et al. Pyraformer: Low-complexity pyramidal attention for long-range time series modeling[J]. arXiv preprint, 2021, 2106.04156.
- [23] 蔡华锋, 唐艺豪. 基于 Pyraformer 网络的短期电力负荷预测 [J]. 武汉大学学报(工学版), 2023, 56(9):1105-1113.
- CAI H F, TANG Y H. Short-term power load forecasting based on Pyraformer network[J]. Engineering Journal of Wuhan University, 2023, 56(9):1105-1113.
- [24] 孙祥瑜, 王超, 杨一昀, 等. CEEMDAN-Pyraformer-LSS 模型在永定河径流预测中的应用 [J/OL]. 南水北调与水利科技(中英文), (2024-12-27). <https://kns.cnki.net/kcms/detail/13.1430.TV.20241226.1539.002.html>.
- SUN X Y, WANG C, YANG Y Y, et al. Application of CEEMDAN-Pyraformer-LSS model in runoff prediction of Yongding River[J/OL]. South-to-North Water Transfers and Water Science & Technology, (2024-12-27). <https://kns.cnki.net/kcms/detail/13.1430.TV.20241226.1539.002.html>.
- [25] 钟燕, 王军, 宋戈, 等. 基于二次重构分解去噪及双向长短时记忆网络的极端天气下超短期电力负荷预测 [J/OL]. 电网技术, (2025-01-06). <https://doi.org/10.13335/j.1000-3673.pst.2024.0935>.
- ZHONG Y, WANG J, SONG G, et al. Ultra-short-term power load forecasting under extreme weather conditions based on secondary reconstruction decomposition denoising and bidirectional long short-term memory network [J/OL]. Power System Technology, (2025-01-06). <https://doi.org/10.13335/j.1000-3673.pst.2024.0935>.

责任编辑:何莉