

强化数据预处理的 BLSTNet-CBAM 短期电力负荷预测^①



陈万志¹, 张思维¹, 王天元²

¹(辽宁工程技术大学 软件学院, 葫芦岛 125105)

²(国网辽宁省电力有限公司 营口供电公司, 营口 115002)

通信作者: 陈万志, E-mail: chenwanzhi@lntu.edu.cn

摘要: 针对负荷数据复杂性、非平稳性以及负荷预测误差较大等问题, 提出一种综合特征构建和模型优化的短期电力负荷预测新方法. 首先采用最大信息系数 (MIC) 分析特征变量的相关性, 选取与电力负荷序列相关的特征变量, 同时, 考虑变分模态分解 (VMD) 方法容易受主观因素的影响, 采用霜冰优化算法 (RIME) 优化 VMD, 完成原始电力负荷序列的分解. 然后改进长短期时间序列网络 (LSTNet) 作为预测模型, 将其递归层 LSTM 更新为 BiLSTM, 并引入卷积块注意力机制 (CBAM) 进行预测. 通过对比实验和消融实验的结果表明: 经 RIME-VMD 优化后, LSTM、GRU、LSTNet 模型预测的均方根误差 (RMSE) 均降低 20% 以上, 显著提高模型预测精度, 且能够适应于不同预测模型. 所提出的 BLSTNet-CBAM 模型与 LSTM、GRU、LSTNet 相比, RMSE 分别降低了 35.54%、6.78%、1.46%, 提高了短期电力负荷预测的准确性.

关键词: 短期电力负荷预测; 霜冰优化算法; 变分模态分解; 长短期时间序列网络; 卷积块注意力机制

引用格式: 陈万志, 张思维, 王天元. 强化数据预处理的 BLSTNet-CBAM 短期电力负荷预测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/9490.html>

Enhanced Data Preprocessing for BLSTNet-CBAM Short-term Power Load Forecasting

CHEN Wan-Zhi¹, ZHANG Si-Wei¹, WANG Tian-Yuan²

¹(School of Software, Liaoning Technical University, Huludao 125105, China)

²(Yingkou Power Supply Company, State Grid Liaoning Electric Power Supply Co. Ltd., Yingkou 115002, China)

Abstract: A new method for short-term power load forecasting is proposed to address issues such as complex and non-stationary load data, as well as large prediction errors. Firstly, this study utilizes the maximum information coefficient (MIC) to analyze the correlation of feature variables and selects relevant variables related to power load sequences. At the same time, as the variational mode decomposition (VMD) method is susceptible to subjective factors, the study employs the rime optimization algorithm (RIME) to optimize VMD and decompose the original power load sequence. Then, the long and short-term time series network (LSTNet) is improved as the prediction model by replacing the recursive LSTM layer with BiLSTM and incorporating the convolutional block attention mechanism (CBAM). Comparative experiments and ablation experiments demonstrate that RIME-VMD reduces the root mean square error (RMSE) of the LSTM, GRU, and LSTNet models by more than 20%, significantly improving the prediction accuracy of the models, and can be adapted to different prediction models. Compared with LSTM, GRU, and LSTNet, the proposed BLSTNet-CBAM model reduces the RMSE by 35.54%, 6.78%, and 1.46% respectively, improving the accuracy of short-term power load forecasting.

Key words: short-term power load forecasting; rime optimization algorithm (RIME); variational mode decomposition

^① 基金项目: 国家重点研发计划 (2018YFB1403303); 辽宁省教育厅高校科研基金 (2021LJKZ0327)

收稿时间: 2023-11-17; 修改时间: 2023-12-22; 采用时间: 2023-12-27; csa 在线出版时间: 2024-03-15

(VMD); long and short-term time series network (LSTNet); convolutional block attention mechanism (CBAM)

当前新能源大规模融入电力系统, 导致影响负荷的因素更加多样化, 电力需求的变化更具随机性, 短期负荷预测研究至关重要^[1]. 考虑到负荷序列的波动性、周期性变化和非线性特征, 因此需要开发有效的模型和方法解决负荷的多样性和不确定性, 确保电力系统的稳定性和可靠性^[2].

针对电力负荷预测方法大致包括两类: 基于数学统计方法和机器学习方法. 基于数学统计方法包括自回归模型^[3]、自回归移动平均模型^[4]、指数平滑法^[5]和卡尔曼滤波^[6]等. 数学统计方法模型更多地依赖于数据的周期性和离群值. 然而, 电力负荷数据与其他参数之间的关系通常是复杂和非线性的, 并且由于这些原因, 很难用统计方法获得准确的预测值. 而机器学习方法主要包括传统机器学习方法和深度学习方法. 传统机器学习方法例如: 支持向量机 (SVM)^[7]、决策树^[8]、随机森林 (RF)^[9]等. 与数学统计方法相比, 传统机器学习方法能够解决相对复杂的数据特征关系等问题, 但无法在较大的数据集上准确预测电力负荷, 同时存在对数据间的相关性考虑不全等问题. 随着机器学习方法的发展, 深度学习方法逐渐被应用于电力负荷预测. 在文献^[10]中, 提出了基于长短期记忆 (LSTM) 的电力负荷预测模型, 结果表明, LSTM 可以产生更好的预测性、实现更高的精度, 但是其训练复杂并且涉及更多参数, 增加了训练和调整模型的复杂性. 作为 LSTM 的变体, 门控循环单元 (GRU) 具有更简单的网络结构. 在文献^[11]中, GRU 被引入用于短期电力负荷预测. 然而, 由于模型中缺乏卷积, 因此仍然有必要提高特征提取的性能. 为了解决这些问题, 卷积神经网络 (CNN) 已被采用其出色的功能提取. 在文献^[12]中, 利用 CNN 进行特征提取与筛选, 减少了运算过程中的过拟合. 然而 CNN 对长期依赖关系的建模能力较弱, 无法满足电力负荷预测的需求. 因此, 在文献^[13]中, 提出了长期和短期时间序列网络 (LSTNet). 其通过卷积层、递归层和递归跳过层提取时间序列的短期局部依赖模式并发现长期趋势. 实验结果表明, LSTNet 成功地捕捉了数据中的短期和长期重复模式, 并结合了线性和非线性模型进行稳健预测. 然而, 单一的深度学习方法面临着诸如计算复杂度和低计算效率等挑战, 为了应对这些问

题, 在文献^[14]中提出了基于 CNN-BiLSTM 的短期电力负荷预测方法. 研究表明, 两种单一模型组合, 不仅弥补各自存在的问题, 而且提高预测精确度.

然而对于电力负荷预测而言, 提高其预测精度, 挖掘电力负荷数据的规律性与预测模型的优化同样关键. 因此, 为降低非线性和波动性的影响, 文献^[15-18]中提出了结合平稳小波变换的短期电力负荷预测方法, 有效提高了预测准确性. 同时, 文献^[19]的经验模式分解 (EMD) 算法为数据分解提供了显著性能改进. 此外, 文献^[20]采用变分模态分解 (VMD) 算法进行数据处理, 解决了 EMD 算法中的模态混叠问题, 并展现出更好的分解效果和较高的鲁棒性. 然而, VMD 算法的二次惩罚因子可能导致内部跳跃过度惩罚, 从而引发预测误差.

综上所述, 现有的短期电力负荷预测方法存在以下不足: (1) 上述文献大多直接利用 VMD 将输入模型数据分解为具有不同特征频率的子序列, 以简化负荷特性, 提高预测精度. 但 VMD 中的参数设置具有主观性, 可能人为影响预测结果, 导致模型的普适性和泛化能力略差. (2) 大部分以改进优化模型来提高预测精度, 却忽略电力负荷数据本身的特征和表达能力.

因此本文提出 RIME-VMD-BLSTNet-CBAM 的短期电力负荷模型方法. 首先, 采用 RIME 算法优化 VMD 方法的参数. 适应度函数选用包络熵, 这使得获得加载子序列具有较低的时间复杂度. 其次, 使用 LSTNet 预测模型, 该模型与其他神经网络相比, LSTNet 可以从长期和短期历史序列中捕捉到不同的特征和趋势. 为了更准确地捕捉到时间序列数据中动态变化和趋势, 将其递归层的 LSTM 更替为 BiLSTM. 之后在 LSTNet 中嵌入 CBAM 注意力机制, 可以减少过拟合和增强模型的泛化能力. 最后用巴拿马地区的实际电力负荷数据集, 通过仿真对比实验来验证本文提出的模型的有效性和准确性.

1 算法原理

1.1 变分模态分解 (VMD)

VMD 是一种用于处理非平稳信号的自适应分解估计方法. VMD 可以将历史电力负荷等序列非递归分解, 得到多个不同频率的特征趋势. 其不仅具有良好的

噪声鲁棒性,而且降低了原始序列的不稳定性.该方法使用维纳滤波进行降噪,对相关参数进行初始化,得到 K 估计的中心频率 ω_k ,然后使用交替方向乘法更新每个模态函数及其中心频率,以最小化总估计带宽.其计算公式为:

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

其中, $\{u_k\} = \{u_1, u_2, \dots, u_k\}$ 是分解出来的 K 个模态 (IMF) 分量; $\{\omega_k\} = \{\omega_1, \omega_2, \dots, \omega_k\}$ 是各模态的中心频率; $\delta(t)$ 是单位脉冲函数. $*$ 是卷积运算.通过在方程 (1) 中引入拉格朗日乘子算子 λ 和二次惩罚因子 α ,并将其转化为无约束变分模型,其计算公式为:

$$\begin{aligned} L(\{u_k\}, \{\omega_k\}, \lambda) = & \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \\ & + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \end{aligned} \quad (2)$$

采用交替方向乘法进行优化,迭代更新 u_k^{n+1} 、 ω_k^{n+1} 和 λ ,在迭代优化序列中寻找方程 (2) 的“鞍点”,进而寻找方程 (1) 的最优解.其计算公式为:

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

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

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

其中, $\hat{u}_k^{n+1}(\omega)$ 是对应于每个模态分量的维纳滤波器; $\hat{f}(\omega)$ 、 $\hat{u}_k(\omega)$ 、 $\hat{\lambda}^n(\omega)$ 分别是 $f(t)$ 、 $u_k(t)$ 、 $\lambda(t)$ 的傅里叶变换; ω_k^{n+1} 为各模态分量的频率中心; τ 为信号的噪声容限; n 为迭代次数.

定义最大迭代次数为 N ,使 n 满足 $n \leq N$. $\forall \varepsilon > 0$,当迭代完成时满足收敛条件式 (6).

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

1.2 霜冰优化算法 (RIME)

RIME^[21]是由 Su 等人在 2023 年 2 月提出的一种基于霜冰物理现象的高效优化算法. RIME 算法通过模拟冰晶的软霜和硬霜生长过程,构造软霜搜索策略和硬霜打孔机制,不断对种群更新从而实现优化,找到最优解. RIME 优化算法过程如下.

(1) 初始化种群: 随机生成一组冰晶 (候选解), 每个冰晶代表一个潜在解.

(2) 受软霜成长的启发,利用霜冰的强随机性和覆盖性,提出软霜搜索策略.针对霜冰粒子的 5 种运动特征,模拟了各个粒子的凝聚过程.其计算公式为:

$$R_{ij}^{\text{new}} = R_{\text{best}j} + r_1 \cdot \cos \theta \cdot \beta \cdot (h \cdot (Ub_{ij} - Lb_{ij}) + Lb_{ij}), r_2 < E \quad (7)$$

其中, R_{ij}^{new} 为更新粒子的新位置, i 和 j 表示第 i 个霜冰因子的第 j 个粒子; $R_{\text{best}j}$ 是霜冰种群 R 中最佳霜冰因子的第 j 个粒子.参数 $r_1 \in (-1, 1)$ 的随机数, r_1 控制粒子运动的方向,同时 $\cos \theta$ 将随着迭代次数而改变如式 (8); β 是环境因子,其遵循迭代次数如式 (9); h 是粘附度,用于控制两个霜冰粒子的中心距离; Ub_{ij} 和 Lb_{ij} 分别是搜索空间的上界和下界,限制了其粒子运动的有效区域; E 是被附着系数,其影响霜冰因子的缩合概率并且随着迭代次数而增加如式 (10),参数 $r_2 \in (0, 1)$ 的随机数,与 E 一起控制粒子是否凝聚,即是否更新粒子位置.

$$\theta = \pi \cdot \frac{t}{10 \cdot T} \quad (8)$$

其中, t 是当前迭代次数, T 是算法的最大迭代次数.

$$\beta = 1 - \left[\frac{w \cdot t}{T} \right] / w \quad (9)$$

其中, w 默认值为 5,用于控制阶跃函数的段数.

$$E = \sqrt{t/T} \quad (10)$$

(3) 通过对硬霜生长过程的模拟,提出硬霜打孔机制.用于霜冰粒子间的置换,其计算公式为:

$$R_{ij}^{\text{new}} = R_{\text{best}j}, r_3 < F^{\text{normr}}(S_i) \quad (11)$$

其中, $r_3 \in (-1, 1)$; $F^{\text{normr}}(S_i)$ 表示当前霜冰因子适应度值的归一化值.

1.3 LSTNet 神经网络

LSTNet 是为了解决电力负荷序列中长期周期性和短期非线性混合的挑战而设计的.其结合卷积层捕获短期模式和递归神经网络层捕获长期模式.然而,这导致了输出不敏感于输入比例的问题.为解决这一问

题, LSTNet 引入了全连接层, 模拟自回归过程, 使输出可以更好地适应输入的尺度变化, 提高了模型的鲁棒性.

1.3.1 卷积层

LSTNet 卷积层是一个没有池化层的卷积神经网络. 该层用于捕获电力负荷数据的短期模式和局部依赖关系. 其包括多个宽度为 ω 和高度为 n 的滤波器. 第 k 个滤波器的公式为:

$$h_k = \text{ReLU}(\omega_k * X + b_k) \quad (12)$$

其中, 卷积运算由 $*$ 表示; h_k 表示输出向量; b_k 为偏移项; 激活函数为 $\text{ReLU}(x) = \max(0, x)$; X 表示输入特征向量.

1.3.2 递归与递归跳跃层

递归层的输出同时输入递归层和递归跳跃层. 递归层部分采用 LSTM 神经网络, 其核心部分是单元状态和门, 作为历史信息的过滤器, 同时可以提供更大的适应性. 其计算公式为:

$$f_t = \sigma(W_f h_{t-1} + U_f x_t + b_f) \quad (13)$$

$$i_t = \sigma(W_i h_{t-1} + U_i x_t + b_i) \quad (14)$$

$$\tilde{c}_t = \text{ReLU}(W_c h_{t-1} + U_c x_t + b_c) \quad (15)$$

$$o_t = \sigma(W_o h_{t-1} + U_o x_t + b_o) \quad (16)$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \tilde{c}_t \quad (17)$$

$$h_t = o_t \otimes \text{ReLU}(c_t) \quad (18)$$

其中, i_t 表示输入门的值, f_t 表示遗忘门的值, o_t 表示输出门的值, c_t 表示存储器单元的值, 以及 b_i 、 b_f 、 b_o 和 b_c 表示各自相应的偏置. 在计算过程中利用权重矩阵 W 和 U ; σ 是 Sigmoid 函数, ReLU 是激活函数, 并且 \otimes 表示元素乘积. h_t 表示输出向量.

为解决 LSTM 在长时间序列上的梯度消失问题, 引入了递归跳跃层, 以增强长期模式的捕获. 递归跳跃层的计算公式为:

$$r_t = \sigma(x_t W_{xr} + h_{t-p} W_{hr} + b_r) \quad (19)$$

$$u_t = \sigma(x_t W_{xu} + h_{t-p} W_{hu} + b_u) \quad (20)$$

$$o_t = \sigma(x_t W_{xo} + h_{t-p} W_{ho} + b_o) \quad (21)$$

$$\tilde{c}_t = \text{ReLU}(x_t W_{xc} + h_{t-p} W_{hc} + b_c) \quad (22)$$

$$c_t = u_t \otimes c_{t-p} + r_t \otimes \tilde{c}_t \quad (23)$$

$$h_t = o_t \otimes \text{ReLU}(c_t) \quad (24)$$

递归跳跃层的输入等于卷积层的输出, 其中, p 表示跳跃的隐藏层状态的数量; h_t^R 表示递归层在 t 时刻的输出; $\{h_{t-p+1}^S, h_{t-p+2}^S, \dots, h_t^S\}$ 表示递归跳跃层在 $t-p+1$ 时刻到 t 时刻的输出.

1.3.3 全连接层

利用全连接层融合递归层和递归跳跃层的输出作为非线性部分的预测结果, 其计算公式为:

$$h_t^D = W^R h_t^R + \sum_{i=0}^{p-1} W_i^S h_{t-i}^S + b \quad (25)$$

其中, h_t^D 表示时间 t 的非线性部分的预测结果.

1.3.4 自回归模型

虽然卷积模块和递归跳跃模块都具有非线性特性, 但其对输入数据的敏感性有限, 会降低预测精度. 因此, LSTNet 模型通过添加自回归 (AR) 模型来解决这个问题. 其计算公式为:

$$h_{t,i}^L = \sum_{k=0}^{q^{\text{ar}}-1} W_k^{\text{ar}} y_{t-k,i} + b^{\text{ar}} \quad (26)$$

其中, AR 模型的预测结果由 $h_{t,i}^L$ 表示, AR 模型的系数是 W^{ar} 和 b^{ar} , 并且 q^{ar} 表示输入窗口的大小.

为了获得 LSTNet 的最终预测结果, 神经网络部分的输出和 AR 模型的结果进行叠加:

$$\hat{Y}_t = h_t^D + h_t^L \quad (27)$$

其中, \hat{Y}_t 表示在 t 时刻的最终预测结果.

1.4 卷积块注意力模块 (CBAM)

CBAM^[22] 具有空间注意力模块和通道注意力模块, 可以捕捉关键特征, 忽略无用特征. 在卷积神经网络生成特征图之后, CBAM 模块通过在通道和空间维度上计算特征图的权重图; 然后, 将其将这个权重图与输入的特征图相乘, 从而实现了自适应学习. 这一轻量级通用模块可以无缝集成到多种卷积神经网络中, 同时支持端到端的训练. 给定特征图 F , 通道注意力模块能推导出 1D 通道注意力向量 M_c , 并且空间注意力模型则能推导出三维空间注意力向量 M_s , 代表了各种注意力的意义位置. 其计算公式为:

$$F' = M_c(F) \otimes F \quad (28)$$

$$F'' = M_s(F') \otimes F' \quad (29)$$

其中, \otimes 表示逐元素乘法. 通道注意力模块关注输入数据中有意义的内容. CBAM 有两个池: 最大池化和平均

池化. 池化可以提取高层特征, 不同的池化方式意味着提取的高层特征更加丰富. 其计算公式为:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (30)$$

空间注意力模块主要关注的是信息中的哪个位置是 有意义的, 也是对通道注意力模块的补充. 其计算公式为:

$$M_s(F) = \sigma(Conv([AvgPool(F); MaxPool(F)])) \quad (31)$$

其中, $\sigma(\cdot)$ 表示 Sigmoid 函数, $MaxPool(\cdot)$ 表示最大池化, $AvgPool(\cdot)$ 表示平均池化, $MLP(\cdot)$ 表示多层感知器 (MLP), $Conv(\cdot)$ 表示三维卷积层.

2 基于 RIME-VMD 与 BLSTNet-CBAM 的预测模型

2.1 基于 RIME 算法优化的 VMD 分解

由于电力负荷数据具有非平稳性, 而且 VMD 分解需要预先定义模态数量 K 和惩罚参数 α , 这可能会使得之后的预测结果有偏差. 为了避免其问题, 采用 RIME 算法优化 VMD 的模态数量参数 K 和惩罚参数 α , 并且适应度函数采用最小包络熵来完成整个优化. 通过 RIME 优化 VMD 的参数, 可以更好地处理电力负荷数据的非平稳性, 有助于提高分解结果的精度, 从而更好地进行预测. RIME 算法优化 VMD 的流程图如图 1 所示.

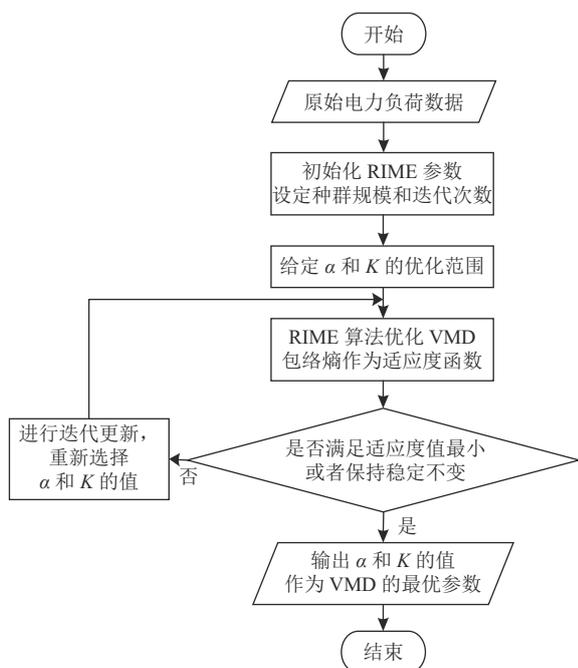


图 1 RIME 算法优化 VMD 流程图

2.2 改进 LSTNet 模型

由于递归层部分采用单一的 LSTM 神经网络, 只考虑对过去的上下文信息进行数据挖掘, 而忽视了未来的上下文信息内容, 因此, 采用双向循环神经网络 (BiLSTM) 来作为递归层的基本单元. 选择 BiLSTM 可以提升模型的建模, 能够基于 LSTM 执行正向和反向双向传播, 从而更有效地捕捉电力负荷数据中的复杂关系. 其计算公式为:

$$h_t = LSTM(x_t, h_{t-1}) \quad (32)$$

$$h_i = LSTM(x_i, h_{i-1}) \quad (33)$$

$$H_t = a_t h_t + b_t h_i + c_t \quad (34)$$

其中, LSTM 表示当方程中的 $p = 1$ 时的模型, 计算公式如式 (14)–式 (19). 正向传播 LSTM 和反向传播 LSTM 计算公式一致, h_t 为正向传播隐藏层状态; h_i 为反向传播隐藏层状态; H_t 为 BiLSTM 的输出结果.

为了有效地处理更多电力负荷数据并捕捉长期依赖关系, 单一的 LSTNet 框架表现出训练速度较慢, 进而影响预测准确性. 因此, 引入卷积块注意力机制模块 (CBAM), 将其置于递归层之后. CBAM 采用三重注意力机制, 有助于更好地交互通道信息和空间信息, 从而更好地捕捉时间序列数据的复杂特征和关系. 这一改进有助于提高模型在短期电力负荷预测中的准确性和鲁棒性. 其改进的 LSTNet 模型框架图如图 2 所示.

2.3 RIME-VMD-BLSTNet-CBAM 模型预测流程

本文旨在提高短期负荷预测的准确性, 结合了经过优化的 VMD 和改进的 LSTNet 模型. 首先, 对原始数据进行预处理, 包括负荷数据和相关特征变量. 利用 VMD 将预处理后的负荷数据分解成多个子序列, 但由于 VMD 分解需要事先定义模态数量 K 和惩罚参数 α , 这可能引入偏差. 因此, 采用 RIME 优化算法对模态数量 K 和惩罚参数 α 进行调优. 鉴于负荷数据的波动性和复杂性, 使用包络熵方法对分解后的多个子序列进行计算和分析, 以提高模型的预测准确度. 相关特征变量经过相关性分析 (MIC) 来反映其之间的关联程度, 以选择更具意义的特征进行进一步分析和建模. 然后, 对处理后的数据进行统一归一化处理并分割为数据集. 使用改进的 LSTNet 模型对每个子序列进行单独预测, 随后采用加权平均方法将预测结果进行融合, 最终输出预测结果. RIME-VMD-BLSTNet-CBAM 模型的流程如图 3 所示.

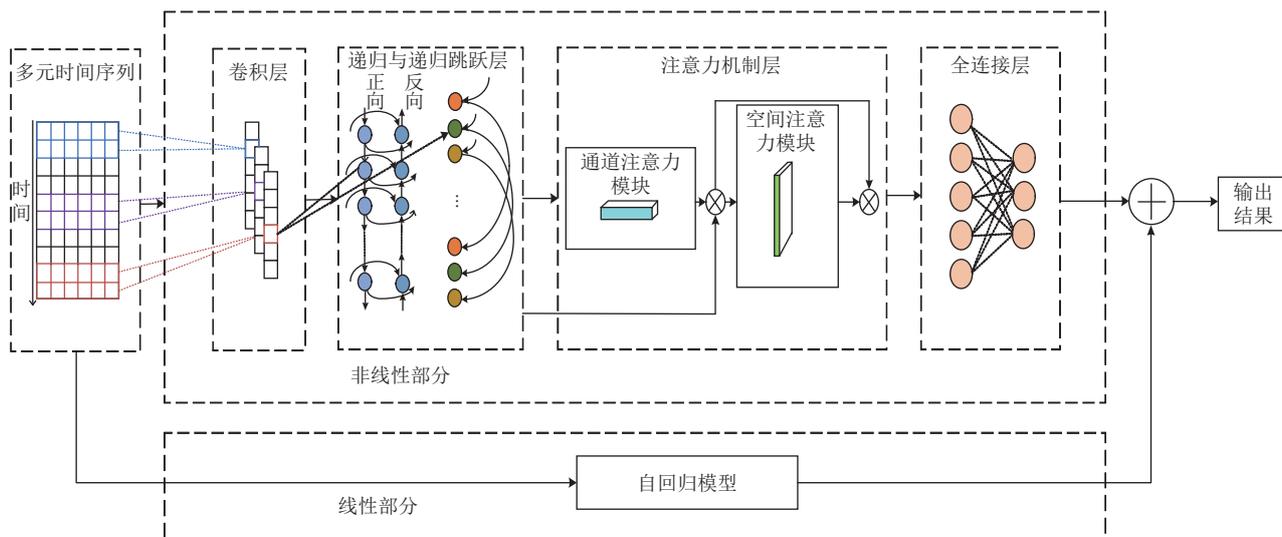


图2 改进的LSTNet模型框架图

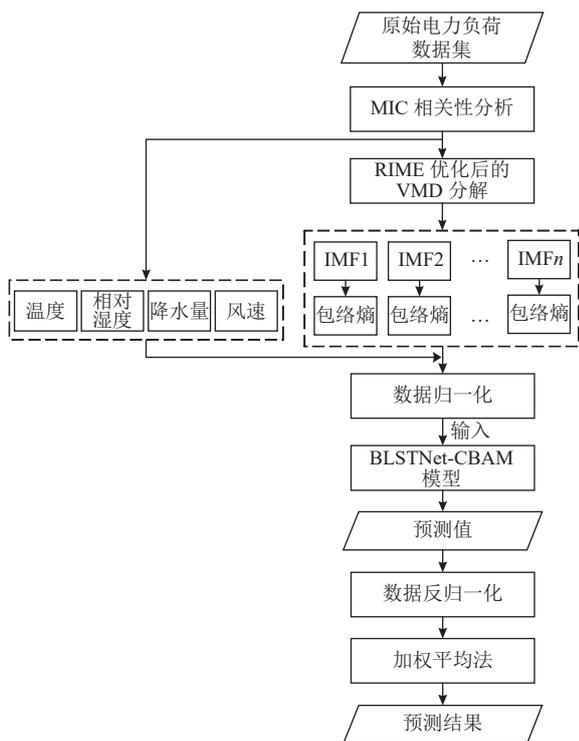


图3 RIME-VMD-BLSTNet-CBAM组合模型流程

3 算例及结果分析

3.1 实验数据选取及数据预处理

本文选取了Kaggle平台上发布的巴拿马在2019年5月1日-2020年5月7日的电力负荷数据集作为实验对象。该数据集涵盖5个特征向量，这些特征的引入有助于更准确地预测电力负荷的变化趋势。在实验过程中，以每1小时为一个采样点，并采用滑动窗口方

法进行负荷预测。具体而言，将滑动窗口的大小设置为24，滑动步长为1，以此来利用过去24个时间点的历史数据进行下一时刻电力负荷的预测。为了全面评估模型的性能，将数据集前8784个样本划分为训练集，其余为测试集，以确保对模型的泛化能力进行有效验证。由于数据集不包含任何缺失值或离群值，同时为了消除不同类型之间的量纲差异，采用归一化方法。归一化是一种数学变换技术，能够将数据压缩至[-1, 1]的范围内，从而使得短期电力负荷预测模型能够更好地适应各种数据类型。具体而言，归一化方法遵循式(35)：

$$x_{i,s}^{norm} = \frac{x_{i,s} - x_s^{min}}{x_s^{max} - x_s^{min}} \quad (35)$$

其中， $x_{i,s}^{norm}$ 表示第*i*个样本的特征向量的第*s*个分量的归一化值， $x_{i,s}$ 是其归一化前的值； x_s^{max} 和 x_s^{min} 分别表示特征向量的第*s*个分量的最大值和最小值。

3.2 电力负荷时间序列模态分解及分析

使用VMD对数据预处理后的电力负荷数据进行分解需要设置多个参数。其中，采用RIME优化算法来确定这些参数的最佳值；采用包络熵作为RIME优化算法的适应度函数。RIME优化VMD参数设置最优值如表1所示。

VMD中 <i>K</i> (模态数目)	VMD中 α (带宽约束)	RIME最佳适应度值
10	2500	8.9146

图4表示预处理后的负荷数据的原始信号，从图4可以看出负荷数据具有很大的波动性，图5表示RIME

优化 VMD 得到的最优迭代次数和最佳适应度值, 图 6 表示 RIME-VMD 的分解结果, 一共分解出 9 个 IMF 分量和 1 个残差量. IMF1 分量具有相对平坦的趋势, 表明其是低频分量, 而其他分量表示高频分量. 这些分量与天气数据、日期数据和其他相关要素相结合, 以创建模型的输入要素. 图 7 表示分解出 10 个分量的频谱图. 从图 7 可以看出每个分量在不同时间尺度上的波动.

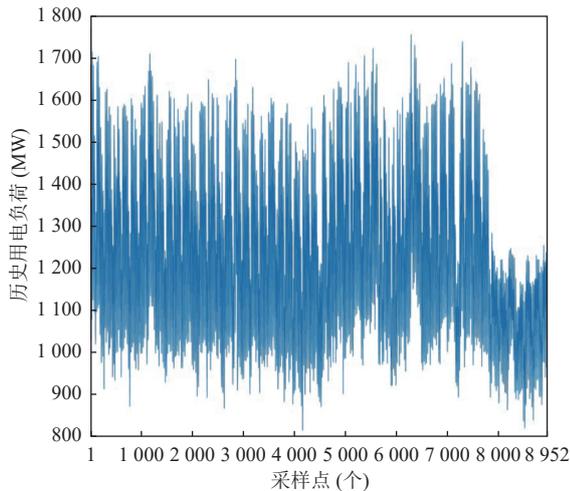


图 4 负荷数据原始信号

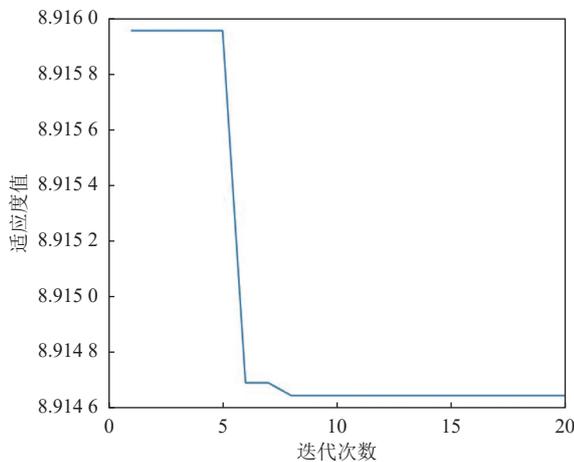


图 5 RIME 优化 VMD 参数的迭代曲线图

3.3 特征选择分析

鉴于各影响因素与电力负荷数据之间存在复杂的非线性关系, 采用 MIC 相关性分析^[23]来检验各变量的显著性. MIC 的值界定在[0, 1]之间. 当 MIC 值越大时, 表示两个变量之间的相关性越强; 反之, 当 MIC 值较小时, 则表明两个变量之间的相关性较弱. 其主要表达式为:

$$M_{MIC}(X, Y) = \max_{xy < B(N)} \{M(D)_{x,y}\} \quad (36)$$

$$M(D)_{x,y} = \frac{I^*(D, x, y)}{\log_2 \min\{x, y\}} \quad (37)$$

其中, $M_{MIC}(X, Y)$ 表示最大信息数, $B(N)$ 表示样本个数函数, 通常 $B(N) = N^{0.6}$, N 表示样本个数 $I^*(D, x, y)$ 为二位变量在数据集 D 下的最大互信息.

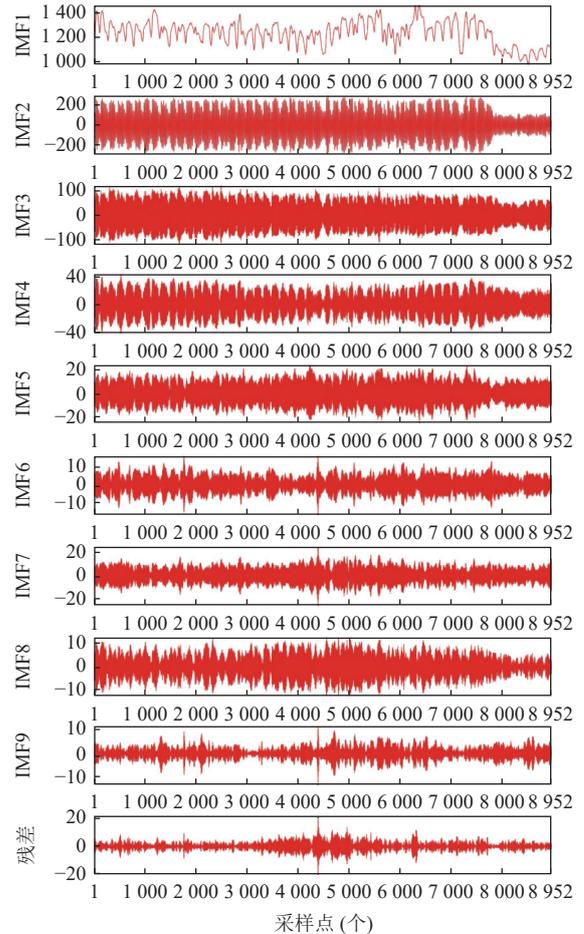


图 6 RIME-VMD 分解结果

图 8 显示了各个影响因素和电力负荷数据之间的相关性. 从图 8 中可以看出, 与历史用电负荷量相关性最强的是温度, 其 MIC 值为 0.38; 而相对湿度、降水量、风速的 MIC 值均为 0.04, 表明温度、相对湿度、降水量、风速与历史用电负荷量之间存在一定的相关性. 其中 nat_demand 代表历史用电负荷量; T2M_roc 代表温度; QV2M_toc 代表相对湿度; TQL_toc 代表降水量; W2M_toc 代表风速.

3.4 实验评估指标

将 RIME-VMD 优化的 BLSTNet-CBAM 的预测

模型与其他模型作对比实验,所有预测模型输入的数据保持一致,确保对比实验的有效性.实验采用多种评估指标对模型的预测性能进行全面评估,这些指标包括平均绝对误差(MAE)、平均百分比误差(MAPE)和均方根误差(RMSE).其中,较低的误差值代表着预测值与实际值之间的更大接近度,这也就意味着模型具有较高的准确度.这些评估指标为模型的性能提供了全面的评估,从而有助于更好地理解和改进模型的性能.其计算公式为:

$$e_{MAE} = \frac{1}{n} \sum_{t=1}^n |\bar{y}_t - \hat{y}_t| \quad (38)$$

$$e_{MAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|\bar{y}_t - \hat{y}_t|}{\bar{y}_t} \times 100\% \quad (39)$$

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (\bar{y}_t - \hat{y}_t)^2} \quad (40)$$

其中, \bar{y}_t 表示真实值, \hat{y}_t 表示预测值, n 表示样本总数.

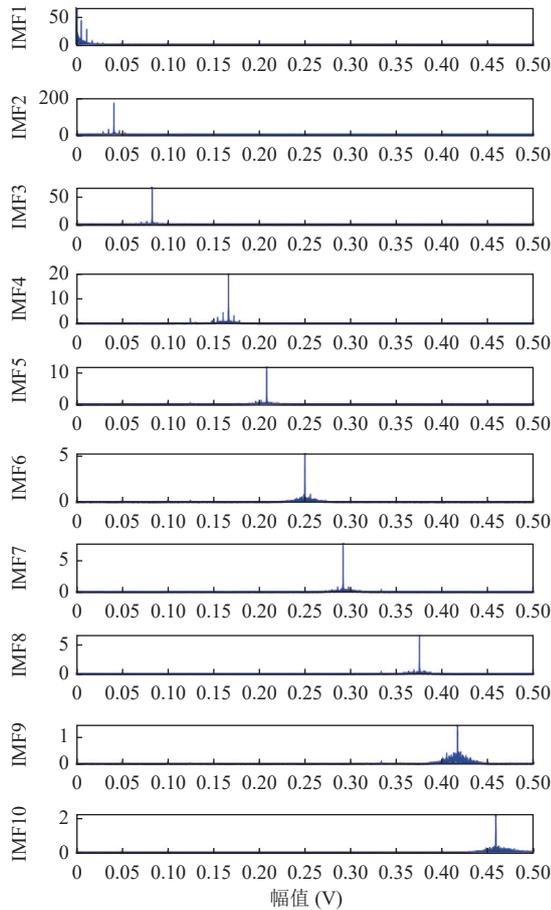


图7 分解出 IMF 子序列的频谱图

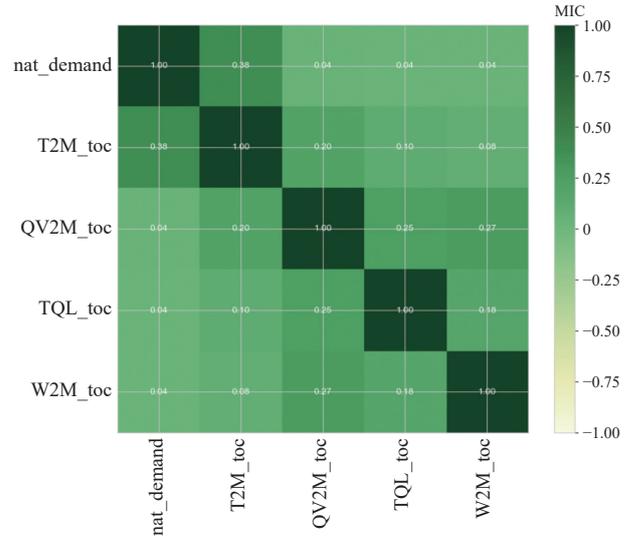


图8 特征相关热图

3.5 实验设置

采用 12th Gen Intel® Core™ i5-12500H 2.50 GHz 处理器和 GPU RTX 3080 服务器作为硬件平台.对于 RIME-VMD 模态分解部分,该部分采用 Matlab 语言编写,而其余部分则使用 Python 3.8 语言编写.整体实验架构基于 PyTorch 框架进行设计与实现.

BLSTNet-CBAM 网络参数包括 64 个神经元的卷积层、64 个神经元的递归层、20 个神经元的递归跳跃层,使用 Adam 算法^[24]的优化以及 64 个训练批量和 100 次迭代.

3.6 基于优化的 VMD-LSTNet 预测结果及对比分析

3.6.1 RIME-VMD 分解对模型预测精度的影响分析

为了评估本文提出的模型的有效性,将 RIME-VMD 优化 3 个单一模型: LSTM、GRU 和 LSTNet 中,将预测模型优化前和优化后的预测性能进行比较.如图 9 和表 2 所示.

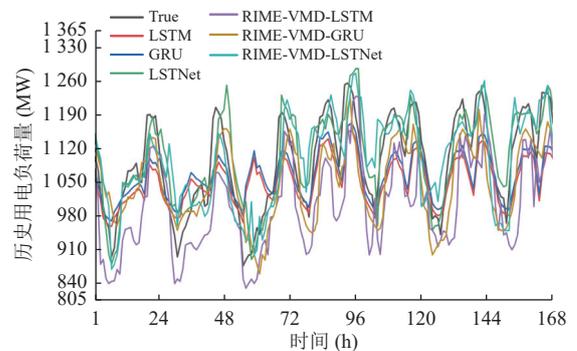


图9 RIME-VMD 优化前后模型预测的比较

表2 RIME-VMD 分解方法优化前后模型评价指标对比

Models	MAE	MAPE	RMSE
LSTM	57.746	0.053	70.767
GRU	62.571	0.057	75.950
LSTNet	32.481	0.030	44.778
RIME-VMD-LSTM	42.357	0.032	53.530
RIME-VMD-GRU	29.528	0.024	37.587
RIME-VMD-LSTNet	28.877	0.022	35.558

从图9中可以看出, LSTM、GRU、LSTNet 在采用 RIME-VMD 优化后, 其预测产生了更好的结果, 表明 RIME-VMD 的分解方法有助于提高模型的预测精度. 同时表2显示在已优化的模型中, RIME-VMD-LSTNet 表现最好, MAE 为 28.877, MAPE 为 0.022, RMSE 为 35.558, 结果表明通过使用 RIME 优化的 VMD 分解负荷序列获得的输入特征提高了模型的预测精度, 同时适应于不同的预测模型.

3.6.2 BLSTNet-CBAM 预测模型比较分析

对于短期电力负荷预测, 本文提出的 BLSTNet-CBAM 模型与其他模型的评估指标如表3所示, 预测对比结果如图10所示.

表3 所提出的模型和其他模型评估指标

Models	MAE	MAPE	RMSE
RIME-VMD-LSTM	42.357	0.032	53.530
RIME-VMD-GRU	29.528	0.024	37.587
RIME-VMD-LSTNet	28.877	0.022	35.558
RIME-VMD-BLSTNet-CBAM	27.131	0.018	35.040

图10显示了短期电力负荷预测中 BLSTNet-CBAM 和其他模型的预测结果曲线. 从图中可以明显看出, BLSTNet 恰当地捕捉了电力负荷数据的非线性和非平稳性, 从而比 LSTM、GRU 和 LSTNet 具有更高的预测精度. 此外还表明了 CBAM 注意力机制可以提高模型的预测精度. 表3显示了 BLSTNet-CBAM 和其他模型的评估结果. 从表中可以清楚地看出, BLSTNet-CBAM 模型均优于其他模型, 与 LSTNet 相比 MAE、MAPE、RMSE 分别降低了 1.746、0.004、0.518. 从而证明本文提出的模型更有效地提取关键特征, 与现有模型相比具有更高的预测精度.

3.6.3 消融实验

为验证所提出的预测模型 RIME-VMD-BLSTNet-CBAM 中各部分的有效性, 进行了一系列的消融实验, 其目的是分析模型中不同组件对预测性能的影响.

首先建立了一个基准模型, 即不包含任何消融组件的 LSTNet 模型. 接着, 逐步引入 VMD 模态分解、

RIME 优化算法优化 VMD、将 LSTNet 的递归层 LSTM 替换为 BiLSTM 层以及 CBAM 注意力机制等组件. 通过逐步添加这些组件, 评估其对模型性能的贡献, 进而有助于揭示各个组件的作用和影响, 因此更全面地理解模型的性能改进. 消融实验结果如表4所示.

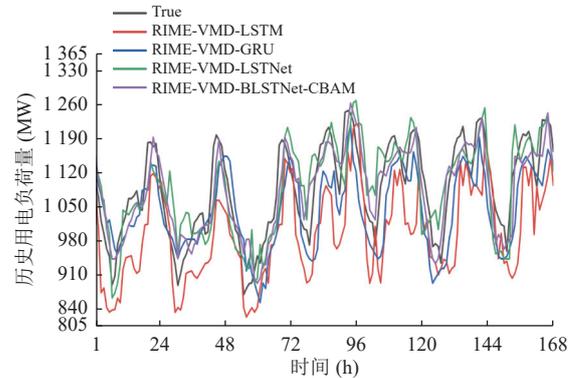


图10 模型预测曲线对比图

表4 消融实验评估误差指标

Models	MAE	MAPE	RMSE
LSTNet	32.481	0.030	44.778
VMD-LSTNet	30.526	0.029	40.437
RIME-VMD-LSTNet	28.877	0.022	35.558
RIME-VMD-BLSTNet	27.658	0.020	35.247
RIME-VMD-BLSTNet-CBAM	27.131	0.018	35.040

通过消融实验的结果表明, 引入 VMD 模态分解和 RIME 优化算法显著提高了模型的预测性能. VMD 模态分解将输入时间序列数据分解成不同频率的子序列, 使得每个子序列在相对较窄的频率范围内更易于建模, 其中与单一的 LSTNet 相比, MAE、MAPE、RMSE 分别降低了 1.955、0.001、4.341. 通过使用 RIME 优化算法对 VMD 分解进行优化, 与直接采用 VMD 分解相比, MAE、MAPE、RMSE 分别降低了 1.649、0.007、4.879. 结果说明 RIME-VMD 分解更好地提取了有用的子序列, 并减少误差噪声的影响, 从而改善了预测准确性.

将 LSTNet 的递归层 LSTM 替换为 BiLSTM 也在一定程度上改善了模型的性能. 相比融合 RIME-VMD 但未改进的 LSTNet 模型, MAE、MAPE、RMSE 分别降低了 1.219、0.002、0.311. 实验结果证明 BiLSTM 具有更好的建模能力, 能够更好地捕捉时间序列数据中的长期依赖关系. 此外, 引入 CBAM 注意力机制进一步提升了模型的性能, 其通过对时间序列数据的通道和空间特征进行有针对性的注意力操作, 能够有效

捕捉关键特征信息. 与未引入 CBAM 注意力机制模型相比, MAE、MAPE、RMSE 分别降低了 0.527、0.002、0.207. 从而证明 CBAM 注意力机制有助于模型更好地聚焦于重要信息, 从而提高短期负荷预测性能.

综上所述, 消融实验的结果验证了 RIME-VMD-BLSTNet-CBAM 模型中各部分的有效性. VMD 模态分解和 RIME 优化、使用 BiLSTM 替换 LSTM 的改进以及嵌入 CBAM 注意力机制, 都对模型的预测性能产生了积极影响, 为短期电力负荷预测任务提供了一种有效的建模方法.

4 结论

为了解决现代电力负荷特征提取和提高预测精度的难题, 本文提出了一种基于 RIME 优化的 VMD 分解方法, 将负荷序列分解作为关键因素, 并结合 BLSTNet 预测模型和嵌入 CBAM 注意力机制, 构建了 RIME-VMD-BLSTNet-CBAM 短期电力负荷预测框架. 为了验证 RIME-VMD 和 BLSTNet-CBAM 算法的有效性, 给出了算法实例, 并得出以下结论.

(1) 通过 RIME 优化后的 VMD 分解方法对电力负荷数据进行分解, 克服了 VMD 分解方法需要人工调整参数的需要. 采用包络熵作为适应度函数, 以最小的时间复杂度分解负荷分量, 从而促进预测模型的学习和特征提取. 所提出的方法使模型的精度得到了提高, 从而验证了计算分析的有效性.

(2) 算例验证了 BLSTNet-CBAM 模型具有高预测精度, 这得益于 RIME 优化的 VMD 负荷序列构造, BLSTNet 对负荷特征的高效学习和准确预测能力, 以及 CBAM 注意力机制对输入序列权重的分配. 通过将两者紧密结合, 解决了电力负荷特征提取的难题, 提高了模型的预测精度, 并最终获得了更高精度的预测结果.

参考文献

- 1 Yildiz B, Bilbao JI, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, 2017, 73: 1104–1122. [doi: 10.1016/j.rser.2017.02.023]
- 2 张大海, 孙锴, 和敬涵. 基于相似日与多模型融合的短期负荷预测. *电网技术*, 2023, 47(5): 1961–1969. [doi: 10.13335/j.1000-3673.pst.2022.1110]
- 3 万昆, 柳瑞禹. 区间时间序列向量自回归模型在短期电力负荷预测中的应用. *电网技术*, 2012, 36(11): 77–81. [doi: 10.13335/j.1000-3673.pst.2012.11.044]
- 4 Alberg D, Last M. Short-term load forecasting in smart meters with sliding window-based ARIMA algorithms. *Vietnam Journal of Computer Science*, 2018, 5(3–4): 241–249. [doi: 10.1007/s40595-018-0119-7]
- 5 杨国华, 郑豪丰, 张鸿皓, 等. 基于 Holt-Winters 指数平滑和时间卷积网络的短期负荷预测. *电力系统自动化*, 2022, 46(6): 73–82. [doi: 10.7500/AEPS20210409003]
- 6 陈培垠, 方彦军. 基于卡尔曼滤波预测节假日逐点增长率的电力系统短期负荷预测. *武汉大学学报(工学版)*, 2020, 53(2): 139–144. [doi: 10.14188/j.1671-8844.2020-02-007]
- 7 万强, 王清亮, 王睿豪, 等. 基于支持向量机的某地区电网短期电力负荷预测. *电网与清洁能源*, 2016, 32(12): 14–20. [doi: 10.3969/j.issn.1674-3814.2016.12.003]
- 8 孙英云, 何光宇, 翟海青, 等. 一种基于决策树技术的短期负荷预测算法. *电工电能新技术*, 2004, 23(3): 55–58, 75. [doi: 10.3969/j.issn.1003-3076.2004.03.013]
- 9 李焱, 贾雅君, 李磊, 等. 基于随机森林算法的短期电力负荷预测. *电力系统保护与控制*, 2020, 48(21): 117–124. [doi: 10.19783/j.cnki.pspc.191594]
- 10 李鹏, 何帅, 韩鹏飞, 等. 基于长短期记忆的实时电价条件下智能电网短期负荷预测. *电网技术*, 2018, 42(12): 4045–4052. [doi: 10.13335/j.1000-3673.pst.2018.0433]
- 11 王增平, 赵兵, 纪维佳, 等. 基于 GRU-NN 模型的短期负荷预测方法. *电力系统自动化*, 2019, 43(5): 53–58. [doi: 10.7500/AEPS20180620003]
- 12 吕志星, 张斌, 王沈征, 等. 基于 K-Means 和 CNN 的用户短期电力负荷预测. *计算机系统应用*, 2020, 29(3): 161–166. [doi: 10.15888/j.cnki.csa.007287]
- 13 Lai GK, Chang WC, Yang YM, et al. Modeling long-and short-term temporal patterns with deep neural networks. *Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*. Ann Arbor: Association for Computing Machinery, 2018. 95–104. [doi: 10.1145/3209978.3210006]
- 14 朱凌建, 荀子涵, 王裕鑫, 等. 基于 CNN-BiLSTM 的短期电力负荷预测. *电网技术*, 2021, 45(11): 4532–4539. [doi: 10.13335/j.1000-3673.pst.2021.0470]
- 15 叶淳铮, 常鲜戎, 顾为国. 基于小波变换和支持向量机的电力系统短期负荷预测. *电力系统保护与控制*, 2009, 37(14): 41–45. [doi: 10.7667/j.issn.1674-3415.2009.14.009]
- 16 赵志强, 陆风怡. 基于小波分解和 ID3 算法的短期日负荷最大值预测方法. *南方电网技术*, 2010, 4(6): 98–102. [doi: 10.13648/j.cnki.issn1674-0629.2010.06.018]

- 17 张平, 潘学萍, 薛文超. 基于小波分解模糊灰色聚类 and BP 神经网络的短期负荷预测. 电力自动化设备, 2012, 32(11): 121–125, 141.
- 18 王玉龙, 崔玉, 李鹏, 等. 基于小波分析改进的神经网络模型电力系统负荷预测. 电网与清洁能源, 2015, 31(2): 16–20, 27. [doi: [10.3969/j.issn.1674-3814.2015.02.003](https://doi.org/10.3969/j.issn.1674-3814.2015.02.003)]
- 19 刘建华, 李锦程, 杨龙月, 等. 基于 EMD-SLSTM 的家庭短期负荷预测. 电力系统保护与控制, 2019, 47(6): 40–47. [doi: [10.7667/PSPC180365](https://doi.org/10.7667/PSPC180365)]
- 20 杨胡萍, 余阳, 汪超, 等. 基于 VMD-CNN-BIGRU 的电力系统短期负荷预测. 中国电力, 2022, 55(10): 71–76.
- 21 Su H, Zhao D, Heidari AA, *et al.* RIME: A physics-based optimization. Neurocomputing, 2023, 532: 183–214. [doi: [10.1016/J.NEUCOM.2023.02.010](https://doi.org/10.1016/J.NEUCOM.2023.02.010)]
- 22 Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 3–19.
- 23 Reshef DN, Reshef YA, Finucane HK, *et al.* Detecting novel associations in large data sets. Science, 2011, 334(6062): 1518–1524. [doi: [10.1126/science.1205438](https://doi.org/10.1126/science.1205438)]
- 24 Kingma DP, Ba J. Adam: A method for stochastic optimization. Proceedings of the 3rd International Conference on Learning Representations. San Diego: ICLR, 2015.

(校对责编: 牛欣悦)