

# 基于 VMD-改进最优加权法的短期负荷变权 组合预测策略\*

李志军<sup>1</sup> 徐博<sup>2</sup> 杨金荣<sup>3</sup> 宁阮浩<sup>2</sup>

(1. 河北工业大学省部共建电工装备可靠性与智能化国家重点实验室 天津 300130;

2. 河北工业大学人工智能与数据科学学院 天津 300130; 3. 河北工业大学电气工程学院 天津 300130)

**摘要:** 为提升短期电力负荷预测精度,提出了一种变权组合预测策略。首先,为了降低负荷数据的不平稳度,使用变分模式分解(variational mode decomposition, VMD)将负荷数据分解成了高频、低频、残差3种特征模式分量。其次,充分计及负荷数据的时序特点,参考指数加权法原理设计自适应误差重要性量化函数,并结合组合模型在时间窗口内的历史负荷数据的均方预测误差设计改进最优加权法的目标函数和约束条件,以完成子模型的准确变权。最后,针对波动较强的高频分量选定极端梯度提升(XGBoost)和卷积神经网络-长短期记忆(CNN-LSTM)模型并使用改进最优加权法进行组合预测,低频分量使用多元线性回归(MLR)模型预测、残差分量使用LSTM模型预测,叠加各模式分量的预测结果,实现了短期负荷数据的准确预测。实验结果表明,使用策略组合模型的平均绝对百分比误差为4.18%。与使用传统组合策略的组合模型相比,平均绝对百分比预测误差平均降低了0.87%。

**关键词:** 短期负荷预测;变分模式分解;改进最优加权法;组合模型

**中图分类号:** TM715 **文献标识码:** A **国家标准学科分类代码:** 510.99

## Short-term load variable weighted combination prediction strategy based on VMD-improved optimal weighting method

Li Zhijun<sup>1</sup> Xu Bo<sup>2</sup> Yang Jinrong<sup>3</sup> Ning Ruanhao<sup>2</sup>

(1. State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300130, China; 2. School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China; 3. School of Electrical Engineering, Hebei University of Technology, Tianjin 300130, China)

**Abstract:** To increase short-term power load forecasting accuracy, this paper proposes a weighted combination prediction strategy. Firstly, in order to reduce the instability of load data, the variational mode decomposition (VMD) is used to decompose the load data into three feature mode components: high-frequency, low-frequency, and residual. Secondly, considering the temporal characteristics of the load data, an adaptive error importance quantification function is designed based on the principle of exponential weighting, and the objective function and constraint conditions of the improved optimal weighting method are designed based on the mean-square prediction error of the historical load data within the time window, in order to achieve accurate weight variation of the submodels. Finally, XGBoost and CNN-LSTM models are selected for the high-frequency components with strong fluctuations, and the improved optimal weighting method is used for combination prediction. The MLR model is used to predict the low-frequency components, and the LSTM model is used to predict the residual components. By superimposing the prediction results of each mode component, accurate prediction of short-term load data is achieved. The experimental results show that the average absolute percentage error of the combined model using this strategy is 4.18%. Compared to the combined model using existing combination strategies, the average absolute percentage prediction error is reduced by 0.87%.

**Keywords:** short-term load forecasting; VMD; improved optimal weighting method; combined model

收稿日期:2023-08-23

\* 基金项目:河北省科技支撑计划(15212105D)项目资助

## 0 引言

“碳达峰”、“碳中和”战略目标的提出使得我国新能源高比例并网的进程不断加快,这意味着电力系统需要消纳更多具有波动性、间歇性的新能源<sup>[1]</sup>。电力负荷预测结果是制定发电、输电和控制系统运行策略的重要参考因素<sup>[2]</sup>。准确地进行短期电力负荷预测,将为发电机组高效利用、新能源有效消纳、电力系统安全稳定运行提供重要支撑<sup>[3]</sup>。

常用的短期电力负荷预测方法主要有统计学和机器学习方法,其中机器学习方法近年来获得了较广泛的应用<sup>[4]</sup>,尤其在深度学习领域发展迅猛。常用的单一模型如长短期记忆网络(long short-term memory network, LSTM)模型<sup>[5]</sup>、卷积神经网络(convolutional neural networks, CNN)模型<sup>[6]</sup>等,在特定场景下均取得了较高的预测精度。但是,受限于算法理论侧重点的不全面性,仅采用单一模型预测易陷入局部最优解,因此其适用的场景较为局限<sup>[7-8]</sup>。

为解决该问题,基于组合模型进行电力负荷预测成为了近期研究的热点<sup>[9-10]</sup>。组合模型的关键在于对子模型权重的分配,包括离线的固定权重分配和在线的变权重分配。文献[11]将极端梯度提升(extreme gradient boosting, XGBoost)模型与LSTM模型离线组合,采用误差倒数法得出固定的子模型权重进行超短期电力负荷的组合预测,取得了良好的效果,但离线组合模型由于固定子模型权重使得组合模型适应能力弱化。因此,能够将子模型在线变权重的组合模型受到研究者的青睐<sup>[12]</sup>。最优加权法是一种经典且较稳定的依赖于时间窗口内历史数据预测误差实现子模型在线变权的方法,其被广泛应用于负荷及其他时序数据的组合预测中<sup>[13]</sup>。文献[14]采用最优加权法进行短期负荷组合预测,精度大幅优于单一子模型。然而,最优加权法将时间窗口内不同位置的预测误差数据对子模型变权的影响视为等同,违背了越旧的负荷数据重要性越低的时序特点。文献[15]基于最优加权法提出了一种基于误差在线更新的组合模型,借鉴AdaBoost调整样本权重的思想,引入一个权重衰减因子使得组合模型在预测误差较大时更为注重最新的负荷数据。文献[16]采用强化学习算法更新子模型权重,设计智能体的奖惩机制时同样忽视了时间窗口内不同位置负荷数据重要性不同的时序特点。文献[17]采用最优加权法并使用改进的遗传算法求解子模型的最优权重,仅仅提升了子模型权重的求解速度,并未对最优加权法加以改进,因此不能准确的获取子模型的权重。

综上所述,传统的电力负荷组合预测模型实现子模型在线变权更新时往往未考虑负荷数据的时序特点,故无法准确地得出子模型权重,干扰了短期负荷预测精度。因此,本文提出了一种基于变分模态分解(VMD)-改进最优加权法的短期负荷变权组合预测策略。首先,将负荷序列

进行VMD分解,以降低负荷数据的非平稳度;参考指数加权法原理设计符合负荷数据特点的自适应误差重要性量化函数,并结合窗口内历史数据的均方预测误差设计改进最优加权法的目标函数和约束条件,以实现子模型的准确变权。使用求解得出的子模型权重矩阵组合XGBoost和卷积长短期记忆网络(CNN-LSTM)模型完成针对高频分量的组合预测。针对低频分量使用多元线性回归(multiple linear regression, MLR)模型预测、残差分量使用LSTM模型预测,叠加各模态分量的预测值得出最终的预测结果。算例分析结果表明,采用本文策略的组合模型的预测精度及效果优于任意单一子模型和采用传统组合策略的组合模型。

## 1 传统组合策略原理及其局限性分析

与单一模型相比较,组合模型可在各子模型之间取长补短,有效提升单一子模型的泛化能力。组合模型的建立通常包含子模型选定和加权求和两个过程。文献[18]指出,选定的子模型不仅需要自身预测精度高,而且不同子模型之间应存在一定的差异性,其本质是在不同的数据空间角度和数据结构角度来观测数据,可有效提升组合模型的泛化性能。常用于负荷预测的组合模型传统加权方法主要有误差倒数法和最优加权法。其中误差倒数法通过量子化模型预测误差得出并固定其权重;最优加权法则使用优化算法求解含有约束的目标函数实现子模型的变权更新。为便于理解本文技术方案,首先对这两种传统加权方法加以简单地分析和讨论,为改进设计提供技术思路。

误差倒数法根据各子模型预测误差及全部子模型预测累计误差获得子模型权重,为便于理解,将子模型按照其预测误差大小进行排序编号,本文以降序排列为例,由式(1)可得出相应子模型的权重,子模型的总数为 $q$ 。

$$\omega_i = \frac{\varepsilon_{q-i+1}}{\sum_{j=1}^q \varepsilon_j} \quad (1)$$

式中: $\omega$ 为子模型的权重; $\varepsilon$ 为预测误差。

组合模型加权求和方法如式(2)所示,将由式(1)获得的子模型权重代入式(2),即可完成各子模型的加权组合并得出基于误差倒数法的组合预测结果。

$$\hat{y} = \omega_1 \cdot M_1 + \omega_2 \cdot M_2 + \dots + \omega_q \cdot M_q \quad (2)$$

式中: $M$ 为子模型预测结果; $\hat{y}$ 为组合模型的预测结果。

由式(1)、(2)可知,采用误差倒数法的组合模型物理意义清晰,算法简单,但由于其仅以子模型预测误差为权重计算依据,子模型的预测误差越大则其在组合预测中的权重越低,权重计算依据较为粗糙且单一。误差倒数组合方法通常采用固定子模型权重模式,无法适应实时变化的运行工况。因此采用误差倒数法的组合模型预测精度、适应能力都受到了限制。

最优加权法可据最新时间窗口内的数据预测误差实现子模型权重的更新,克服了误差倒数法固定权重的

缺点。

假设时间窗口内的负荷数据的样本集合为  $U = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ ,  $x_i$  表示特征值,  $y_i$  表示真实值,  $n$  为时间窗口内样本的总数,  $M$  为子模型预测结果。使用  $q$  个子模型对当前时间窗口  $T$  内的负荷数据进行预测, 预测结果如下:

$$M_T = \begin{bmatrix} M_1(x_1) & \dots & M_q(x_1) \\ \vdots & \ddots & \vdots \\ M_1(x_n) & \dots & M_q(x_n) \end{bmatrix}_{n \times q} \quad (3)$$

由式(3)得到子模型的预测结果后, 对其加权组合计算:

$$\bar{y}_i = \omega_1 M_1(x_i) + \omega_2 M_2(x_i) + \dots + \omega_q M_q(x_i) \quad (4)$$

式中:  $\bar{y}_i$  为组合模型的预测结果。

则组合模型对当前时间窗口内样本点的预测误差计算:

$$e_i = y_i - \bar{y}_i \quad (5)$$

式中:  $e_i$  为相应样本的预测误差。

将全部样本点的预测误差用矩阵进行表示:

$$E = [e_1, e_2, \dots, e_n]^T \quad (6)$$

式(6)可进一步表示为:

$$E = Y_T - M_T \cdot W_T \quad (7)$$

式中:  $Y_T = [y_1, y_2, \dots, y_n]^T$  为时间窗口内所有负荷数据的实际值;  $M_T$  为子模型在时间窗口  $T$  内的预测结果矩阵;  $W_T = [\omega_1, \omega_2, \dots, \omega_q]^T$  为子模型权重矩阵。

为使组合模型的预测误差达到最小, 利用最小二乘法设计待优化的目标函数及约束条件, 如式(8)所示, 其中  $I = [1, \dots, 1]$ 。

$$\begin{cases} \min E^T \cdot E \\ \text{s. t. } I \cdot W_T = 1 \\ \omega_i \geq 0 \end{cases} \quad (8)$$

此时, 子模型变权被转化为了非线性规划问题, 需要借助优化算法进行最新子模型权重求解, 之后通过式(4)得出子模型的变权组合预测结果。同理, 时间窗口滚动到下一个预测周期, 重复上述过程, 即可通过子模型变权完成组合预测。

然而, 采用最优加权法进行负荷数据预测的组合模型存在明显的局限性。图1所示为最优加权法缺陷示意图, 其中红色曲线和蓝色曲线分别为子模型的预测值、实际负荷值图像。结合式(5)~(8)可知, 最优加权法以整个时间窗口内组合模型的预测误差设计待优化的目标函数, 由于负荷数据具有明显的时序性, 历史负荷数据预测误差对变权的影响程度与待预测数据的距离越接近则越高<sup>[15]</sup>。该方法将窗口内不同位置的数据预测误差数据对变权的影响程度视为等同, 违背了负荷数据的时序特点, 难以准确的获得子模型权重。

由以上组合模型的设计原理可知, 子模型权重计算以及负荷数据与待预测值的“距离”都会对组合模型预测性

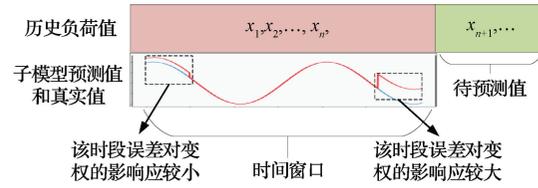


图1 缺陷示意图

能和精度造成影响, 因此, 通过客观评价子模型表现, 同时兼顾负荷数据时序特点进行子模型变权是提升组合模型负荷预测精度和效果的有效途径。

## 2 基于VMD-改进最优加权法的短期负荷变权组合预测策略设计

为进一步提升短期负荷预测精度和效果, 本文提出了一种VMD-改进最优加权法的组合预测策略。该策略流程如图2所示, 数据预处理包括特征提取和数据归一化, 并使用VMD算法将原始负荷序列分解为3个分量。该策略首先针对各模态分量的数据特点选定适宜的模型; 然后, 使用训练集数据训练模型; 之后, 使用验证集数据针对波动性较强的高频分量使用所提改进最优加权法组合XGBoost和CNN-LSTM模型进行预测、低频分量使用MLR模型预测、残差分量使用LSTM模型预测, 叠加各模态分量预测值得出预测结果, 以完成各个模型和改进加权组合策略的参数整定; 最后, 使用测试集数据测试所提策略, 以验证所提策略的有效性。

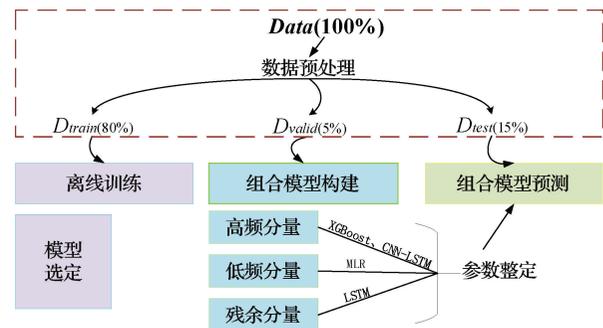


图2 策略流程示意图

### 2.1 VMD分解算法工作原理和效果

VMD是一种非平稳信号自适应分解算法, 其将初始信号分解为  $K$  个本征模态分量 (intrinsic mode portion, IMP)<sup>[19]</sup>, 最终叠加各分量可还原初始信号。

为了降低负荷数据的不平稳度, 本文使用VMD算法将负荷数据分解为了高频、低频、残差3个分量, 并对相应的分量选取适宜的模型分别进行预测, 最终可以叠加各分量预测值得出最终的预测结果。为了直观了解不同模态分量的数据特点, 图3所示为一段原始负荷数据序列及其高频、低频和残差分量的图像。

由图3可知, 低频分量数据较为平滑, 占据了原始负

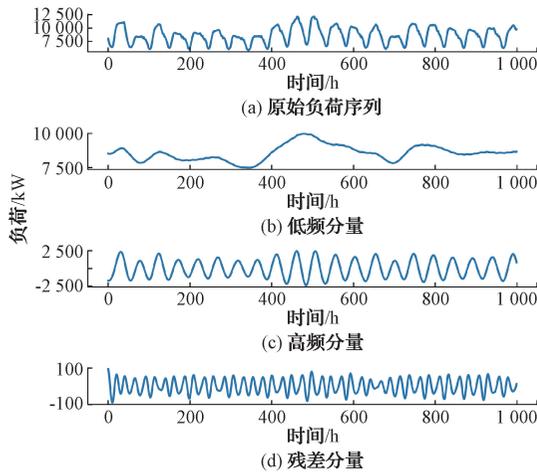


图3 负荷序列分解示意图

荷序列的绝大部分,该部分较容易预测;高频分量数据包含的是波动分量,是负荷序列中重要且又难以预测的部分<sup>[15]</sup>;残差分量数据只占据原始负荷序列较小的比例,其变化频率非常高,具有幅值小和非线性强的特点。

### 2.2 模型选定和组合模型的构建

据 2.1 节所述低频分量数据特点,因高频分量重要且较难预测,为增强单一模型的泛化能力,使用组合模型对其进行预测。为完成组合模型构建,首先要选定子模型,除了要选择精度高的模型外,也要考虑选择理论上具有一定差异性的模型以实现优势互补<sup>[18]</sup>。因此,本文选定在负荷预测中常用且表现较好的 CNN-LSTM、XGBoost 作为构建组合模型的子模型,进行相关分析。

在短期负荷预测中,LSTM 模型使用输入门、遗忘门和输出门来控制输入值、记忆值和输出值,能够有效地捕捉到较长负荷序列的时序联系,同时缓解了梯度消失、爆炸现象。CNN-LSTM 模型在 LSTM 模型的基础上加入了卷积神经网络,能够由 CNN 网络先行提取到负荷数据的深度特征,并输入 LSTM 模型进行预测,可取得较好效果,但 CNN 网络在提取数据的深度特征时也会造成一定信息的丢失。XGBoost 模型属于 Boosting 算法,它通过不断地生成新的决策树模型来拟合上一次的残差,通过不断迭代,精度不断上升,但是它仅凭输入特征预测,无法学习到负荷数据的时序关联。

现有研究中,MLR 模型简单实用且其在预测较平滑数据表现良好<sup>[20]</sup>,选取 MLR 模型对低频分量进行预测;由 2.1 节可知,残差分量非线性特征明显,LSTM 能够发现较长序列的时序关联且能更好地适用于非线性较强的数据的预测之中。因此,对残差分量使用 LSTM 模型进行预测。

### 2.3 改进最优加权法设计

使用组合模型预测最关键的技术环节是变权,2.2 节通过对模型的选定,已经完成组合模型的构建。对所提出

的子模型变权组合方法进行设计。为便于解释,后续以时间窗口  $T$  为例展开分析说明。

为减轻负荷数据“距离”对组合模型短期负荷预测精度的影响,本文参考指数加权法原理,对最优加权法加以改进。如图 4 所示,设计并利用计及负荷数据特点的自适应误差重要性量化函数为组合模型在最新时间窗口内的均方预测误差的重要性进行赋权量化,从而建立计及最新负荷数据特点的目标函数和约束条件,求解建立子模型权重矩阵,并据此将子模型变权,完成负荷组合预测后自适应调整下个时间窗口中量化函数的参数,之后不断重复此过程。

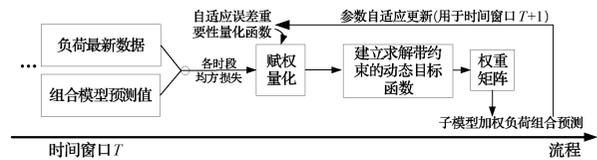


图4 改进最优加权法原理

#### 1) 指数加权法原理

指数加权法是一种经典的时序数据平滑方法,其通过指数函数来分配不同位置时序数据的权重,可以更好的处理负荷数据“距离”的时序特点。该方法基本思想是将每个样本点乘以由指数函数分配的权重,较新样本点的权重较大,而较旧的样本点的权重则较小,然后将所有结果求和,以得出时刻  $t$  待平滑数据的平滑结果。其原理如式(9)和(10)所示,其中超参数  $\tau$  依赖于人为经验设置且属于  $(0,1)$ 。

$$v_t = (1 - \tau) \cdot \theta_t + \tau \cdot v_{t-1} \quad (9)$$

$$v_t = \frac{\theta_t}{1 - \tau^t} \quad (10)$$

式中:  $v_t$  为时刻  $t$  数据的平滑结果;  $\theta_t$  为时刻  $t$  待平滑数据的实际值。

由式(9)和(10)可知,随着  $t$  的增大,  $1 - \tau^t$  逐渐趋近于 1;  $\tau$  设置的越小则较旧数据所占权重越小,即越重视最新的样本点。

#### 2) 自适应误差重要性量化函数设计

为客观地评价量化最新时间窗口内各个位置数据误差对子模型变权的影响程度,将考虑负荷数据的时序“距离”特点,参考指数加权法原理设计了一个计及负荷数据时序特点的指数类型的自适应误差重要性量化函数。如式(11)所示,其中,  $\beta$  是一个在区间  $(0,1)$  自适应变化的浮点型超参数,  $t$  表示时间窗口内的第  $t$  个样本点。该函数得出的值随时间递增且始终为正,考虑了时序负荷数据“距离”对变权的影响。

$$F_{\theta,t}(\beta, t) = \begin{cases} \beta^{n-t} \\ \text{s. t. } 0 < \beta < 1 \end{cases} \quad (11)$$

为便于所提变权方法的说明,将式(11)得出的组合模型的窗口内不同位置的数据预测误差对变权影响的重要

性结果用矩阵  $\beta_T$  表示:

$$\beta_T = [\beta^{n^*}, \beta^{n^*-1}, \dots, \beta^1, \beta^0] \quad (12)$$

由式(12)可知,自适应量化函数的参数  $\beta$  决定了矩阵  $\beta_T$ 。值得注意的是,时间窗口是不断滚动的,负荷数据的特点也在变化,因此不同的时间窗口即便是位于相同位置的均方预测误差对变权的影响程度也在发生变化。因此,本文通过设计式(13)完成参数  $\beta$  的自适应更新,该式基本思想是组合模型在当前预测周期已经完成预测后,通过本预测周期预测误差的反馈来更新参数  $\beta$ 。同时,为防止  $\beta$  的过调节,使用平衡系数  $\lambda$  限制其变化幅度,该系数为一个固定值,其可使用验证集数据进行参数调优得出。

$$\beta' = \begin{cases} \operatorname{argmin}_{\beta'} \frac{1}{N} \cdot (\mathbf{y} - \hat{\mathbf{y}}) \cdot (\mathbf{y} - \hat{\mathbf{y}})^T + \lambda \cdot (\beta' - \beta)^2 \\ \text{s. t. } 0 < \beta < 1 \end{cases} \quad (13)$$

式中:  $\hat{\mathbf{y}} = \mathbf{W}_T \cdot \mathbf{M}_T$ ,  $\mathbf{M}_T$  为子模型在时间窗口  $T$  内的预测结果构成的矩阵,  $\mathbf{W}_T$  为子模型权重矩阵,  $\beta'$  为参数  $\beta$  的更新值;  $\mathbf{y}$  为最新实际负荷数据构成的向量;  $\hat{\mathbf{y}}$  为组合模型的最新预测值构成的向量,  $N$  为预测步长,  $\lambda$  为平衡系数。

本文为最大限度地保证式(13)求解的准确性,在求解时将(0,1)等分为1000份,即  $\beta'$  的取值范围为0.001~1.000。通过穷举法将取不同值的  $\beta'$  代入式(13)计算,以使得式(13)取得最小值,完成该参数的更新。

## 2.4 计及负荷数据特点的改进最优加权短期负荷组合预测方法实现

为实现子模型准确变权,设计改进最优加权法的计及负荷数据时序特点的目标函数和约束条件,如式(14)所示,通过  $\beta_T$  量化了窗口内不同位置均方预测误差对变权的影响程度。参数  $\beta$  的更新使得式(14)构建的目标函数也是动态变化的,可助力组合模型更好的适应最新时间窗口内负荷数据的特点。求解式(14)可得出最新的子模型权重矩阵。

$$\mathbf{W}_T = \begin{cases} \min(\beta_T \cdot \mathbf{E}^T) \cdot (\mathbf{E} \cdot \beta_T^T) \\ \text{s. t. } \mathbf{I} \cdot \mathbf{W}_T = 1 \\ \omega_i \geq 0 \end{cases} \quad (14)$$

式中:  $\mathbf{E}$  为组合模型对时间窗口  $T$  内所有负荷数据样本点的均方预测误差所构成的行向量;  $\mathbf{I} = [1, \dots, 1]$ 。

完成本周期的组合预测后,通过式(13)更新参数  $\beta$ ,之后,时间窗口滚动进入下一预测周期,将式(13)得到的  $\beta'$  应用到最新的时间窗口中并代入式(14)求解出最新子模型权重矩阵,并使用最新子模型权重矩阵组合子模型预测。同理,不断重复上述过程即可实现针对短期负荷数据的变权组合预测。

## 3 算例分析

本文算例采用中国北方某地区2017年1月1日~

2020年1月1日共3年的真实运行电力负荷数据作为数据集,负荷数据采样周期为30 min/次,共有52596条数据。选定CNN-LSTM、XGBoost模型作为子模型进行组合,本文后续组合模型均由上述两个子模型构建。为便于后续说明,将采用本文组合策略的组合模型命名为Online\_ensemble。

为了提升预测效率和精度,首先对原始数据进行了预处理和特征提取。为了验证所提策略的有效性和优越性,针对测试集数据利用本文所提组合策略与传统组合策略进行对比,传统组合策略依次选取负荷组合预测常用的固定子模型权重的误差倒数法、可实现子模型变权的最优加权法。为验证所提策略的优越性及普适性,最后选取了夏季、冬季2个季节中的代表月份进行了不同工况下的仿真预测。为了使子模型权重最大程度地接近相应变权策略所构造目标函数的全局最优解,本文改进最优加权法和最优加权法进行子模型权重求解时均采用穷举法,将子模型权重  $\omega_1$  以步长0.0001从0~1递进变化,则满足等式关系  $\omega_2 = 1 - \omega_1$ ,将  $\omega_1, \omega_2$  分别代入式(8)和(14)求解,得出子模型的最新权重矩阵。

为加快模型的训练收敛速度,使用归一化后数据的80%构造训练集,5%构造验证集。为观察组合策略在组合模型投入较长时间后的使用效果,使用15%的数据构造测试集,将预测结果输出并反归一化,可得到最终的负荷预测结果。

为了排除因参数设置不当对算例分析的干扰,完成组合模型构建后,首先使用验证集数据的高频分量数据并基于网格搜索法进行组合策略参数整定。经反复调优,并根据各组合策略在验证集中的最好表现,将采用最优加权法的组合模型的时间窗口长度设置为51;采用本文改进最优加权法组合策略的Online\_ensemble组合模型的时间窗口长度设置为58,设置式(13)的平衡系数  $\lambda$  为0.85。

在进行特征提取时,采用皮尔逊相关系数选取与负荷数据强相关的特征,最终选取了与负荷强相关的负荷数据  $L$ 、温度  $T$ 、小时  $H$ 。构建样本集合  $D = \{(x_1, y_1), \dots, (x_n, y_n)\}$ ,其中,样本特征  $x_i = \{L_i, \dots, L_{i+7}, T_{i+7}, H_{i+7}\}$ ,  $y_i = \{L_{i+8}, \dots, L_{i+52}\}$ 。

### 3.1 评价指标

本文采用在电力负荷预测中常用的平均绝对百分比误差(mean absolute percentage error, MAPE)和均方根误差(root mean square error, RMSE)作为各模型的评价指标,如式(15)和(16)所示。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left\| \frac{\bar{y}_i - y_i}{y_i} \right\| \times 100\% \quad (15)$$

式中:  $n$  为负荷样本点的个数;  $\bar{y}_i$  为预测值;  $y_i$  为实际值。

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\bar{y}_i - y_i)^2}{n}} \quad (16)$$

### 3.2 负荷预测仿真案例分析

为验证策略的有效性,首先将针对波动较强的高频分量采用 XGBoost 和 CNN-LSTM 模型并使用不同组合策略进行组合预测、低频分量使用同一 MLR 模型预测、残差分量使用同一 LSTM 模型预测,表 1 为使用上述模型分别针对不同模态分量数据的预测评价指标结果。

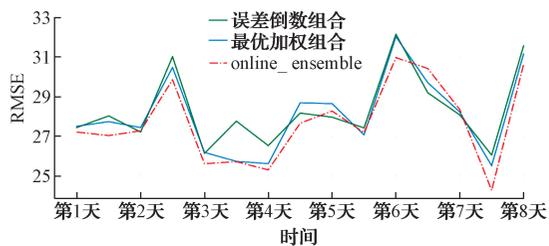
表 1 不同模态分量预测数据的评价指标结果

所用模型/所预测的模态分量	MAPE /%	RMSE /kW
Online_ensemble 组合模型(高频分量)	<b>4.28</b>	<b>30.84</b>
采用最优加权法的组合模型(高频分量)	4.78	34.02
采用误差倒数法的组合模型(高频分量)	5.52	35.45
MLR 模型(低频分量)	3.92	27.13
LSTM 模型(残差分量)	6.53	39.44

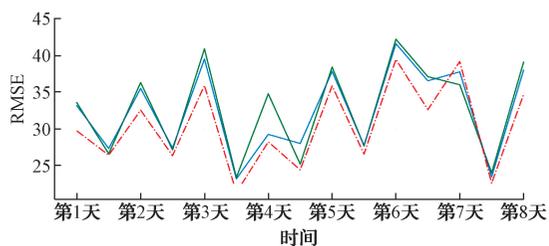
由表 1 可知,采用本文提出的改进最优加权法的 Online\_ensemble 组合模型在高频分量数据的预测中表现出色。MLR 模型在对低频分量数据进行预测时具有较高的准确性。残差分量数据占比较小且具有较强的非线性特征,LSTM 模型能有效学习时序负荷数据的复杂非线性关系,有着良好的预测表现。

本文为了更加明显的观察不同组合策略的预测效果,后续将叠加各模态分量的预测结果,还原预测结果后进行预测对比分析。

第 1 周和最后一周采用不同组合策略的组合模型针对算例测试集数据的 RMSE 预测误差的变化图像如图 5 所示。每隔 12 h 计算一次在此期间所有预测数据的平均 RMSE 误差。时间分别为 2019 年 7 月 22 日~28 日、2019 年 12 月 25 日~31 日。



(a) 第 1 周采用不同组合策略 RMSE 预测误差结果



(b) 最后一周不同组合策略 RMSE 预测误差结果

图 5 组合模型预测结果

表 2 为基于算例测试集数据仿真预测结果获得的组合模型及其子模型的评价指标结果。此外,本文为降低负荷数据的非平稳性,使用了 VMD 算法对原始负荷序列进行了分解,并使用适宜的模型预测了相应的分量。为验证其有效性,使用本文策略未经 VMD 分解直接组合 CNN-LSTM、XGBoost 模型预测原始负荷序列,并计算了其评价指标结果。

表 2 组合模型及其子模型的评价指标结果

是否 VMD 分解/高频分量 所用预测模型	MAPE /%	RMSE /kW
VMD-Online_ensemble 组合模型	4.18	29.11
VMD-采用最优加权法的组合模型	4.67	32.28
VMD-采用误差倒数法的组合模型	5.42	33.75
VMD-子模型 XGBoost	5.74	35.96
VMD-子模型 CNN-LSTM	6.02	37.73
未经 VMD-Online_ensemble 组合模型	4.43	31.01

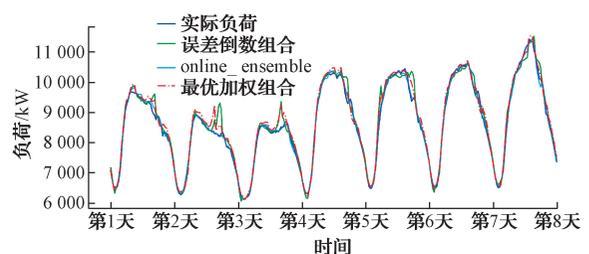
由图 5 可知,在大多数时期采用本文变权组合策略 Online\_ensemble 组合模型的预测结果均具有更小误差,且更为稳定,其能够从始至终完成对负荷数据较为准确的预测。由表 2 可知,组合模型可有效提升单一子模型的预测精度。

综上可知,Online\_ensemble 组合模型的预测精度优于采用其他传统组合策略的组合模型和其任意的单一子模型。通过 VMD 算法降低负荷数据的非平稳性,并选定适宜的模型进行相应分量的预测可有效提升短期负荷的预测精度。

### 3.3 较大差异工况下采用不同组合策略预测效果对比

夏季的 8 月、冬季的 12 月两月的负荷数据特点有着明显差异。为验证所提策略的优越性及普适性,叠加各模态分量还原预测结果后,本文以 2019 年 8 月第 2 周以及 12 月第 2 周的数据为例,对比分析采用不同组合策略的预测精度与效果。得到上述月份第 2 周的负荷预测的结果图像,如图 6 所示。

基于预测结果计算了上述 8 月第 2 周、12 月第 3 周中采用不同组合策略的预测值的评价指标结果,如表 3 所



(a) 8 月第 2 周预测结果

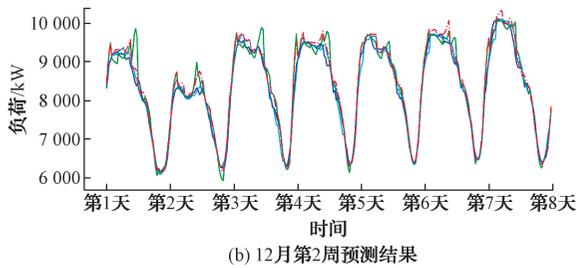


图6 两个月份组合模型预测结果

示。观察表3可知,在数据特征差异明显的各月中,采用本文策略的 Online\_ensemble 模型均具有更高的预测精度,在外界工况环境变化时能够快速适应,具有最佳的预测泛化能力。

表3 两个月份第2周的不同预测模型评价指标结果

月份	是否 VMD 分解/高频分 所用预测模型	MAPE /%	RMSE /kW
8月	VMD-Online_ensemble 组合模型	4.02	28.04
	VMD-采用最优加权法的组合模型	4.51	31.46
	VMD-采用误差倒数法的组合模型	5.13	31.19
	VMD-子模型 XGBoost	5.65	34.71
12月	VMD-子模型 CNN-LSTM	5.94	36.03
	VMD-Online_ensemble 组合模型	4.23	30.02
	VMD-采用最优加权法的组合模型	4.87	33.77
	VMD-采用误差倒数法的组合模型	5.47	34.15
	VMD-子模型 XGBoost	5.83	36.76
	VMD-子模型 CNN-LSTM	6.35	38.03

#### 4 结论

针对传统组合策略不能较好的适应负荷数据特点的问题,为了提升短期电力负荷预测的精度,本文参考指数加权法原理,兼顾负荷数据的时序特点提出了一种基于 VMD-改进最优加权法的短期负荷组合预测策略,并根据该策略测试了组合模型,最终完成对短期负荷的精确预测。

1) 本文所提改进加权组合策略可以在传统负荷组合模型变权更新策略的基础上进一步提升其负荷预测精度。

2) 采用本文所提策略的组合模型可以实现短期负荷的预测,其明显优于任意单一子模型且与传统组合策略比较预测精度高、预测性能更优。

3) 本文所提变权组合策略在运行工况变化时,能够快速调整子模型权重矩阵完成子模型的变权组合,使组合模型仍具有较高的预测精度和良好的泛化能力,可满足短期负荷精确预测的现实需求。

#### 参考文献

- [1] 石文辉,屈姬贤,罗魁,等.高比例新能源并网与运行发展研究[J].中国工程科学,2022,24(6):52-63.
- [2] 朱莉,韩凯萍,朱春强.二次分解策略组合 Informer 的短期电力负荷预测方法[J].国外电子测量技术,2023,42(6):23-32.
- [3] 孟金鑫,黄山,印月.基于特征优选策略和 DLSTMs-FCN 优化的短期负荷预测[J].电子测量技术,2023,46(10):46-52.
- [4] BENTO P M R, POMBO J A N, CALADO M R A, et al. Optimization of neural network with wavelet transform and improved data selection using bat algorithm for short-term load forecasting [J]. Neurocomputing, 2019, 358: 53-71.
- [5] 吴晨,姚菁,薛贵元,等.基于 MMoE 多任务学习和长短时记忆网络的综合能源系统负荷预测[J].电力自动化设备,2022,42(7):33-39.
- [6] 张淑清,李君,姜安琦,等.基于 FPA-VMD 和 BiLSTM 神经网络的新型两阶段短期电力负荷预测[J].电网技术,2022,46(8):3269-3279.
- [7] SHANG C, GAO J, LIU H, et al. Short-term load forecasting based on PSO-KFCM daily load curve clustering and CNN-LSTM model[J]. IEEE Access, 2021, 9: 50344-50357.
- [8] 殷礼胜,吴洋洋.基于改进 VMD-GAT-GRU 的交通流量组合预测模型[J].电子测量与仪器学报,2022,36(7):62-72.
- [9] 崔树银,汪昕杰.基于最大信息系数和多目标 Stacking 集成学习的综合能源系统多元负荷预测[J].电力自动化设备,2022,42(5):32-39,81.
- [10] 邵必林,严义川,曾弄玢.注意力机制下的 VMD-IDBiGRU 负荷预测模型[J].电力系统及其自动化学报,2022,34(10):120-128.
- [11] 陈振宇,刘金波,李晨,等.基于 LSTM 与 XGBoost 组合模型的超短期电力负荷预测[J].电网技术,2020,44(2):614-620.
- [12] 张翔颖,杨永标,徐青山,等.基于多时段相似日理论的光伏功率组合预测方法[J].南方电网技术,2023,17(2):57-65.
- [13] 门东坡,王金力,何超平,等.基于最优加权组合模型的煤炭消费结构预测[J].煤炭工程,2021,53(3):190-196.
- [14] 陈纬楠,胡志坚,岳菁鹏,等.基于长短期记忆网络和 LightGBM 组合模型的短期负荷预测[J].电力系统自动化,2021,45(4):91-97.
- [15] 李俊良,焦润海,王双坤,等.一种基于误差在线更新的集成负荷预测模型[J].中国电机工程学报,2023,43(4):1402-1413.

- [16] 孟安波,许炫淙,陈嘉铭,等. 基于强化学习和组合式深度学习模型的超短期光伏功率预测[J]. 电网技术, 2021,45(12):4721-4728.
- [17] 刘巍,李猛,李秋燕,等. 基于改进遗传算法的电网投资组合预测方法[J]. 电力系统保护与控制, 2020, 48(8):78-85.
- [18] 史佳琪,张建华. 基于多模型融合 Stacking 集成学习方式的负荷预测方法[J]. 中国电机工程学报, 2019, 39(14):4032-4042.
- [19] 江星星,宋秋昱,杜贵府,等. 变分模式分解方法研究与应用综述[J]. 仪器仪表学报, 2023,44(1):55-73.
- [20] LI J, DENG D, ZHAO J, et al. A novel hybrid short-term load forecasting method of smart grid using MLR and LSTM neural network [J]. IEEE

Transactions on Industrial Informatics, 2020, 17(4): 2443-2452.

#### 作者简介

李志军,博士,教授,正高工,主要研究方向为智能电网技术、电力系统稳定与控制等。

E-mail:zhijun\_li@263.net

徐博(通信作者),硕士研究生,主要研究方向为机器学习技术在电力系统中的应用。

E-mail:934489425@qq.com

杨金荣,硕士研究生,主要研究方向为新能源并网稳定性分析与控制、电力系统负荷数据预测。

E-mail:1370191064@qq.com