

基于二次分解和 BiLSTM 的多特征天然气负荷预测

张芷晨, 邵必林

西安建筑科技大学 管理学院, 西安 710055

摘要:目的 针对天然气负荷数据存在噪声多、规律性难以捕捉和特征冗余的典型特点,提出了一种基于二次分解和双向长短期记忆(BiLSTM)模型的多特征天然气负荷预测方法。方法 首先,利用 RF 对特征进行重要性排序,筛选出重要关联特征;其次,利用奇异谱分析(SSA)对原始数据进行分解重构,提取数据的趋势和周期信息并形成分量,分离出高频分量,然后利用变分模态分解(VMD)对高频分量进行二次分解,形成各模态分量,并将各模态分量输入 BiLSTM 模型进行预测;最后融合各分量预测结果得到最终的负荷点预测结果。结果 仿真实验表明:相比于其他传统模型,所提的二次分解和组合模型预测效果更优,预测精度更高。结论 新建模型具有良好的解释性与合理性,能有效应用于天然气负荷预测,为决策者合理规划天然气供应计划提供可靠的理论依据和有效的技术支持。

关键词:随机森林;二次分解;双向长短期记忆网络;组合预测;天然气负荷预测

中图分类号:TP183 **文献标识码:**A **doi:**10.16055/j.issn.1672-058X.2026.0004.018

Multi-Feature Natural Gas Load Forecasting Based on Secondary Decomposition and BiLSTM

ZHANG Zhichen, SHAO Bilin

School of Management, Xi'an University of Architecture and Technology, Xi'an 710055, China

Abstract: Objective Aiming at the typical characteristics of natural gas load data, such as noise, difficulty in capturing regularity, and feature redundancy, a multi-feature natural gas load forecasting method based on secondary decomposition and bidirectional long short-term memory (BiLSTM) model is proposed. **Methods** First, random forest (RF) was used to rank the importance of features, and important associated features were screened out. Second, singular spectrum analysis (SSA) was employed to decompose and reconstruct the original data, extract the trend and periodic information of the data to form components, and separate the high-frequency components. Then, variational mode decomposition (VMD) was used to conduct secondary decomposition on the high-frequency components to generate individual modal components. These modal components were fed into the BiLSTM model for prediction. Finally, the prediction results of each component were fused to obtain the final load point prediction result. **Results** Simulation results showed that compared with other traditional models, the proposed secondary decomposition and combination model had a better prediction effect and higher prediction accuracy. **Conclusion** The newly built model has good interpretability and rationality, and can be effectively applied to natural gas load forecasting, providing a reliable theoretical basis and effective technical support for decision-makers to reasonably make natural gas supply plans.

Keywords: random forest; secondary decomposition; bidirectional long short-term memory network; combined forecasting; natural gas load forecasting

收稿日期:2024-06-26 修回日期:2024-08-27 文章编号:1672-058X(2026)04-0135-07

基金项目:国家自然科学基金面上项目(62072363)资助。

作者简介:张芷晨(1999—),女,四川阆中人,硕士研究生,从事负荷预测研究。Email:17378580953@163.com。

通信作者:邵必林(1965—),男,云南腾冲人,教授,博士生导师,从事数据信息管理与负荷预测研究。Email:sblin0462@163.com。

引用格式:张芷晨,邵必林.基于二次分解和 BiLSTM 的多特征天然气负荷预测[J].重庆工商大学学报(自然科学版),2026,43(4):135-141.

Zhang Zhichen, Shao Bilin. Multi-feature natural gas load forecasting based on secondary decomposition and BiLSTM[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(4): 135-141.

天然气广泛应用于工业生产、家庭供暖和清洁能源发电等领域,凭借着环保清洁、高热效率、灵活便利等优势,已成为全球主要的清洁能源之一^[1]。准确预测天然气负荷可以帮助优化天然气供应和需求之间的平衡,对合理安排天然气供应计划,保障能源供应的稳定性和可靠性具有重要意义。

天然气负荷数据具有波动性强、嘈杂度高、周期性明显的特点,其波动受多因素影响。高红燕等^[2]分析了西安市供暖期、节假日、双休日燃气负荷的变化规律,结果表明西安市日燃气负荷呈单峰型波动变化,供暖期燃气负荷具有双休日、节假日效应,供暖期燃气负荷与最高气温、最低气温、平均气温等气象因子呈显著负相关。为了进一步提高预测效果,特征筛选成了一个关键步骤。杨奕等^[3]指出冗余因素对负荷预测精度及速度会产生影响,并提出了一种结合特征选择 Relief 算法的预测模型,通过筛选关键影响因素显著提高了天然气负荷预测的精度和效率。佟敏等^[4]采用灰色关联度法(GRA)、平均影响值法(MIV)和主成分分析法(PCA)对天然气需求量影响因素进行了特征筛选,提高预测的准确性和精度,然而,这些方法主观性较强,易影响特征筛选的准确性。Veeramsetty 等^[5]使用随机森林算法筛选特征,有效地减少了模型的输入维度,降低了模型计算时间。

天然气负荷预测的方法可分为传统预测方法和基于人工智能预测方法。近年来,深度学习因其捕捉数据特征的优越性,在负荷预测领域等到了广泛应用。Imani M^[6]利用卷积神经网络捕捉时间序列数据中的局部模式,提高了预测精度。孙庆凯等^[7]结合 LSTM 和多任务学习,对综合能源系统中的多元负荷进行预测,有效提升了预测的准确性和效率,但 LSTM 仅使用前向信息进行预测,在捕捉序列信息能力方面仍存在相应不足。此外,由于缺乏对原始数据的有效处理,时间序列数据中的高频成分会限制模型的预测精度^[8]。为减小模型误差,研究者们通过信号分解的方式对原始序列进行处理,常见的有经验模态分解^[9](EMD)、变分模态分解^[10](VMD)和奇异谱分析^[11](SSA)等。陈川等^[12]用 EMD 算法分解残差,再利用 LSTM 模型学习各分量的短期时序规律,但经过 EMD 分解后的高频分量部分通常具有很强的随机性。邓带雨等^[13]使用 EEMD 分解电力负荷数据再进行预测,EEMD 是针对 EMD 存在模态混叠现象而做出的改进,但 EEMD 会产生残余辅助噪声和重构误差。那幸仪等^[14]采用离散小波变换将原始序列数据处理为方差更稳定的分解序列。穆晨宇等^[15]通过 VMD 方法对原始负荷进行分解,分解出的模态分量是一个有带宽限制的调幅-调频函数,具有良好的噪声鲁棒性。然而即使在实施一次分解之后,某些高频分量仍可能给准确预测带来困难。

综上梳理可得,尽管深度学习模型在天然气负荷

预测领域已经得到了广泛应用,但预测主流模型依然为 LSTM 和 GRU 等神经网络模型,这类模型在拟合数据时考虑的信息较为单薄,仅考虑了序列的单向信息。此外,集成模型发展迅速^[16],根据各种模型的优缺点将几种不同的模型组合起来进行预测可以较为明显地改善预测效果,还可以减少复杂模型中的过拟合的概率^[17]。以往的组合模型大多忽略了对原始数据的有效处理,导致预测效果精度不高。即使特征筛选和信号分解已得到了应用,但仍然存在特征筛选准确度不高,信号分解程度不够,且易受主观性影响的问题。

基于此,针对天然气负荷数据复杂性高,单一模型预测准确性低的问题,本文结合特征筛选的思想,融合二次分解模型和 BiLSTM 模型对天然气负荷进行预测。首先,采用随机森林算法(RF)进行特征筛选,解决数据特征冗余的问题;其次,利用 SSA 提取序列趋势,分离序列中的高频分量;再利用 VMD 对高频分量数据进行二次分解,获得多个 IMF 分量,两次分解过程中分别使用 Cao 法和中心频率法确定分解参数;然后将各分量输入 BiLSTM 模型中进行预测,最后通过重构各分量预测结果,实现对天然气负荷的精准预测。

1 研究方法

1.1 随机森林

随机森林是一种集成学习方法,通过构建多个决策树来进行分类或回归。其关键优势是能够评估特征对模型预测能力的贡献,即特征的重要性,通常通过计算平均不纯度减少和平均精度下降这两种方式来衡量。

(1) 平均不纯度减少。这种方法基于决策树构建过程中每个特征所减少的不纯度进行衡量。在决策树的每个节点,分裂操作会根据特征值将数据集划分为不同子集,目的是实现最大化不纯度的减少。不纯度可以通过不同的标准来衡量,如基尼不纯度(Gini impurity)或信息增益(Information Gain)。

(2) 平均精度下降(Permutation Importance)。这种方法通过打乱单个特征的值来评估其对模型精度的影响。首先,使用原始数据集对随机森林模型进行训练,并评估其预测精度;其次,选择一个特征,然后在所有样本中随机打乱该特征的值,这一步的目的是模拟该特征对模型预测能力的影响被完全移除的情况;在特征值被打乱后,重新训练随机森林模型,并评估其预测精度;最后,计算模型在原始数据集上的预测精度与在打乱后数据集上的预测精度之间的差异。这个差异量反映了特征的重要性。

1.2 奇异谱分析

奇异谱分析能直接分离信号的 trend、振荡和噪声信息,分为 4 个过程,其基本算法步骤如下^[18]:

(1) 将序列 $X=(x_1, x_2, x_3, \dots, x_N)$ 转化为如下轨迹矩阵:

$$\mathbf{X} = [x_1 : x_2 : \cdots : x_K]' = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_n \end{bmatrix} \quad (1)$$

式(1)中, $K=n-L+1$ 为选择的滞后向量, L 为所选窗口的长度, $1 < L < N$ 。

(2) 建立 $\mathbf{S}$ 矩阵: $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, $\lambda_1, \lambda_2, \cdots, \lambda_L$ 为矩阵 $\mathbf{S}$ 的特征值, 得到奇异值分解结果:

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_d \quad (2)$$

式(2)中, $\mathbf{X}_i = \sqrt{\lambda_i} \mathbf{U}_i \mathbf{V}_i^T$, $\mathbf{V}_i = \mathbf{X}^T \mathbf{U}_i / \sqrt{\lambda_i}$ ($i = 1, 2, \cdots, d$), $\mathbf{U}_1, \mathbf{U}_2, \cdots, \mathbf{U}_L$ 为对应的标准正交特征矩阵, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq 0$ 。

(3) 将集合 $\{1, 2, \cdots, d\}$ 重构为 m 个不相交的子集, 子集 I 和相应的分组矩阵 $\mathbf{X}_i$ 如下:

$$I = \{i_1, i_2, \cdots, i_p\} \quad (3)$$

$$\mathbf{X} = \mathbf{X}_{i_1} + \mathbf{X}_{i_2} + \cdots + \mathbf{X}_{i_p} \quad (4)$$

式(3)、式(4)中, p 为子集的长度, 由此可得:

$$\mathbf{X} = \mathbf{X}_{i_1} + \mathbf{X}_{i_2} + \cdots + \mathbf{X}_{i_p} \quad (5)$$

(4) 将各个分组矩阵 $\mathbf{Y} = \mathbf{X}_{ij}$ 重组为一维信号序列

$y_1, y_2, \cdots, y_N$:

$$y = \begin{cases} \frac{1}{k} \sum_{m=1}^k y_{m,k-m+1}^*, & 1 \leq k \leq L^* \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+1}^*, & L^* \leq k \leq K^* \\ \frac{1}{N-k+1} \sum_{m=k-K^*+1}^{N-K^*+1} y_{m,k-m+1}^*, & K^* \leq k \leq N^* \end{cases} \quad (6)$$

式(6)中, $1 < i < L, 1 < j < K, K^* = \max(L, K)$, 由此可得一维矩阵:

$$\mathbf{y} = [y_1, y_2, \cdots, y_k, y_{k+1}, y_{k+2}, \cdots, y_L, y_{L+1}, y_{L+2}, \cdots, y_{K+L-2}, y_{K+L-1}] \quad (7)$$

1.3 变分模态分解

变分模态分解(VMD)是一种非递归的自适应信号分解方法,能有效缓解经验模态分解中出现的模态混叠和端点效应问题。

(1) VMD根据固有模态分量估计带宽之和的最小化,通过寻找约束变分模型的最优解来实现原始序列的分解,表达式为

$$\min_{\{u_k\}, \{\omega_k\}} = \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \quad (8)$$

$$\text{s.t.} \sum_{k=1}^K u_k(t) = y(t) \quad (9)$$

式(8)、式(9)中, K 为本征模态函数 IMF 的个数, $\{u_k\}$ 为 IMF 的集合, $\{\omega_k\}$ 为中心频率的集合, $\delta(t)$ 为脉冲函数, $*$ 为卷积算子, j 为虚数单位, $y(t)$ 为原始序列。

(2) 引入二次惩罚因子 σ 和拉格朗日乘子 $\lambda(t)$,

将上式转化为无约束变分问题:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \sigma \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| y(t) - \sum_k u_k(t) \right\|_2^2 + \langle \lambda(t), y(t) - \sum_k u_k(t) \rangle \quad (10)$$

(3) 采用交替乘子方向法迭代更新 $\{u_k\}, \{\omega_k\}, \lambda$, 搜寻上式的鞍点:

$$u_k^{n+1}(\omega) =$$

$$z(\omega) - \sum_{i < k} u_i^{n+1}(\omega) - \sum_{i > k} u_i^n(\omega) + \frac{\lambda^n(\omega)}{2} \quad (11)$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |u_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |u_k^{n+1}(\omega)|^2 d\omega} \quad (12)$$

(4) 满足下列收敛式时,停止更新:

$$\sum_{k=1}^K \frac{\|u_k^{n+1} - u_k^n\|_2^2}{\|u_k^n\|_2^2} < \varepsilon \quad (13)$$

1.4 双向长短期记忆网络

长短期记忆网络引入了输入、输出和遗忘门,能够有效编码和保留与长模式相关的信息,其结构如图1所示,相应的数学表达式如下:

$$i_t = \sigma(X_t W_{xi} + h_{t-1} W_{hi} + B_i) \quad (14)$$

$$f_t = \sigma(X_t W_{xf} + h_{t-1} W_{hf} + B_f) \quad (15)$$

$$o_t = \sigma(X_t W_{xo} + h_{t-1} W_{ho} + B_o) \quad (16)$$

$$\tilde{C}_t = \tanh(X_t W_{xc} + h_{t-1} W_{hc} + B_c) \quad (17)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \tilde{C}_t \quad (18)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (19)$$

式(14)一式(19)中, X_t 为 t 时刻的输入; i_t, f_t, o_t 分别为输入门、遗忘门和输出门; $W_{xi}, W_{hi}, W_{xf}, W_{hf}, W_{xo}, W_{ho}$ 为三个门对应的权重矩阵; B_i, B_f, B_o 为对应的偏置量; W_{xc}, W_{hc}, B_c 分别为候选记忆单元的权重矩阵和偏置项; h_{t-1} 为 $t-1$ 时刻的状态; C_{t-1} 表示 $t-1$ 时刻隐藏层单元状态; σ 为 Sigmoid 激活函数。

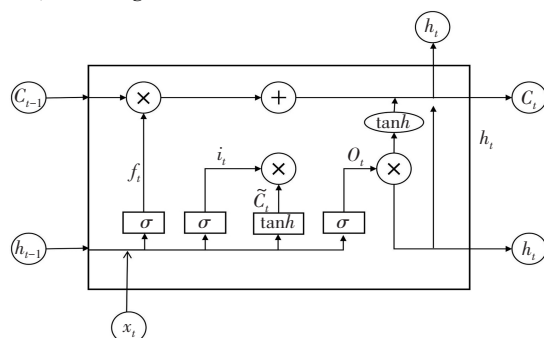


图1 LSTM 结构

Fig. 1 Structure of the LSTM cell

双向长短期记忆网络 (BiLSTM) 是长短期记忆 LSTM 的扩展, 引入了两个独立的 LSTM 层, 一个按照正序处理输入序列, 另一个按逆序处理, 使得模型能够同时捕获到序列中当前时刻的过去和未来信息。这种双向结构能够更全面地理解和表达序列数据, 提高其在序列建模任务中的性能。BiLSTM 结构如图 2 所示, 数学表达式如下:

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}) \quad (20)$$

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t-1}) \quad (21)$$

$$y_t = \sigma(W_y \cdot [\vec{h}_t, \overleftarrow{h}_t] + b_y) \quad (22)$$

式(20)一式(22)中, LSTM 单元为传统 LSTM 网络的计算过程, $\vec{h}_t$ 为前向 LSTM 隐藏层在时刻 t 的状态; $\overleftarrow{h}_t$ 为时刻 t 的后向 LSTM 隐藏层状态; W_y 和 b_y 分别为权矩阵和偏置项; $x_1, x_2, \dots, x_t$ 表示 $t_1 \sim t_i (i \in [1 \sim t])$ 各时刻对应的输入数据; $y_1, y_2, \dots, y_t$ 表示 $t_1 \sim t_i (i \in [1 \sim t])$ 各时刻对应的输出数据。

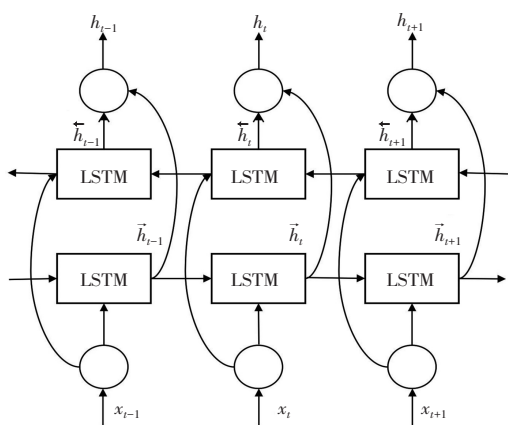


图 2 BiLSTM 结构

Fig. 2 Structure of the BiLSTM network

2 模型构建

2.1 模型架构

本文提出一种基于二次分解和 BiLSTM 的多特征预测模型, 整体框架如图 3 所示。具体步骤如下:

- (1) 利用随机森林法进行特征选择, 筛选出关联性高的特征;
- (2) 用 SSA 对原始负荷数据进行分解重构, 采用 Cao 法^[19]确定分解参数, 形成趋势性分量并分离出高频的随机波动分量;
- (3) 对高频分量进行 VMD 分解, 采用中心频率法确定分解参数, 形成多个模态分量;
- (4) 将重要关联特征序列结合趋势性分量和模态分量输入 BiLSTM 模型中进行预测;
- (5) 叠加各分量预测值得到最终的点预测结果。

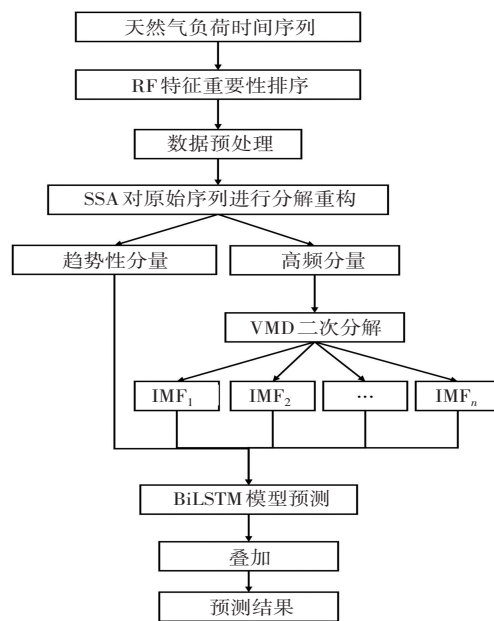


图 3 预测模型框架图

Fig. 3 Framework of the proposed prediction model

2.2 评价指标

选择均方根误差、平均绝对误差和平均绝对误差百分比作为评价预测效果的主要指标。

(1) 均方根误差 (R_{MSE}):

$$R_{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (23)$$

(2) 平均绝对误差 (M_{AE}):

$$M_{AE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (24)$$

(3) 平均绝对百分比误差 (M_{APE}):

$$M_{APE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (25)$$

式(23)一式(25)中, y_i 和 $\hat{y}_i$ 分别表示真实值和预测值; n 表示预测样本数。相应的指标数值越接近 0, 误差越小, 表明模型的预测效果越好。

3 仿真实例与结果分析

3.1 数据预处理

本文选取陕西省汉中市 2018 年 1 月 1 日—2022 年 12 月 30 日某天然气站采集的天然气负荷数据进行仿真实验, 间隔时间为 1 d, 缺失数据用线性插值法填补, 数据共计 1 825 条。同时采集该地区对应时间段每日的最高温度、最低温度、平均温度及风速数据, 并将每日的天气类型、日期类型、节假日类型和季节类型进行无量纲处理, 形成初始特征数据集。

3.2 特征选择

天然气负荷受多方因素影响, 包含线性因素和非线性因素, 选取随机森林法对特征重要性进行排序, 结

果如图 4 所示。结合实际情况,选取季节、平均温度、最高温度、最低温度和天气类型为输入特征。

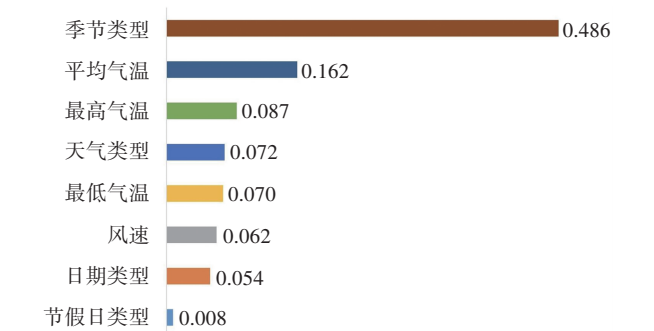


图 4 特征重要性排序图

Fig. 4 Feature importance ranking

3.3 负荷序列的二次分解

对原始序列进行奇异谱分析,SSA 算法需要确定嵌入维数 L ,为避免主观决定所造成的误差,采用 Cao 法确定了时间序列的最佳嵌入维数,最终取值为 $L=7$ 。SSA 分解后的 $IMF_1 \sim IMF_7$ 分量如图 5 所示。

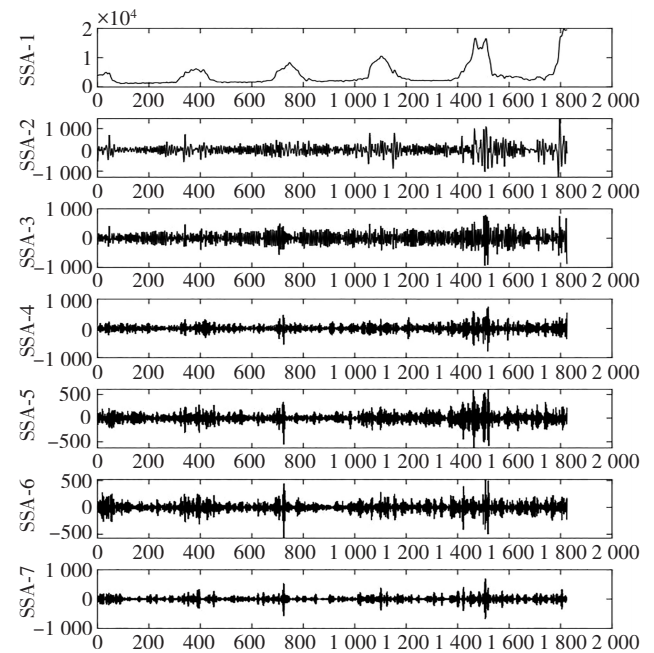


图 5 SSA 分解图

Fig. 5 Decomposition results of SSA

由图 5 可知, IMF_1 包含了原始序列的大部分趋势性信息, $IMF_2 \sim IMF_7$ 复杂度高,噪声含量大,难以直接进行预测。由此可知,在面对高复杂度的序列数据时,单一 SSA 分解难以很好地捕捉序列规律性特征,因此将 $IMF_2 \sim IMF_7$ 进行合并,作为待处理的高频波动分量。

对高频波动分量进行 VMD 二次预测,采用中心频率法确定分解模态数 k 的取值。从 $k=2$ 开始,对随机波动分量进行分解,各 IMF 的中心频率如表 1 所示。

由表 1 可知,当 $k=6$ 时,模态 u_3 和 u_4 、 u_5 和 u_6 的中

心频率相近,很可能出现过度分解的情况,因此取 $k=5$ 为最终的分解模态个数,最终分解效果如图 6 所示。

表 1 各分量中心频率

Table 1 Center frequencies of each component

	IMF ₁	IMF ₂	IMF ₃	IMF ₄	IMF ₅	IMF ₆
2	0.088 0	0.143 0				
3	0.072 7	0.109 8	0.149 0			
4	0.070 3	0.106 5	0.146 5	0.262 1		
5	0.068 2	0.104 2	0.144 6	0.203 5	0.284 0	
6	0.067 5	0.102 8	0.142 1	0.177 2	0.224 6	0.285 1

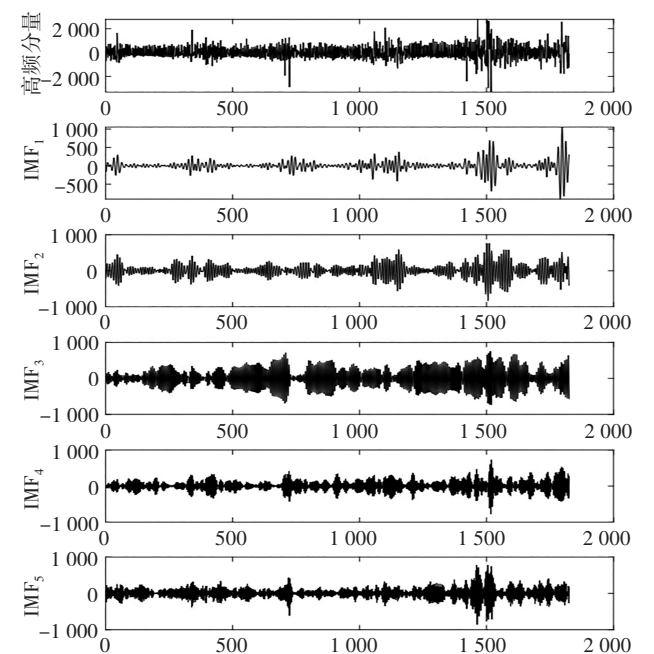


图 6 VMD 分解图

Fig. 6 Decomposition results of VMD

由图 6 可知,SSA 分解出高频分量波动性强,平稳性较差,而经 VMD 二次分解后得到的分量平稳性相对较好,有利于提升后续模型预测准确度。

3.4 预测结果分析

3.4.1 不同模型对比结果

仿真实验设置训练集与测试集的比例为 9 : 1,时间步长设为 7,然后将最终分量输入 BiLSTM 模型中进行预测。

为验证 BiLSTM 模型的有效性,与当下深度学习主流模型 LSTM 和 GRU 进行对比,对比实验均采用相同的数据处理方法和特征选择方法;各模型拟合效果如图 7 所示, M_{AE} 、 M_{APE} 、 R_{MSE} 评价指标结果如表 2 所示。

由图 7 和表 2 可以看出,本文所提模型的拟合效果最好,相比 LSTM 和 GRU, M_{AE} 分别降低了 7.84% 和 6.82%, M_{APE} 分别降低了 2.23% 和 1.89%, R_{MSE} 分别降低了 9.48% 和 5.13%。表明 BiLSTM 模型的预测效果更佳,精度更高。

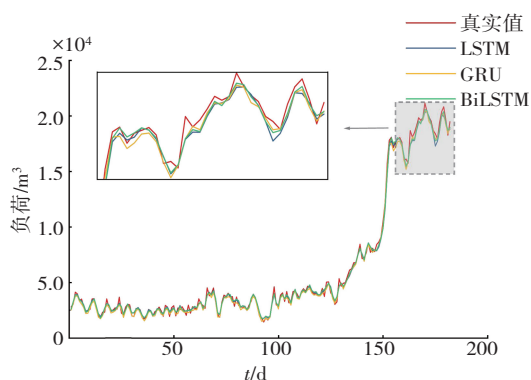


图 7 不同模型对比结果

Fig. 7 Comparison results of different models

表 2 不同模型评价指标对比

Table 2 Comparison of evaluation metrics for different models

预测模型	M_{AE}	M_{APE}	R_{MSE}
LSTM	253.61	5.84	331.69
GRU	250.82	5.82	316.49
BiLSTM	233.72	5.71	300.26

3.4.2 特征选择对比

为验证特征选择的有效性,将其与未经 RF 特征处理的单一模型进行对比;各模型预测效果如图 8 所示,评价指标结果见表 3。由图 8 和表 3 可以看出,相比未进行 RF 特征处理的单一 BiLSTM 模型,本文所提模型 MAE 降低了 13.91%,MAPE 降低了 7.31%,RMSE 降低了 16.14%。表明通过 RF 特征筛选可显著提升模型拟合效果,增加模型有效性和准确性。

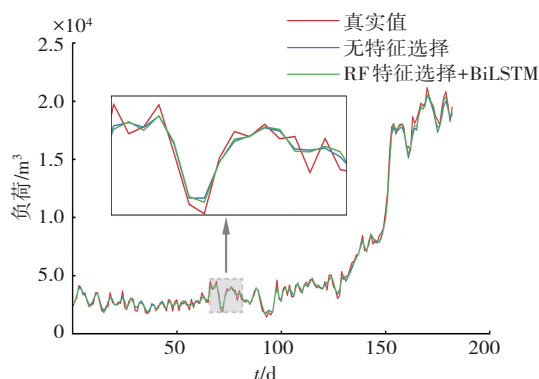


图 8 特征选择对比

Fig. 8 Feature selection comparison

表 3 评价指标对比

Table 3 Comparison of evaluation metrics of different models

预测模型	M_{AE}	M_{APE}	R_{MSE}
无特征选择	271.47	6.16	358.03
RF 特征选择+BiLSTM	233.72	5.71	300.26

3.4.3 分解方式对比

为验证二次分解模块的有效性,对同一组数据分别采用不进行信号分解、进行一次信号分解 (SSA、VMD) 和进行二次信号分解 (SSA-VMD) 的方法进行预测,并观察模型拟合效果和评价指标的变化。实验结果如图 9 和表 4 所示。

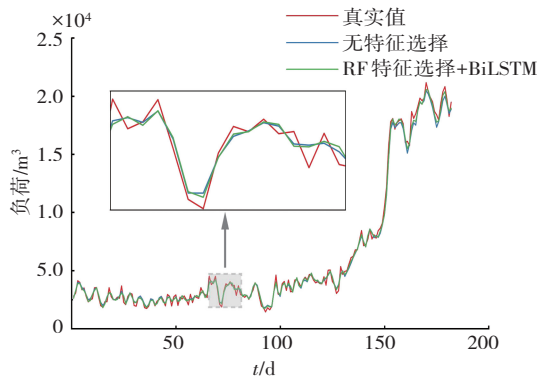


图 9 不同分解方式对比结果

Fig. 9 Comparison results of different decomposition methods

表 4 不同分解方式评价指标对比

Table 4 Evaluation metrics for different decomposition methods

分解模型		M_{AE}	M_{APE}	R_{MSE}
未分解		619.72	13.96	826.30
一次分解	SSA	458.35	10.64	581.53
	VMD	356.88	8.68	464.72
二次分解	SSA-VMD	233.72	5.71	300.26

由图 9 和表 4 可以得出,未进行信号分解时,模型误差最大,拟合效果最差;进行一次分解后,预测精度有一定程度的提高,其中经 SSA 分解的 M_{AE} 、 M_{APE} 、 R_{MSE} 分别降低了 26.04%、23.78% 和 29.62%,经 VMD 分解后 M_{AE} 、 M_{APE} 、 R_{MSE} 分别降低了 42.41%、37.82% 和 43.76%;进行二次分解后,模型预测精度进一步提高,相比 SSA 单一分解模型 M_{AE} 、 M_{APE} 和 R_{MSE} 分别降低了 49.01%、46.33% 和 48.37%,相比 VMD 单一分解模型分别降低了 34.51%、34.22% 和 35.39%。

由此可见,信号分解可通过形成平稳分量的方式降低数据复杂度,但一次分解后存在的高频分量对模型预测效果仍会产生相应影响,二次分解后的分量更加平稳,可有效缓解该问题,显著提高预测精度。

4 结论

针对天然气负荷数据复杂度高、序列规律性难以捕捉所造成的预测精度不高的问题,本文提出了一种基于二次分解和 RF-BiLSTM 模型的预测方法,通过实验分析所得结论如下:

(1) BiLSTM 模型能够同时学习序列前向和后向的细节,相比主流的深度学习模型,在时序预测中表现效果更佳。

(2) 随机森林算法可挖掘时序数据与特征之间的

关键关联,有效改善了特征冗余所造成的预测精度差的问题。

(3) 二次分解模块可减少高频分量对预测结果的影响,进一步减轻了数据复杂度,提高了预测精度。

综上所述,本文提出的基于二次分解和BiLSTM的多特征预测模型具有优异的拟合效果,可实现更高的天然气负荷预测精度,有效拓展了天然气负荷预测思路。

参考文献(References):

- [1] Tian N, Shao B, Bian G, et al. Application of forecasting strategies and techniques to natural gas consumption: a comprehensive review and comparative study[J]. Engineering Applications of Artificial Intelligence, 2024, 129: 107644.
- [2] 高红燕, 杨艳超, 张曦, 等. 西安市供暖期日燃气负荷预测方法[J]. 干旱气象, 2021, 39(5): 857-863, 870.
Gao Hongyan, Yang Yanchao, Zhang Xi, et al. Forecast method of daily gas load during heating period in Xi'an of Shaanxi Province[J]. Journal of Arid Meteorology, 2021, 39(5): 857-863, 870.
- [3] 杨奕, 刘金源, 陈天民, 等. 基于 PSO-SVR 模型的短期天然气负荷预测[J]. 科学技术与工程, 2023, 23(35): 15210-15216.
Yang Yi, Liu Jinyuan, Chen Tianmin, et al. Short-term natural gas load forecasting based on PSO-SVR model[J]. Science Technology and Engineering, 2023, 23(35): 15210-15216.
- [4] 佟敏, 陈忠源, 党乐, 等. 基于特征筛选 BP 神经网络的天然气需求量预测[J]. 天然气技术与经济, 2022, 16(3): 59-65.
Tong Min, Chen Zhongyuan, Dang (Le | Yue), et al. Predicting the natural-gas demand based on feature selection and BP neural network[J]. Natural Gas Technology and Economy, 2022, 16(3): 59-65.
- [5] Veeramsetty V, Reddy K R, Santhosh M, et al. Short-term electric power load forecasting using random forest and gated recurrent unit[J]. Electrical Engineering, 2022, 104(1): 307-329.
- [6] Imani M. Electrical load-temperature CNN for residential load forecasting[J]. Energy, 2021, 227: 120480.
- [7] 孙庆凯, 王小君, 张义志, 等. 基于 LSTM 和多任务学习的综合能源系统多元负荷预测[J]. 电力系统自动化, 2021, 45(5): 63-70.
Sun Qingkai, Wang Xiaojun, Zhang Yizhi, et al. Multiple load prediction of integrated energy system based on long short-term memory and multi-task learning[J]. Automation of Electric Power Systems, 2021, 45(5): 63-70.
- [8] Wen Y, Pan S, Li X, et al. Highly fluctuating short-term load forecasting based on improved secondary decomposition and optimized VMD[J]. Sustainable Energy, Grids and Networks, 2024, 37: 101270.
- [9] Meng Z, Xie Y, Sun J. Short-term load forecasting using neural attention model based on EMD[J]. Electrical Engineering, 2022, 104(3): 1857-1866.
- [10] Shao B, Yao Z, Qiang Y. Point-interval forecasting for electricity load based on regular fluctuation component extraction[J]. Energies, 2023, 16(4): 1988.
- [11] Wei N, Yin L, Li C, et al. Short-term load forecasting using detrend singular spectrum fluctuation analysis[J]. Energy, 2022, 256: 124722.
- [12] 陈川, 陈冬林, 何李凯. 基于 BPNN-EMD-LSTM 组合模型的城市短期燃气负荷预测[J]. 安全与环境工程, 2019, 26(1): 149-154, 169.
Chen Chuan, Chen Donglin, He Likai. Short-term forecast of urban natural gas load based on BPNN-EMD-LSTM combined model[J]. Safety and Environmental Engineering, 2019, 26(1): 149-154, 169.
- [13] 邓带雨, 李坚, 张真源, 等. 基于 EEMD-GRU-MLR 的短期电力负荷预测[J]. 电网技术, 2020, 44(2): 593-602.
Deng Daiyu, Li Jian, Zhang Zhenyuan, et al. Short-term electric load forecasting based on EEMD-GRU-MLR [J]. Power System Technology, 2020, 44(2): 593-602.
- [14] 那幸仪, 贾俊铖, 赵晓筠, 等. 基于小波变换和 LSTM 模型的城市天然气负荷预测[J]. 计算机应用与软件, 2021, 38(12): 61-66.
(Na/Nuo) Xingyi, Jia Juncheng, Zhao Xiao (Jun | Yun), et al. City natural gas load forecasting based on wavelet transform and LSTM model[J]. Computer Applications and Software, 2021, 38(12): 61-66.
- [15] 穆晨宇, 薛文斌, 穆羨瑛, 等. 基于 VMD-LSTM-Attention 模型的短期负荷预测研究[J]. 现代电子技术, 2023, 46(17): 174-178.
Mu Chenyu, Xue Wenbin, Mu Xianying, et al. Research on short-term load forecasting based on VMD- LSTM-Attention model[J]. Modern Electronics Technique, 2023, 46(17): 174-178.
- [16] Liu F, Liang C. Short-term power load forecasting based on AC-BiLSTM model[J]. Energy Reports, 2024, 11: 1570-1579.
- [17] 甘如美江, 傅杰, 汪正, 等. 改进灰狼算法与机器学习混合模型的时间序列预测[J/OL]. 重庆工商大学学报(自然科学版). <https://link.cnki.net/urlid/50.1155.N.20240628.1614.002>.
Gan Rumeijiang, Fu Jie, Wang Zheng, et al. Time series prediction of improved grey wolf algorithm and machine learning hybrid model [J/OL]. Journal of Chongqing Technology and Business University (Natural Science Edition) <https://link.cnki.net/urlid/50.1155.N.20240628.1614.002>.
- [18] Ma W. Short-term prediction of photovoltaic power generation based on LMD permutation entropy and singular spectrum analysis[J]. Energy Engineering, 2023, 120(7): 1685-1699.
- [19] 张立国, 刘婉, 张淑清, 等. 基于相空间重构改进算法的混沌奇异谱分析及应用[J]. 计量学报, 2021, 42(10): 1299-1306.
Zhang Liguang, Liu Wan, Zhang Shuqing, et al. Chaotic singular spectrum analysis based on improved phase space reconstruction algorithm and its application[J]. Acta Metrologica Sinica, 2021, 42(10): 1299-1306.

责任编辑:代小红