

引文: 孙仁金, 邓钰暄, 李慧慧, 等. 基于滑动时间窗和组合模型结合的中国 LNG 现货价格预测方法[J]. 天然气工业, 2025, 45(3): 170-178.
SUN Renjin, DENG Yuxuan, LI Huihui, et al. A prediction method of China's LNG spot price based on sliding time window and combination model[J]. Natural Gas Industry, 2025, 45(3): 170-178.

基于滑动时间窗和组合模型结合的中国 LNG 现货价格预测方法

孙仁金 邓钰暄 李慧慧 刘子越

中国石油大学(北京)经济管理学院

摘要: 近年来, 中国液化天然气(LNG)的生产量和进口量持续攀升, 成为最重要的天然气供给来源之一。由于 LNG 供给灵活、市场参与主体众多、在产业链中市场化程度相对较高, 科学准确预测 LNG 现货价格能够为市场参与者提供决策参考, 降低市场风险。为此, 建立了基于滑动时间窗以及二次分解思想的变分模态分解(VMD)–自适应噪声完全集合经验模态分解(CEEMDAN)–极限学习机模型(ELM)有机组合预测模型, 并以内蒙古自治区的 LNG 价格数据为例进行实证分析。研究结果表明: ①采用滑动时间窗可以有效提取 LNG 价格序列中用于分析建模的部分, 将训练集数据进行分解、建模以及预测等环节, 避免了待预测的 LNG 价格数据混入其中; 同时可以在建模预测步骤完成后, 舍弃最旧日期的 LNG 价格, 将新一期的价格数据纳入其中, 随时间推移有效把握数据规律, 实现了模型的动态更新。②利用 VMD 对 LNG 价格进行初次分解, 再通过 CEEMDAN 对 VMD 的残差序列展开二次分解, 可以充分提取 LNG 价格的数据信息, 提高预测精度。③将二次模态分解 LNG 价格得到的分量模态序列分别带入 ELM 模型中预测, 再将各价格分量预测结果加和集成得到 LNG 价格预测结果, 可以显著提升价格预测的准确度。结论认为, 该模型可以更好提取序列时频信息, 有效规避了数据泄露问题, 充分利用残差数据, 显著提高了预测精度, 是 LNG 现货价格预测的可行方法与有效手段。

关键词: 滑动时间窗; 机器学习; 二次分解; LNG; 价格预测

中图分类号: TE-9 文献标识码: A DOI: 10.3787/j.issn.1000-0976.2025.03.015

A prediction method of China's LNG spot price based on sliding time window and combination model

SUN Renjin, DENG Yuxuan, LI Huihui, LIU Ziyue

(School of Economics and Management, China University of Petroleum, Beijing 102249, China)

Natural Gas Industry, Vol.45, No.3, p.170-178, 3/25/2025. (ISSN 1000-0976; In Chinese)

Abstract: In recent years, the continuous rising production and import volumes of liquefied natural gas (LNG) in China makes LNG one of the most important sources of natural gas supply. In view that LNG has flexible supply, multiple market participants and a high degree of marketization in the industrial chain, the scientific and accurate prediction of LNG spot price can provide market participants with decision-making reference and reduce market risks. In this paper, an organically combined prediction model of variational modal decomposition (VMD), complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) and extreme learning machine (ELM) based on the sliding time window and the idea of quadratic decomposition is established, and then analyzed with the LNG price data from Inner Mongolia as an example. The following results are obtained. First, the sliding time window can effectively extract some parts of LNG price series used for analysis and modeling, and perform the decomposition, modeling and prediction of training set data, so as to prevent the LNG price data to be predicted from mixing with them. After the modeling and prediction is completed, the earliest LNG price is discarded, and the latest price data is introduced to understand the data laws over time effectively, realizing dynamic model update. Second, VMD is adopted to the primary decomposition of LNG price, and CEEMDAN is employed for the quadratic decomposition of VMD residual sequence, which can sufficiently extract the data information of LNG price, so as to improve the prediction accuracy. Third, the component modal sequence obtained from the quadratic modal decomposition of LNG price is introduced into the ELM model for prediction, and the prediction results of each price component are summed and integrated to get the prediction result of LNG price, which can improve the accuracy of price prediction significantly. In conclusion, this model can better extract the time frequency information of the series, effectively avoid data leakage, and make full use of residual data, so as to improve the prediction accuracy greatly. It is a feasible and effective method to predict LNG spot price.

Keywords: Sliding time window; Machine learning; Secondary decomposition; LNG; Price prediction

基金项目: 国家自然科学基金面上项目“能源绿色转型的路径与优化: 基于生态足迹的视角”(编号: 72273151)。

作者简介: 孙仁金, 1965 年生, 教授, 博士研究生导师; 主要从事能源市场与投资决策、低碳经济与投资环境影响评价方面的研究工作。地址: (102249) 北京市昌平区城北街道府学路 18 号。ORCID: 0000-0001-5591-9581。E-mail: sunrenjin@cup.edu.cn

通信作者: 邓钰暄, 1999 年生, 硕士研究生; 主要从事能源经济管理、能源价格预测方面的研究工作。地址: (102249) 北京市昌平区城北街道府学路 18 号。ORCID: 0009-0006-5072-1158。E-mail: 1665913377@qq.com

0 引言

天然气作为低碳的化石能源，在中国构建新型能源体系的过程中占有十分重要的战略地位^[1]。近年来，随着天然气供应来源的多元化发展，中国液化天然气（LNG）的生产量和进口量持续攀升，成为最重要的天然气供给来源之一^[2]。按照“管住中间、放开两头”的天然气市场改革策略^[3]，2013年6月国家发展和改革委员会发布《关于调整天然气价格的通知》，天然气价格管理由现货环节调整为门站环节，放开海上天然气、页岩气、煤层气的现货价格以及 LNG 的气源价格管制，自此 LNG 现货价格实行市场化定价^[4]。当前，中国的 LNG 现货交易通过挂牌或竞价的形式在各石油天然气交易中心进行^[5]，买卖双方协商一致即得到成交价格^[6]。由于 LNG 供给灵活、市场参与主体众多，在产业链中市场化程度相对较高^[7]。因此，有效预测 LNG 现货价格能够指导生产商合理调整生产计划，帮助市场参与者把握价格波动规律，为政府部门制定能源政策提供实时参考依据，助力社会各方有效应对价格变化、降低市场风险^[8]。

已有的研究方法与理论为 LNG 价格预测奠定了良好的基础和可行的条件，自回归移动平均模型（ARIMA）^[9]、K-means^[10]、小波分析^[11]、动态贝叶斯网络^[12]和灰色模型^[13]等数理统计方法，以及支持向量回归（SVR）^[14]、长短期记忆网络（LSTM）^[15]、卷积神经网络（CNN）^[16]模式序列相似性搜索（APSS）^[17]等神经网络算法，为天然气价格预测研究提供了模型基础。基于单一模型构建组合模型能够有效提升算法的预测精度，如粒子群（PSO）^[18]、特征选择（FS）^[19]、麻雀搜索算法（SSA）^[20]、遗传算法（GA）^[21]等优化算法可以进一步提升机器学习的计算效率，经验模态分解（EMD）、集成经验模态分解（EEMD）^[22]、互补集合经验模态分解（CEEMD）^[23]、自适应噪声完全集合经验模态分解（CEEMDAN）^[24]、变分模态分解（VMD）^[25]等模态分解算法可以将原始 LNG 价格分解为若干复杂程度各异的子序列，从而简化数据分析的难度。

通过梳理发现，现有研究的主要预测思路可以概括为：首先针对原始的价格序列展开模态分解，将原序列分解成不同时间尺度下的本征模态序列，再划分训练集与测试集，将各个模态序列带入机器学习模型得到预测值。然而，现有的分解集成预测模型仍存在一定的局限性，例如在数据预处理过程中预先使用了测试集，导致模型预测精度虚高且无法实际投入使

用；或是未能充分利用残差数据，面对预测难度较大的复杂序列选择直接舍弃，导致模型精度不佳。为此，笔者利用 VMD—CEEMDAN 二次模态分解与极限学习机模型（ELM）相结合，构建 LNG 现货价格预测模型，以滑动时间窗的思想划分预测数据集，构建可以实时更新参数的动态分解集成模型，同时确保建模过程中不会利用待预测的数据，防止数据泄露；在构建预测模型时考虑残差，并基于二次模态分解思想，对 VMD 分解后的高频本征模态序列以及残差序列展开 CEEMDAN 二次分解，进一步提取复杂序列的数据，降低预测难度，提高整体预测准确性。

1 算法原理

构建预测 LNG 现货价格的 VMD—CEEMDAN—ELM 组合模型，其算法原理如下所示。

1.1 滑动时间窗

滑动时间窗是一种在处理时间序列数据时常用的数据分析方法，主要用于实时或近似实时的数据处理场景中。其基本思想是在一段连续的时间窗内对数据进行统计分析或者计算，然后随着时间的推移，这个窗口向前“滑动”一个固定的时间间隔，每滑动一次，就对窗口内捕获到的新数据进行处理，并对新的窗口内的数据进行同样的分析或计算^[26]。

将原始时间序列数据记为 X ，在 $t(i)$ 时刻对应的数据记作 x_i ，在原始时间序列样本 X 中， x_{p+i} 可以通过 $x_p \cdots x_{p+i-1}$ 来表示。假设选取原序列中的后 q 个数据作为测试集，则在第一次迭代中将序列中测试集外的 p 个数据置于时间窗内，利用时间窗内的 p 个数据 x_1, \cdots, x_p 构建预测模型，得到 $t(p+1)$ 时刻的预测值 y_{p+1} ；至下一时刻，更新时间窗，舍去 $t(1)$ 时刻的数据 x_1 ，并将 $t(p+1)$ 时刻的数据 x_{p+1} 加入时间窗，保持时间窗内始终有 p 个数据，利用数据 x_2, \cdots, x_{p+i} 重新构建模型得到 $t(p+2)$ 时刻的预测值 y_{p+2} ，多次迭代直至更新出所需的预测结果序列^[27]。

1.2 模态分解

模态分解能够处理时间序列数据，以信号分解的原理挖掘数据中的信息，常用的模态分解方法包括经验 EMD、EEMD、CEEMDAN 以及 VMD 等。EMD 在分解数据时易产生模态混叠现象^[28]；EEMD 虽然改进了 EMD 混频的问题，却存在产生虚假分量的缺点^[29]；CEEMDAN 在分解的过程中增加了有限次数的自适应高斯白噪声^[30]，能够有效规避 EMD 混频

与 EEMD 虚假分量的问题,有效分解复杂数据,但得到的模态分量数会随着数据复杂程度的不同而变化。VMD 在预设合适分解次数 k 的条件下,便可以实现数据的有效分解,得到固定数目的分量序列^[31],但当原始数据频率过高时,分解次数难以准确设定,便会造成分解不完全或过度分解的问题^[32]。因此,笔者结合 VMD 与 CEEMDAN 的优势,利用 VMD 进行初次分解,再通过 CEEMDAN 对 VMD 的残差序列展开二次分解,充分提取序列的数据信息,以提高预测效果。

1.2.1 VMD 分解

VMD 是一种非平稳信号的自适应分解估计方法,能够将时序数据同步分解为若干个具有固定带宽的模态分量,并且能够评估每个分量的中心频率^[33]。对于每个分量序列, VMD 将其定义为调频函数^[34],表示形式如下:

$$u_k(t) = A_k(t) \cos[\varphi_k(t)] \quad (1)$$

式中 t 表示时间; k 表示分量序列的序号; $u_k(t)$ 表示信号; $A_k(t)$ 表示包络幅值; $\varphi_k(t)$ 表示相位函数。求解 VMD 分量模态:

$$\min \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t. } \sum_{k=1}^K u_k(t) = f(t) \quad (2)$$

式中 j 表示虚数; $\partial_t(\cdot)$ 表示对时间进行偏导数运算; $\delta(t)$ 表示狄利克雷函数; $*$ 表示卷积运算; $\|\cdot\|_2^2$ 表示 L2 范数的平方; $f(t)$ 表示原始时间序列数据; $e^{j\omega_k t}$ 表示 $u_k(t)$ 信号的指数项。

1.2.2 CEEMDAN 分解

CEEMDAN 是一种处理非线性、非平稳信号的有效分析方法^[35]。该算法的原理是通过加入白噪声,利用其频谱分布特性进行多次迭代运算,优化 EMD 的分解效果^[36],在确保分解结果的准确性方面具有显著优势。CEEMDAN 的运算步骤如下列步骤所示。

1) 将白噪声加入原始信号 $x(t)$ 后便形成新信号,利用 EMD 对其进行 n 次分解,便可得到分量序列 IMF_1 与余量序列 $r_1(t)$:

$$IMF_{k+1}(t) = \frac{1}{N} \sum_{n=1}^N IMF_1^i(t) \quad (3)$$

$$r_1(t) = x(t) - IMF_1(t) \quad (4)$$

式中 n 表示分解次数, N 表示分解最高次数, i 表示在原序列中添加白噪声的次数。

2) 继续分解,得到第 k 个余量序列 $r_k(t)$ 和第 $k+1$

个分量模态序列 IMF_{k+1} :

$$r_k(t) = r_{k-1}(t) - IMF_k(t) \quad (5)$$

$$IMF_{k+1}(t) = \frac{1}{N} \sum_{n=1}^N E_1[r_k(t)] + \varepsilon_k E_k[u^i(t)] \quad (6)$$

式中 $E_k[\cdot]$ 表示 EMD 分解; ε_k 表示参数。

3) 重复步骤 2), 直到 RES 序列呈单调函数,原始信号最终分解结果 $z_{\text{end}}(t)$ 计算方法为:

$$z_{\text{end}}(t) = x(t) - \sum_{i=1}^k IMF_i(t) \quad (7)$$

1.3 样本熵

样本熵 (SE) 可以用来衡量时间序列数据复杂度,计算过程简单且抗干扰能力强,在众多领域的数学计算均有应用,采用样本熵来衡量模态分解后各个子序列的本征模态函数的复杂度^[37]。具体的计算步骤如下列步骤所示。

1) 将长度为 n 时间序列向量间的绝对值距离定义为向量间最大距离:

$$d[X(i), X(j)] = \max [|X(i+l) - X(j+l)|] \quad (8)$$

式中 $X(\cdot)$ 表示时间序列向量; $l=1, 2, \dots, m-1$; 且 $1 \leq i \neq j \leq n-m+1$ 。

2) 假设序列之间距离 $d[X(i), X(j)]$ 的阈值为 r , 则对于给定阈值 r 的概率可表示为:

$$B_i^m(r) = \frac{B_i}{n-m+1} \quad (9)$$

$$B^m(r) = \frac{1}{n-m} \sum_{i=1}^{n-m} B_i^m(r) \quad (10)$$

式中 B_i 表示序列 $X(i)$ 之间距离不超过阈值条件的数据个数; m 表示 $B(r)$ 的维度。

3) 最终得到序列的样本熵:

$$SE(m, r) = \lim_{n \rightarrow \infty} \left[-\frac{B^m(r)}{B^{m+1}(r)} \right] \quad (11)$$

1.4 ELM 模型

ELM 属于前馈神经网络学习模型,其结构包括输入层、隐藏层和输出层。由于只有单一的隐藏层,ELM 的计算方法相较于多数机器学习模型具有较低的计算复杂度。此外,ELM 不仅展现出卓越的数值模拟性能和鲁棒性,还具备出色的泛化能力^[38]。

假设 ELM 网络的输入层节点为 N , 隐藏层节点为 L , 输出层节点为 M , 训练样本为 $\{(x_i, y_i)\}^N$, 其中 $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ 为输入变量, $y_i = [y_{i1}, y_{i2}, \dots, y_{im}]^T$ 为输出变量。ELM 模型如式 (12) 所示:

$$Y = H\beta \quad (12)$$

式中 $Y = [y_1, y_2, \dots, y_N]^T$ 表示目标矩阵, H 表示隐藏层输出矩阵, β 表示隐藏层与输出层之间的权重矩阵。

机器学习模型在实际预测过程中需要先将数据进行归一化处理，其计算方法为：

$$X_{\text{map}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (13)$$

式中 X_{map} 表示归一化后的序列； X 表示原始序列； X_{max} 表示原始序列中的最大值； X_{min} 表示原始序列中的最小值。

2 预测模型构建

研究构建的 VMD—CEEMDAN—ELM 模型，利用滑动时间窗方法处理数据，以 VMD—CEEMDAN 分解架构对价格序列进行二次模态分解，并将各子序列带入 ELM 模型，各个分量序列的预测结果加总得到 LNG 现货价格预测值。

2.1 组合模型方法

选取滑动时间窗数据处理方法、VMD—CEEMDAN 二次模态分解架构以及 ELM 模型，旨在对现有研究方法的缺陷进行补充完善，同时进一步提升预测的准确度。

2.1.1 滑动时间窗

现有文献进行分解集成预测时，通常先整体分解数据再划分测试集与训练集，但由于待分解序列的数据量与复杂度会影响模态分解的效果，这种“整体分解”方法使测试集的数据参与分解的过程，干扰分解结果，其构建的模型存在数据泄露的问题，难以在实际预测中应用；同时，根据固定时间段数据训练只能得到静态模型，其参数无法根据价格规律变化而更新。

滑动时间窗能够有效规避上述问题，一方面可以有效提取原始序列中用于分析建模的部分，将训练集数据进行分解、建模以及预测等环节，避免测试集数据混入其中；另一方面可以在此次建模预测步骤完成后，舍弃最旧的数据，将新一期的数据纳入其中，随时间推移有效把握数据规律，实现模型的动态更新。

2.1.2 二次分解

VMD 会产生频率各异的本征模态序列，各本征模态序列之和与原序列并不完全相等，二者之间还存在信息丰富的 RES，但由于 RES 序列频率高、预测难度较大，多数学者在预测时直接将其舍弃，影响预测精度^[39]。

VMD—CEEMDAN 二次模态分解架构则可以在

保证模态序列有效分解的情况下，充分利用数据信息。VMD 在确定中心频率的情况下，可以按频率与数据趋势规律，有效分解时间序列数据，避免 EMD、EEMD 等分解方法造成的模态混叠或者过度分解的现象^[40]。CEEMDAN 擅长分解波动性强，频率高的复杂数据，可以用于处理 VMD 产生的残差序列，一方面可以确保数据的充分利用，另一方面也可以有效提取数据信息，从而提高模型预测精度。

2.2 组合模型构建

考虑到 LNG 现货价格具有非线性与随机性的特征，提出基于模态分解与机器学习模型的 VMD—CEEMDAN—ELM 模型，用于 LNG 现货价格的短期预测。其实现步骤（图 1）具体包括：

1) 设置滑动时间窗数据。在原始序列的基础上划分出测试集，选取测试集数据时刻之外的 n 个数据置于时间窗内，用于训练模型。

2) LNG 现货价格的分解。将 LNG 现货价格经过 VMD 初次分解，得到 k 个模态序列 V_1, V_2, \dots, V_k 以及残差序列，并对频率较高的 V_k 和残差序列展开 CEEMDAN 二次分解，降低复杂度，进一步得到 m 个模态序列 C_1, C_2, \dots, C_m 。

3) 序列的预测与集成。利用经过 VMD 与 CEEMDAN 分解得到的分量序列训练 ELM 模型，得到每个分量序列的预测值，并将各个分量序列预测结果线性叠加，计算出 LNG 现货价格的预测结果。

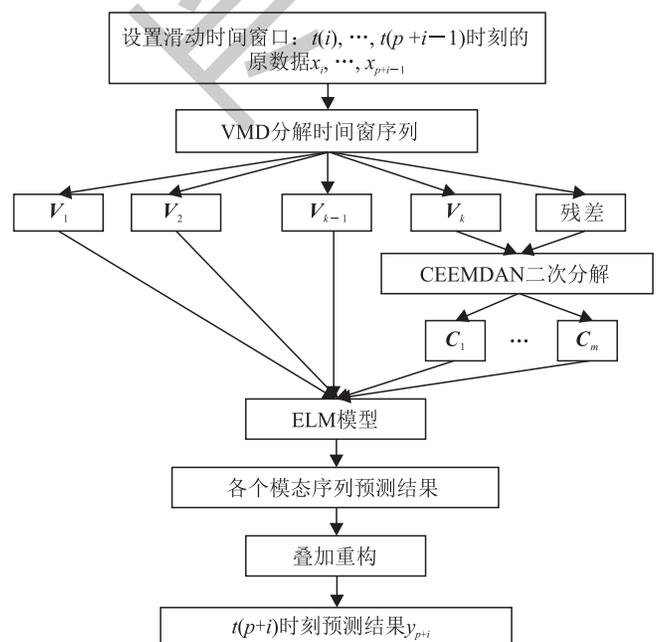


图 1 VMD—CEEMDAN—ELM 模型预测流程图

4) 若需要预测下一时刻的 LNG 现货价格, 则将时间窗向后滑动一个单位 (即在时间窗内加入 1 个新数据, 并舍弃时间最早的 1 个数据, 确保时间窗内的数据量恒为 p), 更新时间窗, 返回步骤 2)。

2.3 误差检验指标

为了评估模型预测的准确性, 采取平均绝对百分比误差 (MAPE)、均方根误差 (RMSE) 以及决定系数 (R^2) 这 3 种指标来综合衡量模型的误差水平, 其计算方法分别见式 (14) ~ (16)。MAPE 是预测误差的无偏统计量, 将离差进行绝对值运算, 并除以数据的真实值, 避免正负相抵的情况, 也消除了量纲对误差大小的干扰, MAPE 越小, 预测误差越小; RMSE 对预测值中的特大或特小误差较为敏感, 能有效体现预测值的精密性, RMSE 越小, 预测误差越小; 决定系数 R^2 可以有效地揭示预测模型的系统误差, 当 R^2 值越趋近于 1 时, 模型的拟合效果越好。

MAPE 的计算公式为:

$$MAPE = \frac{100\%}{N} \times \sum_{i=1}^N \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \quad (14)$$

式中 \hat{Y}_i 表示预测值; Y_i 表示实际值; N 为样本数。

RMSE 的计算公式为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{Y}_i - Y_i)^2} \quad (15)$$

R^2 的计算方法为:

$$R^2 = 1 - \frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y})^2} \quad (16)$$

式中 \bar{Y} 表示实际值的平均值。

3 实证分析

依照上文所构建的模型, 以内蒙古自治区的 LNG 现货价格为例进行预测, 并设置对照试验验证模型的有效性。

3.1 数据来源与描述

内蒙古自治区位于中国北部, 油气资源从东至西遍布全区。其中, 位于鄂尔多斯市境内的中国石

油长庆油田公司苏里格气田是中国目前探明储量规模最大的整装气田, 也是保障京津地区、内蒙古周边区域的主力气源地。由于内蒙古自治区拥有丰富的天然气资源, 其天然气消费受 LNG 进口贸易的影响相对较小, 选取内蒙古自治区可以更加真实地反映国产天然气在国内市场的交易状况。上海石油与天然气交易中心统计了全国范围的 LNG 工厂以及接收站的每日 LNG 成交价格, 其发布的每个区域 LNG 出厂价格指数, 是根据每个省市区内所有 LNG 工厂或者接收站每日 LNG 现货的成交价格以及交易量, 加权计算出的区域平均值, 能够反映各省区市 LNG 现货价格的平均水平^[41], 因此, 选取上海石油与天然气交易中心发布的内蒙古自治区的 LNG 出厂价格指数来代表该区域 LNG 现货价格水平。数据时间范围从 2016 年 1 月 1 日—2023 年 12 月 31 日, 数据来源为 Wind 数据库, 如图 2 所示。

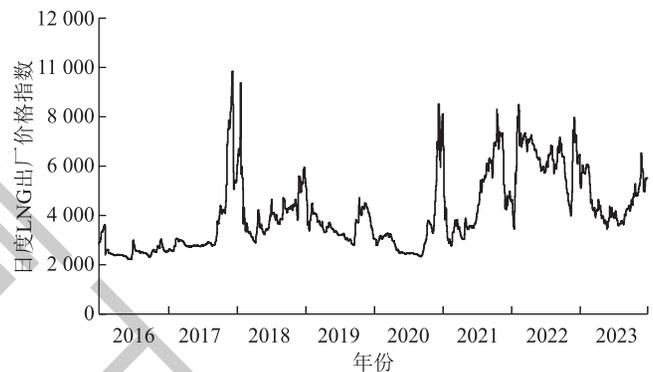


图 2 2016—2023 年内蒙古自治区日度 LNG 出厂价格指数图

研究区域 LNG 现货价格的统计性描述如表 1 所示, P 表示时间序列数据通过 BDS 检验的概率。价格指数极差为 7 158.00, 标准差为 1 507.67, 序列波动剧烈, 且 BDS 检验在 0.05 水平下显著, 说明序列非线性关系较强, 复杂度较高。

在笔者所应用的机器学习预测模型中, 以当日为基期, 设定输入的时间步长为 15, 输入值为包括当日数据在内的前 15 期历史数据, 由于 LNG 现货交易通常每周连续交易 5 d, 预测的滞后期数设为 5。选取 2023 年 6 月 1 日—12 月 29 日的 147 个数据用于测试模型性能, 1 957 个有效数据中除去测试集还剩余 1 810 个数据, 所以滑动时间窗口的长度设为 1 810。

表 1 LNG 出厂价格指数基本统计量描述表

价格指数最大值	价格指数最小值	价格指数极差	均值	标准差	峰度	偏度	BDS 检验 P 值
9 386.00	2 228.00	7 158.00	4 105.39	1 507.67	0.147	0.96	小于 0.01

3.2 序列的分解与重构

VMD 与 CEEMDAN 在数据的模态分解方面各有优势，在设置合适值的情况下，采取 VMD 分解方法可以得到固定数量的模态分量序列，防止时间窗口内数据复杂度差异导致模态分量个数的变化；CEEMDAN 则可以有效分解高频数据，充分提取序列信息，因此本文选择利用 VMD 进行一次分解，再针对高频序列和残差序列进行 CEEMDAN 二次分解。

3.2.1 VMD 分解

首先采用 VMD 对 LNG 现货价格进行一次分解。VMD 在使用过程中需要预设分解个数 k ，为了避免分解不充分或者过度分解的现象，利用中心频率的方法确定 k 值，即在不同的 k 值下测定 VMD 分解得到各分量中心频率的大小（表 2）。

表 2 不同 k 值下模态分类序列对应的各分量中心频率表

模态分量序列	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
V_1	0.04	0.02	0.02	0.02	0.02	0.02	0.02
V_2	13.22	6.34	5.15	4.50	4.20	4.03	3.83
V_3	—	27.39	20.97	15.96	14.29	13.46	12.17
V_4	—	—	181.30	42.67	32.80	28.92	24.59
V_5	—	—	—	201.76	63.97	56.74	47.24
V_6	—	—	—	—	223.92	179.65	87.37
V_7	—	—	—	—	—	225.57	180.72
V_8	—	—	—	—	—	—	225.82

从 $k=7$ 开始，复杂度最高的两序列显示出相近的频率，出现模态混叠现象，所以经过计算中心频率最终确定 $k=6$ ，其余参数均为默认值，分解得到 6 个本征模态序列与一个残差序列，依次记作 V_1, V_2, \dots, V_6 与残差。条件阈值设为序列标准差的 0.2 倍， $m=2$ ，计算 V_1, V_2, \dots, V_6 与残差序列的样本熵结果如表 3 所示。

表 3 VMD 分解得到模态分量的样本熵表

模态分量序列	样本熵
V_1	0.016 8
V_2	0.098 2
V_3	0.224 8
V_4	0.299 8
V_5	0.402 2
V_6	0.596 3
残差	0.479 2

3.2.2 CEEMDAN 二次分解

经过 VMD 分解的序列中， V_6 和残差的值较大，其序列复杂程度也偏高，为了降低序列的复杂程度，提高预测准确性，利用 CEEMDAN 对 V_6 和残差进行二次分解，分解后所有模态序列图像及其样本熵如表 4 所示。经过 CEEMDAN 二次分解后，得到的子序列样本熵均小于 V_6 和残差序列，这也证明了二次分解能够有效降低 V_6 和残差序列的复杂度，从而提高预测的准确性。

表 4 CEEMDAN 分解得到各模态分量的样本熵表

模态分量序列	样本熵	模态分量序列	样本熵
C_1	0.384 6	C_7	0.057 7
C_2	0.273 4	C_8	0.046 1
C_3	0.188 1	C_9	0.025 0
C_4	0.168 8	C_{10}	0.019 6
C_5	0.109 7	C_{11}	0.012 1
C_6	0.045 5	C_{12}	0.002 8

3.3 模型预测结果对比分析

按照前文 2.1 的步骤所示，先后基于 VMD 与 CEEMDAN 模态分解方法对 LNG 现货价格指数序列进行二次分解，以便充分提取数据中的频率信息；然后将分解后的众多模态分量序列代入 ELM 模型预测。同时为了验证本文方法的有效性，另外设置了另外 5 组对照实验，分别是未经模态分解的 ELM 模型，以及经过单次模态分解的 EMD—ELM、EEMD—ELM、CEEMDAN—ELM 与 VMD—ELM 模型。为了衡量模型的稳健性，分别对比预测滞后 1 期、3 期与 5 期情况下，6 种预测方法的预测误差分析结果（表 5）。

经仿真验证，在预测滞后 1 期、3 期与 5 期情况下，对比上述 6 种模型的精度与误差，可以得出以下结果：

1) 对比单一模型与经过模态分解模型的预测结果，单一模型预测误差较大，预测值在真实值的附近波动幅度也更为剧烈，说明机器学习模型虽然能够拟合出序列的非线性波动特征，但在复杂序列预测中容易出现过拟合的现象；而且相比于经过模态分解后的模型，单一模型预测结果表现出显著的滞后性，说明单独使用机器学习模型在提取不同时间尺度的数据信息方面，存在一定的局限性。引入模态分解后可以将原数据按照不同的时间尺度划分，降低预测难度，改进单一模型精度较低的问题。

2) 对比一次分解中，不同模态分解方法对应

表 5 不同模型预测误差分析表

预测滞后期数	模型	R^2	MAPE	RMSE
1 期	ELM	0.878 8	0.041 6	243.421 6
	EMD—ELM	0.926 0	0.027 4	162.501 2
	EEMD—ELM	0.938 7	0.023 3	160.332 5
	CEEMDAN—ELM	0.943 7	0.023 7	159.912 5
	VMD—ELM	0.944 5	0.023 3	157.144 7
	VMD—CEEMDAN—ELM	0.950 2 ^①	0.023 0 ^①	155.982 1 ^①
3 期	ELM	0.849 9	0.045 0	270.926 1
	EMD—ELM	0.890 3	0.039 5	231.577 2
	EEMD—ELM	0.908 5	0.029 5	211.570 0
	CEEMDAN—ELM	0.927 9	0.028 5	187.781 2
	VMD—ELM	0.941 4	0.026 0	169.275 7
	VMD—CEEMDAN—ELM	0.947 0 ^①	0.023 9 ^①	160.917 1 ^①
5 期	ELM	0.749 7	0.056 5	349.822 2
	EMD—ELM	0.874 5	0.041 6	247.770 6
	EEMD—ELM	0.883 3	0.036 4	238.845 2
	CEEMDAN—ELM	0.885 5	0.036 3	236.587 0
	VMD—ELM	0.892 6	0.036 0	229.193 3
	VMD—CEEMDAN—ELM	0.940 8 ^①	0.027 2 ^①	170.166 1 ^①

注：①最优预测结果。

的预测结果，VMD 和 CEEMDAN 的分解误差较小，说明模态分解能够将原始序列分解成不同时间尺度的数据，而不同模态分解方法的计算思路不同，分解效果也各有差异，VMD 和 CEEMDAN 则能够避免模态混叠或者过度分解的现象。选择合适的模态分解方法，能够充分提取数据信息，从而提高预测精度。

3) 对比一次分解与二次分解模型的预测结果，二次分解模型的误差更低，预测结果也更加贴近实际数据。原序列经过模态分解得到复杂度不同的子序列，而当序列复杂度升高，预测难度也随之加大，采用两种不同计算思路的模态分解方法，在对原序列进行初次分解的基础上再对高频序列与残差序列展开二次分解，降低高频序列与残差序列的复杂度，在单次模态分解的基础上进一步优化预测结果。

4) 对比不同预测滞后期数的结果，发现随着预测滞后期数的增加，未经模态分解的单一模型预测效果不断下降，而采取分解技术的模型却能保持较高的预测精度，二次分解的 VMD—CEEMDAN—ELM 模型预测效果在不同预测期数下均表现最佳，说明将原数据分解为不同时间尺度，并充分提取复杂序列中的数据信息能够使预测模型同时具有较好的准确度与稳定性。

4 结论

1) 构建滑动时间窗框架，一是可以有效规避数据泄露问题，建模预测过程中避免使用未来的数据，确保预测结果的真实性；二是可以随时间调整建模数据，动态更新模型参数，使模型和预测结果更加贴近最新的数据规律，提高预测准确度。

2) 原序列经过 VMD 进行初次分解后，再利用 CEEMDAN 对复杂度较高的高频序列与 RES 序列进行二次分解，可以确保数据信息的全面使用，并充分提取其中的重要信息，降低序列复杂度。随着预测滞后期数的增加，单一模型与单次分解模型预测误差明显增加，而 VMD—CEEMDAN—ELM 模型经过二次分解，预测精度最高，且预测效果始终较为稳定。

3) 利用 ELM 模型分别预测 VMD—CEEMDAN 二次分解得到的各个分量序列，再将结果加和得到价格的最终预测结果，预测精度较高，说明 ELM 模型在预测非线性数据方面具有优势。

在天然气价格市场化改革的背景之下，LNG 现货价格的市场化程度将逐渐提高，其变化趋势也将越来越复杂，总体来看，研究提出的预测模型在中短期预测中的稳定性较强，可以为 LNG 现货价格的预测提供参考。

参 考 文 献

- [1] 张颀, 何春蕾, 杜波, 等. 对中国天然气价格改革的回顾与政策建议[J]. 天然气技术与经济, 2023, 17(3): 1-9.
ZHANG Yong, HE Chunlei, DU Bo, et al. Review and policy suggestions on China's natural-gas price reform[J]. Natural Gas Technology and Economy, 2023, 17(3): 1-9.
- [2] 魏恒姝, 马远. “一带一路”倡议对中国天然气价格形成机制的影响[J]. 价格月刊, 2023(8): 18-23.
WEI Hengshu, MA Yuan. The impact of the Belt and Road on the formation mechanism of China's natural gas price[J]. Prices Monthly, 2023(8): 18-23.
- [3] 张颀. 我国天然气产供储销价格体系: 现状、问题与建议[J]. 天然气工业, 2019, 39(1): 131-137.
ZHANG Yong. A price-system network of natural gas production, supply, storage, sales in China: Present status, problems and proposals[J]. Natural Gas Industry, 2019, 39(1): 131-137.
- [4] 周娟, 魏微, 胡奥林, 等. 深化中国天然气价格机制改革的思考[J]. 天然气工业, 2020, 40(5): 134-141.
ZHOU Juan, WEI Wei, HU Aolin, et al. On deepening the reform of China's natural gas price mechanism[J]. Natural Gas Industry, 2020, 40(5): 134-141.
- [5] 白俊. 以竞争为标尺重构中国天然气价格市场化改革[J]. 天然气工业, 2020, 40(5): 117-125.
BAI Jun. Reconstructing China's natural gas price marketization reform based on competition[J]. Natural Gas Industry, 2020, 40(5): 117-125.
- [6] 苏倩, 纪海燕, 万方. 中国海上天然气价格管理及定价机制[J]. 天然气技术与经济, 2021, 15(4): 57-62.
SU Qian, JI Haiyan, WAN Fang. Price management and pricing mechanism for China's offshore natural gas[J]. Natural Gas Technology, 2021, 15(4): 57-62.
- [7] 付俊涛, 李鹏, 李伟. 中国天然气价格市场化改革进展分析[J]. 国际石油经济, 2019, 27(9): 17-24.
FU Juntao, LI Peng, LI Wei. Analysis on the progress of natural gas price marketization reform in China[J]. International Petroleum Economics, 2019, 27(9): 17-24.
- [8] 闫涵, 金乃婧, 刘炳赫. 中国国内天然气价格承受力研究[J]. 当代石油石化, 2022, 30(7): 20-25.
YAN Han, JIN Naijing, LIU Binghe. Research on natural gas price bearing capacity in China[J]. Petroleum & Petrochemical Today, 2022, 30(7): 20-25.
- [9] 李梦祎, 孔盈皓. 基于时间序列模型的天然气价格预测[J]. 油气与新能源, 2023, 35(1): 61-66.
LI Mengyi, KONG Yinghao. Prediction of natural gas price based on time series model[J]. Petroleum and New Energy, 2023, 35(1): 61-66.
- [10] ALAM M S, MURSHED M, MANIGANDAN P, et al. Forecasting oil, coal, and natural gas prices in the pre-and post-COVID scenarios: Contextual evidence from India using time series forecasting tools[J]. Resources Policy, 2023, 81: 103342.
- [11] 姜鑫, 乔佳, 王勋, 等. 基于小波分析的天然气价格预测[J]. 煤气与热力, 2021, 41(6): 34-37.
JIANG Xin, QIAO Jia, WANG Xun, et al. Prediction of natural gas price based on wavelet analysis[J]. Gas & Heat, 2021, 41(6): 34-37.
- [12] 史惠婷, 柴建, 卢全莹, 等. 北美天然气现货价格波动机制分析及波动率预测[J]. 系统工程理论与实践, 2021, 41(12): 3366-3377.
SHI Huiting, CHAI Jian, LU Quanying, et al. Analysis on the fluctuation mechanism of natural gas spot price in North America and its volatility prediction[J]. Systems Engineering—Theory & Practice, 2021, 41(12): 3366-3377.
- [13] 田红丽, 李成群, 闫会强, 等. 基于相关系数和灰色模型的振荡序列预测[J]. 计算机工程与设计, 2020, 41(2): 557-561.
TIAN Hongli, LI Chengqun, YAN Huiqiang, et al. Prediction of oscillating sequence based on correlation coefficient and grey model[J]. Computer Engineering and Design, 2020, 41(2): 557-561.
- [14] WANG Jianliang, LEI Changran, GUO Meiyu. Daily natural gas price forecasting by a weighted hybrid data-driven model[J]. Journal of Petroleum Science and Engineering, 2020, 192: 107240.
- [15] OYEWOLA D O, AKINWUNMI S A, OMOTEHINWA T O. Deep LSTM and LSTM-attention q-learning based reinforcement learning in oil and gas sector prediction[J]. Knowledge-Based Systems, 2024, 284: 111290.
- [16] 裴莹, 李天祥, 王麇清, 等. 基于新闻的国际天然气价格趋势预测方法[J]. 计算机科学, 2021, 48(S1): 235-239.
PEI Ying, LI Tianxiang, WANG Aoqing, et al. Prediction method of international natural gas price trends based on news[J]. Computer Science, 2021, 48(S1): 235-239.
- [17] 王建良, 雷昌然. 基于数据挖掘技术的天然气价格预测方法研究[J]. 中国矿业, 2020, 29(2): 52-58.
WANG Jianliang, LEI Changran. Research on the forecasting method for natural gas price based on the data mining technique[J]. China Mining Magazine, 2020, 29(2): 52-58.
- [18] WANG Jun, CAO Junxing, YUAN Shan, et al. Short-term forecasting of natural gas prices by using a novel hybrid method based on a combination of the CEEMDAN-SE- and the PSO-ALS-optimized GRU network[J]. Energy, 2021, 233: 121082.
- [19] ČEPERIĆ E, ŽIKOVIĆ S, ČEPERIĆ V. Short-term forecasting of natural gas prices using machine learning and feature selection algorithms[J]. Energy, 2017, 140, Part 1: 893-900.
- [20] 胡志瑞, 赵万伏, 宋垠先, 等. 基于改进麻雀搜索算法优化 BP 神经网络的土壤有机质空间分布预测[J]. 环境科学, 2024, 45(5): 2859-2870.
HU Zhirui, ZHAO Wanfu, SONG Yinxian, et al. Prediction spatial distribution of soil organic matter based on improved BP neural network with optimized sparrow search algorithm[J]. Environmental Science, 2024, 45(5): 2859-2870.
- [21] 姜立超, 尚晓兵, 金豹, 等. 基于遗传算法-v 支持向量回归的船舶轨迹预测[J]. 哈尔滨工程大学学报, 2024, 45(10): 2001-2006.
JIANG Lichao, SHANG Xiaobing, JIN Bao, et al. Ship trajectory prediction based on genetic algorithm-v support vector regression[J]. Journal of Harbin Engineering University, 2024, 45(10): 2001-2006.
- [22] 刘克天, 胡新宇, 李想, 等. 基于 EMD 和 TLS-Prony 的次同

- 步振荡模态辨识方法及电压反馈抑制策略[J]. 太阳能学报, 2023, 44(11): 294-302.
- LIU Ketian, HU Xinyu, LI Xiang, et al. Sub-synchronous oscillation modal identification method based on EMD and TLS-Prony algorithm and voltage feedback suppression strategy[J]. *Acta Energetica Solaris Sinica*, 2023, 44(11): 294-302.
- [23] 张金良, 刘子毅, 王明雪. 基于 CEEMD—ELM—ARIMA 的天然气价格预测模型研究[J]. 天然气与石油, 2021, 39(4): 129-136.
- ZHANG Jinliang, LIU Ziyi, WANG Mingxue. Research on natural gas price forecasting model based on CEEMD—ELM—ARIMA[J]. *Natural Gas and Oil*, 2021, 39(4): 129-136.
- [24] NIU Hongli, XU Kunliang, LIU Cheng. A decomposition-ensemble model with regrouping method and attention-based gated recurrent unit network for energy price prediction[J]. *Energy*, 2021, 231: 120941.
- [25] LIN Yu, LU Qin, TAN Bin, et al. Forecasting energy prices using a novel hybrid model with variational mode decomposition[J]. *Energy*, 2022, 246: 123366.
- [26] 郭基联, 张保山, 周章文, 等. 基于双自适应滑动时间窗滚动轴承故障预测模型[J]. 空军工程大学学报, 2023, 24(4): 1-7.
- GUO Jilian, ZHANG Baoshan, ZHOU Zhangwen, et al. A fault prediction model for rolling bearing based on double adaptive sliding time window[J]. *Journal of Air Force Engineering University*, 2023, 24(4): 1-7.
- [27] 范丽伟, 董欢欢, 渐令. 基于滚动时间窗的碳市场价格分解集成预测研究[J]. 中国管理科学, 2023, 31(1): 277-286.
- FAN Liwei, DONG Huanhuan, JIAN Ling. A decomposition ensemble model with sliding time window for forecasting carbon market prices[J]. *Chinese Journal of Management Science*, 2023, 31(1): 277-286.
- [28] QIN Quande, XIE Kangqiang, HE Huangda, et al. An effective and robust decomposition-ensemble energy price forecasting paradigm with local linear prediction[J]. *Energy Economics*, 2019, 83: 402-414.
- [29] 刘聪, 钱坤, 焦淮, 等. 用于损伤状态识别的极值延拓 EMD 和 LS-SVM 算法[J]. 计算机应用, 2023, 43(S2): 256-260.
- LIU Cong, QIAN Kun, JIAO Zhun, et al. Damage state identification algorithm based on extreme value extension empirical mode decomposition and least squares support vector machine[J]. *Computer Applications*, 2023, 43(S2): 256-260.
- [30] 王文川, 杜玉瑾, 和吉, 等. 基于 CEEMDAN—VMD—BP 模型的月径流量预测研究[J]. 华北水利水电大学学报(自然科学版), 2023, 44(1): 32-40.
- WANG Wenchuan, DU Yujin, HE Ji, et al. Research on monthly runoff prediction based on CEEMDAN—VMD—BP model[J]. *Journal of North China University of Water Resources and Electric Power (Natural Science Edition)*, 2023, 44(1): 32-40.
- [31] LI Jinchao, WU Qianqian, TIAN Yu, et al. Monthly henry hub natural gas spot prices forecasting using variational mode decomposition and deep belief network[J]. *Energy*, 2021, 227: 120478.
- [32] FU Wenlong, ZHANG Kai, WANG Kai, et al. A hybrid approach for multi-step wind speed forecasting based on two-layer decomposition, improved hybrid DE—HHO optimization and KELM[J]. *Renewable Energy*, 2021, 164: 211-229.
- [33] 肖烈禧, 张玉, 周辉, 等. 基于 IAOA—VMD—LSTM 的超短期风电功率预测[J]. 太阳能学报, 2023, 44(11): 239-246.
- XIAO Liexi, ZHANG Yu, ZHOU Hui, et al. Ultra short term wind power prediction based on IAOA—VMD—LSTM[J]. *Acta Energetica Solaris Sinica*, 2023, 44(11): 239-246.
- [34] LI Ranran, SONG Xiao. A multi-scale model with feature recognition for the use of energy futures price forecasting[J]. *Expert Systems With Applications*, 2023, 211: 118622.
- [35] 肖成, 曹万鹏, 褚越强, 等. CEEMDAN 联合自适应小波阈值算法的 GA-BP 风电发电机故障预测[J]. 可再生能源, 2024, 42(10): 1332-1340.
- XIAO Cheng, CAO Wanpeng, CHU Yueqiang, et al. CEEMDAN joint adaptive wavelet thresholding algorithm for GA-BP wind turbine fault prediction[J]. *Renewable Energy Resources*, 2024, 42(10): 1332-1340.
- [36] 冬鑫, 陈秋雨, 高俊, 等. 基于 CEEMDAN 分解的风电功率预测方法[J]. 科学技术创新, 2024(22): 89-92.
- DONG Xin, CHEN Qiuyu, GAO Jun, et al. Wind power prediction method based on CEEMDAN decomposition[J]. *Scientific and Technological Innovation*, 2024(22): 89-92.
- [37] 毕贵红, 黄泽, 赵四洪, 等. 基于混合分解和 PCG—BiLSTM 的风速短期预测[J]. 太阳能学报, 2024, 45(1): 159-170.
- BI Guihong, HUANG Ze, ZHAO Sihong, et al. Short-term prediction of wind speed based on hybrid decomposition and PCG—BiLSTM[J]. *Acta Energetica Solaris Sinica*, 2024, 45(1): 159-170.
- [38] 丁同, 傅晓锦. 基于改进粒子群优化算法和极限学习机的锂离子电池健康状态预测[J]. 国外电子测量技术, 2023, 42(5): 163-173.
- DING Tong, FU Xiaojin. Improved particle swarm optimization and an extreme learning machine are used to predict the health state of lithium-ion batteries[J]. *Foreign Electronic Measurement Technology*, 2023, 42(5): 163-173.
- [39] ZHAN Linjie, TANG Zhenpeng. Natural gas price forecasting by a new hybrid model combining quadratic decomposition technology and LSTM model[J]. *Mathematical Problems in Engineering*, 2022, 2022(1): 5488053.
- [40] YANG Kailing, ZHANG Xi, LUO Haojia, et al. Predicting energy prices based on a novel hybrid machine learning: Comprehensive study of multi-step price forecasting[J]. *Energy*, 2024, 298: 131321.
- [41] 肖建忠, 王璇. 中国液化天然气现货价格的传导机制[J]. 天然气工业, 2019, 39(11): 117-125.
- XIAO Jianzhong, WANG Xuan. The analysis of conduction mechanism of China's LNG spot price[J]. *Natural Gas Industry*, 2019, 39(11): 117-125.

(修改回稿日期 2025-01-18 编辑 陈 嵩)



本
文
互
动