

基于 CEEMDAN-VMD 二次分解的原油期货价格预测

姜玉淳

贵州财经大学，贵州贵阳，550025；

摘要：本文聚焦于原油期货价格预测，针对原油期货价格时间序列数据的非平稳性、非结构化及非线性特征，提出了一种基于二次分解技术的预测方法。借鉴深度学习在其他金融领域的成功经验，本文提出 CEEMDAN-SE-VMD-GRU 模型，对上海原油 SC、Brent 原油和 WTI 原油价格日度数据进行深度分解与精细处理，有效捕捉价格波动的微观特征，以实现更高精度的预测，并且模型在原油期货价格的预测精度优于对比模型。

关键词：原油期货价格；金融时间序列预测；二次分解；GRU

DOI：10.69979/3029-2700.24.8.038

传统时间序列预测模型如 ARIMA、VAR 等，受限于线性变换假设，难以准确捕捉原油期货价格等具有非线性、非平稳性及多重周期性特征的数据。随着机器学习与深度学习技术的发展，尤其是与信号分解技术的结合，为提升预测精度提供了新途径。经验模态分解 EMD 及其改进模型 CEEMDAN，以及变分模态分解 VMD，有效解决了信号分解中的模态混叠、重构误差和降噪问题^{[1][2][3][4]}。近几年研究发现通过结合多种分解算法可以进一步降低时间序列的复杂性和非平稳性，显著提高了预测能力。例如，结合 STL 和 VMD、VMD 和 SSA 等两阶段分解策略^[5]，并与深度学习算法如 GWO-SVR、LSTM 和 ELM 等相结合，已在电力需求、风速等非金融领域实现了高精度预测^{[6][7][8]}。这些研究验证了二次分解与深度学习模型融合预测的优势，为金融时间序列预测带来了新的突破。

基于二次分解技术结合深度学习模型在多个领域展现出了卓越的预测能力，为原油期货价格预测提供了新的思路和方法。本文提出 CEEMDAN、SE 和 VMD 的二次分解方法，将原油期货价格时间序列分解为一系列 IMF，以充分降低时间序列的非平稳性，同时避免有效信息的丢失；然后，利用 GRU 神经网络构建预测模型并进行拟合验证预测效果。

1 模型建立

1.1 数据与方法

本文运用数据选取 2018 年 3 月 26 至 2024 年 8 月 19 的上海 SC 原油期货收盘价、Brent 原油期货价格、WTI 源于期货价格的日度价格数据。基于 Tushare 数据平台使用 python 对数据进行下载与处理。然后对原始数据进行 CEEMDAN 分解，根据熵值进行聚类整合，将 IMF

分量划分为高频率序列、中频率序列和低频率序列。对高频数据进行 VMD 二次分解，矩阵输入 GRU 预测模型中，中频率序列和低频率序列向量输入到预测模型中，得到最终预测结果。

1.2 模型构建

1.2.1 分解算法

CEEMDAN 算法是 EMD 的改进算法。CEEMDAN 通过引入自适应噪声和集合平均技术改进 EMD 分解，解决模态混叠、端点效应和分解不稳定问题。CEEMDAN 分解通过引入自适应噪声和迭代更新过程，提高了分解的自适应性。有效抑制模态混叠现象，提高分解精度。

VMD 采用交替方向乘子法优化算法来求解所构建的变分问题。通过不断迭代优化，逐步逼近最优解，从而精确地提取出各个模态函数的中心频率和带宽，实现了对信号的有效分解。

1.2.2 GRU

GRU 在 LSTM 的基础上进行了简化，引入了更少的参数量和结构复杂度，同时保持了较高的性能。GRU 通过使用门控机制有效解决了传统 RNN 存在的梯度消失和梯度爆炸问题，特别适合处理长时间依赖的数据。其核心组件包括更新门和重置门，前者决定当前时间步的隐藏状态需要保留多少前一个时间步的信息。后者决定前一时间步的隐藏状态在多大程度上被忽略。

2 实验设计

2.1 实验流程

CEEMDAN-SE-VMD-GRU 模型对原油期货价格预测的框架包括数据预处理与分解、样本熵计算与聚类整合、

高频序列预测、中低频序列预测、模型评估四个部分。首先，对数据预处理与分解对原油期货价数据进行 CEEMDAN 分解，获得一系列 IMF 分量。然后根据样本熵计算与聚类整合计算上述 IMF 分量的样本熵，并根据熵值进行聚类整合，将 IMF 分量划分为高频数据 Co-IMF0、中频序列 Co-IMF1 以及低频序列 Co-IMF2。对高频数据 Co-IMF0 进行 VMD 分解，以进一步分解其频率成分。得到 VMD 分解结果，对 Co-IMF0 矩阵输入 GRU 预测，得到评价指标。Co-IMF1 和 Co-IMF2 应用 GRU 模型对分解后的低频序列进行预测，得到预测结果和评价指标。最后计算 CEEMDAN-SE-VMD-GRU 模型的评价指标。

2.2 模型参数与评价指标

实验环境在 Windows11 系统下进行，编程语言为 Python 3.11，实验中的算法均在 VScode 中进行。文中机器学习部分选用 Scikit-learn，深度学习部分选用 Keras，GRU 包含三层，每层分别包含 128、64、32 个神经元。所有 GRU 层使用 tanh 作为激活函数。在每层 GRU 之后添加了 Dropout 层，损失率为 0.2。Dense 全连接层使用 tanh 函数。模型初始学习率为 0.001，优化器为 adam，使用 RMSE 作为损失函数。

采用 RMSE、MAPE、MAE 以及 R^2 这四个评估指标考察模型性能。RMSE、MAPE、MAE 数值越小拟合效果越好， R^2 反应模型拟合优度，数值越接近 1 越好。

3 模型分析与对比

3.1 模型预测结果分析

采用 CEEMDAN 方法对 Brent 原油期货价格时间序列进行分解，获得 8 个本征模态函数（IMF0-IMF7）及 1 个残余项（IMF8）。高频 IMF 分量因含较多噪声而预测难度较高。通过 SE 聚类分析，将 IMF 分量分为三类：IMF0-IMF1 为高频（聚类 0），IMF2 为中频（聚类 2），IMF3-IMF7 为低频（聚类 1）。

随后，整合不同频率序列为 co-imf0（高频）、co-imf1（中频）和 co-imf2（低频）。对高频序列 co-imf0 应用 VMD 方法进一步分解降噪，设定参数 $K=10$ 。最后，将处理后的高频序列输入 GRU 模型，中频和低频序列输入 GRU 模型，以获取最终预测结果。

表 1 展示了 CEEMDAN-SE-VMD-GRU 模型预测上海原油 SC、Brent 原油和 WTI 原油价格的指标结果。WTI 原油的 MAPE 和 MAE 最低，预测准确性最高；Brent 原油次

之；SC 原油相对较高，但三者差异均在可接受范围内。RMSE 显示相同趋势，WTI 最低，SC 最高，表明模型预测 WTI 最稳定。所有原油的 R^2 均超 0.90，模型拟合效果好，其中 SC 原油 R^2 最高达 0.9829，模型对其价格解释能力最强。

表 1 模型预测结果

模型	指标	CEEMDAN-SE-VMD-GRU
Brent	MAPE	0.0071
	MAE	0.5614
	RMSE	0.7235
	R2	0.9770
WTI	MAPE	0.0028
	MAE	1.1166
	RMSE	1.4275
	R2	0.9163
SC	MAPE	0.0049
	MAE	2.9896
	RMSE	3.8868
	R2	0.9829

3.2 模型对比分析

为了验证二次分解的深度学习模型的对上海原油 SC、Brent 原油和 WTI 原油价格预测的表现，将 CEEMDAN-LSTM、VMD-LSTM 与 CEEMDAN-SE-VMD-GRU 的预测表现进行对比，采用 R^2 衡量模型的预测效果（图 1）。

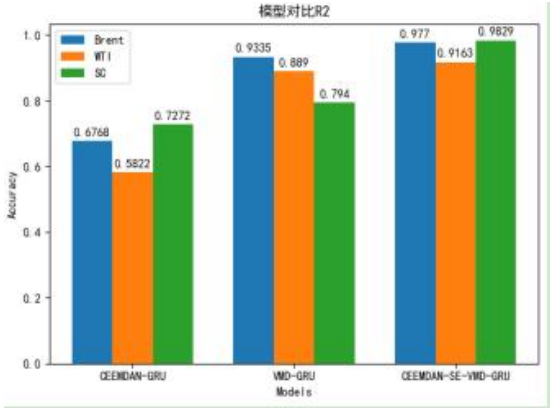


图 1 模型对比结果

CEEMDAN-SE-VMD-GRU 模型在预测 Brent、WTI 和 SC 原油期货价格上表现最佳，优于其他两个模型。结合 CEEMDAN、SE 机制、VMD 二次分解与 GRU 网络，显著提升了预测的准确性和稳定性。VMD-GRU 模型也展现出较强性能，尤其在 Brent 和 WTI 原油预测中， R^2 分别达到 0.9335 和 0.8890，优于 CEEMDAN-GRU 模型，验证了 VMD 在信号分解和特征提取的有效性及 GRU 网络处理时间序列的优势。相比之下，CEEMDAN-GRU 模型预测准确率较低，尤其在 WTI 原油预测中 R^2 仅为 0.5822，表明仅靠

GRU 网络可能不足以准确捕捉和利用时间序列信息。

4 结论

二次分解的预测性能优于一次分解的预测性能。CEEMDAN-SE-VMD-GRU 模型在预测 Brent 原油、WTI 原油和 SC 价格方面具有最高的准确率。CEEMDAN 和 VMD 分解技术的引入,成功地提取了时间序列数据中的有用信息,并有效减少了噪声和冗余成分,为 GRU 模型的精准预测提供了有效的输入特征。同时,二次分解算法与深度学习技术的结合,进一步提升了模型的预测精度。本研究提出的 CEEMDAN-SE-VMD-GRU 模型在原油期货价格预测方面展现出了极高的预测性能和实用价值,为原油期货价格市场的风险管理、投资决策等提供了有力的支持。

参考文献

- [1]Huang N E, Zheng S, Steven R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proceedings: Mathematical, Physical and Engineering Sciences, 1998,454(1971).
- [2]Yeh J R, Shieh J S, Huang N E. Complementar

y ensemble empirical mode decomposition: A novel noise enhanced data analysis method[J]. Advances in adaptive data analysis,2010(2):135-156.

[3]Dragomiretskiy K, Zosso D, Bertozzi A L, et al. Two-Dimensional Compact Variational Mode Decomposition[J]. Math Imaging Vis 58, 294 - 320 (2017).

[4]Liu H, Mi X, Li Y. Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM[J]. Energy Conv Manag, 2018; 159:54 - 64.

[5]张省. 我国碳排放权交易价格组合预测研究——基于二次分解和机器学习方法的分析[J]. 价格理论与实践,2023,(09):142-145+209.

[6]朱菊香,张诗云,张涛,等. 基于二次分解和改进沙猫群优化算法的空气质量预测[J]. 国外电子测量技术,2024,43(05):190-200.

[7]樊冲. 基于 CEEMDAN-VMD-SSA-LSTM 的门诊量预测模型[J]. 微型电脑应用,2024,40(05):214-218+242.