

原油期货价格预测模型 CEEMDAN-PSO-ELM^①



崔金鑫, 邹辉文

(福州大学 经济与管理学院, 福州 350116)
(福州大学 投资与风险管理研究所, 福州 350116)
通讯作者: 邹辉文, E-mail: frenew@fzu.edu.cn

摘要: 为了进一步提升原油期货价格预测的精准性, 本文基于 CEEMDAN 分解算法和 ELM 极限学习机模型, 利用 PSO 粒子群优化算法对机器学习模型进行参数寻优, 进而构建了 CEEMDAN-PSO-ELM 模型用于原油期货价格预测. 先基于 CEEMDAN 算法对原始价格序列进行分解, 然后利用 Lempel-Ziv 复杂度指数对分量进行重构, 得到高频、中频和低频重构分量, 再采用 PSO-ELM 模型对每个重构分量进行预测, 利用 PACF 系数选取模型输入变量, 最终加总集成各分量预测结果. 实证结果表明, 与其他 15 种基准模型相比, CEEMDAN-PSO-ELM 模型的预测性能最佳, MCS 检验和 DM 检验也进一步证实了该模型的稳健性.

关键词: CEEMDAN 分解算法; PSO-ELM 预测模型; 原油期货价格预测

引用格式: 崔金鑫, 邹辉文. 原油期货价格预测模型 CEEMDAN-PSO-ELM. 计算机系统应用, 2020, 29(2): 28-39. <http://www.c-s-a.org.cn/1003-3254/7224.html>

Oil Futures Price Forecasting Model Named CEEMDAN-PSO-ELM

CUI Jin-Xin, ZOU Hui-Wen

(School of Economics and Management, Fuzhou University, Fuzhou 350116, China)
(Institute of Investment and Risk Management, Fuzhou University, Fuzhou 350116, China)

Abstract: In order to further enhance the prediction performance of oil futures price, this study proposes a novel CEEMDAN-PSO-ELM model for oil futures price forecasting based on CEEMDAN decomposition algorithm, extreme learning machine, and particle swarm optimization technology. Firstly, the original oil futures price series is decomposed by CEEMDAN algorithm into several intrinsic mode functions and a residual. Secondly, all the intrinsic mode functions and the residual are reconstructed based on Lempel-Ziv value. Then, the high, medium, and low frequency component are obtained respectively. Thirdly, the extreme learning machine optimized by particle swarm optimization algorithm is employed to predict each component and three component prediction results are obtained. Finally, integrate the prediction results of three components. The empirical research demonstrates that the CEEMDAN-PSO-ELM model proposed in this study has the best prediction performance compared with other 15 benchmark forecasting models. Moreover, the model confidence set and Diebold-Mariano test results further confirm the robustness of the proposed model.

Key words: CEEMDAN decomposition algorithm; PSO-ELM prediction model; oil futures price forecasting

引言

原油期货作为一种非常关键的能源金融衍生品, 其价格变化幅度较大, 且具有较高的时变特征, 原油期

货价格的变化对全球经济都产生着显著的影响^[1]. 因此, 对其价格作出准确的预测, 不仅可以为期货投资者以及市场监管者提供决策参考, 同时也利于原油期货

① 基金项目: 国家自然科学基金 (71573042); 福建省自然科学基金 (2017J01794)

Foundation item: National Natural Science Foundation of China (71573042); Natural Science Foundation of Fujian Province (2017J01794)

收稿时间: 2019-06-05; 修改时间: 2019-07-05; 采用时间: 2019-07-10; csa 在线出版时间: 2020-01-16

市场的长期健康发展.然而,与其他金融价格序列相类似,原油期货价格序列具有非平稳、非线性、多频率等典型特征,因此,对其价格作出精准预测较为困难.即便如此,国内外依然有学者对其展开了预测研究.

传统的金融时间序列预测方法主要包括 ARMA (AutoRegressive Moving Average) 模型、ARIMA (AutoRegressive Integrated Moving Average) 模型和 GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) 模型,然而它们无法完全挖掘出金融时间序列的非线性特征,预测效果往往不佳.随着数据统计科学以及计算机技术的发展,神经网络模型受到了学者们的青睐,并在金融时间序列预测领域得到了大量的应用,例如: ANN (Artificial Neural Network) 人工神经网络模型^[2]、SVR (Support Vector Regression) 支持向量机回归模型^[3]、BPNN (Back-Propagation Neural Network) 神经网络模型^[4]、ELM (Extreme Learning Machine) 极限学习机模型^[5]、ANFIS (Adaptive Network-based Fuzzy Inference System) 自适应模糊推理系统^[6]、LSTM (Long Short-Term Memory) 长短记忆神经网络模型^[7]等.随着预测研究的不断深入,学者们发现单一机器学习预测模型的预测性能仍有提升的空间,因此学者们探寻出了一种更加有效的预测模型:数据分解技术与机器学习算法相结合的预测模型.遵循着“分解”、“预测”、“集成”的整体建模架构,国内外学者们针对金融价格时间序列展开了大量的预测研究,例如:将经验模态分解算法与前馈神经网络相结合的 EMD-FNN (Empirical Mode Decomposition-Feed forward Neural Network) 模型^[8];将集合经验模态分解算法与粒子群算法优化的最小二乘支持向量机进行结合的 EEMD-PSO-LSSVM (Ensemble Empirical Mode Decomposition Particle Swarm Optimization Least Square Support Vector Machine) 模型^[9]、将集合经验模态分解方法、长短记忆模型和 Adaboost 算法相结合的 EEMD-LSTM-Adaboost 模型^[10]、将互补集合经验模态分解与改进的极限学习机相结合的 CEEMD-EELM (Complete Ensemble Empirical Mode Decomposition-Extended Extreme Learning Machine) 模型^[11]等.

CEEMDAN 完全自适应集合经验模态分解是由 Colominas 等^[12]提出的,利用一种自适应的噪声生成算法对 EEMD 集合经验模态分解算法进行改进,由于

CEEMDAN (Complete Ensemble Empirical Mode Decomposition with Adaptive Noise) 分解算法所具备的优异性能,因而近些年来被大量学者所采用,并且将其与机器学习预测模型相结合并应用到各个领域,例如:股市预测^[13]、风速预测^[14]、交通流预测^[15]、电力负荷预测^[16].然而,目前尚未有学者将其运用于原油期货价格预测.基于此,本文采用 CEEMDAN 分解算法对原始的原油期货价格序列进行分解,以期更好地挖掘出原始价格序列的内在特征.

PSO-ELM 基于粒子群优化的极限学习机模型目前在二氧化碳排放量预测^[17]、太阳能光伏发电量预测^[18]、交通流量预测^[19]、碳排放权交易价格预测^[20]等领域均已显示出其优异的预测性能.然而,还未曾见到将 PSO-ELM 模型用于原油期货价格预测的文献.因此,本文将 PSO-ELM 模型作为基础预测模型,探究其是否可以提升原油期货价格预测的精准性.

综合已有文献,多数学者针对石油现货价格展开预测研究.然而,随着商品金融化程度的不断加深和商品期货市场规模的不断扩大,商品现货价格逐渐参照期货进行定价,尤其是石油现货价格的定价完全参照期货价格^[21].因此,对原油期货价格序列展开预测研究将更具现实意义.基于以上认识,将 CEEMDAN 数据分解算法和 PSO-ELM 粒子群优化的极限学习机模型进行组合,本文构建出了一种 CEEMDAN-PSO-ELM 预测模型,以期进一步提升原油期货价格预测的精准性和稳健性.

1 相关理论及模型

1.1 CEEMDAN 分解算法

在经典的集合经验模态分解过程中,一定次数的平均并不能完全抵消所添加的白噪声序列,并且重构误差的大小与集成的次数有关.尽管重构误差会随着平均次数的增多而逐渐地减小,但是所耗费的时间也会增加^[16].而在本文采用的 CEEMDAN 分解流程中,在每个阶段添加有限次数的自适应白噪声序列,即可使得重构误差在较少的平均次数逼近于零.

将原始原油期货价格序列记为 $x(n)$, 则第 t ($1, 2, \dots, T$) 次所添加的白噪声序列记为 $\lambda^t(n)$, 因此,第 t 次分解的原油期货价格序列可以被表示为 $x^t(n) = x(n) + \lambda^t(n)$. 将第 i 个由 EMD 分解算法和 CEEMDAN 分解算法产生的模态分量记为: $E_i(\cdot)$ 和 \widehat{IMF}_i . 参照已有

研究^[15,16], CEEMDAN 分解步骤如下:

第 1 步: 先通过 EMD 分解算法对原始原油期货价格序列进行分解, 从而得到第一个模态分量:

$$\widehat{IMF}_1(n) = \frac{1}{T} \sum_{t=1}^T IMF_1^t(n) = \overline{IMF}_1(n) \quad (1)$$

第 2 步: 在第一个阶段 ($i = 1$) 计算第一个唯一的余量信号:

$$r_1(n) = x(n) - \overline{IMF}_1(n) \quad (2)$$

第 3 步: 将实验重复 $t(1, 2, \dots, T)$ 次, 在每次实验过程中, 对信号 $r_1(n) + \varepsilon_1 E_1(\lambda^t(n))$ 进行分解, 当得到第一个 EMD 模态分量时停止. 此时得到第二个模态分量:

$$\widehat{IMF}_2(n) = \frac{1}{T} \sum_{t=1}^T E_1(r_1(n) + \varepsilon_1 E_1(\lambda^t(n))) \quad (3)$$

第 4 步: 在剩下的每个阶段中 ($i = 2, 3, \dots, I$), 计算第 i 个余量信号, 其计算方法与第三步相同. 此时, 可以得到第 $i + 1$ 个模态分量:

$$r_i(n) = r_{i-1}(n) - \widehat{IMF}_i(n) \quad (4)$$

$$\widehat{IMF}_{(i+1)} = \frac{1}{T} \sum_{t=1}^T E_i(r_i(n) + \varepsilon_i E_i(\lambda^t(n))) \quad (5)$$

第 5 步: 判断余量信号的极值点个数是否至多不超过两个, 如果是, 则算法停止, 也即此时的余量序列无法继续进行分解, 则一共得到 I 个模态分量. 最终的余量信号可以表示为:

$$R(n) = x(n) - \sum_{i=1}^I \widehat{IMF}_i \quad (6)$$

因而, 原始的原油期货价格序列 $x(n)$ 最终可以被分解成为式 (7):

$$x(n) = \sum_{i=1}^I \widehat{IMF}_i + R(n) \quad (7)$$

1.2 PSO-ELM 粒子群算法优化的极限学习机模型

极限学习机 (Extreme Learning Machine, ELM) 是 Huang 等^[22]构建的一个只含有单个隐含层的前馈神经网络, 对其输入权重以及偏置进行随机的初始化, 并且得到相应的权值. 另外, 在训练流程中并不需要对参数进行调整, 仅需对隐含层的神经元个数进行设置, 就可以得到唯一的最优解. 具有简洁高效, 无需调整参数, 学习速度快的优点^[23], ELM 极限学习机的基本网络结构如图 1 所示.

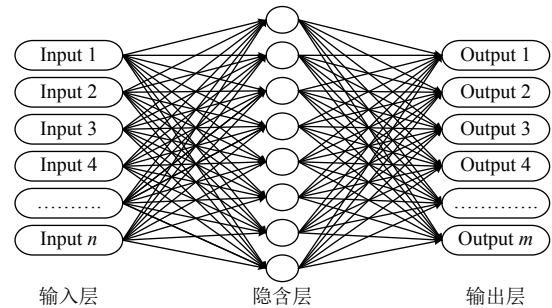


图 1 ELM 模型结构图^[23]

极限学习机模型由输入层、隐含层和输出层构成. 输入层有 n 个神经元, 分别对应着 n 个输入变量; 隐含层有 l 个神经元; 输出层有 m 个神经元, 分别对应着 m 个输出变量.

PSO 参数优化算法主要是对 ELM 模型第 i 个输入层神经元与第 j 个隐层神经元之间的连接权值 ω 以及第 j 层隐层神经元的阈值 b 进行寻优, PSO-ELM 模型的主要步骤包括^[24]: 第 1 步, 初始化粒子群算法的参数以及 ELM 模型的参数, 其中第 i 个粒子的位置可以被表示为 $X_i = [\omega_{11}, \dots, \omega_{h1}, \dots, \omega_{1n}, \dots, \omega_{ln}, b_1, \dots, b_n]$; 第 2 步, 寻找初始极值, 基于粒子的初始位置以及训练集数据对 ELM 模型进行训练, 从而获取 ELM 的输出值, 进一步算出粒子的适应度值. 另外, 还需搜寻到个体及群体极值, 并且记录它们的位置及适应度值; 第 3 步, 选取惯性权重, 在经历一定次数的迭代后, 粒子位置逐步趋于最佳, 此时惯性权重随着迭代次数的增加而不断减小; 第 4 步, 进行迭代寻优操作, 在每一个迭代流程中, 粒子依据个体和群体极值对自己的速度及所处的位置进行更新. 不断地重复迭代流程, 最后即可获得最优个体的粒子位置及适应度值; 第 5 步, 最优个体所对应的粒子位置即为 ELM 模型的权值以及隐含层阈值.

1.3 Lempel-Ziv 复杂度指数

Lempel-Ziv 复杂度指数是由 Lempel 和 Ziv^[25]通过构建算法程序实现的, 由于其良好的性能在非线形科学中得到了广泛的应用^[26-28], 其主要被用来衡量序列的复杂性特征, 序列的 Lempel-Ziv 数值越大, 表明序列的复杂度越高, 其所包含的周期成分越少, 规律性越弱, 隐含的频率信息越丰富; 反之, 序列的复杂度越低, 呈现出一种规律性特征, 序列会按照一定周期波动, 频率相应较低. 本文依据各个分量的 Lempel-Ziv 数值来进行分量的重构操作, 以期提高分量重构的合理性. 由于 Lempel-Ziv 算法流程较为复杂和冗长, 鉴于篇幅限制,

此处不再赘述,详细流程可见文献[25].

2 CEEMDAN-PSO-ELM 模型构建

本文的 CEEMDAN-PSO-ELM 原油期货价格预测模型构建流程如图 2 所示.

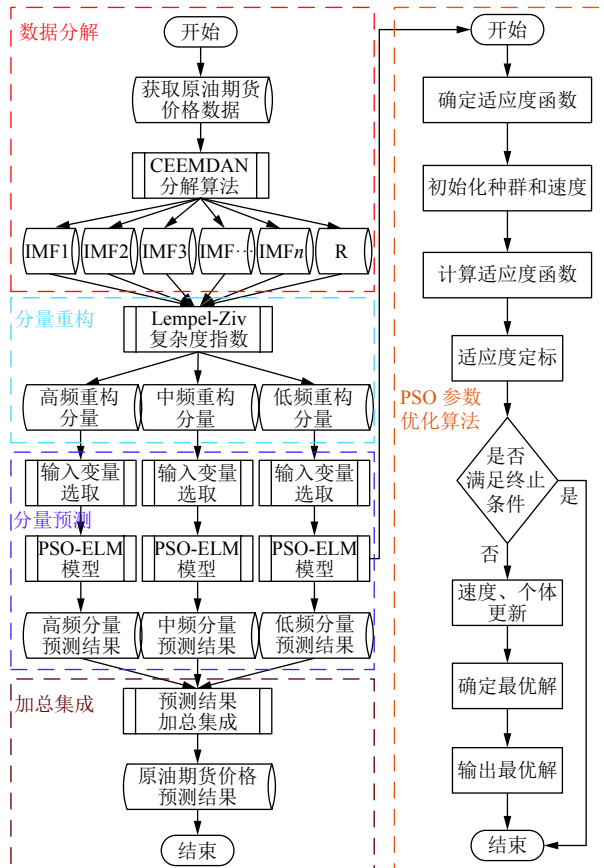


图 2 CEEMDAN-PSO-ELM 预测模型构建

遵循已有的“分解”、“重构”、“预测”和“集成”的整体构建流程. 第 1 步, 数据分解阶段, 先采用鲁棒性更优的 CEEMDAN 分解算法对原油期货原始价格序列进行分解, 进而得到若干个 IMF (Intrinsic Mode Function) 本征模态分量和一个残余项; 第二步, 分量重构阶段, 先计算各个分量的 Lempel-Ziv 复杂度指数, 然后按复杂度数值的高低进行归类, 从而将分解出的分量进行重构, 最终得到高频、中频和低频分量; 第 3 步, 重构分量预测阶段, 利用 PSO-ELM 粒子群优化的极限学习机模型对高频、中频和低频分量分别建立相应预测模型, 采用 PACF (Partial Auto Correlation Function) 偏自相关系数选择预测模型的输入变量, 进而得到 3 个重构分量的预测结果; 第四步, 预测结果集

成阶段, 也即将高频、中频和低频分量的预测结果进行加总集成, 进而得到最终的原油期货价格预测结果.

本文的预测模型在理论上具备以下 3 点优势: (1) 本文采用了更为稳健有效的 CEEMDAN 分解算法, 它在分解流程的每一个步骤增添自适应噪声, 然后计算单一的余量信号从而获取各个模态分量, 可以有效解决 EMD 模态混叠及 EEMD 分解不彻底的问题^[16]; (2) 本文选取 ELM 极限学习机作为基准预测模型, 在其训练过程无需调整参数, 运行效率较高, 预测性能较优. 并且, 采用 PSO 粒子群优化算法对输入层权值及隐含层阈值进行参数寻优, 进一步提升 ELM 预测模型的性能. (3) 本文预测模型遵照已有的“先分解后集成”的经典预测模型建模思路, 将智能分解算法及机器学习预测模型进行组合, 以期充分利用二者的优势, 较经典的计量模型及单一智能预测模型, 本文的预测模型在理论上更具稳健性和有效性.

3 实证分析

3.1 样本选取及描述性统计分析

本文参照龚旭和林伯强^[1]的研究设定, 选取世界上交易量最大, 影响最广泛的原油期货合约之一的 WTI 西德克萨斯中质原油月连续期货合约. 样本区间跨度从 2015 年 3 月 24 日至 2019 年 1 月 31 日, 共包含 1000 个交易日; 参照已有研究^[29,30]的区间划分方法, 将总样本的前 80% (2015-3-24 至 2018-4-24, 800 个交易日) 作为模型的训练集, 用于训练 PSO-ELM 预测模型, 将剩余的 20% (2018-4-25 至 2019-1-31, 200 个交易日) 作为模型的测试集, 用来测试模型的预测性能. 本文的数据均来自 WIND 万德数据库, 文章的实证过程均基于 Eviews10 和 Matlab 2014 (a) 完成.

图 3 给出了 WTI 原油期货价格序列图, 可以明显看出, 原油期货价格序列呈现出明显的波动性和非平稳性. 表 1 为原油期货价格序列描述性统计, 其结果也说明, 原油期货价格序列较离散, ADF (Augmented Dickey-Fuller) 统计值 P 值为 0.4271, 表明原油期货价格序列非平稳.

为了进一步检验原油期货价格序列是否存在非线性特征, 本文对其进行了 BDS (Brock-Decher-Scheikman) 检验, 从表 2 的检验结果可以看出, 嵌入维数从 2 增加到 10, P 值均为 0, 表明原油期货价格序列呈现出典型的非线性特征.

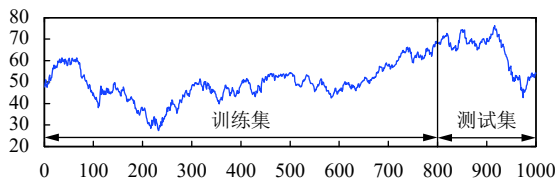


图3 原油期货日度价格序列图

表1 原油期货价格序列描述性统计

平均值	中位数	最小值	最大值	标准差	ADF 值	样本数
52.1829	50.5100	27.3000	76.2200	10.1307	-1.7076	1000

3.2 原油期货价格序列分解及重构

样本的描述性统计和 BDS 检验表明原始的原油期货价格序列呈现出典型的波动性、非平稳性和非线性

特征. 为了进一步挖掘价格序列的内在特征, 本文采用 CEEMDAN 分解算法对原油期货价格序列进行分解, 共得到 9 个 IMF 本征模态分量和 1 个残余项, 如图 4 所示.

表2 原油期货价格序列非线性检验结果

维数	BDS 统计量	标准差	Z 值
2	0.190 267	0.002 057	92.487 30
3	0.323 906	0.003 258	99.420 37
4	0.416 346	0.003 865	107.7280
5	0.479 451	0.004 012	119.4961
6	0.521 830	0.003 854	135.4115
7	0.549 696	0.003 517	156.3074
8	0.567 300	0.003 095	183.2914
9	0.577 710	0.002 651	217.8988
10	0.583 132	0.002 224	262.1965

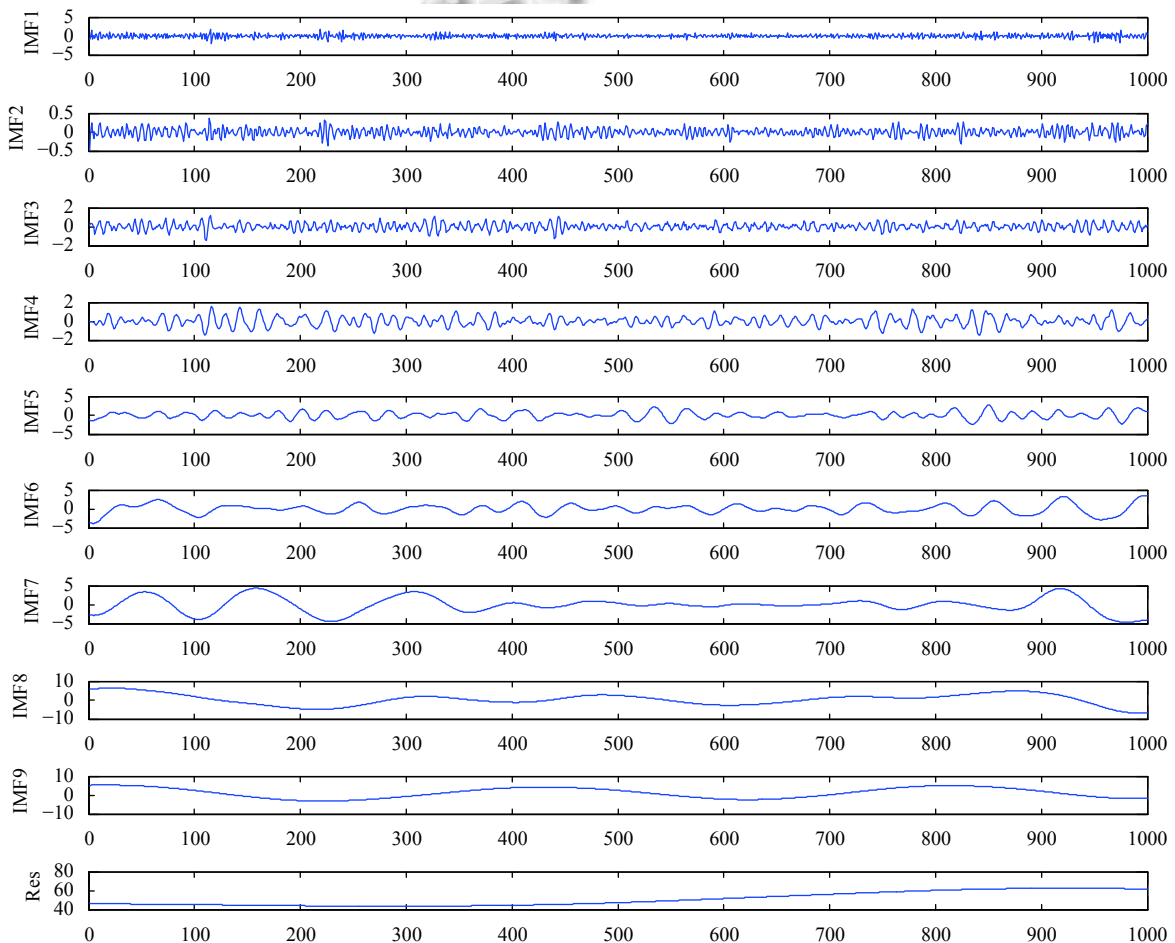


图4 原油期货价格序列 CEEMDAN 分解结果

为了使分量重构过程更加科学, 此处依据 Lempel-Ziv 复杂度指数的数值来对分量进行重构, 从表 3 的复杂度计算结果可以看出, IMF1~IMF4 的复杂度数值较

大, 说明前 4 个分量所包含的周期成分较少, 变化规律性低, 序列趋于一种随机状态; IMF5~IMF7 的复杂度数值处于中等水平, 表明中间 3 个分量存在一定的规律

性,包含的频率成分相对较少; IMF8、IMF9和余项的复杂度数值均小于 0.1, 表明这 3 个分量的周期性和规律性较强, 序列复杂度较低. 因此, 本文将前 4 个分

量 IMF1~IMF4 重组为高频分量, 中间 3 个分量 IMF5~IMF7 重组为中频分量, 剩余的两个分量以及余项重组为低频分量, 分量重构结果如图 5 所示.

表 3 Lempel-Ziv 复杂度指数计算结果

IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	Res
0.9667	0.8272	0.7275	0.5979	0.3488	0.1993	0.1495	0.0698	0.0598	0.0299

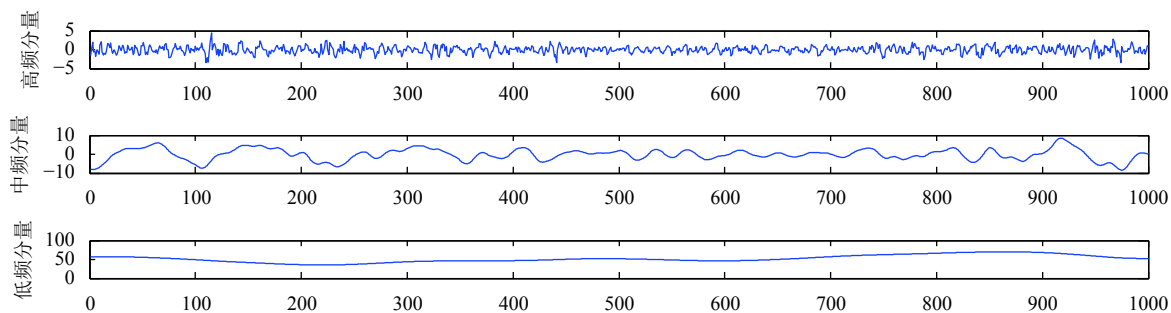


图 5 原油期货价格序列分量重构结果

图 5 所给出的分量重构结果同样具备一定的经济含义^[10], 高频分量呈现出十分复杂的序列特征, 代表着短期的市场波动因素, 这种短期因素所引起的原油期货价格序列的波动是不可持续的, 且方向不定, 频率较高; 中频分量的波动特征相对平缓, 代表着重大事件对原油期货价格序列的影响, 这种因素使得原油期货价格发生大幅度变化, 且持续时间较长, 但是最终会恢复到长期趋势线上; 低频分量呈现出非常平缓的趋势特征, 代表着原油期货价格的长期走向, 不易受到市场及外部环境因素的冲击, 可以通过低频分量预测未来的原油期货价格的大致走势.

3.3 CEEMDAN-PSO-ELM 模型预测

在完成了分量重构以后, 便可以对高频、中频和低频分量分别建立 PSO-ELM 预测模型, 而在机器学习模型初始训练阶段, 参数的设定以及模型输入变量的选取都会对模型的预测效果造成显著影响, 在参考了大量研究后, 本文的预测模型参数设定如下: PSO 参数寻优算法的初始种群大小为 25, 最大迭代次数为 200, 惯性权重 $w_{\min} = 0.8$, $w_{\max} = 1.2$, 加速因子 $c_1 = c_2 = 2.0$. 利用 PSO 对 ELM 极限学习机模型进行参数寻优操作; ELM 模型的隐含节点设定为 30, 隐含层节点输出权值范围为 $[-1, 1]$, 隐含层节点阈值取值范围为 $[0, 1]$, 输入层个数即为输入变量的个数 (需根据具体的预测分量而定). 在预测模型的参数设定好了以后, 还

需要确定高频、中频和低频分量预测模型的输入变量, 本文采用 PACF 偏自相关系数^[29,31]来进行判定. 本文最终的输入变量选择结果如表 4 所示.

表 4 输入变量选择结果

分量	输入变量
高频分量	$H(t-1), H(t-2), H(t-3), H(t-4), H(t-5), H(t-9), H(t-10), H(t-11)$
中频分量	$M(t-1), M(t-2), M(t-3), M(t-4), M(t-5)$
低频分量	$L(t-1), L(t-2), L(t-3), L(t-4), L(t-5), L(t-6)$

图 6~图 8 分别给出了高频、中频和低频分量的预测结果.

从图 6~图 8 可以明显地看出, 中频分量和低频分量的预测效果显著优于高频分量. 原因可能是高频分量的序列复杂度较高, 序列中所包含的规律性特征较少, 因而利用极限学习机模型对高频分量进行预测仍然存在一定的误差. 将高频、中频和低频分量各自的预测结果进行加总集成, 便得到了最终的原油期货价格预测结果.

图 9 显示 CEEMDAN-PSO-ELM 模型的预测结果.

从图 9 可以清晰地看出, CEEMDAN-PSO-ELM 模型的预测值与真实值高度接近. 为了更加科学稳健体现本文所构建模型的优异预测性能, 还需与本文所选取的基准预测模型进行预测效果的对比分析.

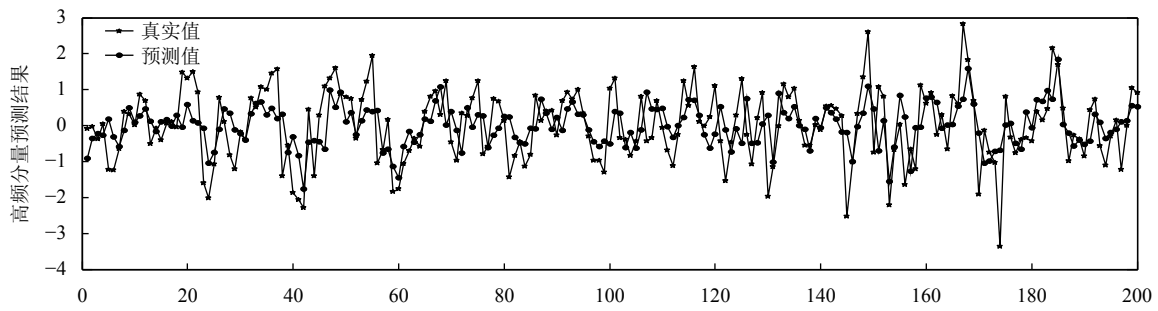


图6 高频分量预测结果

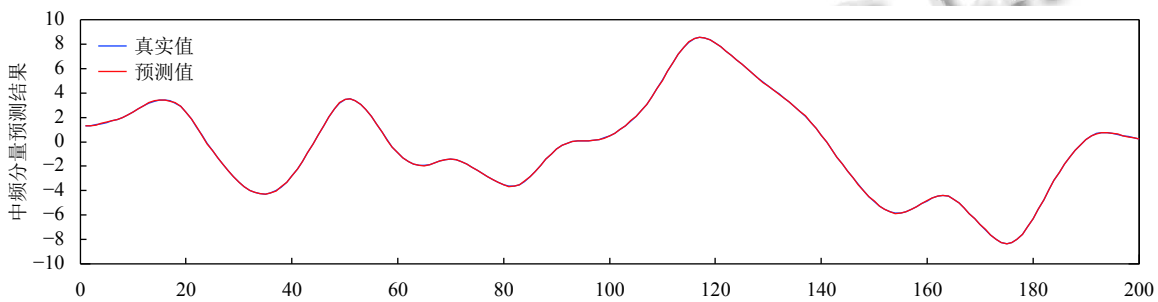


图7 中频分量预测结果

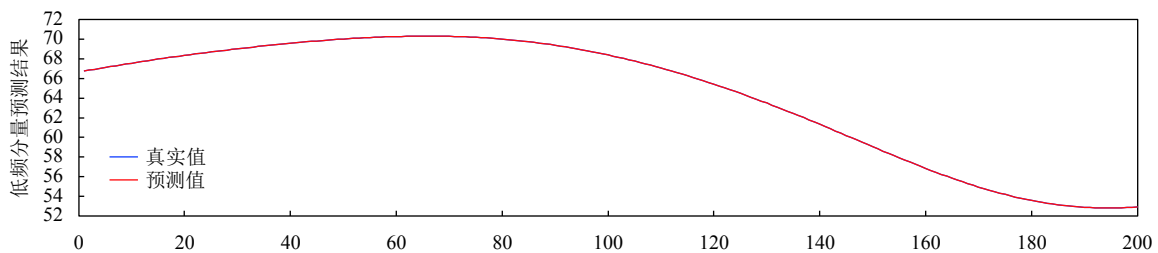


图8 低频分量预测结果

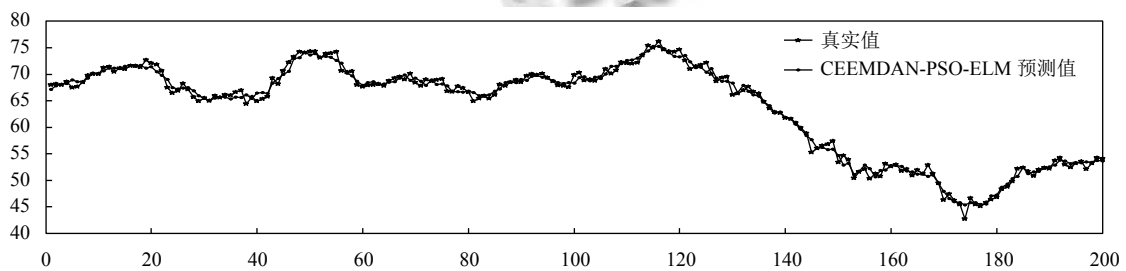


图9 CEEMDAN-PSO-ELM 预测结果

3.4 模型预测效果比较

为了科学稳健地体现出本文所构建的 CEEMDAN-PSO-ELM 模型的优越性, 本文另外选取了 15 个基准对比预测模型. 在对比模型的选取上, 本文主要基于以下 3 个方面的考量: (1) 选取 BPNN、SVR、ANFIS 和

ELM 4 种单一的机器学习预测模型, 利用模型自身内部算法学习能力对原油期货价格进行预测; (2) 采用 PSO 粒子群参数优化算法对 4 种单一预测模型的参数进行寻优, 以期提升它们的预测性能; (3) 分别将 EEMD 和 CEEMDAN 两种分解算法与经过 PSO 参数

优化的4种预测模型相结合。

通过分组对比,可证明本文构建的预测模型性能的优越性,也可体现出CEEMDAN分解算法及PSO粒子群参数优化算法的有效性,还能体现出组合预测较单一预测模型的优势。另外,此处采用均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对误差百分比(MAPE)、Theil不等系数(TIC)以及平均方向精准度(MDA)作为各模型预测效果评价指标,本文相关计算公式如式(8)~式(12)。其中, T 表示预测样本数, Y_t 代表 t 时刻的原油期货价格真实值, P_t 代表 t 时刻的原油期货价格预测值。 $RMSE$ 、 MAE 、 $MAPE$ 和 TIC 都是预测误差衡量指标,值越小表明模型的预测效果越好; MDA 是衡量原油期货价格预测结果趋势与真实原油期货价格趋势一致程度的指标,值越大说明模型预测结果方向精准度越高。

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t - P_t)^2} \quad (8)$$

$$MAE = T^{-1} \sum_{t=1}^T |Y_t - P_t| \quad (9)$$

$$MAPE = T^{-1} \sum_{t=1}^T |(Y_t - P_t)/Y_t| \quad (10)$$

$$TIC = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t - P_t)^2}}{\sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t)^2 + \frac{1}{T} \sum_{t=1}^T (P_t)^2}} \quad (11)$$

$$MDA = \frac{1}{T} \sum_{t=1}^T b_t, b_t = \begin{cases} 1, & (Y_{t+1} - Y_t) \cdot (P_{t+1} - P_t) \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

图10和表5给出了本文16种预测模型的预测效果对比。

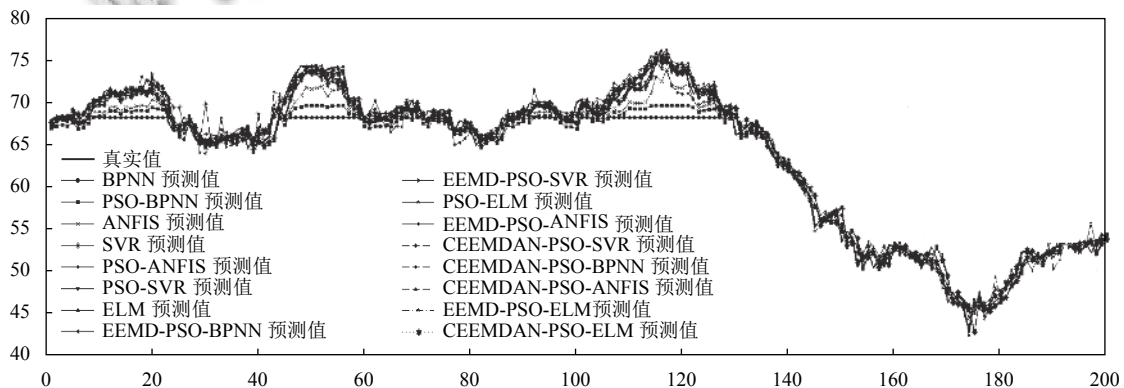


图10 模型预测结果对比图

从图10和表5中可以清晰地看出:各个机器学习预测模型均具备一定的预测性能;PSO粒子群优化算法提升了BPNN、ANFIS、SVR和ELM等4种单一预测模型的预测性能;将机器学习预测模型与数据分解算法相结合可以显著地提升单一预测模型的预测性能;本文所采用的CEEMDAN数据分解算法的表现,在一定程度上优于常用的EEMD数据分解算法;而ELM极限学习机模型在单一预测模型中的预测性能最佳,表明本文选取ELM模型作为基础预测模型具备一定的合理性;本文构建的CEEMDAN-PSO-ELM预测模型不仅预测误差指标数值最低,且方向精准性指标数值最高,表明其预测性能最佳。为了保证本文实证结果的稳健性,还需要进一步进行MCS (Model

Confidence Set)以及DM (Diebold-Mariano)预测性能稳健性检验。

3.5 模型预测效果稳健性检验

MCS检验是由Hansen等^[32]提出的衡量模型预测性能的一种方法,相较SPA (Superior Predictive Ability)更稳健。由于篇幅限制,MCS检验的理论框架及流程详见Hansen等^[32]、Zhang等^[33]。为了得到MCS检验中的各统计量以及 p 值,本文参照雷立坤等^[34]的做法,将MCS检验的参数设置如下: $d=2$ (block length),模拟次数 $B=10\ 000$ 次作为Bootstrap过程的控制参数,MCS检验的显著性水平 α 取值为0.1,也即 p 值大于0.1的模型将幸存下来。

DM检验是由Diebold和Mariano^[35]提出一种衡量

模型预测效果的方法, DM 检验的原假设是目标模型 A 的预期预测精准性与基准模型 B 的预测精准性一致, 因此原假设可以表示为:

$$H_0: E[F(e_t^A)] = E[F(e_t^B)] \quad (13)$$

式中, e_t^A 和 e_t^B 分别代表模型 A 和 B 的预测误差, 误差函数 F 通常被设定为均方误差 MSE , 因此 DM 统计量可以被定义成式 (14), 式中, γ_0 代表 g_t 的方差, $x_{A,t}$ 和 $x_{B,t}$ 分别表示模型 A 和 B 在 t 时刻的预测值, T 代表测试集观测样本的总数量.

$$S_{DM} = \frac{\bar{g}}{\sqrt{(\widehat{V}_{\bar{g}}/T)}} \quad (14)$$

式中,

$$\bar{g} = \frac{1}{T} \sum_{t=1}^T g_t, g_t = (x_t - x_{A,t})^2 - (x_t - x_{B,t})^2 \quad (15)$$

$$\widehat{V}_{\bar{g}} = \gamma_0 + 2 \sum_{t=1}^{\infty} \gamma_t, \gamma_t = Cov(g_{t+1}, g_t) \quad (16)$$

表 6 和表 7 分别给出了模型预测结果的 MCS 稳健性检验和 DM 稳健性检验结果, 从中可以看出, 本文所构建的 CEEMDAN-PSO-ELM 原油期货价格预测模型相较于其他 15 种基准模型预测性能最佳, EEMD-PSO-ELM 模型的预测性能仅次于 CEEMDAN-PSO-ELM 模型, 其余结论与 3.4 节一致, 此处不再赘述.

表 5 模型预测效果对比表

模型	MAE	RMSE	MAPE	TIC	MDA
BPNN	1.768 889	2.479 229	0.026 736	0.019 403	0.495 000
PSO-BPNN	1.622 738	2.106 095	0.024 807	0.016 476	0.465 000
ANFIS	1.247 432	1.578 698	0.019 608	0.012 297	0.475 000
SVR	0.997 767	1.322 584	0.016 277	0.010 252	0.485 000
PSO-ANFIS	0.992 391	1.271 113	0.016 122	0.009 869	0.475 000
PSO-SVR	0.968 254	1.249 564	0.015 793	0.009 696	0.485 000
ELM	0.916 268	1.228 059	0.015 090	0.009 512	0.520 000
EEMD-PSO-BPNN	0.894 177	1.143 111	0.014 448	0.008 845	0.655 000
EEMD-PSO-SVR	0.825 342	1.040 992	0.013 426	0.008 072	0.695 000
PSO-ELM	0.816 069	1.081 749	0.013 164	0.008 383	0.580 000
EEMD-PSO-ANFIS	0.791 419	1.033 675	0.012 731	0.008 009	0.710 000
CEEMDAN-PSO-SVR	0.794 451	1.020 273	0.012 641	0.007 920	0.705 000
CEEMDAN-PSO-BPNN	0.782 455	1.003 332	0.012 792	0.007 782	0.695 000
CEEMDAN-PSO-ANFI	0.703 795	0.900 891	0.011 401	0.006 982	0.725 000
EEMD-PSO-ELM	0.695 086	0.896 206	0.011 165	0.006 946	0.715 000
CEEMDAN-PSO-ELM	0.664 024	0.838 313	0.010 641	0.006 497	0.735 000

4 结论

原油期货价格序列具有典型的非平稳、非线性以及多频率特征, 鉴于原油期货价格序列预测的复杂性, 本文构建了一种 CEEMDAN-PSO-ELM 预测模型, 在本文的样本区间内, 得出以下实证结论: (1) 利用 PSO 粒子群参数优化算法对 SVR、ANFIS、BPNN 和 ELM 等 4 种单一预测模型参数进行优化, 可以显著提升单一预测模型的预测性能; (2) ELM 极限学习机的预测性能优于 SVR、ANFIS 和 BPNN 等 3 种单一预测模型; (3) 将数据分解算法与机器学习预测模型相结合可以显著提升其预测性能; (4) 本文所构建的 CEEMDAN-PSO-ELM 原油期货价格预测模型与其余基准预测模型相比预测性能最佳, 并且具备一定稳健性. 本文的研

究成果可以为期货市场投资者及监管者提供有力的决策参考, 同时为原油期货价格预测领域提供一种新的视角和方法.

不可否认的是, 本文的预测模型也存在一定的不足之处: 第一, PSO 粒子群参数优化算法在运行过程中可能会陷入局部极值, 导致参数寻优效果不佳; 第二, 在输入变量的选取上, 本文没有考虑到现实中影响原油期货价格变动的因素. 因而, 可以从以下两方面改进预测模型: 第一, 将粒子群算法与其他的参数优化算法相结合 (例如粒子群算法与引力搜索算法结合-PSO-GSA 算法) 对机器学习预测模型进行参数寻优; 第二, 优化预测模型输入变量的选取方法, 可以在模型输入变量选取上考虑更多的现实影响因素.

表6 MCS 检验结果

模型	MSE		MSPE		MAE		MAPE	
	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}	T_R	T_{SQ}
BPNN	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
PSO-BPNN	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ANFIS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
SVR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
PSO-ANFIS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
PSO-SVR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ELM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EEMD-PSO-BPNN	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EEMD-PSO-SVR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
PSO-ELM	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EEMD-PSO-ANFIS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CEEMDAN-PSO-SVR	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CEEMDAN-PSO-BPNN	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
CEEMDAN-PSO-ANFIS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
EEMD-PSO-ELM	0.3250	0.3250	0.0320	0.0320	0.4742	0.4742	0.8986	0.8986
CEEMDAN-PSO-ELM	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

注:表中的数字是进行了10 000次Bootstrap模拟的MCS检验的P值,P值大于0.1(加粗的数字),代表在MCS检验中幸存下来的模型,即预测能力较好的模型;P值为1(加粗和加下划线的数字)表示在对应的损失函数下,该模型预测精度最高。

表7 DM 检验结果

M8	M9	M10	M11	M12	M13	M14	M15	M16	
6.1355(***)	6.4381(***)	6.2893(***)	6.4204(***)	6.8826(***)	6.5659(***)	6.8033(***)	6.7711(***)	6.9332(***)	M1
6.3164(***)	6.8254(***)	6.5886(***)	6.7630(***)	7.5296(***)	6.9915(***)	7.4050(***)	7.3320(***)	7.5963(***)	M2
4.9381(***)	6.0615(***)	5.6442(***)	5.7084(***)	6.5995(***)	6.1800(***)	7.2553(***)	7.0431(***)	7.5951(***)	M3
2.2947(**)	3.5574(***)	3.3085(***)	3.2817(***)	3.2256(***)	3.4923(***)	4.9074(***)	4.7239(***)	5.2159(***)	M4
1.8894(**)	3.3163(***)	3.1582(***)	3.0179(***)	3.1381(***)	3.4307(***)	5.1174(***)	4.7659(***)	5.4221(***)	M5
1.5712(*)	2.9582(***)	2.9057(***)	2.7526(***)	2.8034(***)	3.1110(***)	4.7787(***)	4.4782(***)	5.0689(***)	M6
1.2429	2.5701(***)	2.6350(***)	2.4383(***)	2.4063(***)	2.7489(***)	4.4024(***)	4.1322(***)	4.6732(***)	M7
-	2.0412(**)	0.8547	1.7204(**)	1.9205(**)	2.5773(***)	4.8926(***)	4.5705(***)	5.2694(***)	M8
-	-	0.5948	0.1376	0.3757	0.7871	3.7711(***)	3.2047(***)	4.4988(***)	M9
-	-	-	0.6919	0.7655	1.0937	2.7446(***)	2.7674(***)	3.6576(***)	M10
-	-	-	-	0.2039	0.5222	2.6310(***)	2.8444(***)	3.5327(***)	M11
-	-	-	-	-	0.3149	2.5363(***)	2.1581(**)	3.5672(***)	M12
-	-	-	-	-	-	2.7486(***)	2.0765(**)	3.6886(***)	M13
-	-	-	-	-	-	-	0.1345	2.3004(**)	M14
-	-	-	-	-	-	-	-	1.6779(**)	M15

注:M1代表BPNN,M2代表PSO-BPNN,M3代表ANFIS,M4代表SVR,M5代表PSO-ANFIS,M6代表PSO-SVR,M7代表ELM,M8代表EEMD-PSO-BPNN,M9代表EEMD-PSO-SVR,M10代表PSO-ELM,M11代表EEMD-PSO-ANFIS,M12代表CEEMDAN-PSO-SVR,M13代表CEEMDAN-PSO-BPNN,M14代表CEEMDAN-PSO-ANFIS,M15代表EEMD-PSO-ELM,M16代表CEEMDAN-PSO-ELM。在每一行中,上面的数值代表DM检验统计量的值,下面圆括号中的星号个数分别代表在不同的显著性水平下显著,其中,***代表在1%的水平下显著,**代表5%的水平下显著,*代表10%的水平下显著。

参考文献

- 1 龚旭,林伯强.跳跃风险、结构突变与原原油期货价格波动预测.中国管理科学,2018,26(11):11-21.[doi:10.16381/j.cnki.issn1003-207x.2018.11.002]
- 2 Moshiri S, Foroutan F. Forecasting nonlinear crude oil futures prices. The Energy Journal, 2006, 27(4): 81-95.
- 3 Kazem A, Sharifi E, Hussain FK, et al. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. Applied Soft Computing, 2013, 13(2): 947-958. [doi: 10.1016/j.asoc.2012.09.024]
- 4 Jammazi R, Aloui C. Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling. Energy Economics, 2012, 34(3): 828-841. [doi: 10.1016/j.eneco.2011.07.018]
- 5 Wang J, Athanasopoulos G, Hyndman RJ, et al. Crude oil price forecasting based on internet concern using an extreme learning machine. International Journal of Forecasting, 2018, 34(4): 665-677. [doi: 10.1016/j.ijforecast.2018.03.009]

- 6 Svalina I, Galzina V, Lujic R, *et al.* An adaptive network-based fuzzy inference system (ANFIS) for the forecasting: The case of close price indices. *Expert Systems with Applications*, 2013, 40(15): 6055–6063. [doi: [10.1016/j.eswa.2013.05.029](https://doi.org/10.1016/j.eswa.2013.05.029)]
- 7 Kim HY, Won CH. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, 2018, 103: 25–37. [doi: [10.1016/j.eswa.2018.03.002](https://doi.org/10.1016/j.eswa.2018.03.002)]
- 8 Yu LA, Wang SY, Lai KK. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, 2008, 30(5): 2623–2635. [doi: [10.1016/j.eneco.2008.05.003](https://doi.org/10.1016/j.eneco.2008.05.003)]
- 9 Zhang JL, Zhang YJ, Zhang L. A novel hybrid method for crude oil price forecasting. *Energy Economics*, 2015, 49: 649–659. [doi: [10.1016/j.eneco.2015.02.018](https://doi.org/10.1016/j.eneco.2015.02.018)]
- 10 邸浩, 赵学军, 张自力. 基于 EEMD-LSTM-Adaboost 的商品价格预测. *统计与决策*, 2018, 34(13): 72–76. [doi: [10.13546/j.cnki.tjyc.2018.13.016](https://doi.org/10.13546/j.cnki.tjyc.2018.13.016)]
- 11 Tang L, Dai W, Yu LA, *et al.* A novel CEEMD-based EELM ensemble learning paradigm for crude oil price forecasting. *International Journal of Information Technology & Decision Making*, 2014, 14(1): 141–169. [doi: [10.1142/S0219622015400015](https://doi.org/10.1142/S0219622015400015)]
- 12 Colominas MA, Schlotthauer G, Torres ME, *et al.* Noise-assisted EMD methods in action. *Advances in Adaptive Data Analysis*, 2012, 4(4): 1250025. [doi: [10.1142/S1793536912500252](https://doi.org/10.1142/S1793536912500252)]
- 13 Cao J, Li Z, Li J. Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 2019, 519: 127–139. [doi: [10.1016/j.physa.2018.11.061](https://doi.org/10.1016/j.physa.2018.11.061)]
- 14 Zhang WY, Qu ZX, Zhang KQ, *et al.* A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting. *Energy Conversion and Management*, 2017, 136: 439–451. [doi: [10.1016/j.enconman.2017.01.022](https://doi.org/10.1016/j.enconman.2017.01.022)]
- 15 田秀娟, 于德新, 邢雪, 等. 交叉口短时流量 CEEMDAN-PE-OSELM 预测模型. *哈尔滨工业大学学报*, 2018, 50(3): 83–89. [doi: [10.11918/j.issn.0367-6234.201703102](https://doi.org/10.11918/j.issn.0367-6234.201703102)]
- 16 李军, 李青. 基于 CEEMDAN-排列熵和泄漏积分 ESN 的中期电力负荷预测研究. *电机与控制学报*, 2015, 19(8): 70–80. [doi: [10.15938/j.emc.2015.08.011](https://doi.org/10.15938/j.emc.2015.08.011)]
- 17 Sun W, Wang CF, Zhang CC. Factor analysis and forecasting of CO₂ emissions in Hebei, using extreme learning machine based on particle swarm optimization. *Journal of Cleaner Production*, 2017, 162: 1095–1101. [doi: [10.1016/j.jclepro.2017.06.016](https://doi.org/10.1016/j.jclepro.2017.06.016)]
- 18 Behera MK, Majumder I, Nayak N. Solar photovoltaic power forecasting using optimized modified extreme learning machine technique. *Engineering Science and Technology, an International Journal*, 2018, 21(3): 428–438. [doi: [10.1016/j.jestch.2018.04.013](https://doi.org/10.1016/j.jestch.2018.04.013)]
- 19 Lin L, Handley JC, Gu YM, *et al.* Quantifying uncertainty in short-term traffic prediction and its application to optimal staffing plan development. *Transportation Research Part C: Emerging Technologies*, 2018, 92: 323–348. [doi: [10.1016/j.trc.2018.05.012](https://doi.org/10.1016/j.trc.2018.05.012)]
- 20 Zhou JG, Yu XC, Yuan XL. Predicting the carbon price sequence in the Shenzhen emissions exchange using a multiscale ensemble forecasting model based on ensemble empirical mode decomposition. *Energies*, 2018, 11(7): 1907. [doi: [10.3390/en11071907](https://doi.org/10.3390/en11071907)]
- 21 李智, 林伯强, 许嘉峻. 基于 MSVAR 的国际原油期货价格变动研究. *金融研究*, 2014, (1): 99–109.
- 22 Huang GB, Zhu QY, Siew CK. Extreme learning machine: A new learning scheme of feedforward neural networks. *Proceedings of 2004 IEEE International Joint Conference on Neural Networks*. Budapest, Hungary. 2004. 985–990. [doi: [10.1109/IJCNN.2004.1380068](https://doi.org/10.1109/IJCNN.2004.1380068)]
- 23 罗宏远, 王德运, 刘艳玲, 等. 基于二层分解技术和改进极限学习机模型的 PM_{2.5} 浓度预测研究. *系统工程理论与实践*, 2018, 38(5): 1321–1330. [doi: [10.12011/1000-6788\(2018\)05-1321-10](https://doi.org/10.12011/1000-6788(2018)05-1321-10)]
- 24 冯禹铭, 董秀成, 金滔. 基于 PSO-ELM 的机器人精度补偿方法研究. *计算机应用研究*, 2019, 36(10): 3000–3003. [doi: [10.19734/j.issn.1001-3695.2018.03.0210](https://doi.org/10.19734/j.issn.1001-3695.2018.03.0210)]
- 25 Lempel A, Ziv J. On the complexity of finite sequences. *IEEE Transactions on Information Theory*, 1976, 22(1): 75–81. [doi: [10.1109/TIT.1976.1055501](https://doi.org/10.1109/TIT.1976.1055501)]
- 26 Zhu BZ, Ye SX, Wang P, *et al.* A novel multiscale nonlinear ensemble learning paradigm for carbon price forecasting. *Energy Economics*, 2018, 70: 143–157. [doi: [10.1016/j.eneco.2017.12.030](https://doi.org/10.1016/j.eneco.2017.12.030)]
- 27 Wang GF, Liu C. Fault diagnosis of rolling element bearings based on complexity measure and v support vector machine.

- Insight-Non-Destructive Testing and Condition Monitoring, 2013, 55(3): 142–146. [doi: [10.1784/insi.2012.55.3.142](https://doi.org/10.1784/insi.2012.55.3.142)]
- 28 Cui LL, Li BB, Ma JF, *et al.* Quantitative trend fault diagnosis of a rolling bearing based on Sparsogram and Lempel-Ziv. *Measurement*, 2018, 128: 410–418. [doi: [10.1016/j.measurement.2018.06.051](https://doi.org/10.1016/j.measurement.2018.06.051)]
- 29 Sun W, Zhang CC. Analysis and forecasting of the carbon price using multi-resolution singular value decomposition and extreme learning machine optimized by adaptive whale optimization algorithm. *Applied Energy*, 2018, 231: 1354–1371. [doi: [10.1016/j.apenergy.2018.09.118](https://doi.org/10.1016/j.apenergy.2018.09.118)]
- 30 Zhu JM, Wu P, Chen HY, *et al.* Carbon price forecasting with variational mode decomposition and optimal combined model. *Physica A: Statistical Mechanics and its Applications*, 2019, 519: 140–158. [doi: [10.1016/j.physa.2018.12.017](https://doi.org/10.1016/j.physa.2018.12.017)]
- 31 Sun W, Zhang CC, Sun CP. Carbon pricing prediction based on wavelet transform and K-ELM optimized by bat optimization algorithm in China ETS: The case of Shanghai and Hubei carbon markets. *Carbon Management*, 2018, 9(6): 605–617. [doi: [10.1080/17583004.2018.1522095](https://doi.org/10.1080/17583004.2018.1522095)]
- 32 Hansen PR, Lunde A, Nason JM. The model confidence set. *Econometrica*, 2011, 79(2): 453–497. [doi: [10.3982/ECTA5771](https://doi.org/10.3982/ECTA5771)]
- 33 Zhang YJ, Yao T, He LY, *et al.* Volatility forecasting of crude oil market: Can the regime switching GARCH model beat the single-regime GARCH models? *International Review of Economics & Finance*, 2019, 59: 302–317. [doi: [10.1016/j.iref.2018.09.006](https://doi.org/10.1016/j.iref.2018.09.006)]
- 34 雷立坤, 余江, 魏宇, 等. 经济政策不确定性与我国股市波动率预测研究. *管理科学学报*, 2018, 21(6): 88–98. [doi: [10.3969/j.issn.1007-9807.2018.06.007](https://doi.org/10.3969/j.issn.1007-9807.2018.06.007)]
- 35 Diebold FX, Mariano RS. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 1995, 13(3): 253–263. [doi: [10.1080/07350015.1995.10524599](https://doi.org/10.1080/07350015.1995.10524599)]