

Crude Oil Price Forecasting: Leveraging CELMDAN and Intelligent Optimization Algorithm for Informed Environmental Policy-Making

Yuxiang Xiao

School of Mathematics, East China University of Science and Technology

22012573@mail.ecust.edu.cn

Abstract

Reliable crude oil price forecasting is crucial for informing environmental policy decisions. This paper proposes a novel framework, CEAS, which integrates Complete Ensemble Local Mean Decomposition with Adaptive Noise (CELM DAN) and an Enhanced Grey Wolf Optimizer (EGWO) for decomposition and prediction. CELMDAN first decomposes crude oil price series into multiple Product Functions (PFs) and a residual. Each component is classified by its Hurst exponent: low-Hurst components are predicted using EGWO-optimized ARMA, and high-Hurst components with EGWO-optimized SVR. EGWO enhances traditional GWO by introducing two random exploration strategies, reducing the risk of local optima. Comparative and ablation experiments on Dubai, Brent, and WTI datasets demonstrate that CEAS achieves superior forecasting performance.

Keywords: Crude oil price forecasting; Decomposition; Enhanced grey wolf optimizer; Rescaled range analysis.

1 Introduction

The accelerating industrialization in developing countries has continued to drive a surge in global energy demand, with fossil fuels—particularly crude oil—remaining a cornerstone of the global energy system ([Moghani and Loni, 2025](#)). However, this heavy dependence on crude oil has resulted in huge carbon emissions, abnormal climate change and a cascade of environmental crises, including rising sea levels and an increasing frequency of extreme weather events. These consequences pose significant threats to the sustainable development of humanity ([Esmaeili et al., 2024](#); [Nyashina et al., 2018](#)).

In response to these challenges, the global transition toward clean and low-carbon energy has become an urgent objective ([Javaid et al., 2024](#)). Many countries have set carbon neutrality targets and are accelerating investments in renewable energy sources ([Fu et al., 2025](#)). Nevertheless, crude oil continues to occupy a central position in the current global economy and energy landscape ([Alkathery and Chaudhuri, 2021](#)). Fluctuations in crude oil prices not only

significantly impact patterns of energy consumption and investment ([Xu et al., 2021](#)) but also influence the pace and trajectory of the green transition ([Kumar and Mallick, 2024](#); [Benlemlih et al., 2024](#)).

Specifically, rising crude oil prices can reduce the economic attractiveness of fossil fuels, thereby encouraging both firms and households to shift toward clean energy. Conversely, declining crude oil prices may weaken investment enthusiasm for green energy, delaying the development of low-carbon technologies and undermining the effectiveness of environmental policies. Additionally, fluctuations in crude oil prices can significantly influence the design of environmental policies, particularly in countries or regions with energy subsidies ([Arouri et al., 2025](#)) or where carbon pricing mechanisms are linked to market prices. In this context, trends in crude oil prices serve as a key source of information for policymakers in assessing environmental risks.

Therefore, improving the accuracy of crude oil price forecasting is of vital significance for scientific environmental policymaking. A reliable forecasting model not only supports governments in formulating more forward-looking carbon reduction policies but also contributes to advancing the sustainable development of humans.

At present, crude oil price forecasting models can be broadly classified into three categories: traditional econometric models, machine learning models, and hybrid models. Traditional econometric models are relatively simple to implement and offer stronger interpretability, making them widely adopted in crude oil price forecasting. Traditional econometric models include autoregressive integrated moving average (ARIMA) ([Zhang and Liu, 2024](#)), generalized autoregressive conditional heteroscedasticity (GARCH) ([Rao et al., 2025](#)), vector autoregressive regression (VAR) ([Drachal, 2021](#)), and error correction model (ECM) ([Lanza et.al, 2005](#)). For example, [Zhang et al. \(2024\)](#) proposed an ARIMA-based model with fuzzy information granulation to make forecasting. [Rao et al. \(2025\)](#) used GARCH(1,1) to predict Brent crude oil prices at different time frequencies and found that it works well when mixed with GRU. [Drachal \(2021\)](#) developed a TV-VAR model to forecast real crude oil prices and demonstrated improved accuracy. However, these methods cannot capture the nonlinear relationship between independent and dependent variables, thus limiting the predictive effect.

With the development of computer science, machine learning models have begun to be widely used in crude oil price prediction ([Jabeur et al., 2021](#)). Common machine learning models include SVR ([Jha et al., 2024](#)), ELM ([Wang et al., 2018](#)), LSTM ([Wu et al., 2019](#)), GRU([Guo et al., 2023](#)), etc. For example, [Guo et al. \(2023\)](#) demonstrated the superiority of the GRU model in the prediction accuracy of China's crude oil prices through comparison. [Wu et al. \(2019\)](#) used LSTM to predict WTI crude oil prices based on EEMD decomposition and proved that the model had good robustness. However, these methods are easily affected by noise and prone to overfitting, and do not work well in practice.

Currently, more researchers have begun to pay attention to hybrid models ([Dutta et al., 2023](#)). Hybrid models can be mainly divided into two categories: hybrid models based on decomposition algorithms and hybrid models based on swarm intelligence algorithms. In terms of hybrid models based on decomposition algorithms, [Ouyang et al. \(2024\)](#) proposed a model

based on extreme symmetric empirical mode decomposition (ESMD) to decompose and reconstruct crude oil price data, and proved its superiority. Regarding hybrid models based on swarm intelligence algorithms, Wang et al. (2022) used a Multi-objective Grey Wolf optimizer (MOGWO) to make interval predictions of crude oil prices and achieved good generalization ability. However, current decomposition algorithms suffer from endpoint effects, and swarm intelligence algorithms are prone to falling into local optimality, which limits their prediction performance.

Based on the above, although many crude oil price forecasting models have been proposed and have certain effects, there are still some defects and space for improvement. The current main defects are summarized as follows:

(a) Limitation of single models. Whether statistical or machine learning-based, single models often suffer from inherent limitations, such as susceptibility to overfitting or a tendency to become trapped in local optima.

(b) Decomposition performance suffers from endpoint effects and mode mixing. Forecasting models based on decomposition algorithms are frequently constrained by suboptimal decomposition performance; traditional methods tend to produce endpoint effects and mode mixing, which in turn hinder effective model learning.

(c) Optimizers easily converge to local optima. Models based on optimization algorithms are prone to convergence to local optima, making it difficult at times to identify truly optimal model parameters.

Aiming to overcome these shortcomings and provide a reliable forecasting model for the government to formulate correct environmental policies, we propose CELMDAN-EGWO-ARMA-SVR (CELM DAN) for crude oil price forecasting. Our main contributions are as follows:

(a) A hybrid forecasting model is proposed by integrating two basic methods. To integrate the strengths of statistical and machine learning models, our proposed CEAS adapts ARMA to predict the high-frequency components, while SVR is used for mid- and low-frequency components. This design mitigates the limitations of single models and significantly improves prediction accuracy.

(b) A noise-assisted decomposition algorithm, CELMDAN, is introduced to achieve better decomposition performance. To address the issues of endpoint effects and mode mixing in traditional decomposition methods, CELMDAN extends the original signal, injects controlled noise, and averages the decomposition results over multiple iterations, thereby achieving more robust and accurate component separation.

(c) An enhanced optimization, EGWO, is developed by incorporating two early-stage exploration strategies aiming to escape from local optima. In response to the tendency of conventional optimizers to fall into local optima, the proposed EGWO integrates diversified exploration mechanisms during the early phase of search, guiding the wolves toward global optima and enabling the model to obtain better parameters.

The rest of this paper is organized as follows: Section 2 details the CEAS model structure and methodology. Section 3 describes the dataset and error indicator. Section 4 presents and

discusses the experiments and results. Finally, Section 5 concludes with key findings and future research directions.

2 Methodology

2.1 CELMDAN-EGWO-ARMA-SVR

We propose the CELMDAN-EGWO-ARMA-SVR (CEAS) model for crude oil price forecasting. This model is specifically designed to cope with the non-stationary, highly volatile nature of crude-oil price series (Ma et al, 2023). We propose Complete Ensemble Local Mean Decomposition with Adaptive Noise (CELM DAN), a novel hybrid decomposition method, based on Local Mean Decomposition (LMD), to decompose historical crude oil price data. This new decomposition technology is able to deliver superior decomposition performance, owing to its effective mitigation of the endpoint effect and mode mixing issues inherent in LMD. After decomposition, the Hurst exponent of each decomposed component signal is calculated using rescaled range analysis (R/S analysis), providing us with a crucial basis for distinguishing high-frequency short-term memory signals from medium-and-low-frequency long-term memory signals.

For short-term memory signals ($\text{Hurst} < 0.5$), we employ ARMA, utilizing the Akaike Information Criterion (AIC) to automatically select the optimal order (p, q) of ARMA. The Enhanced Grey Wolf Optimizer (EGWO) is then used to optimize the ARMA parameters ϕ and θ . Then, based on the optimized parameters, time series forecasting of ARMA is conducted. For long-term memory signals ($\text{Hurst} \geq 0.5$), we use Support Vector Regression (SVR), introducing EGWO to optimize its penalty parameter c and the kernel function parameter g , to achieve better prediction performance, minimizing the cross-validation error on the training set. Then we use SVR with the optimized c^* and g^* to make predictions.

The framework of the proposed CEAS model is shown in Figure 1. We will introduce each module in detail below.

CEAS Model (CELM DAN-EGWO-ARMA-SVR)

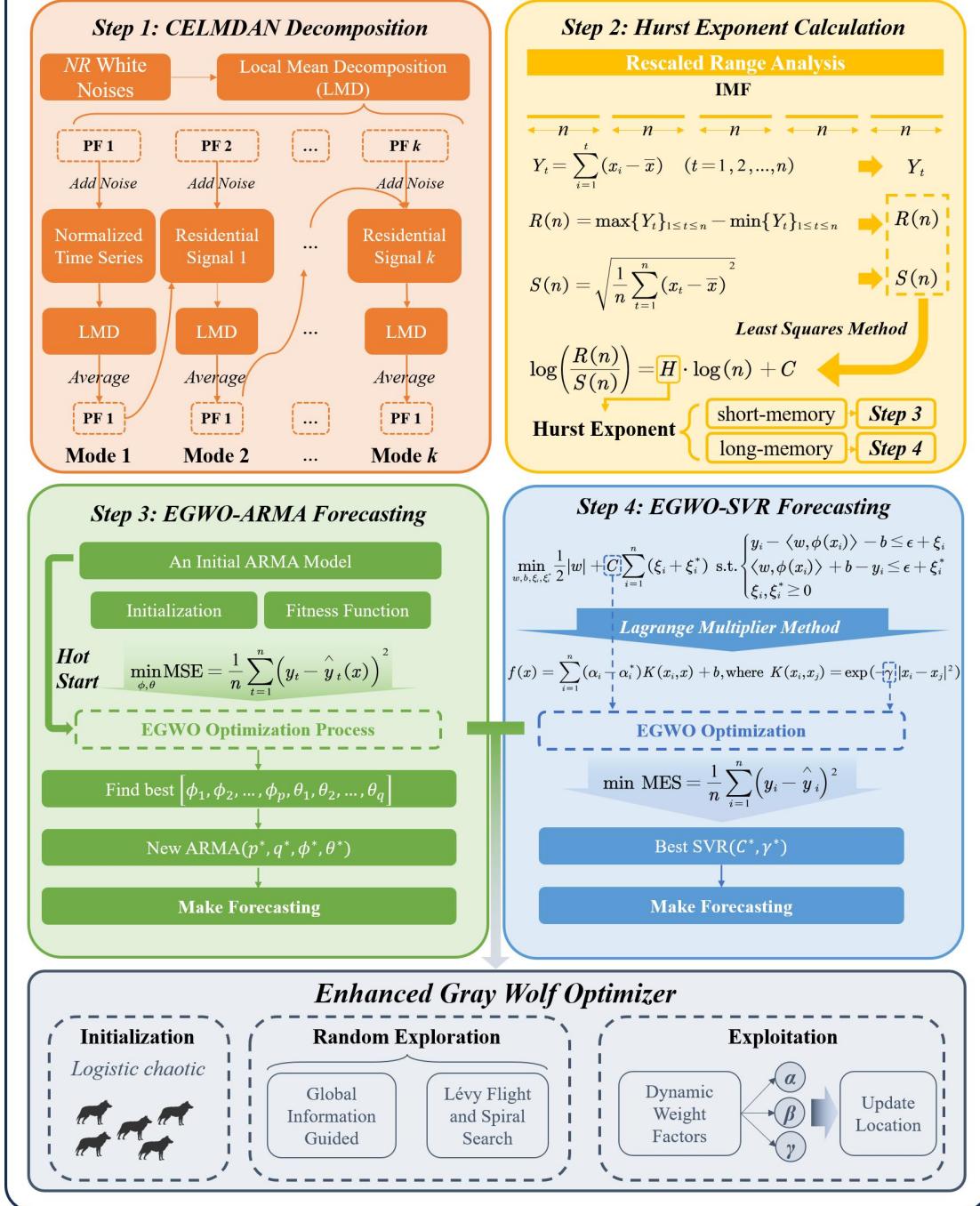


Figure 1: Framework of the proposed CEAS model for crude oil price forecasting.
Illustration of the proposed CEAS model, consisting of key steps: signal decomposition using CELMDAN, Hurst exponent calculation, forecasting with EGWO-optimized ARMA and EGWO-optimized SVR

2.2 Complete Ensemble LMD with Adaptive Noise

2.2.1 Local Mean Decomposition

Local Mean Decomposition (LMD) is an adaptive signal processing method designed to address the analysis of nonlinear and non-stationary signals (Jonathan, 2005). LMD decomposes complex signals into a series of Product Function (PF) components, where each PF is the product of an envelope signal and a pure frequency modulation and amplitude modulation signal. The main process is:

(a) Finding extreme points: For a given signal $x(t)$, identify all local maxima and minima points.

(b) Local mean and local amplitude calculation: Compute the local mean m_i and local amplitude a_i for any adjacent extremum points n_i and n_{i+1} :

$$m_i = \frac{x(n_i) + x(n_{i+1})}{2}, \quad a_i = \frac{|x(n_i) - x(n_{i+1})|}{2}. \quad (1)$$

(c) Interpolation: Using spline interpolation of local means and amplitudes, construct local mean function $m_{11}(t)$ and local amplitude function $a_{11}(t)$:

$$m_{11}(t) = \text{spline}(m_i, t), \quad a_{11}(t) = \text{spline}(a_i, t) \quad (2)$$

(d) Derive the modulation signal $h_{11}(t)$:

$$h_{11}(t) = \frac{x(t) - m_{11}(t)}{a_{11}(t)}. \quad (3)$$

(e) Determine the modulation signal: If $h_{11}(t)$ is not a pure modulation signal (i.e., close to the form $\cos(\omega(t))$), repeat the steps above, processing $h_{11}(t)$ to obtain new $m_{12}(t)$ and $a_{12}(t)$:

$$h_{12}(t) = \frac{h_{11}(t) - m_{12}(t)}{a_{12}(t)}. \quad (4)$$

Continue iterating until $h_{1n}(t)$ satisfies:

$$\lim_{n \rightarrow \infty} m_{1n}(t) \approx 0, \lim_{n \rightarrow \infty} a_{1n}(t) \approx 1. \quad (5)$$

(f) Instantaneous amplitude function and final modulation function calculation: Multiply all envelope functions to derive the instantaneous amplitude function $a_1(t)$: $a_1(t) = a_{11}(t) \cdot a_{12}(t) \cdot \dots \cdot a_{1n}(t)$. The final modulation function is $s_1(t) = h_{1n}(t)$.

(g) PF₁ calculation:

$$PF_1(t) = a_1(t) \cdot s_1(t). \quad (6)$$

(h) Residual signal calculation: Subtract the first PF from the original signal to obtain the residual signal $u_1(t)$:

$$u_1(t) = x(t) - PF_1(t). \quad (7)$$

Use $u_1(t)$ as the new input signal and repeat the above steps until the residual component contains extremum points fewer than a predefined threshold.

(i) Signal reconstruction: $x(t)$ can be expressed as:

$$x(t) = \sum_{i=1}^k PF_i(t) + u_k(t), \quad (8)$$

where k is the total number of PFs, and $u_k(t)$ is the final residual component.

2.2.2 Complete Ensemble LMD with Adaptive Noise

The Complete Ensemble Local Mean Decomposition with Adaptive Noise (CELM DAN) is an improved algorithm developed to address issues such as mode mixing, incomplete decomposition, and noise sensitivity encountered by LMD when processing complex signals. Although LMD exhibits superior performance compared to EMD in handling nonlinear and non-stationary signals, it still faces challenges like mode mixing and endpoint effects (Liang et al., 2020). CELMDAN effectively mitigates these problems by incorporating ensemble techniques and adaptive noise strategies. The framework of CELMDAN is shown in Figure 2.

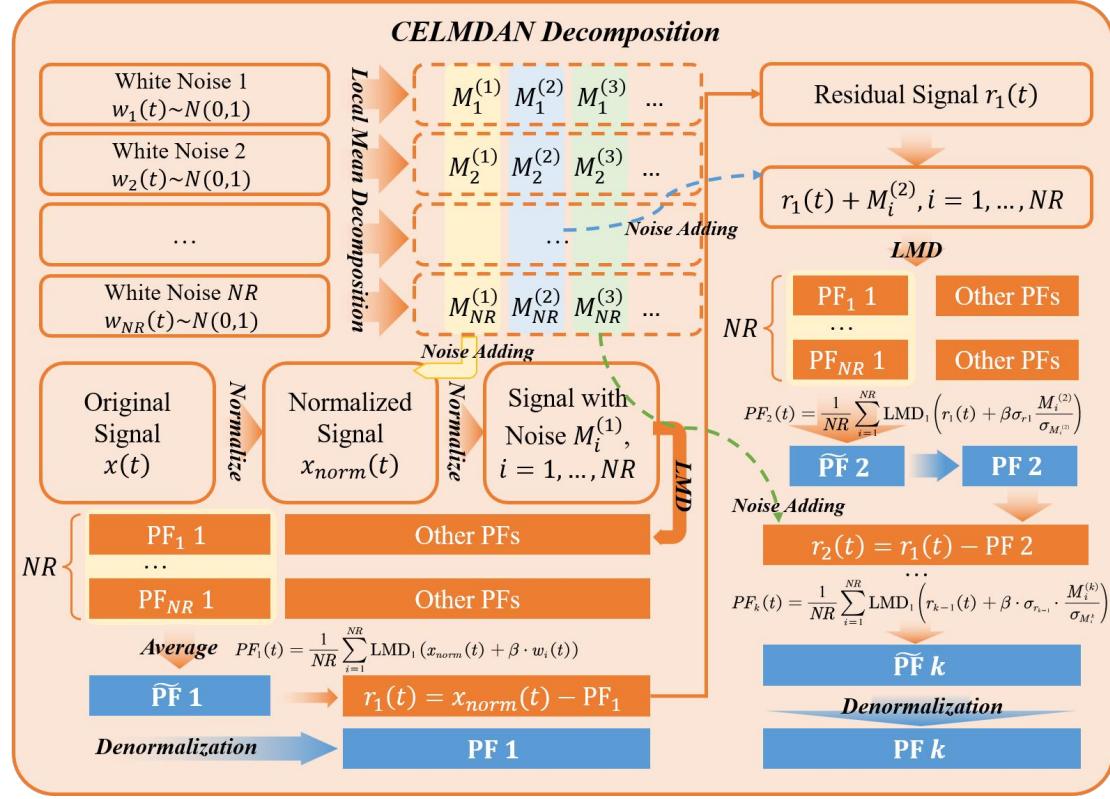


Figure 2: The framework of CELMDAN, including signal normalization.

Detailed flowchart of the CELMDAN decomposition process, which enhances LMD by integrating ensemble averaging and adaptive white noise injection, enabling robust extraction of Product Functions (PFs) through iterative noise-assisted decomposition and residual updating.

The main process of CELMDAN are as follows:

(a) **Signal Preprocessing:** Normalize the signal $x(t)$ to obtain the standardized signal:

$$x_{norm}(t) = \frac{x(t)}{\sigma_x}, \quad (9)$$

where σ_x is the standard deviation of the signal $x(t)$.

(b) **Noise Ensemble Generation:** Generate NR sets of white noise sequences:

$$w_i(t) \sim N(0,1), \quad i = 1, 2, \dots, NR.$$

(c) **Noise Mode Library Construction:** Decompose each set of white noise sequences using LMD to construct the noise mode library:

$$\{M_i^j\} = \text{LMD}(w_i(t)), \quad (10)$$

where M_i^j denotes the j -th mode of the i -th noise set.

(d) Extraction of the First Mode Component: Compute the first PF by integrating NR noise-added LMD results:

$$\text{PF}_1(t) = \frac{1}{NR} \sum_{i=1}^{NR} \text{LMD}_1(x_{norm}(t) + \beta \cdot w_i(t)), \quad (11)$$

where LMD_1 indicates the extraction of the first mode from LMD, and β is the noise control parameter. In this way, the impact of accidental errors can be eliminated and the robustness of the decomposition can be improved.

(e) Residual Signal Calculation: Calculate the residual signal after removing the first mode:

$$r_1(t) = x_{norm}(t) - \text{PF}_1(t). \quad (12)$$

(f) Subsequent Mode Iterative Extraction: Using an adaptive noise strategy, for the k -th mode:

$$\text{PF}_k(t) = \frac{1}{NR} \sum_{i=1}^{NR} \text{LMD}_1 \left(r_{k-1}(t) + \beta \cdot \sigma_{r_{k-1}} \cdot \frac{M_i^k}{\sigma_{M_i^k}} \right), \quad (13)$$

where $r_{k-1}(t)$ is the residual signal after extracting the $(k-1)$ -th mode, $\sigma_{r_{k-1}}$ is its standard deviation, and $\frac{M_i^k}{\sigma_{M_i^k}}$ is the normalized noise mode.

(g) Residual Signal Update: After extracting the k -th mode, update the residual signal:

$$r_k(t) = r_{k-1}(t) - \text{PF}_k(t). \quad (14)$$

(h) Termination Condition Check: Evaluate the number of local extrema in the residual signal. Terminate the decomposition process when the count is below a predetermined threshold (typically 2):

$$\text{if } N_{extrema}(r_k(t)) \leq 2 \text{ then stop.} \quad (15)$$

(i) Final Residual as the Last Mode: Add the final residual signal as the last mode in the decomposition results:

$$\text{PF}_{k+1}(t) = r_k(t). \quad (16)$$

(j) Result Denormalization: Denormalize all extracted modes to restore the original signal scale:

$$\text{PF}_j^{final}(t) = \text{PF}_j(t) \cdot \sigma_x, \quad j = 1, 2, \dots, k+1. \quad (17)$$

In LMD, to alleviate endpoint effects, the signal is symmetrically extended as :

$$x_{ext}(t) = [\text{fliplr}(x(1:L_{pad})), x(t), \text{fliplr}(x(\text{end} - L_{pad} + 1:\text{end}))], \quad (18)$$

where L_{pad} represents the padding length (such as 20% of the length). After performing LMD on $x_{ext}(t)$, the extended portions are then removed.

[Figure 3](#) presents the decomposition results of Dubai crude oil price data (from May 20, 2020, to February 18, 2025) using EMD, LMD, and CELMDAN, respectively.

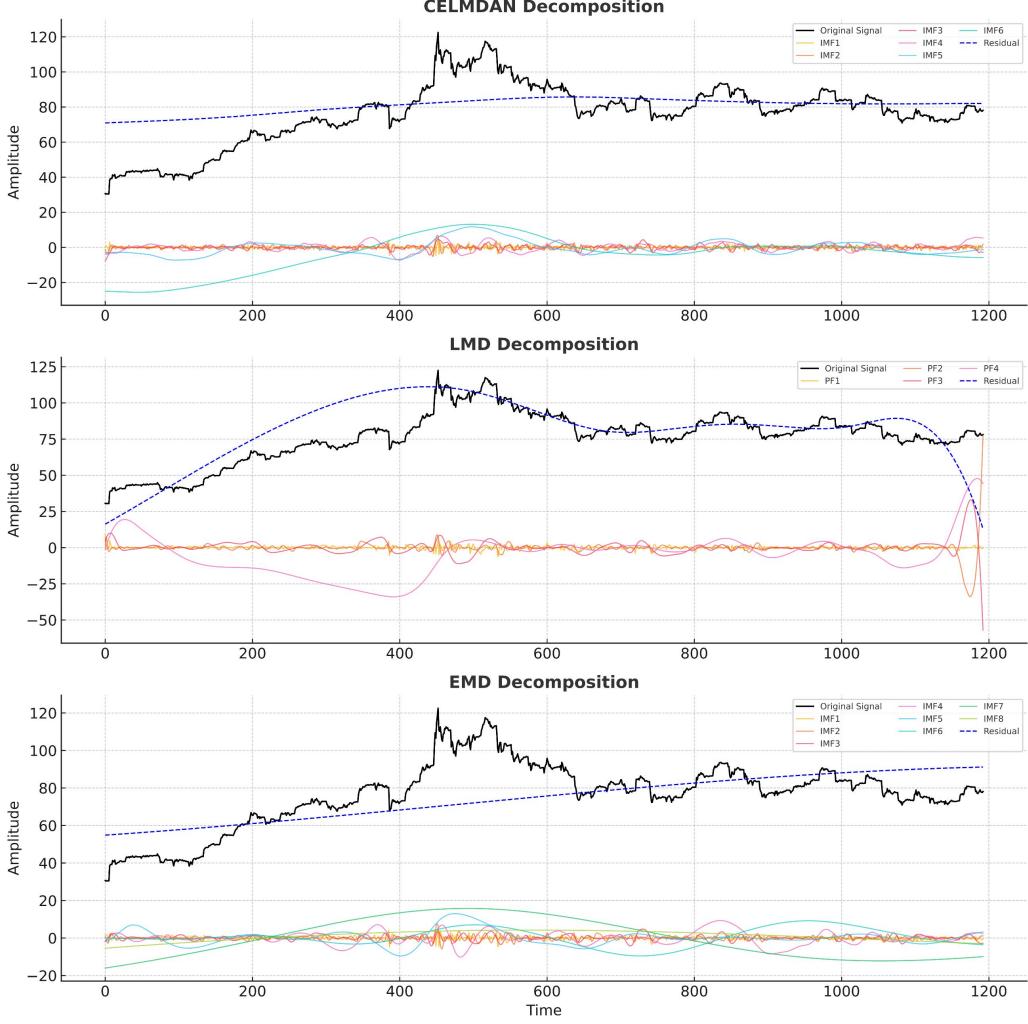


Figure 3: Comparison of signal decomposition results using CELMDAN, LMD, and EMD.

Comparison of signal decomposition results using CELMDAN, LMD, and EMD on crude oil price data. CELMDAN exhibits superior endpoint smoothness and clearer mode separation, while LMD shows significant boundary distortion and EMD presents evident mode mixing, demonstrating the robustness of CELMDAN against endpoint effects and frequency overlap.

As we can see from [Figure 3](#):

(a) Superior Alleviation of Endpoint Effects: In the LMD decomposition results, endpoint effects are significant, as evidenced by the sharp rise at the tail ends of PF2 and PF3. However, this issue in CELMDAN is mitigated. This advantage is attributed to the noise perturbation averaging strategy and the symmetric extension mechanism introduced.

(b) Significant Mitigation of Mode Mixing: In the LMD and EMD decomposition results, such as PF3 and IMF3~IMF5, significant frequency band overlap is observed. Conversely, the modes obtained from CELMDAN exhibit clear band-pass characteristics, making them more conducive for analysis.

2.3 Hurst Exponent Calculation

The Hurst exponent (denoted as H) is a statistical measure used to quantify the long-term memory and self-similarity of a time series (Florido et al., 2025). It serves as a fundamental indicator of the persistence characteristics of stochastic processes.

In this study, the Hurst exponent is estimated using the Rescaled Range (R/S) Analysis method. The procedure is as follows:

(a) Division of the Time Series. Given the time series $\{x_1, x_2, \dots, x_N\}$ of length N , we divide it into $m = \left\lfloor \frac{N}{n} \right\rfloor$ non-overlapping sub-series of equal length n , for multiple values of n (e.g., $n = 10, 20, 40, \dots$).

(b) Calculation of the mean. For each sub-series $X^{(i)} = \{x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}\}$, compute the mean as:

$$\bar{X}^{(i)} = \frac{1}{n} \sum_{k=1}^n x_k^{(i)}. \quad (19)$$

(b) Construction of the cumulative deviation series: For i , construct the cumulative deviation series as:

$$Y_k^{(i)} = \sum_{j=1}^k (x_j^{(i)} - \bar{X}^{(i)}), \quad k = 1, 2, \dots, n. \quad (20)$$

(c) Calculation of the Range. For i , calculate the range of cumulative deviation series as:

$$R^{(i)}(n) = \max_{1 \leq k \leq n} Y_k^{(i)} - \min_{1 \leq k \leq n} Y_k^{(i)}. \quad (21)$$

(d) Calculation of the Standard Deviation: For i , compute the standard deviation of the original series as:

$$S^{(i)}(n) = \sqrt{\frac{1}{n} \sum_{k=1}^n (x_k^{(i)} - \bar{X}^{(i)})^2}. \quad (22)$$

(e) Computation of the Rescaled Range. For i , calculate R/S as:

$$(R/S)^{(i)} = \frac{R^{(i)}(n)}{S^{(i)}(n)}. \quad (23)$$

(f) Estimation of the hurst exponent.

Take logarithms of the segment sizes n and their corresponding average rescaled ranges. Fit a linear model between $\log(n)$ and $\log(R/S(n))$, where $R/S(n)$ is calculated as the average of $(R/S)^{(i)}$ in Equation (23). The slope of the regression line corresponds to the Hurst exponent H :

$$\log(R/S(n)) = H \log(n) + C \quad (24)$$

This classification provides a theoretically grounded basis for assigning suitable forecasting models to each component.

2.4 Enhanced Grey Wolf Optimizer

2.4.1 Grey Wolf Optimizer

The Grey Wolf Optimizer (GWO) is a swarm intelligence optimization algorithm ([Mirjalili et al., 2014](#)). It emulates the four-level hierarchy within grey wolf packs: Alpha (α), Beta (β), Delta (δ), and Omega (ω) wolves, where the Alpha represents the current best solution, and Beta and Delta are the second and third best solutions, respectively. The core of GWO revolves around modeling the behaviors of “encircling, tracking, and attacking prey”. The position of an individual represents a potential solution to the problem, with the encircling process simulated to iteratively update and refine the solutions, thereby gradually approaching the global optimum. The main equations involved are:

Distance calculation:

$$D = |C \cdot X_p(t) - X(t)|. \quad (25)$$

Position update:

$$X_p(t+1) = X_p(t) - A \cdot D. \quad (26)$$

Here, X_p represents the prey (i.e., the position of α , β , or δ), X is the current position of an individual, $A = 2a \cdot r_1 - a$, $C = 2r_2$, and a is a parameter linearly decreased from 2 to 0. Both r_1 and r_2 are random vectors in $[0,1]$.

Finally, the positions of all individuals are updated based on the positions of α , β , and δ as follows:

$$X(t+1) = \frac{1}{3}(X_1 + X_2 + X_3), \quad (27)$$

where X_1 , X_2 , and X_3 are candidate positions calculated based on α , β , and δ wolves, as determined in [Equation \(25\)](#).

2.4.2 Enhanced Grey Wolf Optimizer

The traditional Grey Wolf Optimizer (GWO) faces several challenges:

(a) Unreliable random initialization strategy: This random initialization method can lead to an uneven distribution of solutions, potentially causing the model to miss better solutions.

(b) Insufficient exploration capability in early stages: Due to the leadership of the Alpha wolf in GWO, the algorithm lacks sufficient exploration ability and may fall into a local optimum.

(c) Unrealistic equal-weight allocation mechanism: The weights given to Alpha, Beta, and Delta wolves are equal, disregarding the differential contributions of optimal solutions,

which do not align with the natural characteristics of wolf packs. This effect further causes the model to continue to approach the local optimal solution.

Based on the above, Enhanced Grey Wolf Optimizer (EGWO) was proposed to overcome these problems. The framework of EGWO is shown in Figure 4. The main improvements are as follows.

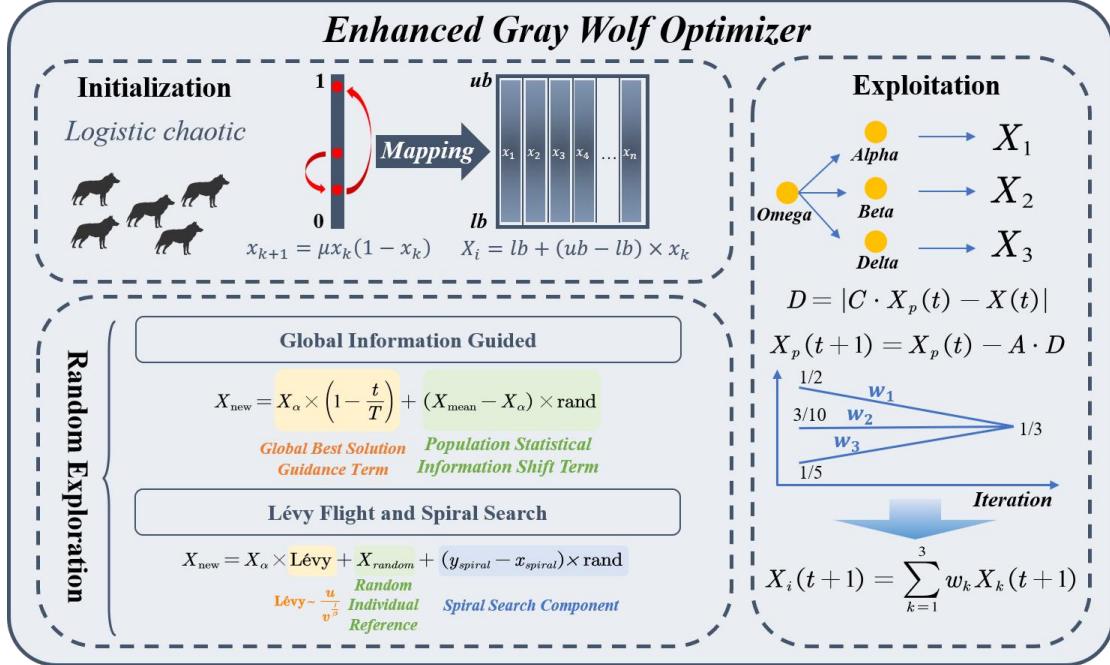


Figure 4: The framework of Enhanced Grey Wolf Optimizer.

Schematic overview of the Enhanced Grey Wolf Optimizer (EGWO), highlighting its improvements over traditional GWO: logistic chaotic initialization for population diversity, dual random exploration strategies (global information guidance and Lévy flight with spiral search), and dynamic weighting in the exploitation phase to enhance convergence and escape local optima.

(a) Population Initialization

The Enhanced Grey Wolf Optimizer (EGWO) utilizes a Logistic chaotic map to generate the initial population:

$$x_{k+1} = \mu x_k (1 - x_k), \quad k = 0, 1, 2, \dots, \quad (28)$$

where $\mu = 4.0$ is the control parameter.

Chaotic Initialization Process:

- Generate random initial values $x_0 \in (0,1)$ for each grey wolf.
- Iterate the Logistic map sufficient times to eliminate transients and achieve a stable chaotic state.
- Map the chaotic sequence to the search space:

$$X_i = lb + (ub - lb) \times x_k, \quad (29)$$

where lb and ub are the lower and upper bounds of the parameter search space, respectively.

Initializing the population using the Logistic chaotic map helps enhance the population diversity and improve the algorithm's global search capability.

(b) Introduction of Random Exploration Modes

In the early stages of the algorithm (the first 2/3 of iterations), we introduce two random exploration modes to enhance the wolves' exploratory capability and prevent them from getting

trapped in local optima. The wolf pack randomly selects between the following two exploration strategies:

(i) Global Information Guided Exploration

Individual wolves update their positions according to the following formula:

$$X_{\text{new}} = X_{\alpha} \times \left(1 - \frac{t}{T}\right) + (X_{\text{mean}} - X_{\alpha}) \times \text{rand}, \quad (30)$$

where X_{α} is the current global best solution, X_{mean} is the current population mean position, t is the current iteration number, T is the total number of iterations, and rand is a random number in $[0, 1]$. This strategy comprises two key components:

(1) Global Best Solution Guidance Term:

$$X_{\alpha} \times \left(1 - \frac{t}{T}\right), \quad (31)$$

with its influence gradually decreasing as iterations progress.

(2) Population Statistical Information Shift Term:

$$(X_{\text{mean}} - X_{\alpha}) \times \text{rand}, \quad (32)$$

introducing random perturbations to explore unknown areas. When the global best solution is distant from the population mean, this strategy facilitates broader search capabilities, increasing the likelihood of discovering potential high-quality solutions.

(ii) Lévy Flight and Spiral Search

Individual wolves update their positions using the following formula:

$$X_{\text{new}} = X_{\alpha} \times \text{Lévy}_d + X_{\text{random}} + (y_{\text{spiral}} - x_{\text{spiral}}) \times \text{rand}, \quad (33)$$

where the step length for Lévy flight is generated using the Mantegna algorithm ([Tarkhaneh and Shen, 2019](#)):

$$\text{Lévy}_d \sim \frac{u}{v^{1/\beta}}, \quad (34)$$

with $u \sim N(0, \sigma_u^2)$, $\sigma_u = \left(\frac{\Gamma(1+\beta) \sin(\pi\beta/2)}{\Gamma((1+\beta)/2)\beta 2^{(\beta-1)/2}}\right)^{1/\beta}$, $v \sim N(0, 1)$, and $\beta = 1.5$, which controls the heavy-tailed nature of the distribution.

The spiral search component is defined by the parametric equations:

$$r(t) = r_0 + u \times t, \quad (35)$$

$$\phi(t) = -\omega \times t + \phi_0, \quad (36)$$

$$x_{\text{spiral}}(t) = r(t) \times \sin(\phi(t)), \quad (37)$$

$$y_{\text{spiral}}(t) = r(t) \times \cos(\phi(t)), \quad (38)$$

where $r_0 = 10$ is the initial radius, $u = 0.0265$ is the spiral expansion rate, $\omega = 0.005$ is the angular velocity, and $\phi_0 = \frac{3}{2}\pi$ is the initial phase.

This strategy comprises three key components:

(i) Lévy Flight: $X_{\alpha} \times \text{Lévy}_d$, utilizing a heavy-tailed distribution to achieve both local fine-tuning and occasional long-distance jumps.

(ii) Random Individual Reference: X_{random} , selecting a random individual from the population as a reference point.

(iii) Spiral Search Component: $(y_{\text{spiral}} - x_{\text{spiral}}) \times \text{rand}$, introducing a deterministic spiral motion pattern.

Through this approach, the wolf pack can explore a wide range of search spaces while also conducting detailed searches in promising regions.

(c) Dynamic Weight Factors in the Exploitation Phase

During the final third of the iterations, the wolf pack focuses on promising regions. This phase utilizes the following position update mechanism:

$$X_i(t+1) = w_1 X_1 + w_2 X_2 + w_3 X_3, \quad (39)$$

where the dynamic weight factors are defined as:

$$w_1 = 0.5 - \frac{t}{6T}, \quad (40)$$

$$w_2 = 0.3 + \frac{t}{30T}, \quad (41)$$

$$w_3 = 0.2 + \frac{2t}{15T}. \quad (42)$$

Figure 5 illustrates the case of $T = 100$.

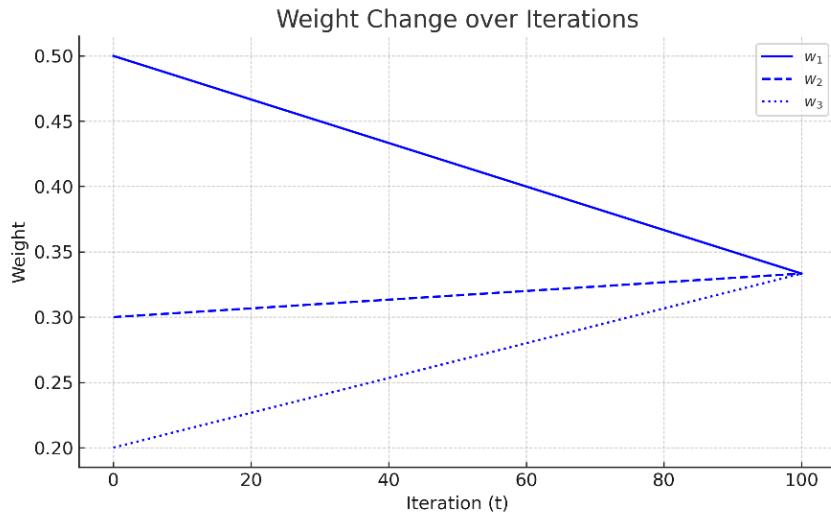


Figure 5: Dynamic weight adjustment for α , β , and δ wolves over iterations in EGWO. Dynamic adjustment of weights over iterations in the EGWO algorithm, where the α wolf's dominance gradually decreases while the contributions of β and δ wolves increase to achieve balanced exploitation in the later optimization stages.

Unlike the simple average used in GWO, EGWO introduces dynamic weights to reflect the distinct roles of the three leading wolves:

(a) The Alpha wolf always dominates. The Alpha wolf, representing the global best solution, is always assigned the highest weight w_1 .

(b) The three weights tend to be average. As the iterations progress, the reliability of the Beta and Delta wolves' solutions increases, and the weight differences among the three wolves gradually diminish. Finally, their weights are the same as $1/3$.

Through the above mechanisms, EGWO establishes an optimization framework that balances global exploration with local exploitation, aiming to enhance search efficiency and solution quality.

2.5 EGWO-ARMA

2.5.1 Autoregressive Moving Average

The Autoregressive Moving Average (ARMA) model is a linear model widely used in time series modeling and forecasting (Box and Jenkins, 1970). This model combines the strengths of both Autoregressive (AR) and Moving Average (MA) components, making it suitable for modeling stationary time series (Li et al., 2021). Mathematically, given a stationary time series $\{y_t\}$, the ARMA(p, q) model is defined as follows:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t, \quad (43)$$

where y_t is the observation at time t , c is constant term, $\phi_1, \phi_2, \dots, \phi_p$ are coefficients of the AR component, $\theta_1, \theta_2, \dots, \theta_q$ are coefficients of the MA component, ε_t is the white noise term, which follows an independent and identically distributed (i.i.d.) process with mean 0 and variance σ^2 .

The core idea of the ARMA is that the value at the current time can be represented as a linear combination of values from several previous time periods (AR) and the errors from several previous periods (MA).

The modeling of ARMA usually involves two steps:

(a) Selection of Model Order:

In this study, by fixing the differencing order $d = 0$ (Stationarity tests indicate that PFs components with a Hurst exponent less than 0.5 do not require differencing.), we first explore the candidate combinations of (p, q) using a search method, employing the Akaike Information Criterion (AIC) as the criterion for evaluating model performance: $AIC = -2 \log L + 2k$, where $\log L$ is the log-likelihood value of the model, and k is the number of parameters in the model. The smaller the AIC value, the better and more parsimonious the model fit.

(b) Parameter Estimation and Prediction:

Once the optimal order is determined, the maximum likelihood estimation method is used to fit the AR and MA coefficients, which are then used to forecast future samples.

2.5.2 EGWO-ARMA

Although traditional methods can estimate the parameters of an ARMA model, the results are often susceptible to initial values, extremum problems, and sample complexity. To improve the accuracy and robustness of the model, this paper introduces the Enhanced Grey Wolf Optimizer (EGWO) to jointly optimize the AR coefficients ϕ_i and MA coefficients θ_j in ARMA.

The objective of the optimizer is to find an optimal parameter vector $[\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q]$ for ARMA, minimizing the mean squared error (MSE) of the prediction residuals on the training set:

$$\min_{\phi, \theta} \text{MSE} = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t(x))^2. \quad (44)$$

The main optimization process is as follows:

(a) Initialization:

An initial ARMA model is fitted using maximum likelihood estimation, and its coefficients are used as one of the initial solutions for the EGWO population. Other individuals are initialized using a logistic chaotic map to enhance population diversity, as described in Equations (28)-(29). The dimension of search space is set as $p + q$.

(b) Fitness Function Construction:

An error sequence is constructed based on the prediction residuals on the training set, and the MSE is calculated as the fitness value in Equation (48).

(c) EGWO Optimization Process:

EGWO optimizes the parameter combination through a two-stage strategy (exploration and exploitation):

Exploration Phase (first 2/3 iterations): According to Equations (30)-(38), the algorithm utilizes two strategies to enhance global search capability: global guidance updates based on the population mean and the Alpha wolf; random perturbation search combining Lévy flight and spiral motion mechanisms.

Exploitation Phase (last 1/3 iterations): The algorithm employs the traditional GWO update strategy, introducing a dynamic weighting factor to assign differentiated weights to the guidance influence of Alpha, Beta, and Delta wolves as detailed in Equations (40)-(42).

(d) Application of Optimal Parameters and Model Prediction:

The optimal parameter combination (ϕ^*, θ^*) obtained through EGWO is used to build an ARMA forecast model on the test set and make predictions as described in Equation (43).

2.6 EGWO-SVR

2.6.1 Support Vector Regression

Support Vector Regression (SVR) (Fan et al., 2016) is a regression method developed as an extension of Support Vector Machines (SVM) (Zhang et al., 2025), with the core idea of finding an optimal function to perform nonlinear fitting of sample data while maintaining generalization capability. Unlike traditional least squares regression, SVR constructs an " ϵ -insensitive loss function" that ignores deviations within an ϵ range and penalizes only errors that exceed this range, thus enhancing the robustness of the model.

For a given training sample set $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n$, $x_i \in \mathbb{R}^d$, $y_i \in \mathbb{R}$, SVR aims to find a function $f(x) = \langle w, \phi(x) \rangle + b$, where $\phi(x)$ is a nonlinear mapping function that maps the input space to a high-dimensional feature space, and w is the weight vector, and b is the bias term. The objective is to ensure that the prediction error of the function for all training samples does not exceed ϵ , while keeping the model's complexity as low as possible. The formulation of its optimization problem is:

$$\min_{w,b,\xi_i,\xi_i^*} \frac{\|w\|}{2} + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad \text{s.t.} \begin{cases} y_i - \langle w, \phi(x_i) \rangle - b \leq \epsilon + \xi_i \\ \langle w, \phi(x_i) \rangle + b - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (45)$$

where $C > 0$ is the penalty parameter that controls the trade-off between fitting accuracy and model complexity; ξ_i, ξ_i^* are slack variables that allow for certain errors.

By introducing Lagrange multipliers and dualizing, the SVR solution can be expressed through a kernel function:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(x_i, x) + b, \quad (46)$$

where $K(x_i, x) = \langle \phi(x_i), \phi(x) \rangle$ is the kernel function. Common kernel functions include linear, polynomial, and Radial Basis Function (RBF) kernels. In this study, the RBF kernel is used, as given by:

$$K(x_i, x_j) = \exp(-\gamma|x_i - x_j|^2), \quad (47)$$

where γ is a kernel parameter that controls the distance of the mapping for samples. The performance of the SVR model in practical applications is heavily dependent on the setting of parameters C and γ .

2.6.2 EGWO-SVR

The two critical hyperparameters in the SVR model, the penalty factor C in Equation (45) and the RBF kernel parameter γ in Equation (47), significantly impact the model's predictive

performance (Frénay and Verleysen, 2011). Traditional grid search methods have high computational costs and may get trapped in local optima, making them unsuitable for modeling high-dimensional complex data (Wu et al., 2009). Therefore, we employ an improved EGWO to jointly optimize the SVR model parameters, enhancing the model's generalization ability and prediction accuracy.

The parameter optimization process is as follows:

(a) Define the optimization objective function: Use the prediction error of the SVR model on the training set, such as root mean square error (RMSE), as the fitness function:

$$\text{Fitness} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}. \quad (48)$$

(b) Search space construction: Set the search space for the penalty factor $C \in [0.01, 100]$ and the kernel parameter $\gamma \in [0.01, 100]$.

(c) Population initialization: Use Logistic chaotic mapping for initial population generation as formulated in Equations (28)-(29), to enhance the diversity and global search capability of the population, preventing it from getting trapped in local optima.

(d) Optimization phase division and strategy selection: As detailed in Equations (30)-(38), during the first two-thirds of the algorithm, the algorithm employs a random exploration based on global mean guidance, combined with Lévy flight and spiral search strategies to effectively expand the search range. In the final third of the algorithm, switch to a development strategy by introducing dynamic weight mechanisms that assign different updating weights to the Alpha, Beta, and Delta wolves ,determined by Equations (40)-(42), enhancing precise search performance.

(e) Obtaining optimal parameters: After the iterations, select the individual with the minimum fitness value, whose position represents the optimal (C^*, γ^*) combination.

(f) Model construction and prediction: Using the optimal parameters, train the SVR model and make predictions on the test set.

3 Error Indicators and Datasets

3.1 Error Indicators

In order to measure the predictive capability of the models, the following indicators were selected to evaluate their forecasting performance.

Firstly, the accuracy of the selected model's numerical predictions is assessed by RMSE as the square root of the mean of the squared errors, and MAPE as the average of the absolute errors divided by the actual values (Xie et al., 2013).

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (x(t) - \hat{x}(t))^2}, \quad (49)$$

$$\text{MAPE} = \frac{1}{N} \sum_{t=1}^N \left| \frac{x(t) - \hat{x}(t)}{x(t)} \right|, \quad (50)$$

where $x(t)$ is the actual value, $\hat{x}(t)$ is the predicted value, and N is the number of observations.

Moreover, when predicting crude oil prices, the accuracy of direction predictions is critical for many decisions. Thus, incorporating Direction Accuracy (DA) is important, as it targets the agreement in change direction between predicted and observed values (Zheng et al., 2023).

$$\text{DA} = \frac{1}{N-1} \sum_{t=2}^N I((x(t) - x(t-1))(\hat{x}(t) - \hat{x}(t-1)) > 0), \quad (51)$$

where $I(\cdot)$ refers to the indicator function, yielding a value of 1 when the specified condition is met and 0 when it is not.

Furthermore, we introduce R^2 to represent the proportion of variance in the input variables explained by the model. A larger R^2 indicates that the model accounts for a higher proportion of the variance (He et al., 2021).

$$R^2 = 1 - \frac{\sum_{t=1}^n (x(t) - \hat{x}(t))^2}{\sum_{t=1}^n (x(t) - \bar{x})^2}, \quad (52)$$

3.2 Datasets

To verify the effectiveness of the proposed CELMDAN-EGWO-ARMA-SVR forecasting model, we selected three major crude oil price datasets—Dubai, Brent, and WTI—as experimental samples.

The historical crude oil price data were sourced from the U.S. Energy Information Administration (EIA) (<https://www.eia.gov/>). The data collection period spans from May 20, 2020, to February 18, 2025, post-COVID-19, in order to mitigate the issues of anomalous fluctuations in data caused by the impact of COVID-19 on crude oil prices.

To evaluate the forecasting performance of the model, the selected time period is divided into a training set (70%) and a testing set (30%) (Liang et al., 2025). The model is trained exclusively on the training set—including both decomposition and prediction—and then used to forecast the testing set, with the predicted values compared against the actual values.

4 Experiment

4.1 Experiment I: Comparison with classic baseline models

In this section, we select some mainstream prediction models from previous references as baseline models. These models can be categorized into three groups: traditional econometric models, classic machine learning models, and commonly used deep learning models (Chen and Tian, 2025). As a representative of traditional econometric models, ARIMA is selected as the baseline model (Mati et al., 2023). ELM, KELM, and SVR are selected as representatives of classic machine learning models (Yu et al., 2025). Finally, BPNN and LSTM represent common deep learning models (Fan et al., 2016). The corresponding experimental results are listed in Figure 6 and Table 1. By these results, we can conclude:

(a) Compared with single baseline models, the prediction performance of the proposed CEAS is significantly improved. From Figure 4 and Table 1, it is evident that our proposed model achieves the best performance in terms of RMSE, MAPE, and R², and also demonstrates highly competitive results on DA. Specifically, regarding MAPE, ARIMA achieves 0.0089, 0.0131, and 0.0139 on Dubai, Brent, and WTI datasets, respectively, ranking as the second-best performing baseline model. In contrast, the proposed CEAS significantly outperforms ARIMA, obtaining much lower MAPE values of 0.0070, 0.0082, and 0.0087 on Dubai, Brent, and WTI datasets, respectively.

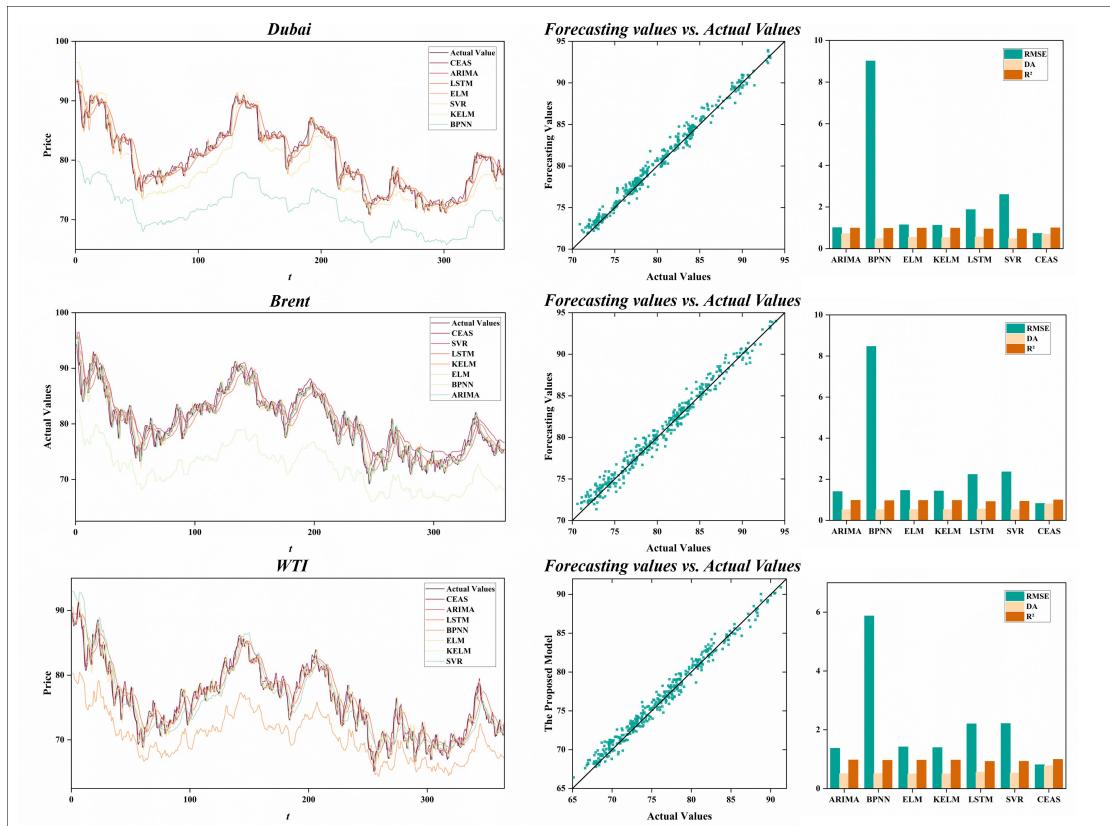
(b) Single models generally fail to match the predictive accuracy achieved by hybrid models. Typically, individual models suffer from inherent limitations that are difficult to overcome. For example, traditional econometric models often struggle to capture the nonlinear relationship among variables. However, hybrid models are able to combine the strengths of different individual models and compensate for their inherent weaknesses. Both SVR and ARIMA are integral components (or a variant) of our proposed model. However, their predictive performance as baseline models is relatively limited. Specifically, compared to ARIMA, our proposed CEAS achieves improvements of (0.0089-0.007)/0.0089=21.35%, (0.0131-0.0082)/0.0131=37.40%, and (0.0139-0.0087)/0.0139=37.41% in MAPE on Dubai, Brent, and WTI datasets, respectively. Similarly, compared to SVR, CEAS achieves improvement of (0.0275-0.0070)/0.0275=74.55%, (0.0233-0.0082)/0.0233=64.81%, and (0.0229-0.0087)/0.0229=62.01% in MAPE on these datasets.

Based on the above comparisons and analyses, it can be concluded that individual models suffer from inherent limitations and perform unsatisfactorily in practical applications. Therefore, integrating multiple models to enhance forecasting performance is necessary.

Table 1

Comparison of prediction performance between the proposed CEAS model and classic baseline models.

Datasets	Indicators	ARIMA	BPNN	ELM	KELM	LSTM	SVR	CEAS
Dubai	RMSE	1.0046	9.0056	1.1392	1.1177	1.8716	2.5949	0.7235
	MAPE	0.0089	0.1078	0.0087	0.0084	0.0152	0.0275	0.0070
	DA	0.7147	0.4741	0.5374	0.5259	0.5575	0.4741	0.6893
	R ²	0.9853	0.9692	0.9778	0.9794	0.9394	0.9390	0.9934
Brent	RMSE	1.3992	8.4603	1.4514	1.4254	2.2323	2.3581	0.8184
	MAPE	0.0131	0.1010	0.0137	0.0134	0.0221	0.0233	0.0082
	DA	0.5190	0.5112	0.5224	0.5196	0.5419	0.5112	0.7912
	R ²	0.9709	0.9555	0.9639	0.9666	0.9115	0.9265	0.9903
WTI	RMSE	1.3673	5.8675	1.4162	1.3915	2.1996	2.2099	0.8088
	MAPE	0.0139	0.0706	0.0147	0.0143	0.0231	0.0229	0.0087
	DA	0.5027	0.5082	0.4973	0.5000	0.5577	0.5275	0.7703
	R ²	0.9710	0.9598	0.9651	0.9672	0.9212	0.9247	0.9911

**Figure 6:** Visual comparison of forecasting performance between CEAS and baseline models.

Forecasting performance comparison across Dubai, Brent, and WTI datasets using various models. The left panels show time series predictions, the middle panels depict actual vs. predicted scatter plots, and the right panels present RMSE, DA, and R² metrics, demonstrating the superior accuracy and stability of the proposed CEAS model over baseline methods.

4.2 Experiment II: Comparison with models based on different decomposition technologies

In previous studies, EMD (Fang et al., 2023), ESMD (He et al., 2023) are often used in data decomposition to forecast crude oil price. In this section, we select them respectively as decomposition methods and compare them with the proposed CEAS model. The results are shown in Figure 7 and Table 2. By analyzing the results, we can draw the following conclusions:

(a) CELMDAN significantly improves prediction compared to EMD and ESMD: The CEAS model demonstrates the best forecasting performance in terms of RMSE, MAPE, and R². In particular, for RMSE, CEAS outperforms the EMD-based model by $(0.9528 - 0.7235)/0.9528 = 24.07\%$, $(0.9234 - 0.8184)/0.9234 = 11.37\%$, $(0.9089 - 0.8088)/0.9089 = 11.01\%$ on Dubai, Brent, and WTI datasets. Compared to the ESMD-based model, the improvements are $(7.0568 - 0.7235)/7.0568 = 0.8975$, $(1.3557 - 0.8184)/1.3557 = 0.3963$, $(7.7246 - 0.8088)/7.7246 = 0.8953$. The improvements indicate that both EMD and ESMD struggle to overcome their weakness, such as the endpoint effect. However, CELMDAN, by taking signal extension and noise strategies, achieves more effective signal decomposition and enhanced predictive accuracy.

(b) Inappropriate decomposition methods may degrade the forecasting performance. As evidenced by the results, ESMD-based model exhibit poor predictive capabilities, especially for the Dubai and WTI datasets—even worse than some single models used in Experiment I. For instance, on the Dubai dataset, the ESMD-based model achieves an R² of 0.8064, whereas ARIMA attains 0.9853, representing a relative improvement of $(0.9853 - 0.8064)/0.8064 = 22.19\%$. Similarly, for the WTI dataset, the ESMD-based model records an R² of 0.5057, while ARIMA reaches 0.9710, yielding a relative gain of $(0.9710 - 0.5057)/0.5057 = 92.01\%$. These suggest that an inappropriate decomposition strategy may significantly compromise prediction accuracy, to the extent that it performs worse than conventional single models.

Based on the analysis, it can be concluded that, due to the endpoint effects and mode mixing, traditional decomposition methods fail to achieve satisfactory performance. CELMDAN alleviates these issues by extending the signal and introducing noise, thereby improving the quality of decomposition and significantly enhancing the predictive accuracy of the model. Moreover, we find that if the employed decomposition algorithm performs poorly, suffering from extreme endpoint effects, the model’s accuracy may decline, even worse than some single models.

Table 2
Comparison of prediction performance of models using different signal decomposition methods (EMD, ESMD, and CELMDAN).

Datasets	Indicators	Using EMD	Using ESMD	CEAS
Dubai	RMSE	0.9528	7.0568	0.7235
	MAPE	0.0099	0.0798	0.007
	DA	0.5678	0.5311	0.6893

Datasets	Indicators	Using EMD	Using ESMD	CEAS
Brent	R ²	0.9899	0.8064	0.9934
	RMSE	0.9234	1.3557	0.8184
	MAPE	0.0087	0.0149	0.0082
	DA	0.7802	0.6429	0.7912
WTI	R ²	0.9876	0.9895	0.9903
	RMSE	0.9089	7.7246	0.8088
	MAPE	0.0095	0.0694	0.0087
	DA	0.7838	0.5378	0.7703
R ²		0.9871	0.5057	0.9911

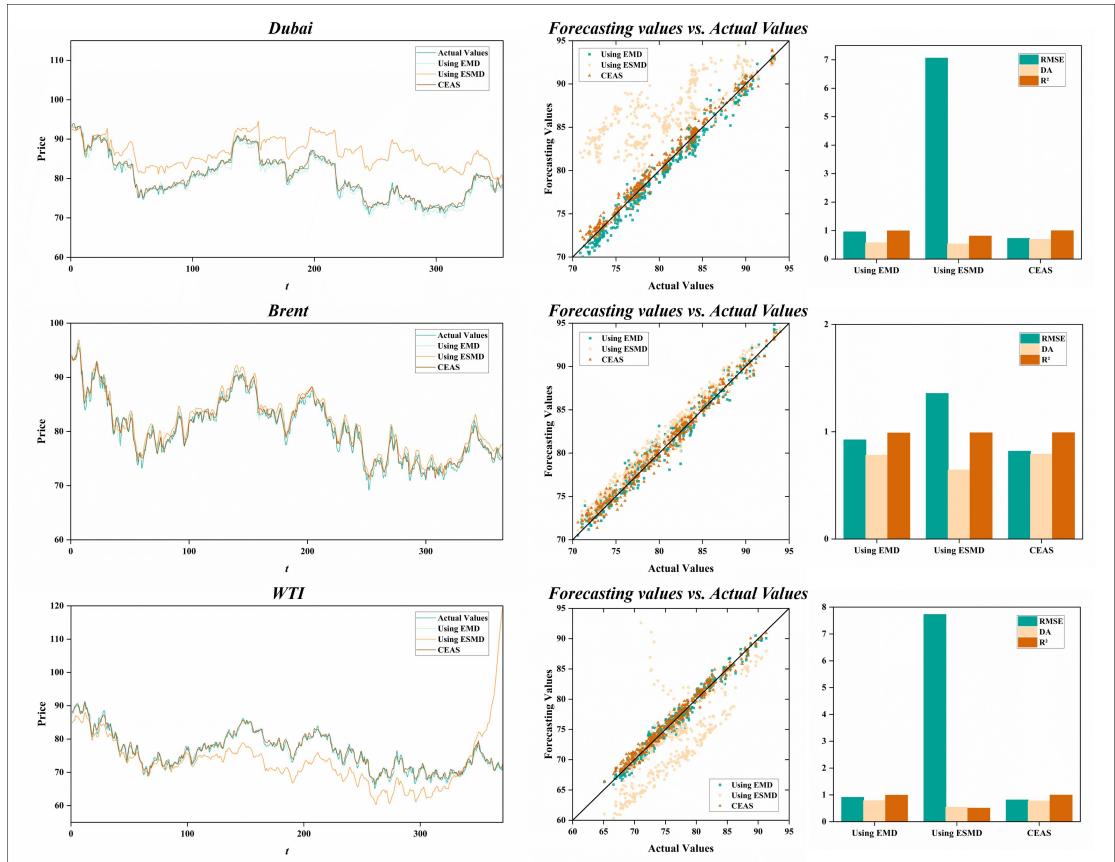


Figure 7: Visual comparison of prediction results using different decomposition methods. Comparison of crude oil price forecasting performance using different decomposition techniques (EMD, ESMD, and CELMDAN-based CEAS) on Dubai, Brent, and WTI datasets. The CEAS model demonstrates more accurate tracking in time series plots (left), tighter alignment in actual vs. predicted scatter plots (middle), and superior performance in RMSE, DA, and R² metrics (right), confirming the advantage of CELMDAN in decomposition quality.

4.3 Experiment III: Comparison with models based on different metaheuristic algorithms

In previous studies, numerous metaheuristic algorithms have been employed for parameter optimization. Ant Colony Optimization (ACO), a classical metaheuristic inspired by the foraging behavior of ants (Pan et al., 2020), is selected for comparison in this section. Exponential-Trigonometric Optimization (ETO), a recent algorithm inspired by a complex combination of

exponential and trigonometric functions (Luan et al., 2024), is also included as a representative of emerging optimization techniques. In this section, only the optimizer for finding the parameters of ARMA and SVR is replaced, while all other components are kept unchanged. The results are presented in [Table 4](#) and [Figure 8](#). Based on the results, the following conclusions can be drawn:

(a) EGWO exhibits superior compatibility with the proposed model. As shown in the results, CEAS, based on EGWO, achieves the best performance across all datasets, outperforming both ACO and ETO. For instance, on the Dubai dataset, EGWO achieves a MAPE of 0.0070, representing a 34.58% and 51.05% improvement over ACO (0.0107) and ETO (0.0143), respectively. Additionally, the DA and R² values of CEAS are consistently higher, indicating more accurate directional prediction and better model fit. These improvements can be attributed to the chaotic initialization and adaptive exploration-exploitation strategies introduced in EGWO, which enhance global search ability and reduce the risk of local optima.

(b) CEAS based on EGWO performs more stably on different datasets. By calculating the variance of each indicator of the three methods based on ACO, ETO, and EGWO on different data sets, we find that the proposed model based on EGWO is the most stable in each indicator. The calculation results of the variance are shown in [Table 3](#). For DA, although the model based on ETO performs better than that based on ACO in terms of variance, CEAS is still lower by $|0.004150 - 0.001931| / 0.004150 = 0.4616\%$. This shows that EGWO can provide more stable and accurate results, making the model have a stronger generalization ability.

Based on the above analysis, we conclude that the proposed CEAS model based on EGWO is superior to the other two models based on other algorithms in both prediction accuracy and stability, so that it is more suitable for crude oil price prediction.

Table 3
Comparison of prediction performance variance of models using different metaheuristic algorithms.

	RMSE	MAPE	DA	R ²
ACO	0.022497	0.000003	0.005575	0.000028
ETO	0.057964	0.000006	0.004150	0.000069
CEAS	0.001819	0.000001	0.001931	0.000002

Table 4
Comparison of prediction performance of models using different metaheuristic algorithms.

Datasets	Indicators	Using ACO	Using ETO	CEAS
Dubai	RMSE	1.1455	1.4291	0.7235
	MAPE	0.0107	0.0143	0.0070
	DA	0.5932	0.6215	0.6893
	R ²	0.9805	0.9702	0.9934
Brent	RMSE	0.9097	0.8664	0.8184
	MAPE	0.0092	0.0085	0.0082
	DA	0.7692	0.7720	0.7912
	R ²	0.9880	0.9894	0.9903
WTI	RMSE	1.2716	0.9949	0.8088
	MAPE	0.0136	0.0104	0.0087
	DA	0.7243	0.7378	0.7703

Datasets	Indicators	Using ACO	Using ETO	CEAS
	R^2	0.9750	0.9856	0.9911

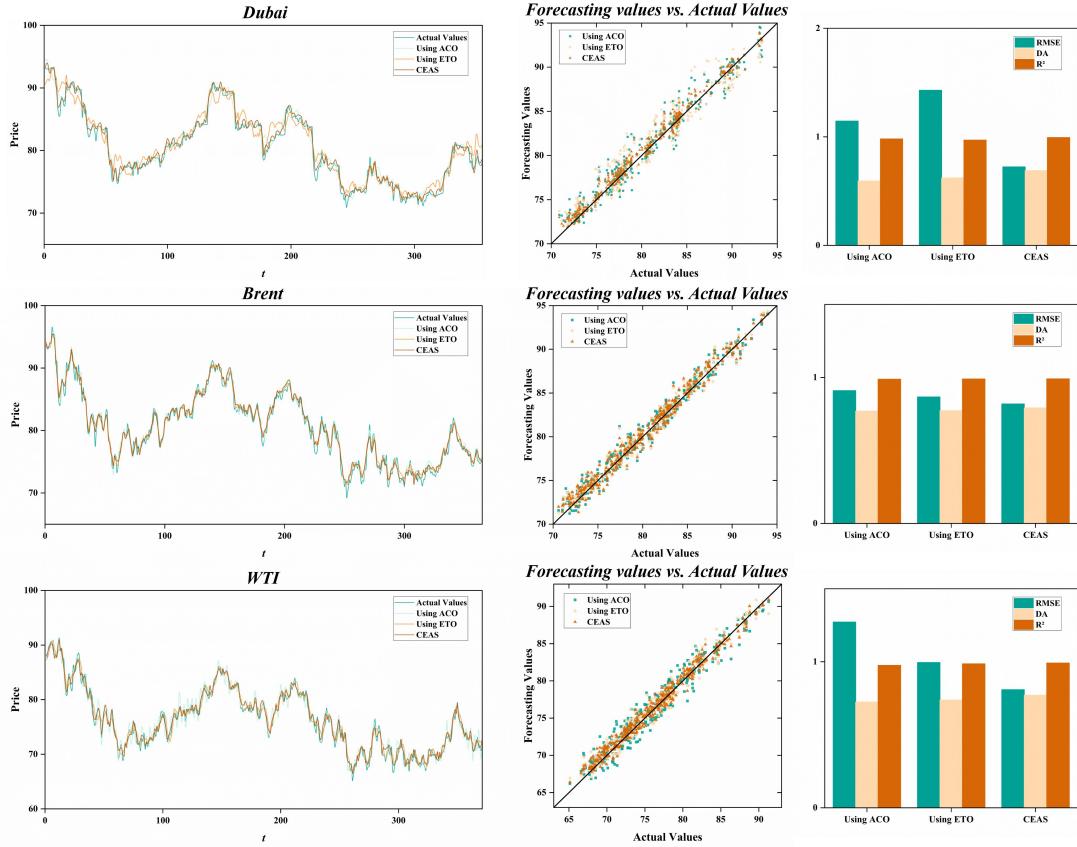


Figure 8: Visual comparison of CEAS model performance using different metaheuristic optimizers (ACO, ETO, and EGWO).

Performance comparison of different metaheuristic optimization algorithms—ACO, ETO, and the proposed EGWO—on crude oil price prediction across Dubai, Brent, and WTI datasets. CEAS, powered by EGWO, achieves the closest fit to actual values (left), higher alignment in forecast vs. actual scatter plots (middle), and better RMSE, DA, and R^2 scores (right), demonstrating superior accuracy and stability.

4.4 Ablation Study

4.4.1 Experiment IV: Comparison with model based on LMD

To demonstrate the efficacy of our improvement operations on Local Mean Decomposition (LMD), including continuation, noise addition, and integration, we conducted the following ablation study (Gong et al., 2024): the improvements to LMD were removed, reverting the decomposition method to the traditional LMD, while keeping other parts of the model unchanged. The experimental results are presented in Table 5 and Figure 10. By analyzing the results, we can conclude:

(a) The improvements of CELMDAN are effective and enhance the prediction accuracy. It is evident that CEAS outperforms the LMD-based model across all evaluation metrics. This advantage is particularly pronounced in terms of R^2 . On the Dubai, Brent, and WTI datasets, CEAS achieves improvements of $|0.5576 - 0.9934| / 0.5576 = 78.15\%$, $|0.6931 - 0.9903| / 0.6931 = 42.88\%$, $|0.7574 - 0.9911| / 0.7574 = 30.86\%$, respectively, compared to the LMD-

based model. These improvements can be attributed to the better robustness of the decomposition process achieved by CELMDAN through signal extension and noise-assisted techniques. This demonstrates that the improvements on LMD are indeed effective, leading to better decomposition quality and, consequently, superior predictive performance.

(b) Solving the endpoint effect plays a significant role in improving prediction accuracy. As shown in Figure 9, due to the inherent limitations of LMD, its decomposition results exhibit instability and shock at the beginning and end of the signal. This issue hinders the model's ability to capture the patterns of individual modes, thereby affecting prediction accuracy. In contrast, CELMDAN effectively alleviates the endpoint effect through taking signal extension, leading to more stable decompositions and improved predictive performance. These results highlight the significant impact of the endpoint effect on model accuracy and show the importance of addressing this issue.

In summary, CELMDAN's improvement on LMD is effective. It significantly alleviates the endpoint effect and mode mixing of LMD, thereby improving the prediction performance.

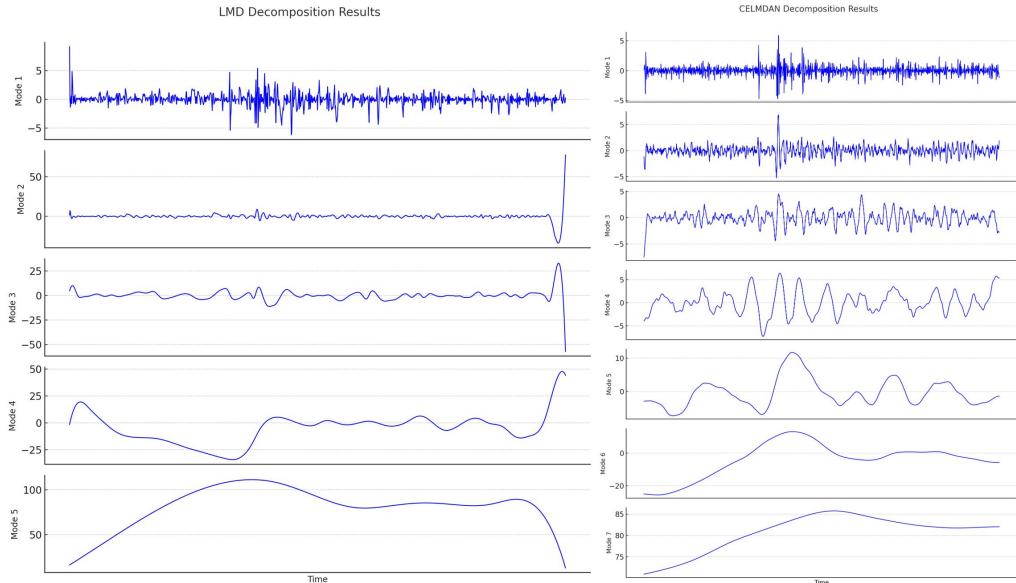


Figure 9: Visual comparison of mode decomposition results using traditional LMD (left) and CELMDAN (right). CELMDAN yields more stable, structured, and frequency-distinct modes with reduced endpoint distortion, highlighting its enhanced capability to separate meaningful oscillatory components from noise.

Table 5
Comparison of prediction performance of models using LMD and CELMDAN decomposition.

	Dubai		Brent		WTI	
	Using LMD	CEAS	Using LMD	CEAS	Using LMD	CEAS
RMSE	7.9307	0.7235	4.7370	0.8184	4.5818	0.8088
MAPE	0.0371	0.0070	0.0308	0.0082	0.0281	0.0087
DA	0.5932	0.6893	0.6841	0.7912	0.6027	0.7703
R ²	0.5576	0.9934	0.6931	0.9903	0.7574	0.9911

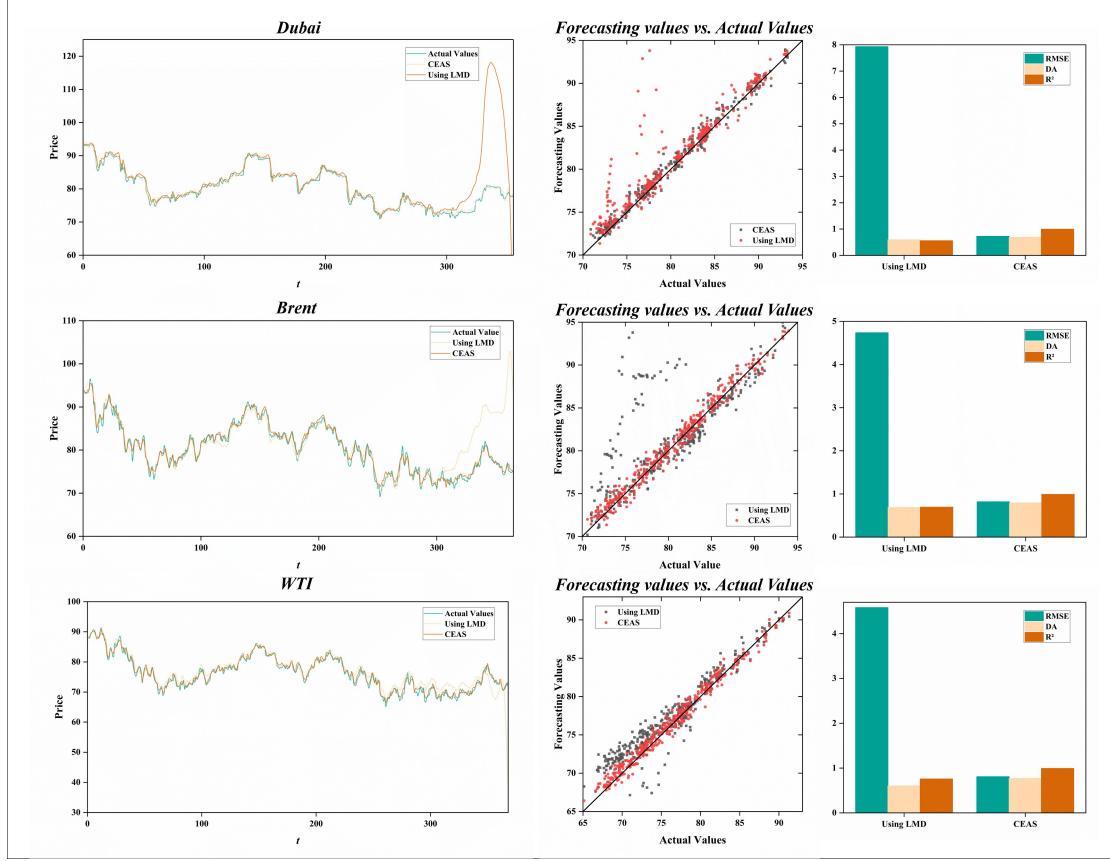


Figure 10: Comparison of prediction results between the CEAS model and the LMD-based model across Dubai, Brent, and WTI datasets.

The CEAS model achieves more accurate forecasts (left), better alignment with actual values in scatter plots (middle), and lower RMSE along with higher DA and R² scores (right), confirming its superiority over traditional LMD-based approaches.

4.4.2 Experiment V: Comparison with model based on GWO

To verify the effectiveness of the improvements in the EGWO compared to the traditional GWO (Wang et al., 2020), we conducted this ablation study. In this experiment, the method of optimizing parameters in ARMA and SVR was replaced with GWO. The experimental results are shown in Table 6 and Figure 12.

By analyzing the experimental results, the following conclusions can be drawn:

(a) The improvement of GWO proves to be effective. It is obvious that CEAS outperforms the GWO-based model across all metrics—RMSE, MAPE, DA, and R²—on the Dubai, Brent, and WTI datasets. Although Section 4.3 indicates that GWO is more proper than other types of algorithms in this task, its performance is not as good as EGWO. For example, in terms of MAPE, CEAS achieves higher accuracy than the GWO-based model by $(0.0103 - 0.0070)/0.0103 = 32.03\%$, $(0.0137 - 0.0082)/0.0137 = 40.15\%$, $(0.0145 - 0.0087)/0.0145 = 40.00\%$ on these datasets, respectively. This demonstrates that the proposed improvements to GWO are valuable, enabling the model to discover better parameter settings and thereby achieve better predictive performance.

(b) EGWO gains a greater chance of escaping local optima and discovering better solutions. Figure 11 illustrates the evolution of the objective function during the optimization

of SVR parameters on the Dubai dataset using EGWO and GWO, respectively. It can be observed that, for GWO, the objective function decreases rapidly during the early iterations, but then plateaus in the mid-to-late stages, indicating premature convergence. In contrast, while EGWO also shows a rapid initial decline, a distinct advantage emerges around 2/3 of the total iterations —corresponding to a change from exploration to exploitation, where the objective function experiences a further drop. This ultimately leads to the discovery of better parameters compared to GWO. These findings suggest that incorporating early exploration allows the algorithm to better avoid local optima and obtain superior parameter values.

Therefore, it can be concluded that, EGWO, by introducing exploration strategies, effectively alleviates the tendency of GWO to become trapped in local optima. This enhancement enables EGWO to identify better model parameters, thereby improving the predictive accuracy of the model.

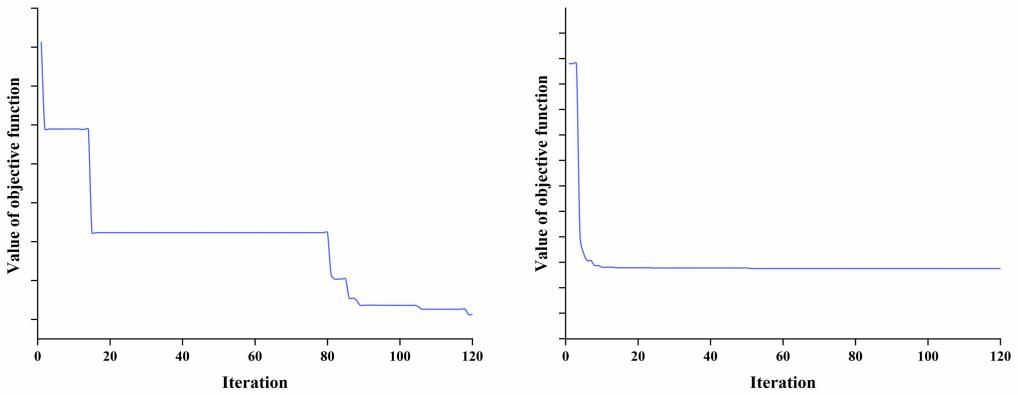


Figure 11: Left: Objective function convergence curve during SVR parameter optimization using GWO (on the Dubai dataset). Right: Objective function convergence curve during SVR parameter optimization using EGWO (on the Dubai dataset).

Table 6
Comparison of prediction performance of models using GWO and EGWO (CEAS).

	Dubai		Brent		WTI	
	Using GWO	CEAS	Using GWO	CEAS	Using GWO	CEAS
RMSE	1.1115	0.7235	1.3832	0.8184	1.3960	0.8088
MAPE	0.0103	0.0070	0.0137	0.0082	0.0145	0.0087
DA	0.6384	0.6893	0.6648	0.7912	0.6568	0.7703
R ²	0.9803	0.9934	0.9708	0.9903	0.9695	0.9911

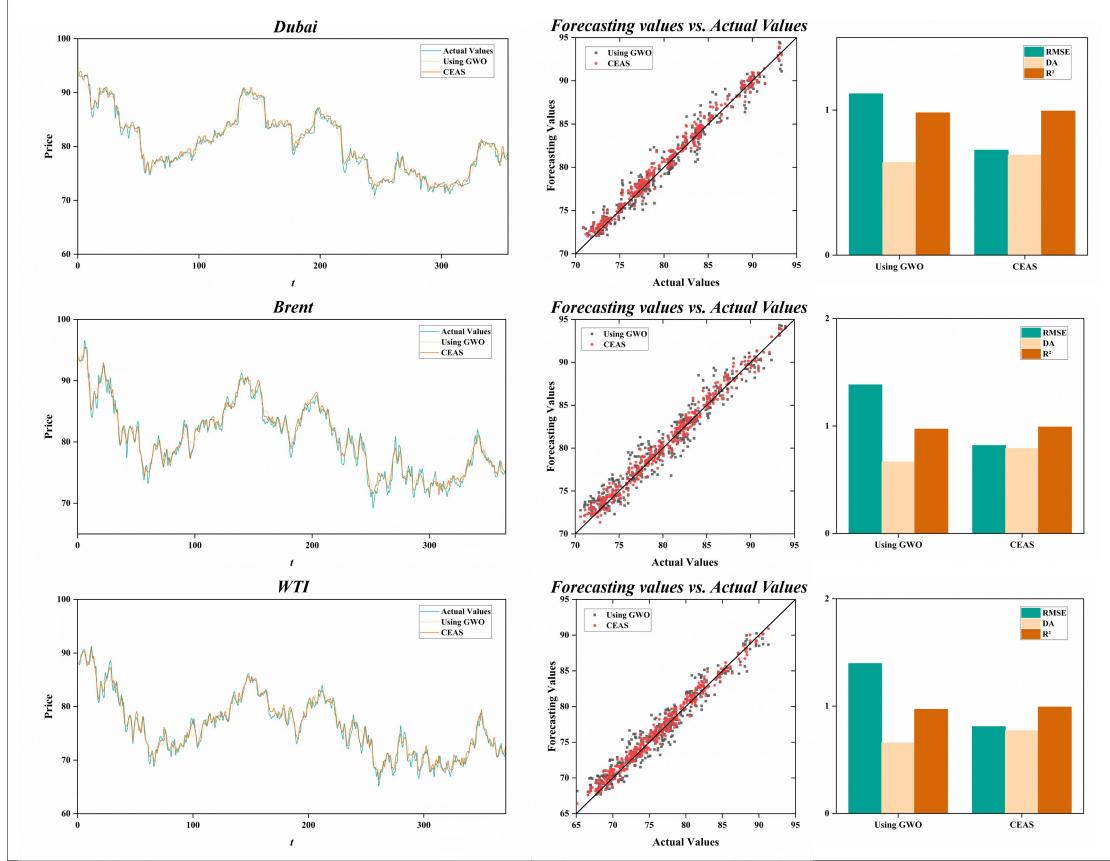


Figure 12. Ablation comparison between the proposed CEAS model and the traditional GWO-based model on Dubai, Brent, and WTI datasets. CEAS achieves tighter fits in time series plots (left), more concentrated scatter distributions (middle), and improved RMSE, DA, and R^2 scores (right), validating the effectiveness of EGWO enhancements over standard GWO.

5 Conclusion and Future Work

In this study, a novel hybrid model CEAS (CELM DAN-EGWO-ARMA-SVR), was proposed to enhance the accuracy and robustness of crude oil price forecasting. By integrating the adaptive and noise-resilient decomposition capability of CELMDAN with the global optimization strength of EGWO, the model effectively captures both high-frequency fluctuations and low-frequency trends inherent in crude oil price series. Specifically, CELMDAN was employed to decompose the original price series into interpretable components, which were subsequently classified by their Hurst exponent values. This classification enabled the adoption of component-specific forecasting strategies: ARMA for short-memory signals and SVR for long-memory signals, both optimized by EGWO. Comprehensive experiments on multiple benchmark datasets (WTI, Brent, Dubai) confirmed that the proposed CEAS model consistently outperforms traditional models and recent decomposition-ensemble frameworks in terms of prediction accuracy, directional correctness, and model stability.

Compared with conventional EMD and LMD-based decomposition schemes, CELMDAN demonstrated superior resistance to mode mixing and endpoint distortion, thereby preserving the structural integrity of IMFs for downstream modeling. Likewise, the proposed EGWO algorithm, incorporating chaotic initialization, stochastic exploration strategies, and dynamically

adaptive leadership weights, showed significant improvements over standard GWO in avoiding local optima and enhancing convergence speed. The ablation studies further validated the individual contributions of CELMDAN and EGWO to the overall model performance, underlining the necessity of each component in the CEAS pipeline.

However, several aspects of this work need future exploration. For example, the parameter of EGWO (e.g., number of wolves, chaos map parameters) is crucial for the model's performance and needs to be chosen in some method. Additionally, CEAS currently relies on historical price series as the sole source of input. Incorporating external features, such as geopolitical events, macroeconomic indicators, or sentiment signals, may further enrich the predictive power.

References

- M. Moghani and R. Loni, "Review on energy governance and demand security in oil-rich countries," *Energy Strategy Reviews*, vol. 57, p. 101625, 2025/01/01/ 2025, doi: <https://doi.org/10.1016/j.esr.2024.101625>.
- Javaid, M. Sajid, E. Uddin, A. Waqas, and Y. Ayaz, "Sustainable urban energy solutions: Forecasting energy production for hybrid solar-wind systems," *Energy Conversion and Management*, vol. 302, p. 118120, 2024/02/15/ 2024, doi: <https://doi.org/10.1016/j.enconman.2024.118120>.
- Kumar and S. Mallick, "Oil price dynamics in times of uncertainty: Revisiting the role of demand and supply shocks," *Energy Economics*, vol. 129, p. 107152, 2024/01/01/ 2024, doi: <https://doi.org/10.1016/j.eneco.2023.107152>.
- Rao, G. D. Sharma, A. K. Tiwari, M. R. Hossain, and D. Dev, "Crude oil Price forecasting: Leveraging machine learning for global economic stability," *Technological Forecasting and Social Change*, vol. 216, p. 124133, 2025/07/01/ 2025, doi: <https://doi.org/10.1016/j.techfore.2025.124133>.
- Frénay and M. Verleysen, "Parameter-insensitive kernel in extreme learning for non-linear support vector regression," *Neurocomputing*, vol. 74, no. 16, pp. 2526-2531, 2011/09/01/ 2011, doi: <https://doi.org/10.1016/j.neucom.2010.11.037>.
- Box, George; Jenkins, Gwilym (1970). Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.
- Xu, R. Fu, and C. K. M. Lau, "Energy market uncertainty and the impact on the crude oil prices," *Journal of Environmental Management*, vol. 298, p. 113403, 2021/11/15/ 2021, doi: <https://doi.org/10.1016/j.jenvman.2021.113403>.
- C.-H. Wu, G.-H. Tzeng, and R.-H. Lin, "A Novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression," *Expert Systems with Applications*, vol. 36, no. 3, Part 1, pp. 4725-4735, 2009/04/01/ 2009, doi: <https://doi.org/10.1016/j.eswa.2008.06.046>.
- G. Fu et al., "Energy, economic and environmental impact analysis of adjusting oil refinery capacity: An interregional optimization approach for resource allocation," *Energy*, vol. 320, p. 135275, 2025/04/01/ 2025, doi: <https://doi.org/10.1016/j.energy.2025.135275>.
- G. Xie, S. Wang, Y. Zhao, and K. K. Lai, "Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study," *Applied Soft Computing*, vol. 13, no. 5, pp. 2232-2241, 2013/05/01/ 2013, doi: <https://doi.org/10.1016/j.asoc.2013.02.002>.
- H. Gong, H. Xing, Y. Yu, and Y. Liang, "A combined model using secondary decomposition for crude oil futures price and volatility forecasting: Analysis based on comparison and ablation experiments," *Expert Systems with Applications*, vol. 252, p. 124196, 2024/10/15/ 2024, doi: <https://doi.org/10.1016/j.eswa.2024.124196>.
- J. B. Florindo, R. R. Lima, F. A. dos Santos, and J. L. Alves, "GHENet: Attention-based Hurst exponents for the forecasting of stock market indexes," *Physica A: Statistical Mechanics*

and its Applications, vol. 667, p. 130540, 2025/06/01/ 2025, doi: <https://doi.org/10.1016/j.physa.2025.130540>.

- J. Zhang and Z. Liu, "Interval prediction of crude oil spot price volatility: An improved hybrid model integrating decomposition strategy, IESN and ARIMA," *Expert Systems with Applications*, vol. 252, p. 124195, 2024/10/15/ 2024, doi: <https://doi.org/10.1016/j.eswa.2024.124195>.
- J. Liang, L. Wang, J. Wu, Z. Liu, and G. Yu, "Elimination of end effects in LMD by Bi-LSTM regression network and applications for rolling element bearings characteristic extraction under different loading conditions," *Digital Signal Processing*, vol. 107, p. 102881, 2020/12/01/ 2020, doi: <https://doi.org/10.1016/j.dsp.2020.102881>.
- J. E, K. He, H. Liu, and Q. Ji, "A novel separation-ensemble analyzing and forecasting method for the gold price forecasting based on RLS-type independent component analysis," *Expert Systems with Applications*, vol. 232, p. 120852, 2023/12/01/ 2023, doi: <https://doi.org/10.1016/j.eswa.2023.120852>.
- J. Wang, X. Niu, Z. Liu, and L. Zhang, "Analysis of the influence of international benchmark oil price on China's real exchange rate forecasting," *Engineering Applications of Artificial Intelligence*, vol. 94, p. 103783, 2020/09/01/ 2020, doi: <https://doi.org/10.1016/j.engappai.2020.103783>.
- K. Drachal, "Forecasting crude oil real prices with averaging time-varying VAR models," *Resources Policy*, vol. 74, p. 102244, 2021/12/01/ 2021, doi: <https://doi.org/10.1016/j.resourpol.2021.102244>.
- K. Yu, X.-J. He, X. Han, X. Luo, and S. Lin, "Forecasting crude oil option prices with dynamic factors using integrated machine learning models," *Communications in Nonlinear Science and Numerical Simulation*, p. 108879, 2025/04/15/ 2025, doi: <https://doi.org/10.1016/j.cnsns.2025.108879>.
- L. Guo, X. Huang, Y. Li, and H. Li, "Forecasting crude oil futures price using machine learning methods: Evidence from China," *Energy Economics*, vol. 127, p. 107089, 2023/11/01/ 2023, doi: <https://doi.org/10.1016/j.eneco.2023.107089>.
- L. Fan, S. Pan, Z. Li, and H. Li, "An ICA-based support vector regression scheme for forecasting crude oil prices," *Technological Forecasting and Social Change*, vol. 112, pp. 245-253, 2016/11/01/ 2016, doi: <https://doi.org/10.1016/j.techfore.2016.04.027>.
- L. Zheng, Y. Sun, and S. Wang, "A novel interval-based hybrid framework for crude oil price forecasting and trading," *Energy Economics*, vol. 130, p. 107266, 2024/02/01/ 2024, doi: <https://doi.org/10.1016/j.eneco.2023.107266>.
- M. A. Alkathery and K. Chaudhuri, "Co-movement between oil price, CO₂ emission,renewable energy and energy equities: Evidence from GCC countries," *Journal of Environmental Management*, vol. 297, p. 113350, 2021/11/01/ 2021, doi: <https://doi.org/10.1016/j.jenvman.2021.113350>.

- M. Arouri, M. Gomes, and G. Pijourlet, "Does climate policy uncertainty shape the response of stock markets to oil price changes? Evidence from GCC stock markets," *Journal of Environmental Management*, vol. 375, p. 124229, 2025/02/01/ 2025, doi: <https://doi.org/10.1016/j.jenvman.2025.124229>.
- M. Benlemlih, I. El Ouadghiri, J. Peillex, and Ç. V. Yavaş, "Crude oil price volatility and environmental performance," *Journal of Environmental Management*, vol. 367, p. 121938, 2024/09/01/ 2024, doi: <https://doi.org/10.1016/j.jenvman.2024.121938>.
- M. Wang et al., "A novel hybrid method of forecasting crude oil prices using complex network science and artificial intelligence algorithms," *Applied Energy*, vol. 220, pp. 480-495, 2018/06/15/ 2018, doi: <https://doi.org/10.1016/j.apenergy.2018.03.148>.
- M. He, Y. Zhang, D. Wen, and Y. Wang, "Forecasting crude oil prices: A scaled PCA approach," *Energy Economics*, vol. 97, p. 105189, 2021/05/01/ 2021, doi: <https://doi.org/10.1016/j.eneco.2021.105189>.
- M. Pan et al., "Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization," *Journal of Cleaner Production*, vol. 277, p. 123948, 2020/12/20/ 2020, doi: <https://doi.org/10.1016/j.jclepro.2020.123948>.
- N. Jha, H. Kumar Tanneru, S. Palla, and I. Hussain Mafat, "Multivariate analysis and forecasting of the crude oil prices: Part I – Classical machine learning approaches," *Energy*, vol. 296, p. 131185, 2024/06/01/ 2024, doi: <https://doi.org/10.1016/j.energy.2024.131185>.
- Tarkhaneh and H. Shen, "Training of feedforward neural networks for data classification using hybrid particle swarm optimization, Mantegna Lévy flight and neighborhood search," *Heliyon*, vol. 5, no. 4, p. e01275, 2019/04/01/ 2019, doi: <https://doi.org/10.1016/j.heliyon.2019.e01275>.
- P. Esmaeili, M. Rafei, M. Salari, and D. Balsalobre-Lorente, "From oil surges to renewable shifts: Unveiling the dynamic impact of supply and demand shocks in global crude oil market on U.S. clean energy trends," *Energy Policy*, vol. 192, p. 114252, 2024/09/01/ 2024, doi: <https://doi.org/10.1016/j.enpol.2024.114252>.
- Q. Liang, Q. Lin, M. Guo, Q. Lu, and D. Zhang, "Forecasting crude oil prices: A Gated Recurrent Unit-based nonlinear Granger Causality model," *International Review of Financial Analysis*, vol. 102, p. 104124, 2025/06/01/ 2025, doi: <https://doi.org/10.1016/j.irfa.2025.104124>.
- R. Li, Y. Hu, J. Heng, and X. Chen, "A novel multiscale forecasting model for crude oil price time series," *Technological Forecasting and Social Change*, vol. 173, p. 121181, 2021/12/01/ 2021, doi: <https://doi.org/10.1016/j.techfore.2021.121181>.
- S. Ben Jabeur, R. Khalfaoui, and W. Ben Arfi, "The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning," *Journal of Environmental Management*, vol. 298, p. 113511, 2021/11/15/ 2021, doi: <https://doi.org/10.1016/j.jenvman.2021.113511>.
- S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

- S. Mati, M. Radulescu, N. Saqib, A. Samour, G. Y. Ismael, and N. Aliyu, "Incorporating Russo-Ukrainian war in Brent crude oil price forecasting: A comparative analysis of ARIMA, TARMA and ENNReg models," *Heliyon*, vol. 9, no. 11, p. e21439, 2023/11/01/ 2023, doi: <https://doi.org/10.1016/j.heliyon.2023.e21439>.
- Smith Jonathan S. 2005 The local mean decomposition and its application to EEG perception data. *J. R. Soc. Interface.* 2: 443–454. <https://doi.org/10.1098/rsif.2005.0058>
- T. Fang, C. Zheng, and D. Wang, "Forecasting the crude oil prices with an EMD-ISBM-FNN model," *Energy*, vol. 263, p. 125407, 2023/01/15/ 2023, doi: <https://doi.org/10.1016/j.energy.2022.125407>.
- T. M. Luan, S. Khatir, M. T. Tran, B. De Baets, and T. Cuong-Le, "Exponential-trigonometric optimization algorithm for solving complicated engineering problems," *Computer Methods in Applied Mechanics and Engineering*, vol. 432, p. 117411, 2024/12/01/ 2024, doi: <https://doi.org/10.1016/j.cma.2024.117411>.
- W. Zhang, S. Yan, H. Tang, X. Li, and D. Lei, "An improved shuffled frog leaping algorithm based on support vector machine for hybrid flow shop rescheduling with disturbance prediction," *Swarm and Evolutionary Computation*, vol. 93, p. 101843, 2025/03/01/ 2025, doi: <https://doi.org/10.1016/j.swevo.2025.101843>.
- X. Ma, T. Yu, and Q. Jiang, "Does geopolitical risk matter in carbon and crude oil markets from a multi-timescale perspective?," *Journal of Environmental Management*, vol. 346, p. 119021, 2023/11/15/ 2023, doi: <https://doi.org/10.1016/j.jenvman.2023.119021>.
- X. Wang, X. Li, and S. Li, "Point and interval forecasting system for crude oil price based on complete ensemble extreme-point symmetric mode decomposition with adaptive noise and intelligent optimization algorithm," *Applied Energy*, vol. 328, p. 120194, 2022/12/15/ 2022, doi: <https://doi.org/10.1016/j.apenergy.2022.120194>.
- Y. Chen and Z. Tian, "A crude oil price forecasting framework based on Constraint Guarantee and Pareto Fronts Shrinking Strategy," *Applied Soft Computing*, vol. 174, p. 112996, 2025/04/01/ 2025, doi: <https://doi.org/10.1016/j.asoc.2025.112996>.
- Y.-X. Wu, Q.-B. Wu, and J.-Q. Zhu, "Improved EEMD-based crude oil price forecasting using LSTM networks," *Physica A: Statistical Mechanics and its Applications*, vol. 516, pp. 114-124, 2019/02/15/ 2019, doi: <https://doi.org/10.1016/j.physa.2018.09.120>.
- Z. Ouyang, M. Lu, Z. Ouyang, X. Zhou, and R. Wang, "A novel integrated method for improving the forecasting accuracy of crude oil: ESMD-CFastICA-BiLSTM-Attention," *Energy Economics*, vol. 138, p. 107851, 2024/10/01/ 2024, doi: <https://doi.org/10.1016/j.eneco.2024.107851>.