Pol. J. Environ. Stud. Vol. 34, No. X (2025), 1-15
DOI: *10.15244/pjoes/187592* ONLINE PUBLICATION DATE:

Original Research

A Study on the Impact Mechanism of WTI Futures Price Forecasting Considering Mutation Factors

Shenhui Song, Haisheng Yu*

School of Mathematics and Statistics, Ludong University,Yantai 264001, China

Received: 9 March 2024
Accepted: 18 April 2024

Abstract

Crude oil futures price forecasting plays a crucial role in assessing energy demand, promoting the development of renewable energy, formulating environmental protection policies, and achieving a balance between economic development and environmental protection. Limited studies have focused on exploring this phenomenon from the perspective of establishing or explaining the influence mechanism. This paper conducts three sets of experiments based on existing theoretical studies. Specifically, the control group solely employs price for prediction, the conventional group integrates conventional factors onto this basis, and the mutation group further incorporates the influence mechanism of mutation factors based on the conventional group. Comparative analysis of the experimental results between the control and conventional groups reveals the underlying principles of how conventional factors influence price trends, and the experiments between the conventional and mutation groups simulate price directions during unexpected situations. The results demonstrate that prediction accuracy follows the order of the mutation group, conventional group, and control group, thus validating the hypothesis proposed in this paper. These research findings hold great significance for futures price prediction and provide valuable insights for related theories.

Keywords: impact mechanism, mutation factors, LSTM, WTI Futures Price Forecasting

Introduction

With the changes in the global economy, crude oil futures prices have become an important factor affecting the global energy market and macroeconomics, and accurate prediction of crude oil futures prices is of great significance to investors, governments, and enterprises. However, the prediction of crude oil futures prices has always been a challenge, and most of the existing research focuses on the traditional machine learning field or combinatorial modeling, and in recent years, some scholars have taken the influencing factors into consideration. Mutations in the influencing factors may lead to drastic fluctuations in crude oil futures prices thus affecting the accuracy of price prediction. Therefore, the purpose of this paper is to analyze the impact of mutating factors on crude oil futures prices to establish an impact mechanism to correct the magnitude of the trend of predicting the movement of crude oil futures prices. This paper will provide investors with more accurate crude oil futures price forecasts, enrich the understanding of the price formation mechanism of the crude oil market, and provide a useful reference for future related research.

*e-mail: haisenltn@163.com

There have been many recent studies on crude oil price forecasting, and most scholars have used composite models. The use of hybrid models to forecast crude oil prices can improve the forecasting accuracy. Hybrid models combine multiple forecasting methods, which can make full use of the advantages of different models and reduce the errors that may exist in a single model. By integrating the prediction results of multiple models, a more accurate and stable prediction value can be obtained. Some scholars have used hybrid models for feature selection on the processing of historical time series data of crude oil prices and factors affecting oil prices [1]. Some scholars use hybrid models to forecast the long-term trend and short-term fluctuations of prices separately before integrating the results [2]. Some scholars have used combinatorial models to capture specific features such as nonlinearities, lag effects, and market interrelationships in oil price time series [3-6]. Wu et al. [7] introduced a self-optimizing integrated learning model, which obtains accurate forecasts by decomposing the subsequences, optimizing the parameters, and aggregating the predicted values. Lu [8] discussed the release time, model structure, prediction accuracy, prediction range, and input variables for energy price forecasting. Chai et al. [9] proposed a methodology to capture various volatility features in crude oil data series, including change points, state switching, time-varying determinants, trend decomposition of high-frequency series, and potential nonlinearities in the model setup. Guliyev et al. [10] utilized the U.S. macroeconomic and financial factors, as well as the global crises and collapses, to forecast the price of West Texas Intermediate (WTI) crude oil price dynamics. Tang et al. [11] used the procedure of multi-factor data processing, multi-scale analysis, and typical forecasting techniques application for oil price forecasting. In summary, hybrid models are adaptable, flexible, and risk-reducing, which can synthesize complex factors, make more accurate and timely forecasts in the face of different situations, and formulate more reasonable investment strategies. In addition, these models mainly use subsequence decomposition for forecasting, which solves data characteristics such as nonlinearity, high noise, and lag effect, but they do not address price fluctuations under sudden shocks, and they do not propose measures to solve such sudden events in practical applications.

In order to improve the prediction accuracy, some scholars added various influencing factors in the analysis, which helps to improve the prediction accuracy, enhance the reliability of the prediction, better respond to the market changes, provide decision support, etc. Zhou et al. [12] proposed a carbon price prediction model based on the Sparrow search algorithm, and considered the structural and non-structural influencing factors in the model. Lu et al. [13] utilized the amplitude modulation method to capture the effect of different past temporal feature states on the closing price of a stock to improve the prediction accuracy. Nwulu [14] used key economic indicators as inputs and resultant data as outputs. Regarding the factors affecting the movement of crude oil prices, some scholars identified supply, demand, and financial markets as the main influencing factors [15]. Some scholars used self-focused mechanisms and spatio-temporal graph neural networks to demonstrate that the US dollar index, LIBOR, and VIX have surpassed supply and demand as the most influential predictors of WTI futures prices [16]. Some scholars have also provided new evidence that the US dollar index has significant out-of-sample predictive power for oil prices [17]. Some scholars have comprehensively explored the various factors affecting the volatility of crude oil prices in terms of commodity properties, macroeconomic factors, geopolitical events, and alternative energy sources [18]. Others have thoroughly investigated the impact of emergencies such as the COVID-19 global pandemic on the low-carbon transition in the post-pandemic era [19]. Liu and team [20] investigated whether oil investor attention (OA), as measured by Google search volume, contains incremental information that predicts crude oil futures volatility. Miao et al. [21] considered six categories of factors (supply, demand, financial markets, commodity markets, speculation, and geopolitics) and tested their significance using various forecasting models. Ma and colleagues [22] used a generalized dynamic factor model to construct three predictors of crude oil price volatility: a fundamental (physical) predictor, a financial predictor, and a macroeconomic uncertainty predictor. Wang et al. [23] investigated the relationship between oil price volatility, inflation rate and economic growth. Incorporating the influence factors into forecasting was used to improve the forecasting accuracy, which was not only applied to forecasting crude oil prices, but also applied to forecasting carbon prices and other aspects. Zhang et al. [24] dynamically predicted the carbon peaking paths by using LSTM neural networks, and analyzed the appropriate paths based on the carbon emission intensity, cumulative emissions, and peaking time of each province. Hao et al. [25] proposed a hybrid prediction framework for carbon pricing that considers multiple influencing factors, incorporating the advantages of new algorithms and successfully solving the problem of carbon price prediction based on multiple factors. Wu et al. [26] demonstrated that the factors of population, urbanization rate, per capita GDP, the proportion of coal consumption, automobile ownership, the intensity of energy consumption, and the proportion of tertiary industry have a facilitating or inhibiting effect on carbon emission. Zhang et al. [27] explored the impacts of economic development on carbon emission in Sichuan Province between 1990 and 2020 and analyzed the impacts of economic development on carbon emission. The impact of economic development on carbon emissions in Sichuan Province and put forward suggestions to accelerate the development of a low-carbon economy. Zhang et al. [28] analyzed the spatial distribution pattern of green total factor productivity in Chinese agriculture and its dynamic evolution pattern in different regions and explored the factors affecting its spatial differentiation. However, most of the studies only extracted the characteristics of the influencing factors for price prediction, and did not describe the specific influencing process and the degree of influence, nor did they consider the adjustments when emerging factors play a major role in price changes. The failure to differentiate the influencing factors according to different situations is not conducive to predicting and judging sudden price changes.

There are also scholars on the social significance and environmental impacts of energy forecasting. D'Orazio [29] further deepens the understanding of how high energy prices affect one of the most vulnerable groups in society. Azubike [30] presents and discusses some of the factors affecting energy security and energy transition in Libya.

The main contributions of this paper can be summarized in two parts. Firstly, based on existing predictive models and the characteristics of the data in this paper, a higher-order multivariate input and single output framework is constructed. It not only verifies that incorporating regular influencing factors leads to more accurate crude oil futures price predictions but also explains the specific mechanisms and degrees of influence of each factor on the price changes. Secondly, considering the corresponding sudden-change influencing factors, a mechanism for these factors is designed. This mechanism can be used not only to explain historical price changes but also to predict the trend of crude oil price changes when a sudden-change factor arises in the future.

The rest of this paper is organized as follows: Theoretical Study describes the research framework and the Long Short-Term Memory (LSTM) network model; The modeling section includes an introduction and explanation of regular and sudden-change influencing factors, the setting of the sudden-change impact mechanism, and symbolic formulas; Experimental Results introduce data sources, price predictions for regular group experiments, and price predictions for sudden-change group experiments; Conclusion summarizes the research.

Material and Methods

Research Framework

The research framework of this paper, as shown in Fig. 1, begins with variable selection. We chose 14 factors that influence crude oil price changes, including 10 regular factors and 4 sudden-change factors. The 10 regular factors are crude oil production, imports, exports, global economy, exchange rates, interest rates, oil inventories, Brent crude futures prices, natural gas futures prices, and crude oil trading volume; the 4 sudden-change factors are political events, economic crises, technological revolutions, and global pandemics. Each influencing factor is introduced, and their impact results and degrees on prices are estimated.

The second stage involves establishing predictive models and impact mechanisms. In the control group, input variables are previous prices of WTI futures, and the predictive model is the LSTM model. In the regular group, previous data of regular influencing factors are added to the input variables, and the predictive model is the LSTM model. In the sudden-change group, the results of the regular group are adjusted using the impact mechanism. For sudden-change influencing factors, data from periods when these factors predominantly influenced prices are extracted, features are identified, and key parameters are used to construct the framework.

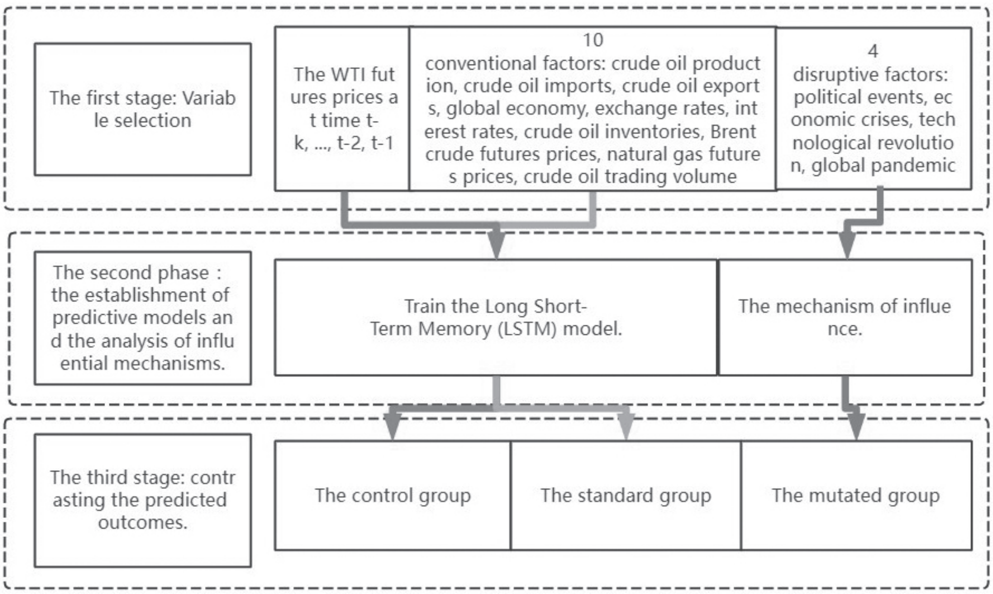

Fig. 1. Research Framework.

The third stage involves comparing the prediction results and errors of the control group, regular group, and sudden-change group to verify the hypotheses.

Long Short-Term Memory Network Model

This article adopts the Long Short-Term Memory (LSTM) network model for analysis and prediction. The LSTM network has attracted attention for its excellent performance in handling time series data, making it widely applicable in fields such as crude oil futures price prediction. The LSTM network consists of a unit containing input, forget, and output gates. These gates control the entry, forgetting, and output of information, thereby enabling the learning and memorization of long-term dependencies. The input gate decides which information is allowed through, the forget gate determines which information needs to be forgotten, and finally, the output gate decides what information will be passed to the next layer of the network. This design allows the LSTM network to better handle long-term dependencies in time series data, making it suitable for complex temporal data such as crude oil futures prices.

Below are the mathematical formulas for the internal computational process of an LSTM unit:

$$i_t = \sigma\left(W_{xi}x_t + W_{ui}h_{t-1} + b_i\right) \tag{1}$$

$$\tilde{C}_t = \tanh\left(W_{xc}x_t + W_{uc}h_{t-1} + b_c\right) \tag{2}$$

$$o_t = \sigma\left(W_{xo}x_t + W_{uo}h_{t-1} + b_o\right) \tag{3}$$

$$f_t = \sigma\left(W_{xf}x_t + W_{uf}h_{t-1} + b_f\right) \tag{4}$$

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{5}$$

$$h_t = o_t \cdot \tanh\left(C_t\right) \tag{6}$$

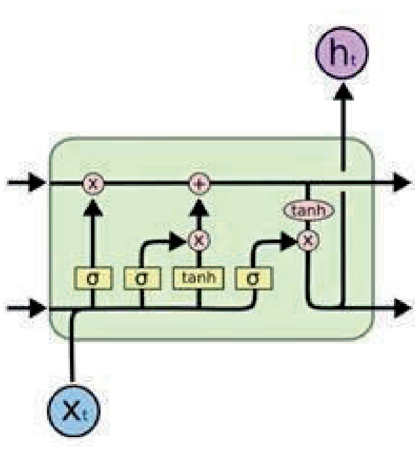

Fig. 2. LSTM Cell.

where f_t represents the forget gate, i_t represents the input gate, o_t represents the output gate, $\tilde{C}_t$ represents the value of the candidate memory cell, C_t represents the updated value of the memory cell, x_t represents the input value, h_t represents the output value, σ,tanh represents the activation function, W represents the weight matrix, and b represents the error.

In addition to the formulas above, the following Fig. 2 is an illustration of an LSTM cell:

Model Building

Explanation of Influencing Factors

Regular Factors

Factors that have a long-term and consistent influence on the price fluctuations of WTI crude oil futures are referred to as conventional factors. Through selection, the following ten factors are identified as conventional: crude oil production, crude oil imports, crude oil exports, global economy, exchange rates, interest rates, crude oil inventories, Brent crude oil futures prices, natural gas futures prices, and crude oil trading volume. These conventional factors are characterized by their longevity and stability. Fig. 3 presents a comparison chart of these ten conventional influencing factors against the price movements of WTI, where the red line represents the trend in changes of the ten conventional influencing factors, and the blue line represents the trend in WTI price changes. The following is an analysis of the specific principles and data selection process for each conventional factor's impact on prices, in conjunction with Fig. 3.

Crude Oil Production. This paper selects data on the production of West Texas Intermediate (WTI) crude oil, measured in thousands of barrels per month. Generally, an increase in crude oil production, which boosts supply, tends to lead to a decrease in prices; conversely, a decrease in production, which reduces supply, tends to lead to an increase in prices. As shown in Fig. 3a), when production trends upward, prices correspondingly trend downward, and vice versa. However, the degree of price change is not always the same due to the influence of other factors.

Crude Oil Imports. This paper selects data on the total volume of U.S. crude oil imports, measured in thousands of barrels per day. Generally, an increase in crude oil imports by a country or region often leads to an increase in global supply, which might be perceived as oversupply, exerting downward pressure on crude oil prices. Conversely, a decrease in imports might lead to a reduction in global supply, perceived as a shortage, thereby supporting crude oil prices and causing them to rise. However, in the early stages, domestic production was far less than consumption, so crude oil relied heavily on imports. As illustrated in Fig. 3b), crude oil imports remained high in the early stages, only decreasing in recent years.

Crude Oil Exports. This paper selects data on the total volume of U.S. crude oil exports, measured in thousands of barrels per day. Generally, an increase in the exports of a country or region tends to increase global supply, which might be seen as oversupply, putting downward pressure on crude oil prices. Conversely, a decrease in exports might reduce global supply, perceived as a shortage, thereby supporting crude oil prices and causing them to rise. However, actual export volumes are limited by domestic production and consumption levels. In the early stages, domestic consumption far exceeded production, so there were almost no exports, but exports have increased in recent years due to technological advancements. As shown in Fig. 3c), crude oil export volumes surged after 2014.

Global Economy. This paper uses annual global GDP values as an indicator to quantify the global economy. Generally, when the global economy prospers, accelerated production development leads to increased demand for crude oil, driving up its price. In contrast, during economic recessions, a general decline in consumption slows down production, leading to a decrease in crude oil demand and, thus, a decrease in prices. Overall, global economic development tends to cause more severe fluctuations in crude oil prices, with greater economic growth corresponding to larger price fluctuations. As shown in Fig. 3d), the price changes become more intense as the economy grows year after year.

Exchange Rates. This paper uses the U.S. Dollar Index as an indicator of exchange rate conditions. Exchange rate changes also have a certain impact on crude oil prices. Firstly, when the domestic currency depreciates, the cost of importing crude oil increases, which might lead to a rise in crude oil prices as importers have to pay more in their domestic currency to buy crude oil. Conversely, when the domestic currency appreciates, the cost of importing crude oil decreases, potentially leading to a fall in crude oil prices as less domestic currency is needed for purchase. Additionally, changes in exchange rates can affect international trade and cross-border investments, potentially influencing global economic activity and crude oil demand. Overall, exchange rate changes can impact crude oil prices, but their effect is usually assessed in conjunction with other factors. As shown in Fig. 3e), over the past 20 years, the trend in exchange rate changes has generally been in sync with price changes.

Interest Rates. This paper uses the data on the effective federal funds rate. The relationship between interest rates and prices might not be direct but is influenced by a variety of factors indirectly affecting each other. When interest rates rise, corporate production costs increase, reducing investment in the crude oil industry; when interest rates fall, it stimulates economic development and promotes production and consumption. As shown in Fig. 3f), the direct relationship between the two is not clearly evident.

Crude Oil Inventories. This study utilizes weekly U.S. crude oil inventory data. Crude oil inventories refer to the quantity of unused oil reserves, and their increase or decrease can impact market supply and demand, thereby influencing crude oil prices. An increase in crude oil inventories usually indicates an oversupply, as there are more reserves available in the market. Oversupply leads to a supply-demand imbalance, driving down crude oil prices. Conversely, a decrease in inventories typically signifies a supply shortage due to reduced reserves in the market, which can drive up prices. As shown in Fig. 3g), due to government market regulation, crude oil inventories generally maintain a certain level over an extended period. There is an inverse relationship between local inventory levels and price changes, with the magnitude of changes being fairly similar.

Brent Crude Oil Futures Prices. Generally, if Brent crude oil futures prices rise, WTI crude oil futures prices are also likely to increase, and vice versa. This is because supply and demand factors in the international oil market typically affect global crude oil prices, and both Brent and WTI crude oils are important benchmarks in the international oil market. As depicted in Fig. 3h), the price movement curves of both oils trend similarly, almost converging in the later stages.

Natural Gas Futures Prices. This paper examines the prices of U.S. natural gas futures contracts, measured in dollars per million British thermal units. The price trends of natural gas and crude oil may be correlated to some extent. For instance, if natural gas prices rise, it may lead some energy consumers to switch to crude oil as an alternative energy source, thereby increasing demand for crude oil. On the other hand, the production and supply of natural gas and crude oil may also affect each other. As shown in Fig. 3i), the prices of crude oil and natural gas display a trend of long-term similarity and short-term complementarity.

Crude Oil Trading Volume. This analysis uses data on WTI crude oil futures trading volume. An increase in the trading volume of the crude oil futures market usually signifies an increased demand for crude oil or speculative activities, causing market fluctuations. Therefore, an increase in trading volume can lead to increased price volatility. Conversely, a decrease in trading volume might indicate lower market demand for crude oil or a lack of confidence among market participants about future price trends, potentially leading to a fall in crude oil futures prices. As illustrated in Fig. 3j), when trading volumes are low, price changes tend to be more moderate; with higher volumes, price fluctuations become more intense.

WTI Crude Oil Futures Prices. Historical price trends can provide some reference for investors. Through technical and trend analysis, investors can observe historical patterns of price fluctuations, as well as support and resistance levels. This information can help investors better understand market trends and assist them to some extent in predicting future price movements, as shown in Fig. 3k).

Mutating Factors

Through selection, the following four factors are identified as mutating: political events, economic crises, technological revolutions, and global pandemics. Mutating factors are characterized by their suddenness, unpredictability, wide-ranging impact, urgency, and relative rarity. Considering factors such as the time frame of the events and data availability, specific events corresponding to mutating factors are selected, as shown in Table 1. Fig. 3 presents a chart of WTI price fluctuations during periods of these events, where prices during political events are represented by a red solid line, economic crises by a blue solid line, technological revolutions by a purple solid line, and global pandemics by a black solid line. The following is an analysis of the specific principles and data selection process for each mutating factor's impact on prices, in conjunction with Fig. 3.

Political Events: Political events include geopolitics, international policies, and local conflicts. Prices show short-term fluctuations before the occurrence of political events, severe fluctuations during the events, and long-term fluctuations after. Considering factors like the duration of the event and data availability, this paper selects data from periods like the Gulf War, the Arab Spring, and the Russo-Ukrainian conflict to construct a mechanism for the impact of political events. As shown in Fig. 4a), b), and c), prices first surge rapidly within a short period and then plummet to levels near those at the onset of the political event.

Economic Crisis: Economic crises often lead to a slowdown in global economic activities, triggering investors' concerns about future economic prospects

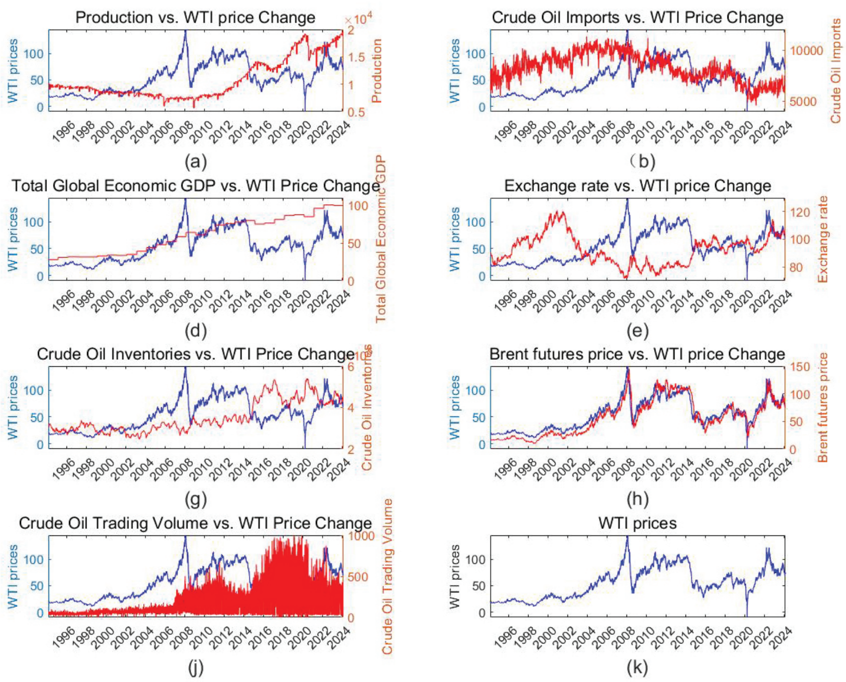

Fig. 3. Comparison of Conventional Influencing Factors with WTI Price Fluctuations.

Table 1. Specific Events and Timelines of Mutating Factors.

Category	Specific Event	Time Frame
Political Events	Gulf War	Aug 1990 - Feb 1991
	Arab Spring	Feb 2010 - Nov 2011
	Russia-Ukraine Conflict	Jan 2022 - Dec 2022
Economic Crisis	Financial Crisis	Jan 2007 - Feb 2009
Technological Revolution	Electric Vehicle Popularization	Jan 1997 - Dec 1999
	Shale Oil Revolution	Jun 2014 - Jan 2015
Global Pandemic	COVID-19 Pandemic	Dec 2019 - Dec 2022

and reducing investments in crude oil, thereby causing a decline in oil prices. It can also lead to political instability in some countries or regions, which may affect oil supply and cause price fluctuations. This paper selects data from the 2008 financial crisis to construct the impact mechanism of economic crises. As shown in Fig. 4d), before the financial crisis, the economic situation was favorable and oil prices rose. However, following the outbreak of the crisis, there was a steep decline in oil prices.

Technological Revolution: The technological revolution includes advancements in oil exploration, extraction, and transportation technologies. This paper selects data from periods of electric vehicle popularity and the shale oil revolution to construct the impact mechanism of technological revolutions. As shown in Figs 4e) and f), electric vehicles are substitutes for fuel-powered cars. The rise in popularity of electric vehicles correspondingly squeezed the market for fuel vehicles, leading to a fall in oil prices. During the shale oil revolution, a breakthrough in extraction technology led to a significant increase in production and supply, causing a decrease in oil prices.

Pandemic Outbreak: The outbreak of a global pandemic leads to a decline in productivity and transportation bottlenecks. As agriculture, industry, and tertiary industries rely on oil resources, the demand for crude oil decreases. This paper selects data from the COVID-19 pandemic period to construct the global pandemic impact mechanism. As shown in Fig. 4g), during the COVID-19 pandemic, a sharp decline in productivity led to a crash in oil prices, which gradually recovered as the pandemic situation improved.

Design of Impact Mechanism for Mutation Factors

Due to the characteristics of mutation factors, such as their suddenness, unpredictability, wide impact range, urgency, and relative rarity, accurately predicting future prices during periods influenced by these factors is extremely challenging. Therefore, leveraging key parameters extracted from data during these unexpected events, combined with specific circumstances for auxiliary prediction, is advisable. The relevant parameters are listed in Table 2, where multiple examples of political events and technological revolutions are selected, and the key parameters are averaged across these instances. Fig. 5 illustrates these key parameters for each event in a graphical format.

The mechanism of political event impact is designated as Z with its corresponding influence coefficient denoted as α. This includes a key parameter Z comprising the maximum value, minimum value, volatility, time span, increase magnitude, and decrease magnitude. The value of α is set within $\alpha \in [0, +\infty]$. When $\alpha = 0$ it signifies an absence of mutational factors due to political events; whereas $\alpha \in (0,1)$ indicates varying degrees of interference from political event mutational factors, with increasing values representing progressively greater impact. At $\alpha = 1$, the interference from political event mutational factors reaches a historical peak; exceeding $\alpha>1$ suggests an interference level surpassing historical peaks. Similar mechanisms apply to other mutational factors: the economic crisis impact mechanism is represented as J with its influence coefficient β; the technological revolution impact mechanism as S with its coefficient γ; and the global

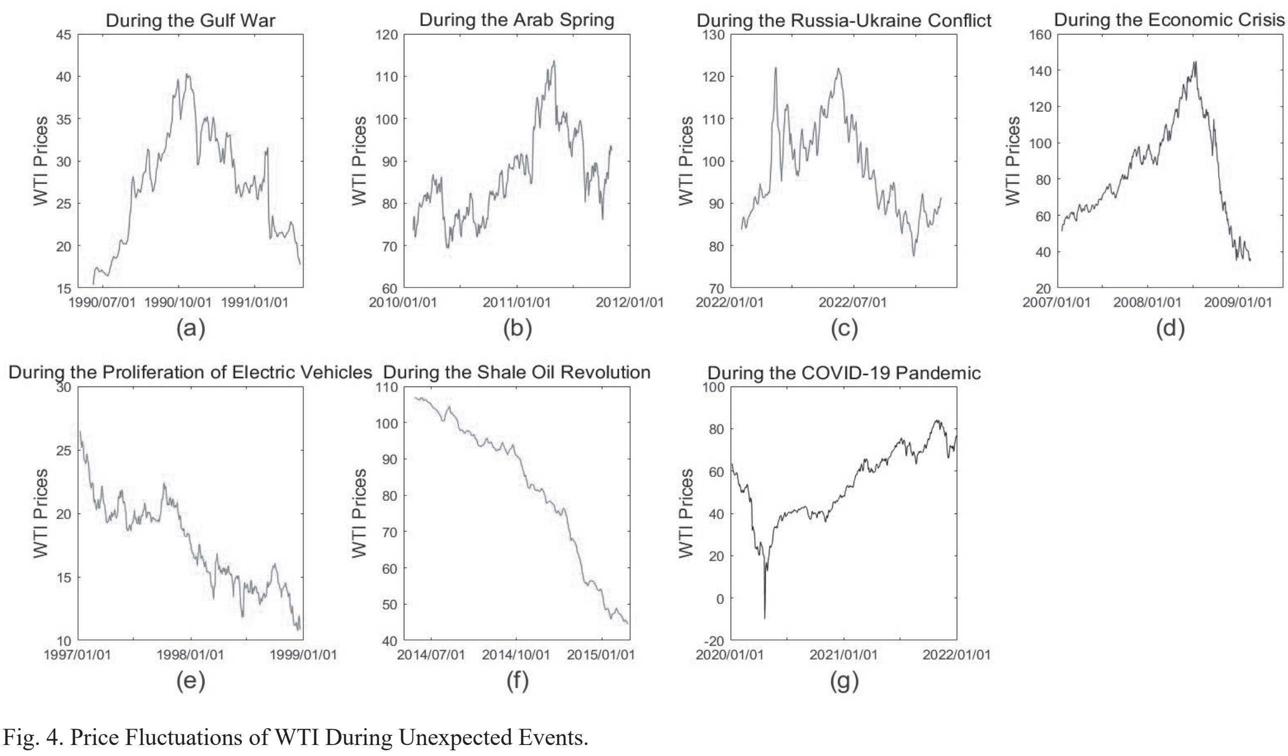

Fig. 4. Price Fluctuations of WTI During Unexpected Events.

Table 2. Relevant Parameters of Data During the Periods of Mutation Factor Events.

Category	Event	Time Span (months)	Maximum Value	Minimum Value	Volatility	Increase	Decrease
Political Events	Gulf War	8.19	40.34	15.35	23.84%	81.71%	-96.64%
	Arab Spring	21.14	113.71	69.33	11.79%	101.97%	-98.48%
	Russia-Ukraine Conflict	10.24	122.19	77.49	11.21%	167.47%	-190.45%
	Political Average	13.19	92.08	54.05	15.61%	117.05%	-128.52%
Economic Crisis	Financial Crisis	25.00	144.94	34.82	32.43%	114.68%	-135.25%
Technological Revolution	Electric Vehicle Proliferation	23.38	26.49	10.77	19.36%	21.06%	-23.73%
	Shale Oil Revolution	7.90	106.93	44.48	23.91%	55.14%	-78.23%
	Technology Average	15.64	66.71	27.62	21.63%	38.10%	-50.98%
Global Pandemic	COVID-19 Pandemic	24.95	84.19	-9.95	31.95%	79.98%	-95.71%

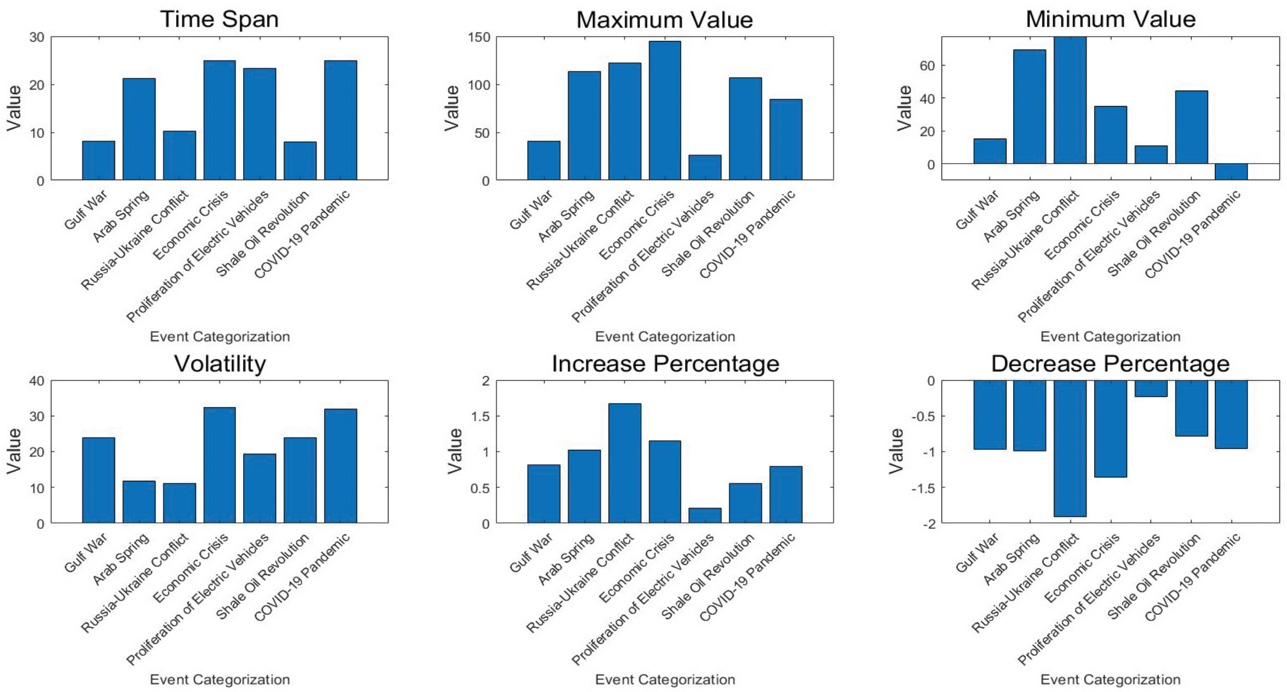

Fig. 5. Bar Graph of Data-Related Parameters for the Occurrence Period of Mutational Factor Events.

pandemic impact mechanism as Y with its coefficient η. The range for these influence coefficients is consistent with the aforementioned.

Three-Group Experimental Symbol Formulas

Symbols

Table 3 contains the symbols representing the variables used in the three experimental groups of this study.

Formula

In the control group experiment, the WTI crude oil futures price at the t-th time point is predicted using the prices of the previous t–1-th, t–2-th, ... t–k-th time points. y_1 represents the predicted WTI crude oil futures price for the t-th time point in the first experimental group, as shown in Equation (7).

In the conventional group experiment, based on the control group experiment, the values of conventional influencing factors from the previous moment

Table 3.Variable Symbols.

Symbol	Meaning	Symbol	Meaning	Symbol	Meaning
n	Total number of moments selected	i_t	Natural gas futures price at time t	$\varepsilon_{9,k}$	Order coefficient of natural gas futures price at time t
t	Any given moment	j_t	Trading volume at time t	$\varepsilon_{10,k}$	Order coefficient of trading volume at time t
k	Order	y_t	WTI futures price at moments prior to moment t	$\varepsilon_{11,k}$	Order coefficients of WTI futures prices for the moments before moment t
a_t	Crude oil production at time t	$\varepsilon_{1,k}$	Order coefficient of crude oil production at time t	Z	Political events
b_t	Crude oil imports at time t	$\varepsilon_{2,k}$	Order coefficient of crude oil imports at time t	J	Economic crisis
c_t	Crude oil exports at time t	$\varepsilon_{3,k}$	Order coefficient of crude oil exports at time t	S	Technological revolution
d_t	Global economic development at time t	$\varepsilon_{4,k}$	Order coefficient of global economic development at time t	Y	Influence of global epidemic
e_t	Exchange rate at time t	$\varepsilon_{5,k}$	Order coefficient of exchange rate at time t	α	Coefficient of political events
f_t	Interest rate at time t	$\varepsilon_{6,k}$	Order coefficient of interest rate at time t	β	Economic Crisis Correspondence Coefficient
g_t	Crude Oil Inventory at time t	$\varepsilon_{7,k}$	Order coefficient of crude oil inventory at time t	η	Correspondence coefficient of technological revolution
h_t	Brent oil futures price at time t	$\varepsilon_{8,k}$	Order coefficient of the Brent oil futures price at time t	η	Correspondence coefficient of global epidemic

are utilized as inputs. The corresponding coefficients are employed to adjust the weightings of each influencing factor, resulting in y_1 as the predicted WTI crude oil futures price for the t-th time point in the conventional group experiment, as indicated in Equation (8).

In the mutation group experiment, building upon the conventional group experiment, the impact mechanism of four mutation factors is incorporated. The corresponding coefficients are used to represent the degree of influence of these mutation factors, yielding y_3 as the predicted WTI crude oil futures price for the t-th time point in the mutation group experiment, as shown in Equation (9).

$$y_1 = \varepsilon_{11,1} y_{t-1} + \varepsilon_{11,2} y_{t-2} + \ldots + \varepsilon_{11,k} y_{t-k} \tag{7}$$

$$y_2 = y_1 + \varepsilon_{1,1} a_{t-1} + \varepsilon_{2,1} b_{t-1} + \ldots + \varepsilon_{10,1} j_{t-1} \tag{8}$$

$$y_3 = y_2 + \alpha \cdot Z + \beta \cdot J + \gamma \cdot S + \eta \cdot Y \tag{9}$$

Results and Discussion

Data Sources

Considering data availability, this study selects the values of input variables for various influencing factors from January 1994 to January 2024. The dataset is divided into a training set, test set, and validation set in the ratio of 7:2:1. The model's predictive performance is measured using metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Accuracy Rate (AR). The formulas for these metrics are as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^{n} \left| \hat{y}_t - y_t \right| \tag{10}$$

$$MAPE = \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{\hat{y}_t - y_t}{y_t} \right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} \left(\hat{y}_t - y_t \right)^2} \tag{12}$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_t \right)^2}{\sum_{t=1}^{n} \left(y_t - \overline{y}_t \right)^2} \tag{13}$$

$$AR = 1 - \frac{100\%}{n} \sum_{t=1}^{n} \left| \frac{\hat{y}_t - y_t}{y_t} \right| \tag{14}$$

Three sets of experiments are designed, with the prediction model being the LSTM model in each case. The first group of experiments serves as the control group, where the input variables are the historical data of higher-order WTI crude oil futures prices. The second group is the conventional group, which includes the input variables from the first group plus 10 standard influencing factors. The third group is the anomaly group, where the input variables from the conventional group are used, and the model is adjusted to incorporate the impact mechanisms of 4 anomalous factors.

Conventional Prediction

In this study, the parameter enumeration method is used to determine the parameters of the LSTM model. The number of input nodes corresponds to the number of features in the input variables, and the number of output nodes, set to 1, matches the features of the output variable.

Table 4 compares the training, testing, and validation set errors between the control and conventional groups. The comparison across all three datasets shows that the error in the conventional group is lower than that in the control group, indicating that incorporating influencing factors provides a calibration effect on the predictive data.

Fig. 6, 7, and 8 respectively present the predictive results and errors for the training, validation, and testing sets of the conventional group. Each Fig. comprises two charts: the upper chart displays the data fitting, and the lower chart shows the error analysis.

In Fig. 6, the initial segment of the upper image, denoted by red blocks, exhibits a notable deviation between the predicted outcomes and the actual results, which are represented by the gray dotted lines. This discrepancy is particularly pronounced in comparison to other instances, as reflected in the corresponding error graph below. This phenomenon can be attributed to the initial misalignment of the model parameters with the data characteristics. In the latter half of the graph, the error variation is less conspicuous, suggesting an adaptation of the model structure to the dataset features, achieved through the incorporation of a substantial amount of training set data. At sample point 3500, a minor peak in the error graph is observable, corresponding to a significant change in the actual data which the model failed to predict with high precision. This indicates that during abrupt data variations, minor error peaks still occur. The red fitting curve closely follows the gray actual data curve in most parts, with the overall error ranging between 0 and 1, indicating a reasonable predictive capability of the model, albeit with certain deviations in specific areas.

Fig. 7 reveals a significant error peak exceeding -5 between samples 1200 and 1400, where the predicted results substantially underperform the actual data. This deviation is attributed to the impact of an external sudden event on price movements, leading to a severe divergence between the predicted trajectory and the actual outcomes during this period. The blue block line, representing the predicted results, closely aligns with the black dotted line of actual results in most areas, with nearly all errors having an absolute value of less than 0.1, barring the aforementioned peak. This demonstrates the model's efficacy in simulating or predicting actual data during the validation phase. Overall, the Figure indicates that while the LSTM model exhibits commendable fitting accuracy on most validation samples with minimal error, it registers significant negative deviations on a minority of samples. These observations from Fig. 7 offer insights into the model's robustness and reliability under specific conditions.

In Fig. 8, the purple triangular line, representing the predicted results, generally aligns slightly below the green cross line of actual results, suggesting a broadly consistent prediction quality. However, in certain segments, such as between samples 200 to 300 and 500 to 600, the predicted data demonstrates notable fluctuations, indicating potential lapses in predictive accuracy in these intervals. This could be due to external shocks causing abrupt short-term increases and decreases in the actual prices, to which the model responds with delayed and more subdued fluctuations. While most errors cluster around zero, some error bars extend beyond -0.2, indicating underpredictions at certain sample points.

From Figs 7 and 8, it is evident that prices exhibit sudden rises or falls in response to external shocks, with the model requiring time to adjust, thereby resulting in subdued predictive changes compared to the actual

Table 4. Error Comparison between Control and Conventional Groups.

Dataset	Experimental Group	MAE	MAPE	RMSE	R^2	AR
Training Set	Control Group Error	7.4718	18.9932%	10.4002	0.9743	81.0068%
	Conventional Group Error	5.9408	15.6315%	8.8844	0.9817	84.3684%
Validation Set	Control Group Error	3.5918	9.2209%	4.3260	0.9772	90.7791%
	Conventional Group Error	2.6472	7.2211%	3.4837	0.9646	92.7789%
Test Set	Control Group Error	6.2899	6.7130%	10.4335	0.7836	93.2870%
	Conventional Group Error	6.2326	6.8807%	9.4877	0.7699	93.1193%

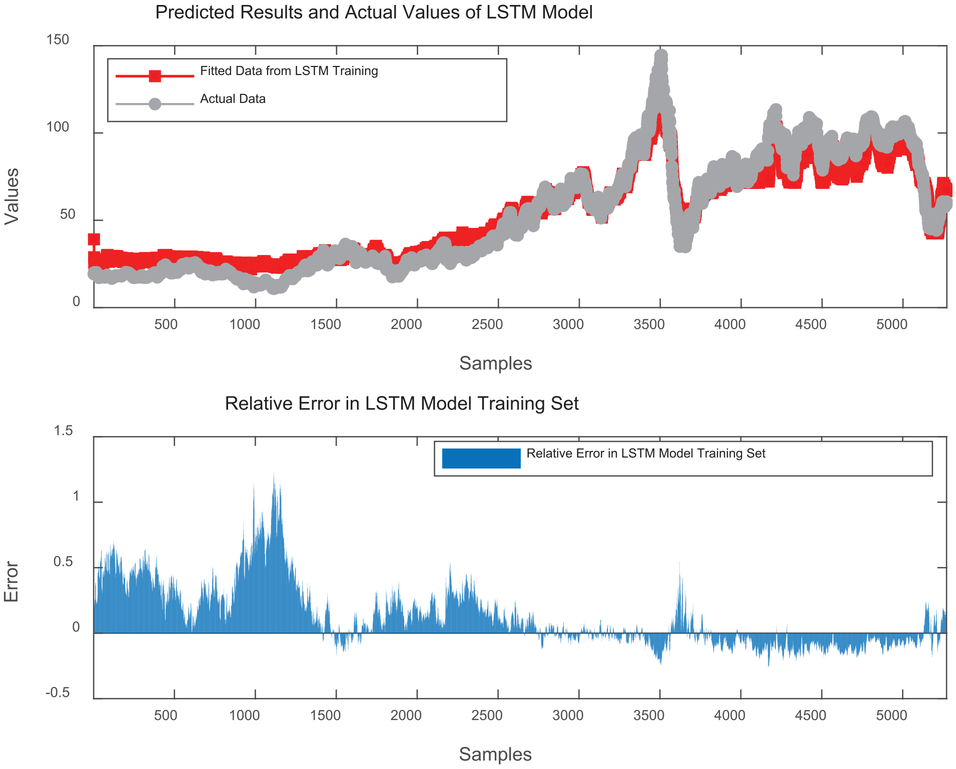

Fig. 6. Error in Training Set Results for the Control Group Experiment.

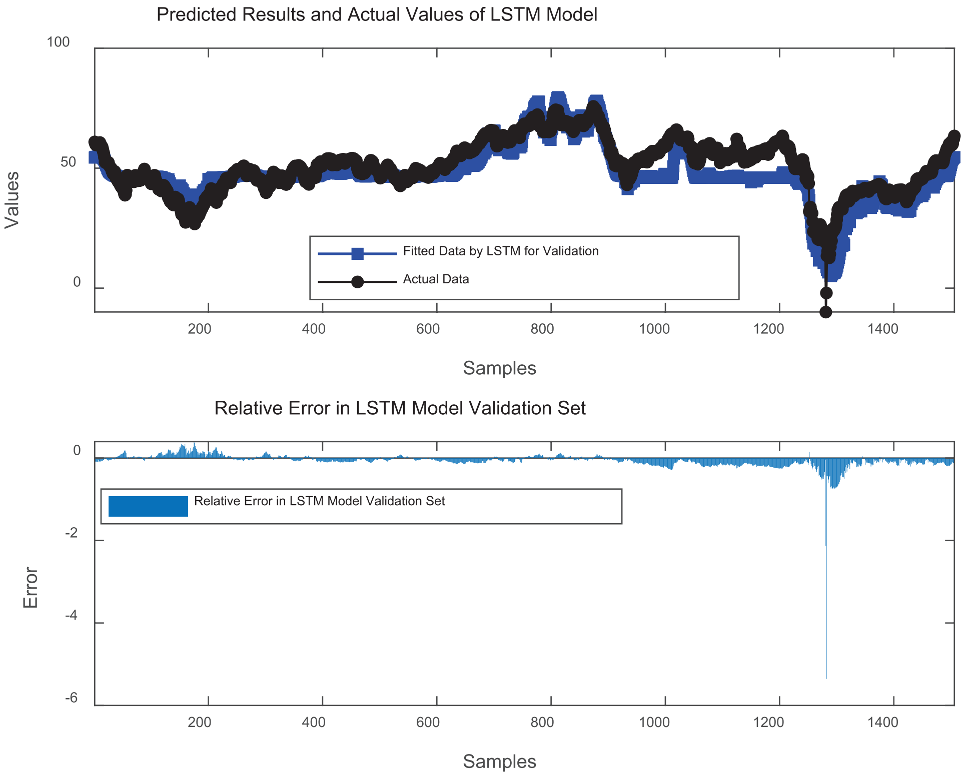

Fig. 7. Error in Validation Set Results for the Control Group Experiment.

data. This leads to implausible predictions and errors under certain conditions, laying the groundwork for the subsequent section on anomaly prediction.

Anomaly Prediction

Building on the previous section's analysis of regular influencing factors, this section focuses on adjusting

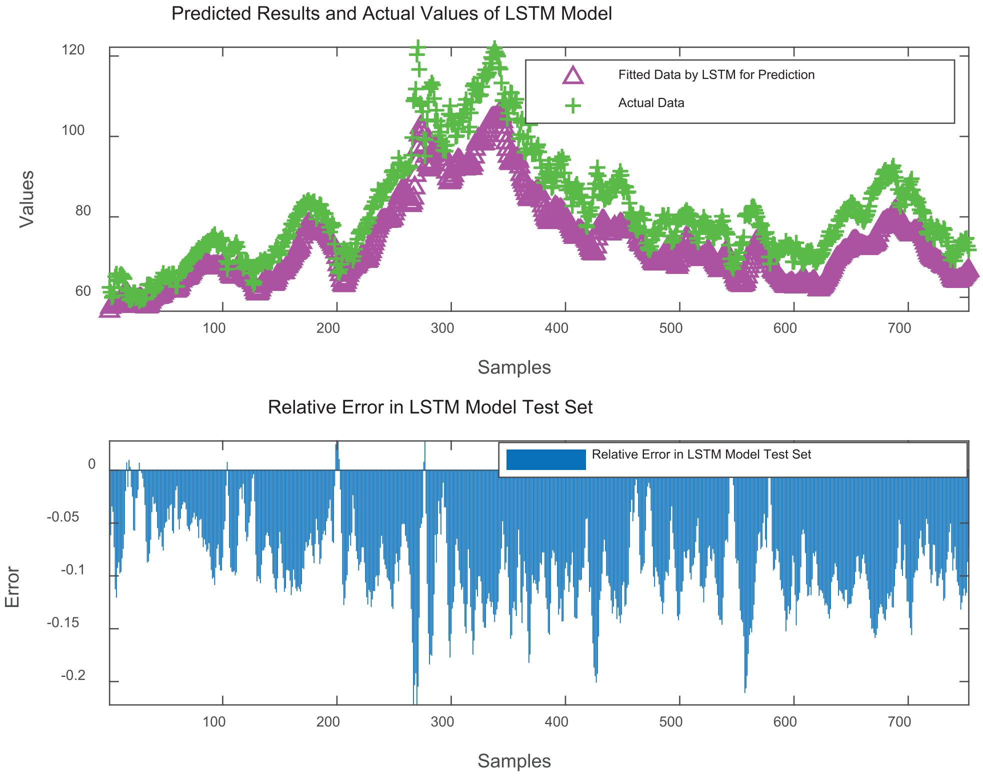

Fig. 8. Error in Test Set Results for the Control Group Experiment.

the predictions deviating from actual outcomes due to anomalous factors.

The recent Russia-Ukraine conflict and the COVID-19 pandemic serve as examples, where prices during these periods are adjusted using corresponding impact mechanisms. Table 4 details the key parameters of these mechanisms. The coefficients for the political event and global pandemic mechanisms are determined by comparing the timelines of these events with those in Table 5, and are then applied to key data parameters like maximum values, minimum values, volatility, and the magnitudes of rises and falls. These adjusted parameters are used to refine the standard predictions, and the actual outcomes, standard predictions, and modified predictions are then compared. Figs 9 and 10 illustrate the comparisons of the actual, standard, and modified predictions for the WTI price movements during the Russia-Ukraine conflict and the COVID-19 pandemic, respectively. Table 5 compares the errors between the standard and modified predictions for these periods.

Fig. 9 shows the fluctuation of WTI prices during the Russia-Ukraine conflict, with the selected time being 2022. The red solid line represents the actual results, the blue dashed line represents the regular group experimental prediction results, and the black dotted line represents the anomaly group experimental prediction results. In the anomaly group prediction experiment, the political event impact mechanism was used to predict that during this period, the maximum price could reach 71.47, the minimum price could reach 41.95, and the volatility rate was approximately 15.61%. These predictions were made by combining other parameters. As shown in Fig. 9, the overall trend of the three datasets is similar, but there are significant differences in values. The regular group's prediction results are much lower than the actual results, and the inclusion of the impact mechanism in the anomaly group adjusted the overall data values closer to the actual results. Although not extremely precise, it significantly reduced the error. Table 6 shows that after using

Table 5. Key Parameters of Impact Mechanisms.

Event \ Parameters	Time Span (months)	Maximum Value	Minimum Value	Volatility	Increase	Decrease
Political Event	13.19	92.08	54.05	15.61%	117.05%	128.52%
Economic Crisis	25.00	144.94	34.82	32.43%	114.68%	135.25%
Technological Revolution	15.64	66.71	27.62	21.63%	38.10%	-50.98%
Global Pandemic	24.95	84.19	-9.95	31.95%	79.98%	-95.71%

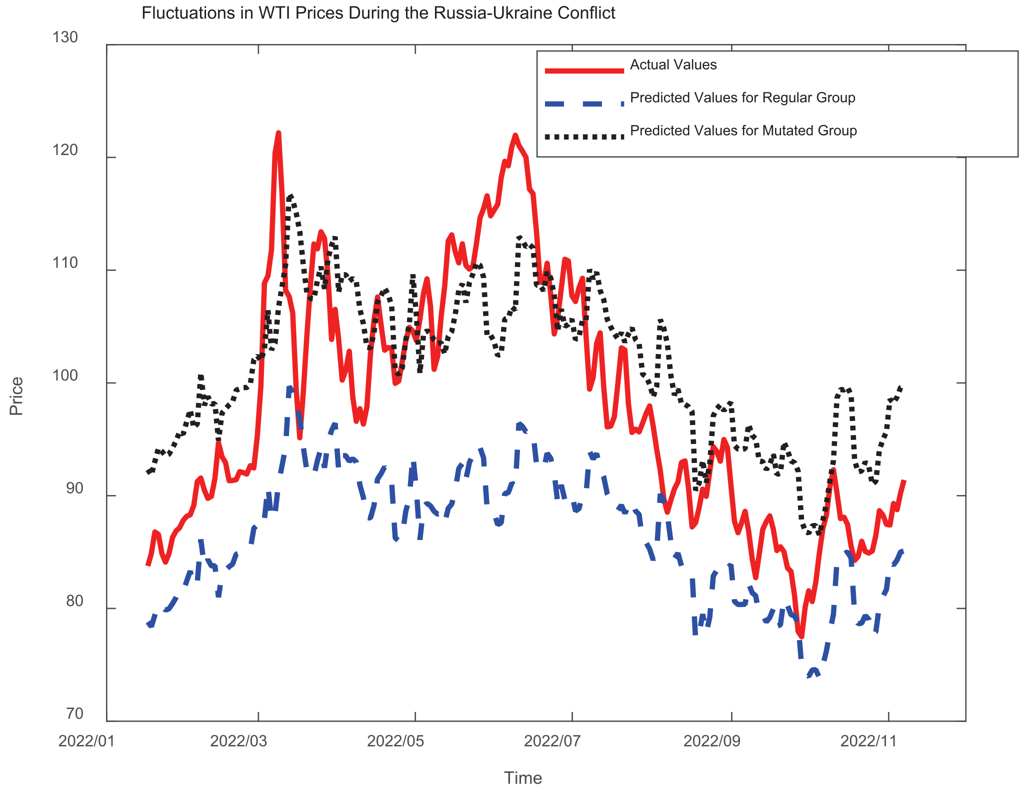

Fig. 9. Data Comparison During the Russia-Ukraine Conflict Period.

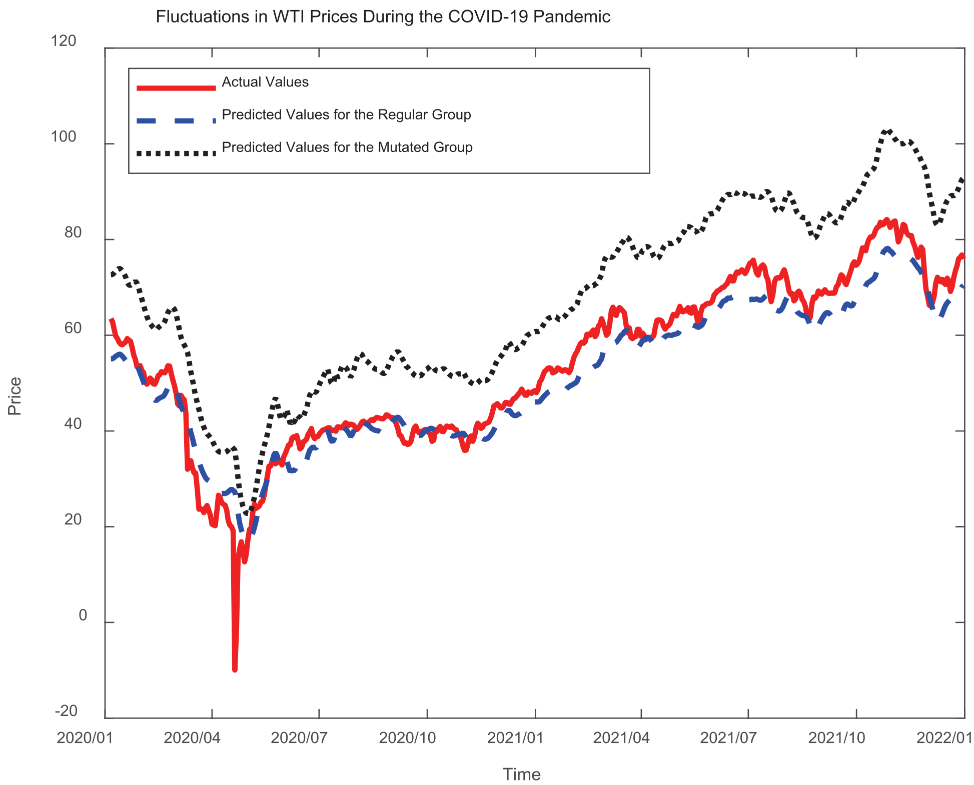

Fig. 10. Data Comparison During the COVID-19 Pandemic Period.

the impact mechanism, the mean absolute error (MAE), mean relative error (MAPE), and root mean square error (RMSE) all decreased significantly, while the coefficient of determination remained largely unchanged, and the accuracy rate (AR) improved slightly.

Table 6. Error Comparison of Regular and Anomaly Group Prediction Results During the Russia-Ukraine Conflict and COVID-19 Pandemic.

Anomalous Events	Experimental Group	MAE	MAPE	RMSE	R^2	AR
Russia-Ukraine Conflict	Regular	11.49763	11.20%	13.51495	0.81013	88.80%
	Anomaly	6.47890	6.77%	7.55504	0.81013	93.23%
COVID-19 Pandemic	Regular	15.30071	33.75%	16.13800	0.95719	66.25%
	Anomaly	8.60056	20.44%	10.15033	0.95719	79.56%

Fig. 10 shows the fluctuation of WTI prices during the COVID-19 pandemic period, with the selected time being 2020 and 2021. The red solid line represents the actual results, the blue dashed line represents the regular group experimental prediction results, and the black dotted line represents the anomaly group experimental prediction results. In the anomaly group prediction results, the global pandemic impact mechanism was used to predict that during this period, the maximum price could reach 84.19, and the minimum price could go down to -9.9, with a volatility rate of about 31.95%. These predictions were made by combining other parameters. As shown in Fig. 10, the overall trend of the data is similar, except for a drastic drop in actual results in April 2020, which was not simulated by either the regular or anomaly group experiments. In terms of numerical predictions, most of the anomaly group results were closer to the actual values than the regular group results. Table 6 shows that after the adjustment with the impact mechanism, the mean absolute error (MAE), mean relative error (MAPE), and root mean square error (RMSE) all decreased to varying extents, while the coefficient of determination remained largely unchanged, and the accuracy rate (AR) improved significantly.

A vertical comparison of the two examples in Table 6 reveals that the error during the COVID-19 period was larger than during the Russia-Ukraine conflict. The former had a longer duration and greater fluctuations, making predictions more challenging. The prediction results and error data from these two examples demonstrate that the impact mechanism significantly enhances prediction accuracy and reduces errors, especially when anomalous factors predominantly influence price movements, thereby substantially narrowing the gap in numerical predictions and compensating for the shortcomings of model predictions.

Conclusion

This paper provides an in-depth analysis of the changing law, prediction accuracy, and stability of crude oil futures prices by exploring a crude oil futures price prediction model that considers the influence mechanism of mutation factors. The results of the study show that the consideration of mutation factors plays an important role in the influence of crude oil futures prices. First, mutating factors play a crucial role in the crude oil futures market. These mutating factors may include political conflicts, natural disasters, supply disruptions, etc., which can directly affect the supply and demand relationship and price fluctuations of crude oil. Through the comprehensive consideration of mutation factors, we can more accurately predict the movement of crude oil futures prices, provide investors and traders with more accurate market information, and help them make scientific decisions. Second, exploring the influence mechanism of mutation factors is important for understanding the volatility of crude oil futures prices and market behavior. Our findings reveal the complex relationship between mutating factors and crude oil futures prices and confirm that the multi-series crude oil futures price prediction model that considers the influencing mechanism of mutating factors has high prediction accuracy and stability. This finding provides investors and traders in the crude oil futures market with a more reliable price prediction tool, which helps to reduce trading risks and improve investment returns. The research in this paper also provides important insights into environmental protection. The extraction, transportation, and use of crude oil have a negative impact on the environment, leading to problems such as climate change and resource waste. By accurately predicting crude oil futures prices, we can assess future energy demand and supply, providing a basis for policymaking for sustainable development and environmental protection. At the same time, promoting the development and utilization of renewable energy can also reduce the dependence on crude oil and ease the environmental burden.

In summary, the findings of this paper highlight the importance of forecasting crude oil futures prices for the environment. Accurately predicting crude oil futures prices can help investors and traders make informed decisions, while providing a decision-making reference for environmental protection and sustainable development. By better understanding the mechanism of the influence of mutation factors on crude oil futures prices, we can better balance the relationship between economic development and environmental protection, and promote the process of green energy transition and sustainable development.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by Shandong Province Social Science Planning Research Project of China (Grant no.22CSDJ13).

References

1. UROLAGIN S., SHARMA N., DATTA T.K. A combined architecture of multivariate LSTM with Mahalanobis and Z-Score transformations for oil price forecasting. Energy, **231**, 120963, **2021**.
2. ZHAO L.T., MIAO J., QU S., CHEN X.H. A multi-factor integrated model for carbon price forecasting: market interaction promoting carbon emission reduction. Science of The Total Environment, **796**, 149110, **2021**.
3. ABDOLLAHI H., EBRAHIMI S.B. A new hybrid model for forecasting Brent crude oil price. Energy, **200**, 117520, **2020**.
4. ABDOLLAHI H. A novel hybrid model for forecasting crude oil price based on time series decomposition. Applied energy, **267**, 115035, **2020**.
5. JIANG H., HU W., XIAO L., DONG Y. A decomposition ensemble based deep learning approach for crude oil price forecasting. Resources Policy, **78**, 102855, **2022.**
6. FENG J., LU X., LIU Y., WU J., HE J., CHEN Z., ZHAO Z. Foundation Settlement Prediction of High-Plateau Airport Based on Modified LSTM Model and BP Neural Network Model. Polish Journal of Environmental Studies, **33** (3), 2037, **2024**.
7. WU J., MIU F., LI T. Daily crude oil price forecasting based on improved CEEMDAN, SCA, and RVFL: a case study in WTI oil market. Energies, **13** (7), 1852, **2020**.
8. LU H., MA X., MA M., ZHU S. Energy price prediction using data-driven models: A decade review. Computer Science Review, **39**, 100356, **2021**.
9. CHAI J., XING L.M., ZHOU X.Y., ZHANG Z.G., LI J. Forecasting the WTI crude oil price by a hybrid-refined method. Energy Economics, **71**, 114, **2018**.
10. GULIYEV H., MUSTAFAYEV E. Predicting the changes in the WTI crude oil price dynamics using machine learning models. Resources Policy, **77**, 102664, **2022**.
11. TANG L., ZHANG C., LI L., WANG S. A multi-scale method for forecasting oil price with multi-factor search engine data. Applied Energy, **257**, 114033, **2020.**
12. ZHOU J., WANG S. A carbon price prediction model based on the secondary decomposition algorithm and influencing factors. Energies, **14** (5), 1328, **2021**.
13. LU W., LI J., WANG J., QIN L. A CNN-BiLSTM-AM method for stock price prediction. Neural Computing and Applications, **33** (10), 4741, **2021.**
14. NWULU N.I. A decision trees approach to oil price prediction. International Artificial Intelligence and Data Processing Symposium, 1, **2017.**
15. LU Q., LI Y., CHAI J. Crude oil price analysis and forecasting: A perspective of "new triangle". Energy Economics, **87**, 104721, **2020**.
16. ZHAO G., XUE M., CHENG L. A new hybrid model for multi-step WTI futures price forecasting based on self-attention mechanism and spatial–temporal graph neural network. Resources Policy, **85**, 103956, **2023**.
17. DAI Z., KANG J., HU Y. Efficient predictability of oil price: The role of number of IPOs and US dollar index. Resources Policy, **74**, 102297, **2021**.
18. ZHAO J. Exploring the influence of the main factors on the crude oil price volatility: An analysis based on GARCH-MIDAS model with Lasso approach. Resources Policy, **79**, 103031, **2022**.
19. MINION L., BANERJEE A. "I can feel the money going out of the window": How high energy prices evoke negative emotions in people with previous experience of homelessness. Energy Research & Social Science, **108**, 103387, **2024**.
20. LIU Y., NIU Z., SULEMAN M.T., YIN L., ZHANG H. Forecasting the volatility of crude oil futures: The role of oil investor attention and its regime switching characteristics under a high-frequency framework. Energy, **238**, 121779, **2022**.
21. MIAO H., RAMCHANDER S., WANG T., YANG D. Influential factors in crude oil price forecasting. Energy Economics, **68**, 77, **2017**.
22. MA Y., JI Q., PAN J. Oil financialization and volatility forecast: Evidence from multidimensional predictors. Journal of Forecasting, **38** (6), 564, **2019**.
23. WANG G., SHARMA P., JAIN V., SHUKLA A., SHABBIR M.S., TABASH M.I., CHAWLA C. The relationship among oil prices volatility, inflation rate, and sustainable economic growth: Evidence from top oil importer and exporter countries. Resources Policy, **77**, 102674, **2022.**
24. ZHANG G., WANG T., LOU Y., GUAN Z., ZHENG H., LI Q., WU J. Analysis of China's Provincial-Level Carbon Peaking Pathways Based on LSTM Neural Networks. Chinese Journal of Management Science, 1, **2024**.
25. HAO Y., TIAN C.A Hybrid Framework for Carbon Trading Price Forecasting: Incorporating the Role of Multiple Influence Factors. Journal of Cleaner Production, **262**, 120378, **2022**.
26. WU Z., SHI J. Analysis and Trend Prediction of Energy Carbon Emissions Influencing Factors in Beijing Based on the STIRPAT and GM (1,1) Models. Chinese Journal of Management Science, **20** (S2), 803, **2012**.
27. ZHANG B., MING T. An Analysis of the Factors Influencing China's Provincial Economy on Land Use Carbon Emissions Based on a Decoupling Model: a Case Study of Sichuan Province from 1990 to 2020. Polish Journal of Environmental Studies, **33** (3), 2457, **2024**.
28. ZHANG J. Spatial Distribution of Green Total Factor Productivity in Chinese Agriculture and Analysis of Its Influencing Factors. Polish Journal of Environmental Studies, **33** (3), 2473, **2024**.
29. D'ORAZIO P. Charting the complexities of a post-COVID energy transition: emerging research frontiers for a sustainable future. Energy Research & Social Science, **108**, 103365, **2024**.
30. AZUBIKE V.C., GATIESH M.M. The intricate goal of energy security and energy transition: Considerations for Libya. Energy Policy, **187**, 114005, **2024**.