Research Article

Selection of Machine Learning Models for Oil Price Forecasting: Based on the Dual Attributes of Oil

Lei Yan [ID],[1] **Yuting Zhu**,[2] **and Haiyan Wang** [ID][1]

[1]*Business School, Zhejiang Wanli University, Ningbo 315101, China*
[2]*School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China*

Correspondence should be addressed to Haiyan Wang; wanghaiyan@zwu.edu.cn

Since the commodity and financial attributes of crude oil will have a long-term or short-term impact on crude oil prices, we propose a de-dimension machine learning model approach to forecast the international crude oil prices. First, we use principal component analysis (PCA), multidimensional scale (MDS), and locally linear embedding (LLE) methods to reduce the dimensions of the data. Then, based on the recurrent neural network (RNN) and long-term and short-term memory (LSTM) models, we build eight models for predicting the future and spot prices of international crude oil. From the analysis and comparison of the prediction results, we find that reducing the dimension of the data can improve the accuracy of the model and the applicability of RNN and LSTM models. In addition, the LLE-RNN/LSTM models can most successfully capture the nonlinear characteristics of crude oil prices. When the moving window size is twenty, that is, when crude oil price data are lagging by almost a month, each model can minimize its error, and the LLE-RNN /LSTM models have the best robustness.

1. Introduction

Forecasting the price of crude oil is of great significance for energy policy-makers, market participants, portfolio diversification, and energy risk management. There are many factors influencing the crude oil price, and the influence period of each factor on the crude oil prices is not consistent, so the crude oil prices have nonlinear characteristics [1, 2]. However, identifying the formation process of crude oil prices is of significance for accurate prediction, but this process is complicated. Therefore, we try to use the machine learning methods to deal with the vague influence among various factors. The formation process of crude oil prices can lead the traditional econometrics model to have a large error in crude oil price prediction, but the RNN and LSTM models can fit well. In particular, we have considered financialization of crude oil markets. The commodity attributes form the long-term trend of crude oil prices, and its financial attributes cause short-term fluctuation. In this paper, we try to forecast the price of crude oil from both spatial perspective and historical perspective.

From a spatial perspective, there are many factors that affect crude oil prices, such as the number of oil fields, mining technology, and crude oil endowments. These factors are closely related to the commodity attributes of oil. From a historical perspective, the formation process of international crude oil prices is complicated. Identifying the formation process of crude oil prices is of great significance for accurate prediction of crude oil prices. The formation of crude oil prices is not directly affected by certain specific factors. For example, crude oil, as the main support energy for maintaining a country's economic development, determines the country's economic development process. Therefore, the imbalance between the demand for and supply of crude oil has become international politics. The key to the game is that this game will inevitably have an impact on crude oil prices. In addition, the influence cycles of various factors on crude oil prices are not consistent, so crude oil prices themselves have obvious nonlinear characteristics.

It can be seen that crude oil itself has both commodity attributes and financial attributes [3]. This makes crude oil prices to exhibit a long-term trend, but also local

fluctuations. First, the relationship between supply and demand of crude oil is the main driving force behind the formation of crude oil prices. Therefore, the impact of crude oil supply and demand on oil prices is long-term and is also a key factor in stabilizing oil prices. Second, international crude oil uses the U.S. dollar as the unit of settlement. Therefore, the trend of the US economy and related political and economic activities will inevitably bring about fluctuations in crude oil prices. Third, the impact of these financial factors on crude oil can only last for a short period of time, causing crude oil prices to fluctuate in the short term in the future, but cannot affect the long-term trend of crude oil prices. Finally, the sudden change in the international economic situation has also caused a certain impact on crude oil prices, causing temporary fluctuations in crude oil prices and bringing the trend of crude oil prices to the unknown. Long-term factors and short-term factors work together to make crude oil prices have significant nonlinear characteristics.

Due to the nonlinear characteristics of oil, forecasting oil price is different. However, the research on the prediction of international crude oil prices never stops. For example, constructing the cointegration model [4, 5], ARIMA model [6], and GARCH model [7] is used to seek better results. Yu et al. [8] hold that SVM performed better in predicting the WTI crude oil prices and Brent crude oil prices. However, they cannot capture the nonlinear relation of crude oil prices because all of the methods only focused on short-term effects, or ignored the interaction of short-term factors and long-term factors and assumed all factors as long-term factors. With the development of machine learning, neural networks have shown good performance in capturing nonlinear relationships among variables. In addition, compared with traditional econometrics, the neural network has a better performance [9], so relevant studies have emerged accordingly. For example, in the study of crude oil price prediction, Moshiri and Foroutan [10] thought that the neural network performs better in prediction. SenGupta et al. [11] improves the BN-S model with the implementation of various machine learning algorithms, and validate the efficacy of the proposed model for long-range dependence. Wang et al. [12] proposed a multigranularity heterogeneous combination approach to enhance the forecasting accuracy, and their results demonstrate that this approach can outperform not only individual competitive benchmarks but also single-granularity heterogeneous and multigranularity homogenous approaches. In addition to the application of the above single model, the composite model is used to predict the price of the crude oil [13].

The nonlinear characteristic of crude oil price is confirmed by the application of the neural network. However, the above method still assumed the relationship between crude oil prices and various factors as a short-term influence in essence and ignored the joint action of long-term and short-term factors of crude oil prices. The emergence of the recurrent neural network (RNN) solves this problem. By setting the recurrent unit in the hidden layer, RNN can keep the memory of the historical data, which overcomes the defect that the BP neural network cannot make the prediction based on the historical data. When making a prediction, it combines historical data with current input to jointly predict the output. However, RNN has obvious defects as an exploding gradient and vanishing gradient. To solve this problem, a long-term and short-term memory model (LSTM) as a variant of RNN has been proposed. The LSTM model controls the memory and discard the information by adding gate structure. Among them, cell state, as the most important structure of LSTM, stores long-term memory and solves the problems of exploding gradient and vanishing gradient [14]. At present, LSTM is mainly applied in the research of speech recognition [15, 16], stock-price prediction [17, 18], and other fields. Based on the above facts, RNN and LSTM models can take into account the combined effects of long-term and short-term factors in crude oil prediction, so they are suitable for crude oil price prediction. The RNN model also was employed on the crude oil price prediction. Cen and Wang [19] collected three crude oil price data and used LSTM to predict WTI and Brent crude oil prices. At the same time, the authors employed the ensemble empirical mode decomposition model to decompose the price of WTI and used the LSTM model to predict the oil prices with the modal function as the input. Wu et al. [20] proposed the EEMD-LSTM mode to decompose WTI crude oil spot prices and used the modal function as input to predict crude oil prices. However, all the above methods take the historical data of crude oil prices as the input to predict the oil prices, which is insufficient in exploring the factors and nonlinear formation process of crude oil prices.

This research makes the following three contributions. First, by analysing the dual attributes of crude oil, this paper owes the formation process of crude oil prices to the commodity attribute that has a long-term impact on it and the financial attribute that has a short-term impact on it. Second, RNN and LSTM models are used as the benchmark model to predict the crude oil prices. Third, considering that too many influencing factors are likely to cause overfitting, the de-dimension method is adopted to improve the accuracy of the model. That is, we propose a de-dimension machine learning model approach to forecast the international crude oil prices by taking account of the dual attributes of oil.

The structure of the paper is as follows: in Section 2, we discuss the key factors affecting the price of crude oil by analysing its attributes; in Section 3, relevant dimension reduction methods and RNN and LSTM benchmark models are introduced, and the "dimension reduction prediction" model is constructed; Section 4 forecasts the future and spot price of crude oil through empirical analysis and verifies the robustness of each model through comparative analysis; and the conclusion is provided in Section 5.

2. Influencing Factors of Crude Oil Prices

The first step in predicting the price of crude oil is to determine the influencing factors. However, because there are so many factors that influence the crude oil prices, the scholars' opinions are inconsistent. In this paper, the

nonlinear characteristics of crude oil prices are attributed to the dual attributes of crude oil, namely, commodity attributes and financial attributes. The factors that affect the change in crude oil prices are illustrated through the analysis of the two attributes.

Crude oil is an essential energy, and its commodity attribute is reflected by the relation between supply and demand and the relation between oil prices and the price of alternative products. The relation between supply and demand, as a reflection factor of commodities, determines the trend of crude oil prices. Gallo et al.'s research showed that supply is the main factor leading to crude oil price change, but the change in demand is the result of price change [21]. Because supply and demand cannot dominate the oil prices all the time, until the 20th century, the impact of supply and demand increased [22]. In addition to supply and demand factors, petroleum, as the main product of crude oil, affects the price of oil. In addition, changes in the price of other commodities also affect the price of oil, which in turn affects the price of crude oil. Mensi et al. [23] showed that there is a correlation between silver and oil prices. Bildirici and Turkmen [24] found the bidirectional causality of oil prices to silver prices and gold prices, but the effect of precious metal prices on crude oil price shocks is nonlinear. Zhang and Sun [25] confirmed the synergy between the European coal market and the energy (coal, gas, and Brent oil) market. In particular, natural gas and oil prices have not decoupled, and there is a long-term cointegration relation [26]. Therefore, in terms of commodity attributes of crude oil, the relationship between supply and demand for crude oil and the price of alternative products are the main reasons that affect the fluctuation of crude oil prices.

With the establishment of the crude oil financial market, the financial attributes of crude oil is further strengthened. Crude oil is settled in US dollars, so the trade activity of using crude oil to replace currency for payment makes crude oil act as a currency with a value scale. At the same time, the future contract of crude oil makes crude oil become the carrier of capital flow, so there is a significant negative correlation between crude oil and dollar exchange rate [27], and this negative correlation has strengthened in the last decade [28]. Kilian and Vigfusson [29] refuted this and argued that the relation between crude oil prices and the dollar exchange rate is nonlinear. Chen et al. [30] showed that this nonlinear relationship is not significant, but under the supply and demand driven price shocks, the dollar exchange rate has caused different impacts on the crude oil prices. Therefore, the dollar exchange rate is not the only factor that influences the crude oil prices. For example, the research results of Bouoiyour et al. [31] indicated that the crude oil prices are positively correlated with the Russian exchange rate. Blokhina et al. [32] confirmed the close interrelation between oil prices and rouble exchange rate. The price of crude oil is the result of a combination of the world economy, mainly because crude oil has replaced some of the functions of gold. Under the Jamaican system, the dollar was decoupled from gold. However, investors still want to find physical protection against the risks posed by fluctuations in the value of the dollar. Crude oil is used by investors in asset allocation, and crude oil financial markets came into being. Like gold, crude oil enters the safe-haven currency system as a physical asset, so there is a correlation between the crude oil market and the gold market [33, 34]. Zhang and Wei [35] found that the relation between oil and gold prices is positive. Although the correlation only exists in the short term [36], interest rates influence the US dollar, which in turn influences international crude oil prices.

From the above discussions, there are many factors influencing the fluctuation of crude oil prices, including crude oil supply and demand, substitute prices, dollar exchange rate, money supply, and gold prices. In this paper, the above variables are selected as the factors affecting the international crude oil prices and substituted into the model for empirical analysis.

3. Methods

3.1. De-Dimension. There are various factors that affect the price of crude oil. Too many factors will lead to the slow calculation speed of the model and inaccurate prediction results. Therefore, dimension reduction is required for the data. The method of dimension reduction mainly includes two kinds. One is feature selection, that is, through some method, find out the important factors, and discard other factors, to achieve the purpose of dimension reduction. The second is feature extraction, that is, the data points in the original high-dimensional space are mapped to the low-dimensional space through some mapping method. We hope to find a suitable mapping function $f: x \longrightarrow y$, mapping the original high-dimensional spatial data x to low-dimensional data y, and f can be linear and nonlinear. Since the feature selection method will discard some factors and lead to the loss of information of a certain factor, which is not in line with the original intention of this study. This paper chooses the feature extraction method to reduce the dimension of data.

There are a lot of feature extraction dimension reduction methods of related research, mainly including principal component analysis (PCA), linear discriminant analysis (LDA), multidimensional scale (MDS), isometric embedding (Isomap), and locally linear embedding (LLE) method. PCA and LDA are the linear mapping methods, while MDS, Isomap, and LLE are nonlinear mapping methods. PCA is an unsupervised dimension reduction method. LDA is a supervised dimension reduction method, and it is necessary to calculate the dimension reduction result according to the target value. Therefore, PCA is selected as the linear dimension reduction method for analysis in this paper. Considering the possible nonlinear characteristics among the data of various factors, this paper also employed the nonlinear method to reduce the dimension of the data. In the selection of nonlinear methods, MDS is used as the most classical method. However, the Isomap method, due to its idea of global optimization, slows down the operation speed when the amount of data is large, while LLE only focuses on local optimization. Therefore, the LLE method is also employed.

3.1.1. PCA. PCA, as a dimension reduction method based on feature extraction, can extract key information in data by a linear combination of original variables to form new variables that are independent of each other and contain original variable information as much as possible.

For matrix $X_{n \times m}$ with n characteristic values and row mean of 0, reduce to k dimension data, and the calculation process of PCA is as follows:

(1) Calculate the covariance matrix $\mathrm{Cov} = (1/m)XX^T$, where Cov is an $n \times n$ orthogonal matrix.

(2) Calculate the characteristic values λ_i of Cov and the corresponding characteristic vector p_i, assuming $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k \cdots \geq \lambda_n$, arrange p_i into a matrix P according to characteristic values from the largest to the smallest, taking the first k rows in P to form the matrix K.

(3) Calculate $Y = KX$, Y is the data after dimension reduction.

In the above calculation process, the contribution rate of the jth principal component is $(\lambda_j / \sum_{i=1}^n \lambda_i)$, and the cumulative contribution rate of the k principal component is $(\sum_{i=1}^k \lambda_i / \sum_{i=1}^n \lambda_i)$.

PCA can effectively extract the main information between variables and achieve the purpose of dimension reduction.

3.1.2. MDS. MDS map high-dimensional data to the low-dimensional space and keep the distance between the samples in the high-dimensional space as unchanged as possible. For matrix $X_{n \times m}$ with n characteristic values, the element d_{ij} in matrix D represents the distance between the sample x_i and x_j in X, which is $d_{ij} = \|x_i - x_j\|$. Assuming that the dimension after mapping is k dimension, Y represents the $k \times m$ matrix after mapping, then $\|y_i - y_j\| = d_{ij}$. Since the Y cannot be directly solved, it is solved with $B = Y^T Y$. Simplifying, assuming that the Y satisfies $\sum_{i=1}^m y_i = 0$, the solution process of Y is followed.

Get $b_{ij} = y_i^T y_j$, $d_{ij}^2 = \|y_i - y_j\|^2 = \|y_i\|^2 - 2y_i^T y_j + \|y_j\|^2 = b_{ii} - 2b_{ij} + b_{jj}$ from $B = Y^T Y$, that is, $d_{ij}^2 = b_{ii} - 2b_{ij} + b_{jj}$. The following can be obtained since d_{ij} is known:

$$
\sum_i d_{ij}^2 = \sum_i \|y_i\|^2 - 2 \left(\sum_i y_i^T \right) y_j + m \|y_j\|^2 = \mathrm{tr}(B) + mb_{jj},
$$
$$
\sum_j d_{ij}^2 = m \|y_i\|^2 - 2y_i^T \left(\sum_i y_j \right) + \sum_j \|y_j\|^2 = \mathrm{tr}(B) + mb_{ii}, \tag{1}
$$
$$
\sum_i \sum_j d_{ij}^2 = \sum_i \left(m \|y_i\|^2 + \sum_j \|y_j\|^2 \right) = m \left(\sum_i \|y_i\|^2 + \sum_j \|y_j\|^2 \right) = 2m\,\mathrm{tr}(B),
$$

where $\mathrm{tr}(*)$ represents the trace of the matrix. According to the above equation, we obtain

$$
\mathrm{tr}(B) = \frac{1}{m} \sum_i \sum_j d_{ij}^2, \tag{2}
$$

$$
b_{ii} = \frac{1}{m} \left(\sum_j d_{ij}^2 - \mathrm{tr}(B) \right) = \frac{1}{m} \sum_j d_{ij}^2 - \frac{1}{2m^2} \sum_i \sum_j d_{ij}^2, \tag{3}
$$

$$
b_{jj} = \frac{1}{m} \left(\sum_i d_{ij}^2 - \mathrm{tr}(B) \right) = \frac{1}{m} \sum_i d_{ij}^2 - \frac{1}{2m^2} \sum_i \sum_j d_{ij}^2, \tag{4}
$$

$$
b_{ij} = \frac{1}{2} \left(b_{ii} + b_{jj} - d_{ij}^2 \right) = \frac{1}{2m} \left(\sum_i d_{ij}^2 + \sum_j d_{ij}^2 \right) - \frac{1}{m^2} \sum_i \sum_j d_{ij}^2 - \frac{1}{2} d_{ij}^2. \tag{5}
$$

B can be obtained from equations (3)–(5). B is decomposed into $B = V \Lambda V^T$, Λ is the diagonal matrix composed of the characteristic values of the B from large to small, and V is the matrix spanned by the characteristic vectors.

Take the first k characteristic values to form the matrix Λ_1 and the corresponding characteristic vector V_1, and get $Y = \Lambda_1^{1/2} V_1^T$, where Y is the data after dimension reduction.

3.1.3. LLE. Different from the Euclidean distance measurement method of the matrix in the MDS method, Isomap changes the Euclidean distance to geodesic distance and finds the low-dimensional mapping relation. Isomap changes the Euclidean distance to the distance from the nearest neighbour value, and changes to the infinite value from other nonnearest neighbour values. However, this method tries to find the global optimal solution of all samples. In the face of a large amount of data and high dimension, the calculation is time-consuming, so LLE was proposed. LLE only pays attention to the optimal solution of the local space to reduce the computation and achieve the purpose of dimension reduction.

For a sample x_1, samples x_2, x_3, x_4 of the nearest neighbours can be found. Therefore, x_1 can be expressed linearly as $x_1 = w_{12}x_2 + w_{13}x_{13} + w_{14}x_4$ through x_2, x_3, x_4, where w is the weight coefficient. We hope that the dimension reduction data x_1', x_2', x_3', x_4' still maintain the relation of $x_1' = w_{12}x_2' + w_{13}x_3' + w_{14}x_4'$, where w remains unchanged or only makes the smallest change. For the above objectives, the solution process of LLE is followed.

For matrix $X_{n \times m}$, with n characteristic values, after linear representation of each sample with q nearest neighbour samples, the error between its real value and the expressed value is expected to be minimal, that is, the optimization function is as follows:

$$\varepsilon(w) = \sum_{i=1}^{m} \left\| x_i - \sum_{j \in Q(j)} w_{ij}x_j \right\|^2, \tag{6}$$

$$\sum_{j \in Q(j)} w_{ij} = 1. \tag{7}$$

Let $Z_i = (x_i - x_j)(x_i - x_j)^T$, $W_i = (w_{i1}, w_{i2}, \ldots, w_{iq})^T$, then (6) and (7) can be expressed as

$$\varepsilon(w) = \sum_{i=1}^{m} W_i^T Z_i W_i$$
$$\sum_{j \in Q(j)} w_{ij} = W_i^T 1_q = 1, \tag{8}$$

where 1_q is a q dimension vector of all 1. The weight coefficient W_i can be obtained from the following equation:

$$W_i = \frac{Z_i^{-1} 1_q}{1_q^T Z_i^{-1} 1_q}. \tag{9}$$

Assuming that the dimension after mapping is k, and Y is $k \times m$ order matrix after mapping, Y satisfies

$$\varepsilon(y) = \sum_{i=1}^{m} \left\| y_i - \sum_{j \in Q(j)} w_{ij}y_j \right\|^2, \tag{10}$$

$$\sum_{i=1}^{m} y_i = 0, \tag{11}$$

$$\frac{1}{m} \sum_{i=1}^{m} y_i y_i^T = 1. \tag{12}$$

Simplifying (10),

$$\varepsilon(y) = \sum_{i=1}^{m} \left\| YI_i - YW_i \right\|^2 = \text{tr}\left(Y(I - W)(I - W)^T Y^T\right). \tag{13}$$

Let $M = (I - W)(I - W)^T$, simplifying (12) and (13),

$$\varepsilon(y) = \text{tr}\left(YMY^T\right),$$
$$YY^T = mI. \tag{14}$$

The solution of the optimal solution is the first $k + 1$ characteristic values of matrix M and their corresponding characteristic vectors $V = (v_1, v_2, \ldots, v_{k+1})$, then $Y = (v_1, v_2, \ldots, v_{k+1})^T$ is the data after dimension reduction.

3.2. Benchmark Model

3.2.1. RNN. Compared with the general neural network, the RNN model can fit the time series of variable length by establishing the connection weight in the hidden layer [37]. RNN realized the memory storage of historical information through the connection weight between the hidden layers. In the next prediction, it can make a joint prediction according to the historical information and the current input, thus realizing the prediction of time series. Figure 1 shows the basic structure of the hidden layer.

X_t is the input at time step t, h_t is the output of the hidden layer, o_t is the output, f_1 and f_2 are activation functions, W, U, and V represent the weight coefficient matrices. At time t, the hidden layer information is updated as $h_t = f_1(Ux_t + Wh_{t-1})$, and the output information is $o_t = f_2(Vh_t)$. h_t is the process by the historical data accumulation, which is thus able to store information for a long time. However, as the sequence continues to grow, the impact of long-term information on the current is exponentially decreasing or increasing as the activation function continues to multiply, thus causing the well-known problem of "gradient disappearance" or "gradient explosion" [38, 39]. The emergence of the LSTM has solved the problem.

3.2.2. LSTM. LSTM, as a variant of RNN realized the long-term and short-term information storage in the network by adding gates to control the input and output of data. Like basic neural networks, the LSTM model includes an input layer, a hidden layer, and an output layer. The hidden layer is composed of memory cells. Each memory cell is equipped with three units, namely, forget, input, and output gates. The three gates control the memory input and output of information. The cell state can store history information. Figure 2 shows the basic structure of a memory cell.

X_t denotes input at time step t, h_t is output, C_t denotes the cell state at time step t, $\overline{C}_t$ is the update value of cell state, f_t, i_t, and o_t represent the forget, input, and output gate information, respectively. The output of each gate is
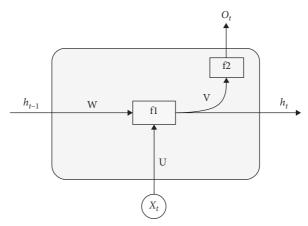
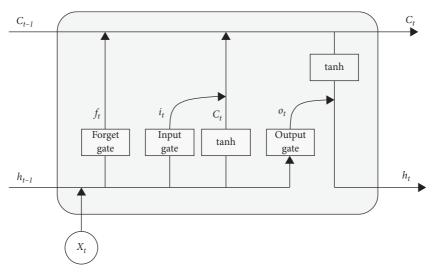Figure 1: Basic structure of the RNN hidden layer.

Figure 2: Basic structure of a memory cell with three gates.

controlled by sigmoid function. $\sigma(*)$ is sigmoid function, W_* denotes weight matrix, and b_* denotes bias. At time t, the initial cell state is C_{t-1}, the cell state update and output information updates are as follows: (1) determining the forget information of the cell state through the forget gate, (2) determining the update value of the cell state according to the input gate, (3) updating the cell state, and (4) the output information is determined through the output gate. The final output is

$$h_t = o_t \times \tanh(C_t), \qquad (15)$$

$\sigma(x)$ and $\tanh(x)$ are activation function, and they are defined by the following equations, respectively,

$$\sigma(x) = \frac{1}{1 + e^{-x}},$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \qquad (16)$$

In the above process, the LSTM model can realize memory storage of the input state through updating the cell state in the memory cell to predict the crude oil prices

through historical data. To evaluate the accuracy of the prediction results of the model, this paper evaluates the model through mean absolute error (MAE) and root mean square error (RMSE). The calculation formula is as follows:

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^{m} |\hat{y}_i - y_i|,$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (\hat{y}_i - y_i)^2}, \qquad (17)$$

where $\hat{y}_i$ denotes the predicted value and y_i denotes the true value.

3.3. The Overall Model. Based on the de-dimension machine learning model approach to forecast the international crude oil prices, this paper combines the above-mentioned de-dimension methods and benchmark prediction models to predict the international crude oil prices. Figure 3 illustrates the prediction process. First, the main factors causing the fluctuation of crude oil prices based on the analysis of the influencing factors of crude oil prices are determined.
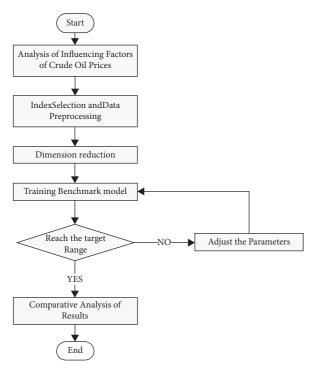
FIGURE 3: Crude oil price prediction process.

Second, indicators and data are collected, and the data are preprocessed and standardized. Third, dimension reduction of data is performed to generate training data. Then, the training data is inputted into the RNN and LSTM models, and the model parameters are adjusted to achieve the optimal training model. Finally, by comparing the results of other models, the performance of the model is determined, and the prediction of crude oil prices is realized.

4. Experimental Analysis

4.1. Sample Generation and Model Design

4.1.1. Data Selection and Preprocessing. There is a strong correlation between the international crude oil future prices and the spot prices. The spot price of crude oil and the future price of crude oil interact with each other. Therefore, this paper predicts the crude oil future prices (FWTI) and spot prices (SWTI), respectively, to test the robustness of the model.

According to the analysis of influencing factors of crude oil prices, the indicators selected in this paper include crude oil future and spot prices, oil supply and demand, natural gas prices, dollar exchange rate, rouble exchange rate, gold prices, and silver prices.

Specific data are selected as follows: West Texas crude oil future (FWTI) and spot (SWTI) crude oil prices are selected, respectively. Natural gas future prices as natural gas prices are selected. World crude oil production as the total supply of oil is selected. Because the total demand for oil does not have access to data, the sum of major consumption by countries is chosen as the demand for oil, including Canada, France, Germany, Italy, Japan, South Korea, the United

Kingdom, the United States, and China. The dollar exchange rate and rouble exchange rate, respectively, are selected and converted with 2010 as the base period. The gold future closing prices as the price of gold and silver future closing prices as the price of silver are selected.

The data were collected from January 2, 1997 to February 28, 2019, totally 5547 pieces. In order to improve the speed and accuracy of the model, the monthly data were converted into daily data by replication, the missing values were deleted, and the above data were normalized to the experimental data with a mean of 0 and a variance of 1.

We select data from 2008 to 2019 for the specific illustration. Figure 4 shows the trend of crude oil future prices, spot prices, and natural gas prices. It can be seen that crude oil future prices and spot prices have the same trend, and there is a weak positive correlation with natural gas prices. Therefore, when analysing the relationship between crude oil prices and variables, we choose crude oil future prices for analysis. Figure 5 shows the relationship between crude oil future prices and crude oil supply and demand. We can find that the relation between crude oil future prices and supply or demand is not significant, and oil supply and demand is increasing in the recent years. Figure 6 shows the relation of the crude oil future prices with the dollar exchange rate and the rouble exchange rate. Positive correlation exits between the crude oil future prices and the rouble exchange rate, but a weak negative correlation with the dollar exchange rate. It can be seen from Figure 7 that there is a certain positive correlation between crude oil future prices and gold prices, and gold prices are more volatile, and crude oil prices with silver prices have the same correlation.

4.1.2. De-Dimension Results

(1) PCA Result. Three principal components were extracted by PCA, and its characteristic value and the rotated variance contribution ratio is shown in Table 1, and the rotation matrix is shown in Table 2. Three of the principal components contain 85.493% of the original information. Among them, silver prices, gold prices, and oil supply contributed the most to the first factor; rouble exchange rate and oil demand contributed the most to the second factor; and the dollar exchange rate and natural gas prices contributed the most to the third factor.

(2) MDS Result. MDS was used to conduct dimension reduction on the experimental data. Kruskal stress [40] is calculated to evaluate the dimension reduction result. The defined equation of the Kruskal stress is as follows:

$$\text{KStress} = \sqrt{\frac{\sum_{i<j}\left(d_{ij} - D_{ij}\right)^2}{\sum_{i<j}D_{ij}^2}}, \qquad (18)$$

where d_{ij} represents the distance between the ith sample and the jth sample after dimension reduction, and D_{ij} represents the distance between the ith sample and the jth sample in the original high-dimensional space. The relation between
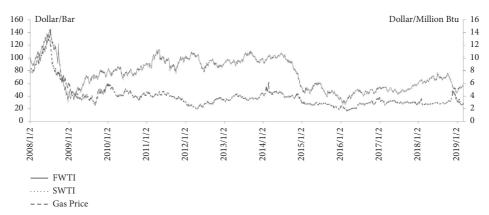
Figure 4: Relation between crude oil future prices, spot prices, and natural gas prices.
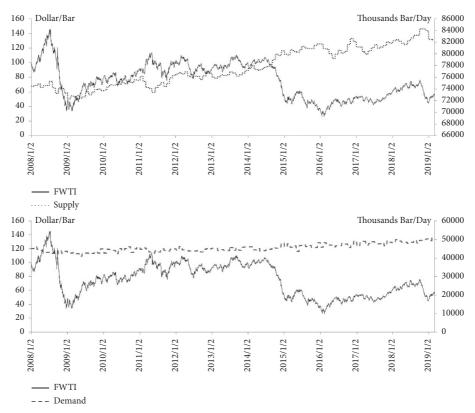

Figure 5: Relation between crude oil future prices, supply, and demand.

dimension and Kruskal stress index is shown in Table 3. According to the stress index and fitness degree (Table 4), when the dimension is 3, the dimension reduction preforms best. Figure 8 shows the relation between the dimension and Kruskal stress index. The stress index drops suddenly when the data of the three dimensions are extracted. As the dimension increases, the reduction range of stress index decreases. Therefore, three components were extracted in this paper for crude oil price prediction.

(3) LLE Result. The dimension reduction method of LLE is related to the number of the nearest neighbour values, therefore, the nearest neighbour values of 5, 10, 20, 30, and 40 are, respectively, selected to reduce the dimension of the experimental

data. The error is shown in Table 5. When the nearest neighbour value is 10, the error is the smallest. Therefore, data with the nearest neighbour value of 10 are selected in this paper to reduce the dimension of the original data.

4.1.3. Sample Selection and Parameter Setting. To test the trained model performance, the data need to be divided into training samples and testing samples. The training samples used for model training and test sample test performance of the model. In this paper, the training sample data from 1997 to 2014 and the test sample data from 2015 to February 2019 were considered. Moving the window approach is employed to generate the sample data.

Figure 6: Relation between crude oil future prices, dollar exchange rate, and ruble exchange rate.
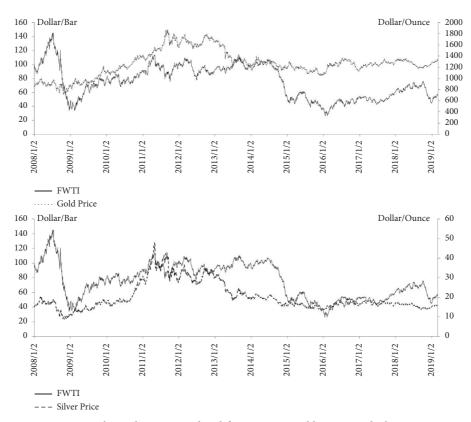

Figure 7: Relation between crude oil future prices, gold prices, and silver prices.

The model can directly take the input sample as lag information by generating the sample through the moving window. The input data of crude oil prices prediction model can be guaranteed to conform to the generation process of time series. The process of generating samples through moving window is as follows: first, the size of the window is selected, and the data consistent with the size of the window are obtained. Second, the crude oil price of the last day in the window data is taken as the output data, and other data are taken as the input data to obtain a training sample. With a fixed step size, repeat the above steps and move the window continuously to build a data unit for model training and prediction. Figure 9 describes the sampling process of the moving window. The crude oil future price data from January to March 2018 were selected and divided into training samples and test samples with a window size of 20 and a step length of 1.

In this study, RNN and LSTM models were, respectively, used to predict crude oil prices. The RNN model consists of one input layer and one output layer, and one RNN hidden layer, which contains 100 nodes. The LSTM model consists of one input layer, one output layer, and two LSTM hiding layers, where the first hiding layer contains 100 nodes, and the second hiding layer contains 100 nodes. The optimizers

TABLE 1: Characteristic values and variance contribution rate.

Principal component	Characteristic values	Variance contribution rate (%)	Accumulated variance contribution rate (%)
1	2.474	35.339	35.339
2	1.782	25.458	60.797
3	1.729	24.696	85.493

TABLE 2: Component matrix.

	Factor 1	Factor 2	Factor 3
Silver prices	0.961	0.078	−0.013
Gold prices	0.933	0.280	0.164
Supply	0.669	0.629	0.161
Ruble exchange rate	−0.056	−0.788	0.139
Demand	0.464	0.761	0.126
Dollar exchange rate	0.056	0.201	0.924
Natural gas prices	−0.102	0.247	−0.887

TABLE 3: Dimension and Kruskal stress index.

Dimension	KStress
1	0.424
2	0.075
3	0.016
4	0.009
5	0.005
6	0.005

TABLE 4: Relation between Kruskal stress and goodness of fit.

Kruskal stress	Goodness of fit
0.200	Poor
0.100	Fair
0.050	Good
0.025	Excellent
0.000	Perfect

of both models selected Adam, the data was iterated 20 times, and the activation function of the output layer is tanh function.

4.2. Result Analysis

4.2.1. Prediction Results. We constructed eight models according to de-dimension methods and benchmark model mentioned and predicted the spot prices and future prices, respectively. All models we constructed are shown in Table 6.

The fixed window size is 60, which predicts crude oil prices by the above models, respectively. The predicted results are shown in Figures 10–13. Figures 10 and 11, respectively, show the prediction results of crude oil future prices and spot prices, where the benchmark model is RNN. It can be seen that PCA-RNN and MDS-RNN do not fit well in the prediction of future oil prices. In addition, when predicting spot oil prices, prediction results of the models after dimension reduction performed better than the single RNN model. Figures 12 and 13, respectively, show the prediction results of crude oil future prices and spot prices,

where the benchmark model is LSTM. It can be seen that LLE-LSTM performed well in fitting the prediction of crude oil future prices and spot prices than other models in prediction of future oil prices. In prediction of spot oil prices, de-dimension improves the fitting effect of the model. RNN does not have the vanishing gradient problem. Therefore, both RNN models and LSTM models perform well and can not only capture the nonlinear characteristics of crude oil prices but also fit the long-term factors and short-term factors in the data.

4.2.2. Comparative Analysis. In order to further compare the prediction results of the model, BP neural network, SVM model, and ARIMA model were, respectively, constructed in this paper for comparison with the above models. According to the calculation of MAE and RMSE, Table 7 shows the results of each benchmark model. Among them, ARIMA has the worst performance, so the influence of various factors should be considered in crude oil prediction. MAE and RMSE of the SVM model are both larger than the results of the neural network models, so there is a nonlinear correlation between crude oil prices and various factors affecting crude oil prices. In neural network models, RNN and LSTM models performed well than the BP network model.

In order to compare the influence of de-dimension on the prediction results, the benchmark model and its variants are, respectively, used to predict the crude oil prices. The error is shown in Table 8. The results of MAE and RMSE of the LLE-RNN model are the smallest, followed by the LLE-LSTM model. In panel A, RMSE of the LLE-RNN model (FWTI) is 0.0442, which is 54.99% lower than that of the RNN model. The LLE-RNN model shows the best result in predicting spot prices, where the RMSE is 56.75% lower than the RNN benchmark model, and less than Cen and Wang's [19] result of 0.0539, and the optimal results in the study of error are reduced by 22.68%. Meanwhile, in the study of Wu et al. [20], the error of the EEMD-LSTM model is 35.75% lower than that of the LSTM model. Therefore, the LLE-RNN model performed better in crude oil price prediction. In panel B, the LLE-LSTM model performed well in the prediction of future and spot oil prices than other models. We get the same result in panel C. Compared with PCA and MDS, LLE realized linear transformation by finding the nearest neighbour of the data and converging the long-term similarity factors and short-term similarity factors. At the same time, the overall data presented the characteristics of nonlinear transformation, so as to achieve the purpose of dimension reduction. Therefore, LLE model fits better than other methods and confirmed the nonlinear correlation among the factors that affect the price of crude oil. With the same dimension reduction method, RNN
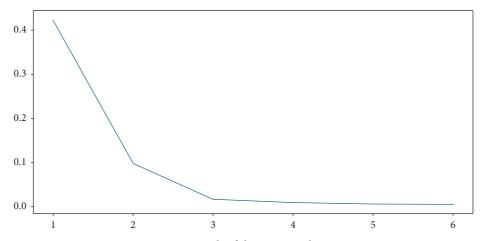
Figure 8: Result of dimension reduction.

Table 5: Errors after dimension reduction of LLE.

Neighbors	Error
5	1.07×10^{-18}
10	-9.01×10^{-19}
20	-1.84×10^{-18}
30	6.13×10^{-18}
40	3.15×10^{-17}

models are better than the LSTM models and BP neural network models. Except for the PCA-LSTM model, other LSTM model errors are less than the BP neural network models. Therefore, the factors that affect the price of crude oil include the impact of long-term factors and the impact of short-term factors. RNN and its variant models are better than the LSTM model; and the effect of the RNN model will lose far memory. Therefore, while the long-term and short-term factors work together in the formation of the price of crude oil, the data far from the current data have little influence on the results of the current. That is, with the passage of time, the impact of long-term factors on the price of crude oil gradually reduced.

The formation of crude oil prices is the result of all factors. Compared with traditional econometrics methods, machine learning methods performed better in crude oil price prediction. In the machine learning methods, the neural network can capture nonlinear characteristics of crude oil prices. In the neural network model, RNN and LSTM can better fit the long-term impact and short-term impact of each factor on the crude oil prices.

In order to verify the influence of different lag lengths on the prediction results and verify the robustness of the model, different window sizes were set in this paper. RNN benchmark models and LSTM benchmark models were employed to predict the crude oil future prices and spot prices, respectively. Models with 40 day (about 2 months), 60 day (about 3 months), and 80 day (about 4 months) window sizes were selected to predict the crude oil prices, and MAE and RMSE of each model were calculated to evaluate the model. The results are shown in Tables 9 and 10.

According to Table 9, under different window sizes, the LLE-RNN model and LLE-LSTM model performed better than other same benchmark models in the prediction of the international crude oil future prices. When the window size is 40, MAE and RMSE of the LLE-RNN model and LLE-LSTM model are smaller than other models with the same benchmark. In panel A, when the window size is 60, the LLE-RNN model takes the second place, and the RMSE only increased by 7.54%. In panel B, when the window size is 60, the LLE-RNN model takes the second place, and the RMSE increased in 0.0145. Therefore, the oil future prices have a lag period of one to two months. In panel A, PCA-RNN fitting is the best when the window size is 80. However, when the window size is 40, the prediction ability of each model is the best, and window size 60 takes the second place. However, the error in different windows is small. Therefore, the RNN benchmark models showed better robustness when the window length changes. In panel B, the RMSE of the LLE-LSTM model is relatively small, and each model shows inconsistent performance in different window sizes. However, compared with RNN benchmark models, the stability of LSTM benchmark models are relatively poor, especially in PCA-LSTM and MDS-LSTM models.

From Table 10, we can find that the model using the LLE dimension reduction method is also superior to the similar models in predicting the spot price of crude oil. In panel A, the LLE-RNN model is superior to the model with the window size of 60 and 80 in turn when the window length is 40, and the RMSE is reduced by 1.45% and 15.98%, respectively. The PCA-RNN model has the worst performance when the window size is 80. However, MDS-RNN and RNN models performed best when the window size is 80. Each RNN model is relatively stable. In panel B, the PCA-LSTM model showed the best performance when the window length is 80. The RMSE of MDS-LSTM at window size 60 is the smallest. Both PCA-LSTM and MDS-LSTM are not stable. The performance of the LLE-LSTM model and the single LSTM model is the best in the window length 40. The LLE-LSTM model is significantly superior to other models with the same benchmark, and under each window size, the difference between MAE and RMSE is small. When
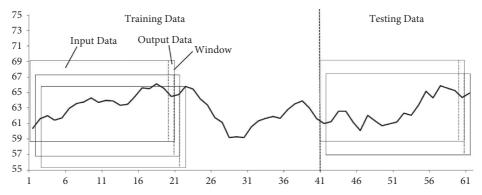
FIGURE 9: Sampling process of the moving window.

TABLE 6: Models.

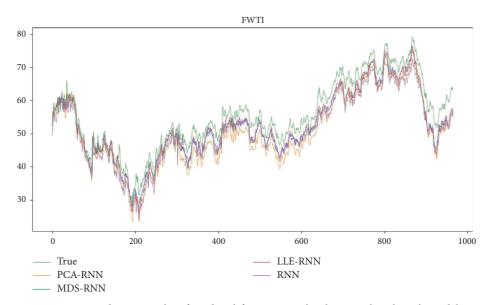|	PCA	MDS	LLE	—
RNN	PCA-RNN	MDS-RNN	LLE-RNN	RNN
LSTM	PCA-LSTM	MDS-LSTM	LLE-LSTM	RNN

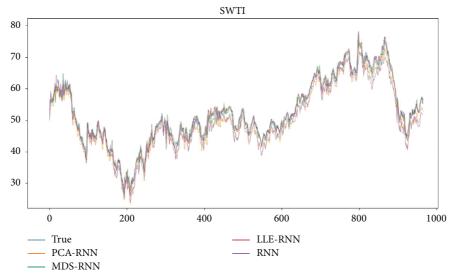FIGURE 10: Prediction results of crude oil future prices by the RNN benchmark model.

FIGURE 11: Prediction results of crude oil spot prices by the RNN benchmark model.
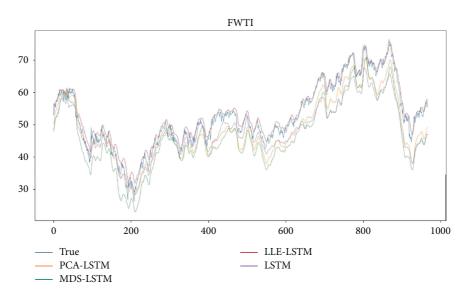
FIGURE 12: Prediction results of crude oil future prices by the LSTM benchmark model.
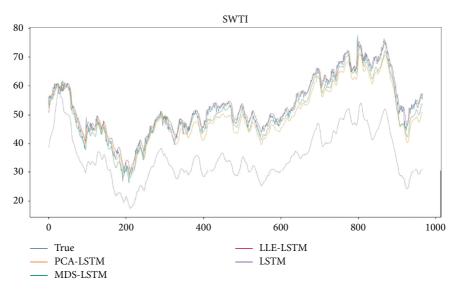

FIGURE 13: Prediction results of crude oil spot prices by the LSTM benchmark model.

TABLE 7: Errors of benchmark model.

	FWTI		SWTI	
	MAE	RMSE	MAE	RMSE
RNN	0.0844	0.0982	0.0845	0.0969
LSTM	0.0905	0.1118	0.0956	0.1157
BP	0.1247	0.1509	0.1499	0.1734
SVM	0.2784	0.3496	0.2807	0.3529
ARIMA	6.9717	8.3986	8.0444	6.6726

the window length is 40, the LLE-LSTM model's RMSE decreases by 10.9% and 22.51%, respectively, when the window size is 60 and 80, and better stability. When the window size of the single LSTM model is 40, RMSE decreases by 21.52% and 37.68%, compared with 60 and 80, respectively.

Table 8: Errors of benchmark and variant model.

	FWTI		SWTI	
	MAE	RMSE	MAE	RMSE
Panel A				
PCA-RNN	0.0595	0.0703	0.0600	0.0723
MDS-RNN	0.0494	0.0626	0.0649	0.0794
LLE-RNN	**0.0344**	**0.0442**	**0.0323**	**0.0419**
RNN	0.0844	0.0982	0.0845	0.0969
Panel B				
PCA-LSTM	0.1428	0.1649	0.1269	0.1461
MDS-LSTM	0.0798	0.1016	0.0929	0.1126
LLE-LSTM	**0.0494**	**0.0636**	**0.0448**	**0.0587**
LSTM	0.0905	0.1118	0.0956	0.1157
Panel C				
PCA-BP	0.0830	0.1028	0.0694	0.0864
MDS-BP	0.1438	0.1858	0.1008	0.1256
LLE-BP	**0.0495**	**0.0621**	**0.0495**	**0.0622**
BP	0.1247	0.1509	0.1499	0.1734
Panel D				
PCA-SVM	0.3071	0.3634	0.2917	0.3578
MDS-SVM	0.3010	0.3602	0.2958	0.3686
LLE-SVM	0.2719	0.3600	0.2755	0.3679
SVM	0.2784	0.3496	0.2807	0.3529

Table 9: Error with different window sizes (future prices).

	40		60		80	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Panel A						
PCA-RNN	0.0471	0.0573	0.0595	0.0703	0.0385	0.0497
MDS-RNN	0.0736	0.0871	0.0494	0.0626	0.1385	0.1529
LLE-RNN	**0.0319**	**0.0411**	**0.0344**	**0.0442**	**0.0337**	**0.0443**
RNN	0.0789	0.0940	0.0844	0.0982	0.1918	0.2217
Panel B						
PCA-LSTM	0.1117	0.1332	0.1428	0.1649	0.0643	0.0798
MDS-LSTM	0.1234	0.1488	0.0798	0.1016	0.1495	0.1747
LLE-LSTM	**0.0378**	**0.0491**	**0.0494**	**0.0636**	**0.0584**	**0.0749**
LSTM	0.1336	0.1500	0.0905	0.1118	0.0715	0.0903

Table 10: Error with different window sizes (spot prices).

	40		60		80	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
Panel A						
PCA-RNN	0.0499	0.0610	0.0600	0.0723	0.1582	0.1741
MDS-RNN	0.0835	0.0971	0.0649	0.0794	0.0486	0.0619
LLE-RNN	**0.0314**	**0.0413**	**0.0323**	**0.0419**	**0.0362**	**0.0479**
RNN	0.0762	0.0879	0.0845	0.0969	0.0439	0.0558
Panel B						
PCA-LSTM	0.0994	0.1217	0.1269	0.1461	0.0621	0.0769
MDS-LSTM	0.2225	0.2579	0.0929	0.1126	0.1660	0.2025
LLE-LSTM	**0.0399**	**0.0523**	**0.0448**	**0.0587**	**0.0524**	**0.0675**
LSTM	0.0740	0.0908	0.0956	0.1157	0.1217	0.1457

5. Conclusion

The formation process of crude oil prices includes factors of horizontal disturbance and the factors of vertical impact. Therefore, this paper analyses the possible factors that affect the international crude oil prices through the dual attributes of crude oil and divides the influencing factors into commodity attributes causing the long-term trend and financial attributes causing the short-term fluctuation of oil prices. Considering the inconsistency of the impact, which various factors have on the crude oil prices, this paper selects RNN and LSTM models to predict the crude oil future prices and spot prices. At the same time, considering the complexity of influencing factors, PCA, MDS, and LLE were selected in this paper to reduce the dimension of data, respectively, so as to improve the generalization ability of the model. Based on the idea of "dimension reduction prediction," this paper constructs eight models to predict the crude oil prices.

Crude oil prices were predicted by RNN benchmark models, LSTM benchmark models, BP benchmark models, SVM benchmark models, and ARIMA model, respectively. According to MAE and RMSE comparison, the following conclusions are drawn. First, the formation process of crude oil prices is the result of the joint action of various factors, and the duration of influence of various factors on the crude oil prices is inconsistent, resulting in a nonlinear process of crude oil prices. Second, compared with the single RNN model or the single LSTM model, the prediction result of dimension-reduction-RNN/LSTM models has the best generalization ability and can fit the real crude oil prices more effectively. Third, because the LLE dimension reduction method has a good performance in capturing the nonlinear correlation between various factors, the performance of LLE-RNN/LSTM models is better than other models with the same benchmark.

The window size of 40 (2 months), 60 (3 months), and 80 (4 months) was set, respectively, and the MAE and RMSE of each model were calculated to analyse the influence of the prediction sequence length on crude oil price prediction and verify the robustness of each model, and the following conclusions were drawn. First, the LLE-RNN/LSTM model has good stability in different window sizes and in the prediction of crude oil future prices and spot prices. Second, the model with PCA and MDS is less robust in crude oil price prediction.

Through the above analysis, compared with other models, RNN and LSTM models can capture the nonlinear factors in the formation process of crude oil prices and fit the formation process of crude oil prices. The models with LLE performed well in crude oil price prediction. Based on a de-dimension machine learning model approach to forecast the international crude oil prices by taking account of the dual attributes of oil, this paper realized the prediction of crude oil prices and provided alternative tools to predict the future trend of crude oil prices. Based on the factors in the formation process of crude oil, it is helpful to distinguish the long-term and short-term factors that affect crude oil prices so as to seize the opportunity and avoid the risks brought by the fluctuation of crude oil prices. With the development of deep learning methods, there is still a lot of room for development in the prediction of time series. Considering the variation of the relevant LSTM model, it is the next research direction to improve the accuracy of the prediction model of crude oil prices.

Data Availability

The data used in this article comes from the Wind database. The Excel data used to support the findings of this study are included within the supplementary information file named "data.xlsx."

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Supplementary Materials

(1) Data source. The data were taken from China Wind Economic Database (https://www.wind.com.cn/en/edb.html). The Wind Economic Database pairs over 1.3 million macroeconomic and industry time series with powerful graphics and data analysis tools to give financial professionals the most comprehensive insights into China's economy. (2) The meaning of each column in the table. "DATE" column represents the date; "FWTI" represents West Texas crude oil future price; "SWTI" represents spot price; "PRO" represents the total supply of oil; "COM" represents total demand for oil; "GAS" represents natural gas prices; "RDL" represents dollar exchange rate; "RUB" represents rouble exchange rate; "GOLD" represents price of gold; and "SIL" represents the price of silver. (*Supplementary Materials*)

References

[1] S. Degiannakis, G. Filis, and V. Arora, "Oil prices and stock markets: a review of the theory and empirical evidence," *Energy Journal*, vol. 39, no. 5, 2018.

[2] K. Lang and B. R. Auer, "The economic and financial properties of crude oil: a review," *The North American Journal of Economics and Finance*, vol. 52, Article ID 100914, 2020.

[3] L. Yanqiong, L. Zhenghui, Y. Yanyan, and H. Dong, "Asymmetry of risk evolution in crude oil market: from the perspective of dual attributes of oil," *Energies*, vol. 4063, no. 13, pp. 2–22, 2021.

[4] S. Abosedra and H. Baghestani, "On the predictive accuracy of crude oil futures prices," *Energy Policy*, vol. 32, no. 12, pp. 1389–1393, 2004.

[5] G. Chevillon and C. Rifflart, "Physical market determinants of the price of crude oil and the market premium," *Energy Economics*, vol. 31, no. 4, pp. 537–549, 2009.

[6] G. K. Abledu and K. Agbodah, "Stochastic forecasting and modeling of volatility of oil prices in Ghana using ARIMA time series model," *European Journal of Business and Management*, vol. 4, no. 16, pp. 122–131, 2012.

[7] W. Kristjanpoller and E. Hernández, "Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors," *Expert Systems with Applications*, vol. 84, pp. 290–300, 2017.

[8] L. Yu, X. Zhang, and S. Wang, "Assessing potentiality of support vector machine method in crude oil price forecasting," *EURASIA Journal of Mathematics, Science and Technology Education*, vol. 13, no. 12, pp. 7893–7904, 2017.

[9] N. Bashiri Behmiri and J. R. Pires Manso, "Crude oil price forecasting techniques: a comprehensive review of literature," *SSRN*, Article ID 2275428, 2013.

[10] S. Moshiri and F. Foroutan, "Forecasting nonlinear crude oil futures prices," *Energy Journal*, vol. 27, no. 4, pp. 81–95, 2006.

[11] I. SenGupta, W. Nganje, and E. Hanson, "Refinements of Barndorff-Nielsen and Shephard model: an analysis of crude oil price with machine learning," *Annals of Data Science*, vol. 8, no. 1, pp. 39–55, 2021.

[12] J. Wang, H. Zhou, T. Hong, X. Li, and S. Wang, "A multi-granularity heterogeneous combination approach to crude oil price forecasting," *Energy Economics*, vol. 91, Article ID 104790, 2020.

[13] W. Kristjanpoller and M. C. Minutolo, "Forecasting volatility of oil price using an artificial neural network-GARCH model," *Expert Systems with Applications*, vol. 65, pp. 233–241, 2016.

[14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[15] A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013.

[16] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modelling," in *Proceedings of the 15th Annual Conference of the International Speech Communication Association*, Singapore, 2014.

[17] H. Y. Kim and C. H. Won, "Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models," *Expert Systems with Applications*, vol. 103, pp. 25–37, 2018.

[18] D. E. Rapach and G. Zhou, "Time-series and cross-sectional stock return forecasting: new machine learning methods," *Machine Learning for Asset Management: New Developments and Financial Applications*, Wiley, Hoboken, NJ, USA, pp. 1–33, 2020.

[19] Z. Cen and J. Wang, "Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer," *Energy*, vol. 169, pp. 160–171, 2019.

[20] 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.

[21] A. Gallo, P. Mason, S. Shapiro, and M. Fabritius, "What is behind the increase in oil prices? Analyzing oil consumption and supply relationship with oil price," *Energy*, vol. 35, no. 10, pp. 4126–4141, 2010.

[22] A. Breitenfellner, J. C. Cuaresma, and C. Keppel, "Determinants of crude oil prices: supply, demand, cartel or speculation," *Monetary Policy & the Economy*, vol. 4, no. 4, pp. 111–136, 2009.

[23] W. Mensi, S. Hammoudeh, and S. H. Kang, "Precious metals, cereal, oil and stock market linkages and portfolio risk management: evidence from Saudi Arabia," *Economic Modelling*, vol. 51, pp. 340–358, 2015.

[24] M. E. Bildirici and C. Turkmen, "Nonlinear causality between oil and precious metals," *Resources Policy*, vol. 46, pp. 202–211, 2015.

[25] Y.-J. Zhang and Y.-F. Sun, "The dynamic volatility spillover between European carbon trading market and fossil energy

market," *Journal of Cleaner Production*, vol. 112, pp. 2654–2663, 2016.

[26] M. Brigida, "The switching relationship between natural gas and crude oil prices," *Energy Economics*, vol. 43, pp. 48–55, 2014.

[27] R. A. Lizardo and A. V. Mollick, "Oil price fluctuations and U.S. dollar exchange rates," *Energy Economics*, vol. 32, no. 2, pp. 399–408, 2010.

[28] M. I. Turhan, A. Sensoy, and E. Hacihasanoglu, "A comparative analysis of the dynamic relationship between oil prices and exchange rates," *Journal of International Financial Markets, Institutions and Money*, vol. 32, pp. 397–414, 2014.

[29] L. Kilian and R. J. Vigfusson, "Nonlinearities in the oil price-output relationship," *Macroeconomic Dynamics*, vol. 15, no. S3, pp. 337–363, 2011.

[30] H. Chen, L. Liu, Y. Wang, and Y. Zhu, "Oil price shocks and U.S. dollar exchange rates," *Energy*, vol. 112, pp. 1036–1048, 2016.

[31] J. Bouoiyour, R. Selmi, A. K. Tiwari, and M. Shahbaz, "The nexus between oil price and Russia's real exchange rate: better paths via unconditional vs. conditional analysis," *Energy Economics*, vol. 51, pp. 54–66, 2015.

[32] T. K. Blokhina, O. A. Karpenko, and A. V. Guirinskiy, "The relationship between oil prices and exchange rate in Russia," *International Journal of Energy Economics and Policy*, vol. 6, no. 4, pp. 721–726, 2016.

[33] M. Shahbaz, M. Balcilar, and Z. Abidin Ozdemir, "Does oil predict gold? A nonparametric causality-in-quantiles approach," *Resources Policy*, vol. 52, pp. 257–265, 2017.

[34] R. Khalfaoui, "Oil-gold time varying nexus: a time-frequency analysis," *Physica A: Statistical Mechanics and Its Applications*, vol. 503, pp. 86–104, 2018.

[35] Y.-J. Zhang and Y.-M. Wei, "The crude oil market and the gold market: evidence for cointegration, causality and price discovery," *Resources Policy*, vol. 35, no. 3, pp. 168–177, 2010.

[36] Y. S. Wang and Y. L. Chueh, "Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices," *Economic Modelling*, vol. 30, pp. 792–798, 2013.

[37] J. L. Elman, "Finding structure in time," *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.

[38] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.

[39] H. Y. Li, W. Xu, Y. Cui, Z. Wang, M. Xiao, and Z. X. Sun, "Preventive maintenance decision model of urban transportation system equipment based on multi-control units," *IEEE Access*, vol. 8, pp. 15851–15869, 2019.

[40] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, no. 1, pp. 1–27, 1964.