

A Study of Crude Oil Price Forecasting Based on Long Short-Term Memory Model

Shihui Wang^{1,a,*}

¹*School of Economics, Henan University, Zhengzhou, 450046, China*

a. sjerry@ldy.edu.rs

**corresponding author*

Abstract: In industrial production, crude oil prices play a pivotal role, influencing economic stability due to their propensity to induce fluctuations. Predicting these prices accurately is thus a crucial task in economics. This study addresses this challenge by employing a Long Short-Term Memory (LSTM) model, trained on data spanning January 2005 to January 2019, to forecast crude oil prices. Compared against expectations from economists, financial markets, and policymakers, the LSTM model demonstrates robust fitting and reliable predictive capability across various time frames. Notably, it outperforms alternative models in terms of accuracy when forecasting crude oil prices over extended periods. However, the model's accuracy diminishes with shorter forecasting intervals, suggesting the need for supplementary variables in shorter-term predictions. Therefore, while the LSTM model proves effective for longer-term forecasts, enhancing its precision for shorter intervals may necessitate integrating additional predictive factors. This research underscores the LSTM model's potential as a valuable tool in navigating the complexities of crude oil price forecasting, contributing insights crucial for informed economic decision-making.

Keywords: Crude oil price forecasting, LSTM model, Artificial neural networks.

1. Introduction

Since the mid-20th century, crude oil has emerged as a pivotal gauge of economic activity, owing to its pivotal role in the global energy landscape. Yet, as alternative sources like wind, hydro, and solar power have gained prominence, crude oil's stature as a primary energy provider has steadily diminished. Nonetheless, oil continues to wield significant influence in the economy, extending its impact across every facet of societal existence [1].

As a result, economists generally agree that there is a strong relationship between a country's economic growth rate and oil price fluctuations. How to understand and respond to this relationship has become a sharp and important issue today. Forty years after the 1973-1974 oil crisis, the global crude oil market is operating under a new regime that allows oil prices to fluctuate freely to a greater extent in response to changes in supply and demand forces [2, 3]. During the crisis, the price of imported oil surged nearly fourfold in a single quarter. The available literature generally agrees that oil price fluctuations have a significant impact on economic activity, and that their effects vary between oil-importing and exporting countries. Higher oil prices are good for oil exporters and bad for oil importers; Falling oil prices have the opposite effect. The transmission mechanism of oil price fluctuations to real economic activities is mainly manifested through two channels: supply and

demand. On the supply side, crude oil is the basic input for production, and rising oil prices lead to rising production costs, which in turn inhibits corporate output. The demand-side effect is expressed in the impact on consumption and investment, which varies with the change in disposable income. In addition, the expectation of the duration of the oil price shock also affects the intensity of its impact on the economy. With the deterioration of global environmental quality and the intensification of climate change, the international community generally agrees on the need to coordinate economic development and environmental protection policies. In the past, the development model has tended to "develop the economy first, then solve the environmental problems", and seek a trade-off between economic and environmental interests. However, more and more countries are turning to the development of new energy sources, although petroleum energy still plays an irreplaceable role. Based on oil price data from 1987 to 2019, this article attempts to make predictions about future oil prices and explore their potential impact on the global economy.

2. Literature Review

Many studies have shown that oil prices often experience unforeseen fluctuations. The magnitude of these fluctuations indicates whether they meet expectations. Research shows that forecasting oil prices can be viewed from the perspectives of economists, policymakers, financial market participants, and consumers [4]. Economists strive to comprehend oil price dynamics using vector autoregressive (VAR) models grounded in economic theory. In the empirical study, it employs the traditional VAR model, which includes real oil prices, global crude oil production, global real economic activity, and changes in global crude oil inventories [5]. Policymakers tend to believe that financial market participants have more information than quantitative economic models. However, financial market participants might ignore, misinterpret, or overlook crucial information about model predictions, especially when obtaining this information is costly [6]. A primary source of market expectations for oil prices is the price of oil futures contracts. Futures contracts allow traders to lock in today's price for buying a set amount of crude oil at a future date [7]. Financial market expectations for oil prices are based on the risk premium estimate from Hamilton and Wu's (2014) term structure model for term h minus the WTI futures price for the current expiration h [8]. Consumer oil price expectation theory focuses on the expectations of businesses or households. This approach extrapolates consumer expectations for the nominal price of oil based on current oil prices and inflation forecasts. None of these prediction methods are effective during an oil price shock. An unexpected change in oil prices is known as an oil price shock [9]. By comparing oil price expectations with actual outcomes, it can assess the magnitude of the oil price shock. The occurrence and extent of an oil price shock depend on how oil price expectations are measured. Therefore, using LSTM models in deep learning for these prediction challenges might yield better results [10].

3. Methodology

3.1. Data and Model

LSTM is a deep learning model derived from the Recurrent Neural Network (RNN). LSTM has played a significant role in the advancement of deep learning. One of the primary benefits of RNNs is their ability to utilize previous information to inform current tasks, meaning they can effectively retain and apply past data to compute the current output for time series analysis and prediction. LSTM addresses the issues of gradient vanishing and gradient explosion that RNNs face. Through its unique cell structure, which includes cell states and three gates (input gate, forget gate, and output gate), LSTM can establish long-term dependencies and accurately predict crude oil prices. LSTM not only fully extracts historical data on crude oil prices but also makes predictions based on current data characteristics, demonstrating its robust short-term and long-term memory capabilities while

avoiding information loss. The model excels in identifying long-term dependencies in crude oil price data and can automatically uncover nonlinear characteristics and complex patterns, leading to superior performance in price prediction. LSTM is a powerful forecasting tool widely applied in various related fields. This study selected the LSTM model for its precision in predicting crude oil prices. The model's cellular structure facilitates selective memory storage and information transmission through the cell state and control gates, with the gating mechanism effectively managing information flow.

This article organizes and reindexes the prices for each date, ensuring that data from January 1, 2005, onwards are included. This results in a time series dataset of crude oil prices, with missing values appropriately handled (e.g. Figure 1 and Figure 2). To enhance the model's performance and stability, the data were normalized to ensure uniform scales and ranges. The dataset was then split into a training set (70% of the total dataset) and a test set (30%). The model's prediction performance on both sets was assessed, with training loss and validation loss visualized to provide a comprehensive understanding of the model's learning process and effectiveness. The mean absolute error (MAE) and root mean square error (RMSE) were calculated on the training and test sets to evaluate the LSTM model's precision and accuracy in forecasting crude oil prices.

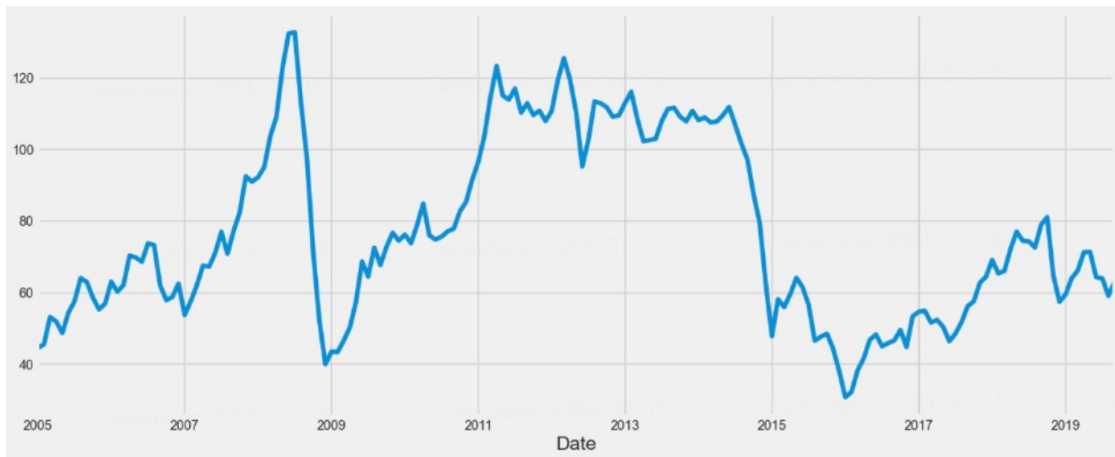


Figure 1: Resample by month and calculation the average for each month
(Photo/Picture credit: Original)

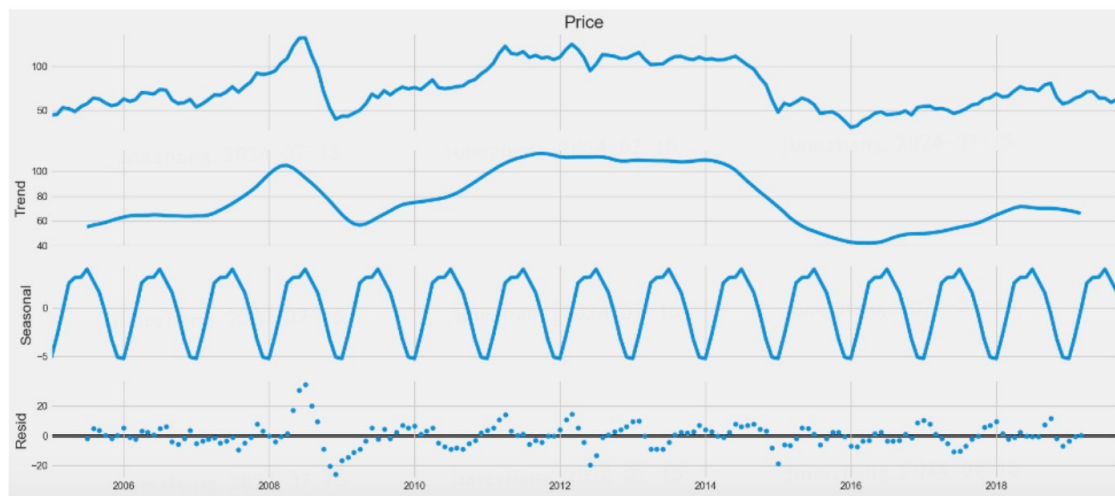


Figure 2: The decomposed trend, seasonality, and residual components
(Photo/Picture credit: Original)

3.2. Evaluation Indicators

In the context of forecasting crude oil prices, various models are available, necessitating the selection of the optimal model based on effective evaluation metrics. To assess prediction accuracy and stability, this study primarily utilizes Mean Squared Error (MSE) and Mean Absolute Error (MAE). Specifically:

MAE quantifies the average of the absolute discrepancies between the predicted and actual values. Smaller MAE values suggest that the model's predictions are closer to the actual values, indicating fewer errors.

RMSE is calculated as the square root of the mean of the squared prediction errors, reflecting the variance of these errors. Lower RMSE values signify that the model's predictions have less deviation from the actual values, indicating better accuracy.

These metrics serve to quantify and compare the predictive performance and reliability of the models under consideration in this research.

4. Results

4.1. Fitting Effect Analysis of LSTM Model

Based on the results of MAE and RMSE, lower MAE values on both the training and test sets indicate a better model performance, as they reflect smaller prediction errors relative to the actual values. Similarly, lower RMSE values suggest that the model's predictions deviate less from the true values. The model's effectiveness is illustrated in Figure 3.

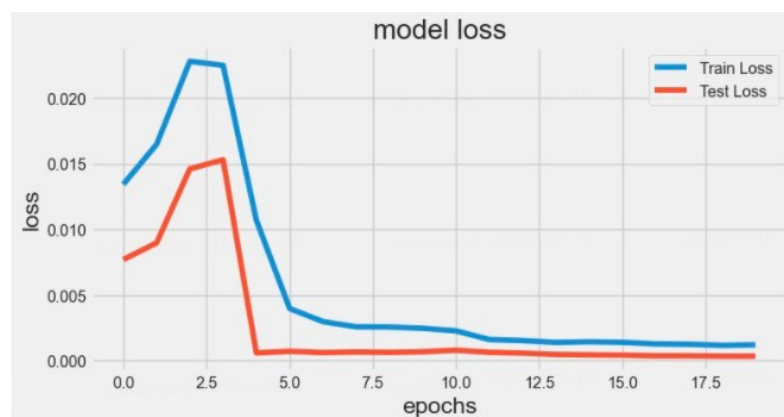


Figure 3: The fitting effect of the LSTM model (Photo/Picture credit: Original)

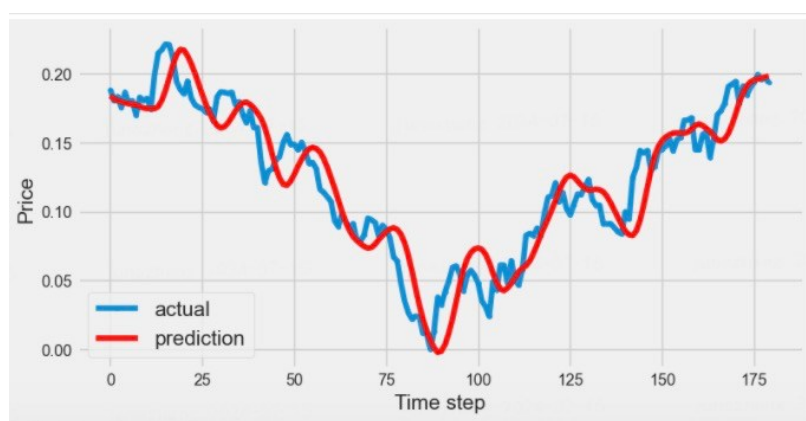


Figure 4: The comparison of actual and predicted values (Photo/Picture credit: Original)

As the epochs increase, the fitting effect of the LSTM model gradually improves. As illustrated in Figure 4, the predicted values mirror the trends observed in the actual data. When the actual prices decline, the predicted prices also decrease, and conversely, when actual prices rise, the predicted values show corresponding increases. Moreover, it is crucial to analyze the disparity between predicted and actual values. Among the three models assessed, the LSTM model distinguishes itself. Particularly noteworthy is its ability to closely approximate actual values across most scenarios, barring outliers. In terms of forecasting accuracy, both the LSTM and ANN models demonstrate proficiency in identifying critical inflection points in crude oil prices, effectively capturing fluctuations with precision.

5. Conclusion

This study utilizes oil price data spanning from January 2005 to December 2019 and employs the LSTM model to forecast the prices of two types of oil. It also considers the viewpoints of economists, policymakers, financial market participants, and consumers to assess the LSTM model's prediction accuracy and stability. The empirical results lead to several conclusions. Firstly, the LSTM model demonstrates strong generalization capabilities and maintains a stable fitting performance for crude oil prices across different time scales. Whether the focus is on short-term or long-term predictions, and whether for WTI or Brent crude oil prices, the LSTM model's fitting effectiveness remains unaffected. Notably, the model performs exceptionally well in the medium term. Secondly, compared to models based on different perspectives, the LSTM model exhibits higher prediction accuracy and greater stability across various time scales. It avoids errors arising from differing perspectives and factors.

The LSTM model's strong predictive capabilities provide significant value to researchers investigating the dynamics of the crude oil market. Results suggest that the LSTM model effectively manages the nonlinear and erratic patterns inherent in oil price movements. Moreover, the accurate forecasts generated by the LSTM model can enhance the understanding of the mechanisms driving fluctuations in crude oil prices, influencing decisions related to production, operations, and investments. This, in turn, supports domestic economic growth and security. While there may exist more advanced techniques for predicting crude oil prices, implementing further enhancements is restricted by limitations in resources and specialized knowledge.

References

- [1] C. Baumeister, L. Kilian, "Forty years of oil price fluctuations: Why the price of oil may still surprise us," *J. Econ. Perspect.*, vol. 30, no. 1, pp. 139-160 (2016).
- [2] K. Zhang, M. Hong, "Forecasting crude oil price using LSTM neural networks," *Volume 2, Issue 3*, pp. 163-180 (2016).
- [3] A. Adelman, "The real oil problem," *Regulation*, vol. 27, no. 1, pp. 16-21 (2004).
- [4] A. K. Patidar, P. Jain, P. Dhasmana, T. Choudhury, "Impact of global events on crude oil economy: A comprehensive review of the geopolitics of energy and economic polarization," *GeoJournal*, vol. 89, no. 2 (2024).
- [5] H. Abdollahi, "A novel hybrid model for forecasting crude oil price based on time series decomposition," *Applied Energy*, vol. 267, 115035 (2020).
- [6] S. Butler, P. Kokoszka, H. Miao, et al., "Neural network prediction of crude oil futures using B-Splines," *Energy Economics*, vol. 94, 105080 (2021).
- [7] T. Fischer, C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654-669 (2018).
- [8] J. Hamilton, J. C. Wu, "Risk premia in crude oil futures prices," *National Bureau of Economic Research* (2013).
- [9] M. B. Karaki, A. Neaimeh, "Do higher global oil and wheat prices matter for the wheat flour price in Lebanon?," *Agricultural Economics* (2024).
- [10] J. E. Hokayem, I. Jamali, A. Hejase, "A forecasting model for oil prices using a large set of economic indicators," *Journal of Forecasting*, vol. 43, no. 5, pp. 1615-1624 (2024).