Quantitative Finance and Economics

http://www.aimspress.com/journal/QFE

QFE, 6(2): 342–358.
DOI: 10.3934/QFE.2022015
Received: 13 May 2022
Revised: 13 June 2022
Accepted: 15 June 2022
Published: 20 June 2022

Research article

Fourier transform based LSTM stock prediction model under oil shocks

Xiaohang Ren[1], **Weixi Xu**[1] **and Kun Duan**[2,*]

[1] School of Business, Central South University, Changsha, 410083, China
[2] School of Economics, Huazhong University of Science and Technology, Wuhan, 430074, China

* **Correspondence:** Email: kunduan@hust.edu.cn.

Abstract: This paper analyses the impact of various oil shocks on the stock volatility prediction by using a Fourier transform-based Long Short-Term Memory (LSTM) model. Oil shocks are decomposed into five components following individual oil price change indicators. By employing a daily dataset involving S&P 500 stock index and WTI oil futures contract, our results show that different oil shocks exert varied impacts on the dynamics of stock price volatility by using gradient descent. Having exploited the role of oil shocks, we further find that the Fourier transform-based LSTM technique improves forecasting accuracy of the stock volatility dynamics from both statistical and economic perspectives. Additional analyses reassure the robustness of our findings. Clear comprehension of the future stock market dynamics possesses important implications for sensible financial risk management.

Keywords: Fourier transform; LSTM; stock prediction; oil shock; stock volatility

JEL Codes: G01, G10, G12, G15

1. Introduction

Crude oil, being termed as "industrial blood", has long been not only an strategic material for operations of the global economy (Aastveit, 2014), but also a key source that fluctuates the dynamics of the financial system (Georgellis, 1994). Since the beginning of the 21st century, the oil market has featured evident characteristics that closely links to industrial production while exerting an important influence on the financial sector, leading to oil-related assets featured with increasingly-evident financial attributes (Ren et al., 2022a,b). Recently, changes in financial market trends are no longer the only important factors affecting stock prices, while globally important commodities like oil also manifest marked influences on fluctuating the stock market dynamics (Ahmed and Huo, 2021; Hu and Ying, 2017). In parallel, along with changes in the demand and supply of oil, the volatility of stock prices varies accordingly (Chiou and Lee, 2009), which would then have an enormous impact on the economic and financial system worldwide and even the way of life of mankind (Zhang, 2017; Zhao et al., 2016). Such changes in the oil market could result in oil shocks, which would then further affect stock market volatility.

It is known that oil can affect stock price volatility which makes the risk of market rise (An et al., 2018). Therefore, it is important to account for the role of oil shocks in driving the stock market dynamics and the associated systemic risk. While existing literature examines the impact of a single oil price shock on the stock market (Shahrestani and Rafei, 2020), but they are questioned by the fact that only one component of the oil price dynamics is considered, leading to the overfitting of the corresponding results. At the same time, there exists related literature that employs various linear models for the prediction analysis such as the quantile approach (Xu et al., 2019), and GARCH-type models (Cunado and de Gracia, 2014). However, the linear framework is often questioned by limited predictive power and a relatively large variance of asymptotic unbiasedness in a large sample (Duan et al., 2021; Barsky and Kilian, 2001).

To this end, we follow the literature Lu et al. (2021) and fill the gap by making a comprehensive decomposition on various types of oil shocks that contains NPI, ANP, SNP, LPI, NPI2, respectively. By using an international daily dataset ranging from 1 January 1989 to 31 January 2020, the prediction of the stock market dynamics is then investigated under a Fourier transform-based LSTM framework, while considering the aforementioned five types of oil shocks. As for the oil shock decomposition, it is worth noting that we extend the literature by testing significance of the five oil shocks' coefficients through gradient descent and thus capture the main elements. Then we apply predictability metrics and conduct a series of robustness tests to examine the performance of different models for oil shocks over a predictable forecast horizon.

Several important findings emerge from our analysis. First, the five oil shocks are found to have either positive effects (e.g., NPI, LPI, SNP and NPI2) or negative effects (i.e., ANP) on the stock market volatility during the in-sample prediction. Accordingly, estimated coefficients of the contemporaneous impact of the five oil shocks on the stock volatility dynamics are 0.0024, 0.0018, 0.0153, 0.0420, −0.0005. In terms of effectiveness of the predictive role of oil shocks, ANP, SNP and NPI2 are shown to be more efficient in learning stock market volatility, while NPI and LPI entail relatively larger training rounds with relatively insignificant predictive power. Moreover, the Fourier transform-based LSTM model is found to be superior to other competing models, indicating that the Fourier transform helps enhance the stock volatility prediction. Robustness of our results is further checked through a battery of additional analyses including the alternative evaluation method, and different forecasting horizons, etc. Our results are important and possess practical implications for various stakeholders.

This paper is organized as follows. We conduct a literature review in section 2. Several oil shock measures are used in section 3 ,along with some mathematical provement and efficiency of comparative models. Section 4 shows the evaluation results of our empirical analysis and conducts related tests. Section 5 makes a conclusion.

2. Literature review

The mystery of international crude oil price volatility is one of the tantalizing problems that economists have been trying to unravel, and it is also one of the world's most difficult problems (Dickey and Fuller, 1979). Since the oil crisis in 1974 , the issue of oil prices has received increasing attention from academics and government agencies at home and abroad (Haugom et al., 2014), and has now become the focus of attention in the world economy, politics and diplomacy. Since oil resources have become urgent and important for the development of the world economy (Liow et al. , 2021; Charfi and Mselmi, 2022), many scholars have developed various forecasting methods to predict oil prices. In 1931, Hotelling proposes the famous depletable resource model to study the mechanism of oil price formation. By assuming different market participants and oil market structures, they establishes various theoretical models to analyze and predict the causes and trends of international crude oil shocks. In addition, since international crude oil prices are time series reflecting the supply and demand information of the oil market (Ren et al., 2022c), scholars at home and abroad usually use time series methods to study and forecast international crude oil prices, such as Granger's causality research method (Rapach et al., 2010), Box and Jenkins' ARIMA (autoregressive integrated sliding average) model method (Wei, 2003), Engle's ARCH (autoregressive conditional (heteroskedasticity) model approach, the co-integration theory and error correction model proposed by Engle and Granger in 1987, etc.

At the same time, there exists an ongoing strand of research for the stock price volatility. In terms of stock market volatility characteristics, Fama (1970), in his seminal study of stock returns, finds that the nth period volatility of returns is affected by prior period volatility, showing initial volatility clustering and the presence of conditional heteroskedasticity effects. Chen and Zhu (2007) in their study of stock volatility clustering in different industries, show that in the presence of major unexpected events, all industries volatility shocks emerge, but the degree of volatility varies. Chen and Zhou (2002) conduct an empirical analysis of China's SSE Composite Index through a GARCH family model and find that, similar to stock markets in Europe and the United States, volatility clustering exists in China's stock market. Chen (2009) conducts a study based on the daily returns of the CSI 300 index through the ARCH family model and conclude that the daily return volatility of the CSI 300 index is not only characterized by clustering and persistence, but also shows a spike and a thick tail. LIU and WANG (2017) analyze the returns of SZSI component index and SSE Composite Index by ARMA-TGARCH-M model. They point out that both indices have the characteristics of volatility agglomeration and spike back tail, and there is a long duration of external shock volatility.

Regarding the stock market dynamics driven by oil shock, Duan and Wang (2016) focuses on emerging market countries and analyzes the relationship between stock prices and oil prices through a VAR-BEEK-GARCH model, and the study shows that international oil shock have a certain degree of directional effect on stock volatility, while oil prices are not significantly influenced by stock market fluctuations. Sun and Yang (2012) use a VAR model to examine the relationship between stock prices and oil prices using a dataset ranging from 2007 to 2011. He (2012) uses a VAR model to analyze the daily data of Brent crude oil spot price and SSE Composite Index during 2007–2011. The results show that although oil price shocks affect the SSE Composite Index, the effect is insignificant. He et al. (2020) analyzes the relationship between stock market and oil price volatility after the stock reform in China, using various GARCH-type models, and concludes that oil price volatilized has an inverse effect on the stock market in normal market conditions. Such the relationship is not evident under extremes conditions. He and Wang (2013) use a VAR model to investigate the relationship between WTI crude oil spot price, Daqing crude oil spot price, and SSE index by using a VAR model. The study shows that the SSE index return is not significantly affected by Daqing crude oil spot price movements, while the SSE index is positively impacted by WTI crude oil prices. Cheng and Chang (2016) employ a VAR model using weekly data from 2011–2015 to analyze whether the share prices of listed oil companies in China are affected by oil price shocks. The empirical test shows that oil price volatility only affects share prices of two subsidiaries of CNPC. It still remains a question that how oil shocks affect stock market since views are multiple, either positive or negative, and the forecasting methods that currently exist are somewhat deficient in terms of accuracy.

3. Method and data

3.1. *Monthly realized volatility*

Following previous researchers, we're inspired by some achievements (Christiansen et al., 2012; Ren et al., 2022c). Due to the various application of the S&P 500 index, we take it as an resource to build the data set. Moreover,It is rather easy to calculate the monthly realized volatility (RV in short) from daily returns. The way of computing this index can be expressed as below:

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2 \tag{1}$$

where M represents the index of samples in month t, $r_{t,j}$ acts as the j_{th} daily return of month t.

3.2. *Five oil shock measurements*

Inspired by Maheu et al. (2020) which establish oil shock measures based on quarter, we adopt five kinds of oil shocks as the input variable, which can be shown more deeply later.

1. Net price increase

According to Maheu et al. (2020), net price increases can be expressed as:

$$d_t^+ = 100max\{0, logO_t/O_t^*\} \tag{2}$$

where O_t shows what WTI oil futures' price is at t period. Additionally, $O_t^* = max\{O_{t-1}, \ldots, O_{t-36}\}$ is calculated on the past three years as research (Hamilton, 2011) depicts that data based on three-year has more contribution to the performance of predictions than that in one-year.

2. Asymmetric net price change

Considering both positive and negative shocks (Qiao et al., 2019; Sukharev, 2020), the asymmetric predictive power should be considered in prediction.The calculation of a positive shock is illustrated before. And a negative shock can be gained as follows:

$$d_t^- = 100min\{0, logO_t/O_t^{**}\} \tag{3}$$

where $O_t^{**} = min\{O_{t-1}, \ldots, O_{t-36}\}$ is calculated by the three-year data as it in NPI.

3. Symmetric net price change

Knowing that it's valid to put the constraint identically on the positive shocks along with negative shocks as a method to improve the performance of forecasting (Kilian and Vigfusson, 2013), we also adopt this kind of measurement:

$$d_t^* = d_t^- + d_t^+ \tag{4}$$

where d_t^+ indicates the net price increase, while the net price was represented as d_t^- decrease. The occasion that a shock of oil prices is from the highest to the lowest means the measurement will be zero since both kinds of shock are dealt with in the same way.

4. Large price increase

Based on the common expectation theory, an unexpected shock will make an enormous influence on the stock price. This index can be used the same as Kilian and Vigfusson (2013),which is:

$$d_t^{large} = r_t I(r_t > std\{r_{t-1}, ..., r_{t-36}\}) \tag{5}$$

where I() is an indicator function, which means only when the argument is true can its result be one while zero in other situations.

5. Net price increase indicator

The indicator's value is 0 or 1 to find out if there is an increase of oil price over O_t^*, which can be formulated as:

$$d_t^I = I(d_t^+ > 0) \tag{6}$$

where d_t^+ represents the net price increase. This indicator can lessen outliers' effect which might be shown the capability of capturing this kind of asymmetric shock.

3.3. The Revised Benchmark AR Model

The result of (Paye, 2012) shows that, the volatility of stock relates to actual economic circumstances a lot with a slight hysteresis. Moreover, current economic conditions are reflected in this phenomenon, meaning that lagged volatility includes necessary information (Wen et al., 2022; Xiao et al., 2019). Based on the information standard of AIC and SC (Barsky and Kilian, 2004), 1 lag of volatility is conducted to carry on the prediction. Following the concept, the autoregressive (AR) containing 1 lag is defined:

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \epsilon_t \tag{7}$$

where RV_{t-1} refers to 1 lag of RV, ϵ_t represents the stochastic component.

Adding five types of oil shocks to this basic model, we can understand the effects of them better, which can make comparasion with each other in forecasting performance:

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \beta_{NPI} NPI_t + \epsilon_t \tag{8}$$

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \beta_{ANP} ANP_t + \epsilon_t \tag{9}$$

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \beta_{SNP} SNP_t + \epsilon_t \tag{10}$$

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \beta_{LPI} LPI_t + \epsilon_t \tag{11}$$

$$Ln(RV_t) = \beta_0 + \beta_1 Ln(RV_{t-1}) + \beta_{NPI2} NPI2_t + \epsilon_t \tag{12}$$

Referring to the conclusion given by Zhang et al. (2019), combining various approaches contributes a lot to the forecasting process which will be better than individual model's average performance considering that actual market conditions keep altering, and there will be many outliers that are not easy to detect immediately. As a consequence, a multi-forecast model is conducted to largely lessen the uncertainty. This is achieved by aligning weighted averages to various models.

$$RV_{c,t} = \sum_{k=1}^{N} w_{k,t-1} RV_{k,t} \tag{13}$$

Assume that each weight is w_1, w_2, w_3, w_4, w_5.

$$\begin{aligned} Ln(RV_t) = \frac{1}{5}((w_1 + w_2 + w_3 + w_4 + w_5)\beta_0 + (w_1 + w_2 + w_3 + w_4 + w_5)\beta_1 Ln(RV_{t-1}) \\ + w_1\beta_{NPI} NPI_t + w_2\beta_{ANP} ANP_t + w_3\beta_{SNP} SNP_t + w_4\beta_{LPI} LPI_t \\ + w_5\beta_{NPI2} NPI2_t + (w_1 + w_2 + w_3 + w_4 + w_5)\epsilon_t) \end{aligned} \tag{14}$$

which can be redenoted as,

$$Ln(RV_t) = \beta'_0 + \beta'_1 Ln(RV_{t-1}) + \beta'_2 NPI_t + \beta'_3 ANP_t + \beta'_3 SNP_t + \beta'_4 LPI_t + \beta'_5 NPI2_t + \epsilon'_t \tag{15}$$

3.4. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a kind of temporal recurrent neural network, which is particularly conducted to fix the long-term dependence issue of general recurrent neural networks, and can achieve the prediction of periodically varying values (Clark and West, 2007). The gates in LSTM include the input gate, the forget gate and the output gate. The mentioned gates contribute to the model differently, acting as multiple functions together. Deciding the refreshment part of the input information is the function of the input gate, which can result of gaining the result iteratively. As its name, the forget gate aims to forget some kind of information, namely kept part of the information in the former memory cell, which helps to reduce the phenomenon of over fitting. At last, the output gate decides what kind of information should be treated as the final outcome. The three gates and formulas are as follows:

$$\begin{cases} f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \\ i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \\ \widetilde{c_t} = tanh(w_c[h_{t-1}, x_t] + b_c), \\ c_t = f_t c_{t-1} + i_t \widetilde{c_t} \\ o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \\ h_t = o_t tanh(c_t) \end{cases} \tag{16}$$

where h_{t-1} is the output of the previous layer, x_t is the current layer input, σ is the sigmoid activation function that maps variables between 0 and 1, w_f, w_i, w_c, w_o is the network weight, b is network bias, c_{t-1} is the memory unit of the previous moment, c_t is the memory unit of the current state, and h_t us the output of LSTM unit.

3.5. Fourier transform

Fourier transform, which represents the ability to indicate a certain function satisfying certain circumstances as a linear combination of trigonometric functions along with their integrals, by which the period of the change data can be found.

Suppose there is a polynomial,

$$f(x) = a_0 + a_1 x + ... + a_{n-2}x^{n-2} + a_{n-1}x^{n-1} \tag{17}$$

Divide the polynomial $f(x)$ into two parts according to the parity of the a subscript.

$$f(x) = (a_0 + a_2x^2 + a_4x^4 + ... + a_{n-2}x^{n-2}) + (a_1x + a_3x^3 + a_5x^5 + ... + a_{n-1}x^{n-1}) \tag{18}$$

Suppose $f_1(x), f_2(x)$,

$$f_1(x) = a_0 + a_2x^1 + a_4x^2 + ... + a_{n-2}x^{\frac{n}{2}-1} \tag{19}$$

$$f_2(x) = a_1 + a_3x + a_5x^2 + ... + a_{n-1}x^{\frac{n}{2}-1} \tag{20}$$

Then we gain,

$$f(x) = f_1(x^2) + xf_2(x^2) \tag{21}$$

And since,

$$w_{rn}^{rk} = cos\frac{2rk\pi}{rn} + isin\frac{2rk\pi}{rn} = w_n^k \tag{22}$$

$$w_n^{k+\frac{n}{2}} = w_n^k(cos\pi + isin\pi) = -w_n^k \tag{23}$$

$$\bar{w_n^k} = cos\frac{2k\pi}{n} - isin\frac{2k\pi}{n} = w_n^{n-k} \tag{24}$$

So when substituting $w_n^k(k<\frac{n}{2})$ into $f(x)$,

$$f(w_n^k) = f_1(w_n^{2k}) + w_n^k f_2(w_n^{2k}) = f_1(w_{\frac{n}{2}}^k) + w_n^k f_2(w_{\frac{n}{2}}^k) \tag{25}$$

And when substituting $w_n^{k+\frac{n}{2}}(k<\frac{n}{2})$ into $f(x)$,

$$f(w_n^{k+\frac{n}{2}}) = f_1(w_n^{2k+n}) + w_n^{k+\frac{n}{2}} f_2(w_n^{2k+n}) = f_1(w_{\frac{n}{2}}^k) - w_n^k f_2(w_{\frac{n}{2}}^k) \tag{26}$$

This translates into a recursive solution of the subproblem.

3.6. Gradient descent

Gradient descent turns out to be one of the most popular methods as a way to solve optimization questions with no constraints (Engle et al., 2013). By step to step refreshment strategy of changing the values of parameters according to the gradient direction of the object function, the results can be gained easily by this kind of method which shows good property since it works simply for the computer to conduct.

According to the bias between prediction and actual value, we can easily find that for regression function,

$$f(x) = \omega x + b \tag{27}$$

in which ω, b denote the vector of coefficient and intercept, its loss function is,

$$E_{\hat{\omega}} = \frac{1}{2n} \sum_{i=1}^{n} (\omega_i x_i + b_i - y_i)^2 \tag{28}$$

in which i denotes the i-th sample of the data set.

For the cost function $E_{\hat{\omega}}$, following the common denotion, we replace $E_{\hat{\omega}}$ with $J(\theta)$, and ω, b with θ_0, θ_1.

Actually, it is relatively easy to write in math form:

Repeat until convergence

$$\{\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)(for \quad j = 0 \quad and \quad j = 1)\} \tag{29}$$

In which α denotes the learning rate, and the parameters should be refreshed simultaneously. The reliance on the loss function to approximate the actual parameters is more sensitive, which makes it a good identification of oil shocks.

As for the multi-linear regression, the algorithm works the same except for additional individual parameters refreshment, which means each parameter in each iteration is updated simultaneously based on the gradient of the previous round, making the contemporaneous parameters without time series correlation.

Actually, different parameters needs different learning rate, RMS and RMSProp can be used to help fix this situation.In RMS condition,

$$\theta_i^{t+1} = \theta_i^t - \frac{\eta}{\sigma_i^t} g_i^t \quad \sigma_i^t = \sqrt{\frac{1}{t+1} \sum_{t=0}^{t} (g_i^t)^2} \tag{30}$$

While in RMSProp condition,

$$\theta_i^{t+1} = \theta_i^t - \frac{\eta}{\sigma_i^t} g_i^t \quad \sigma_i^t = \sqrt{\alpha(\sigma_i^{t-1})^2 + (1-\alpha)(g_i^t)^2} \tag{31}$$

Through the above learning rate adjustment, the recognition of oil shock will be more sensitive, and even solve some unrecognizable problems that cannot be solved by traditional joint cubic equations.In addition, this will also improve the prediction effect by filtering out the main factors influencing the prediction problem in the latter order.

3.7. Data

This paper studies the role of oil shocks on the stock volatility prediction where the data information of the crude oil and stock markets is respectively represented by WTI oil futures and S&P 500 index.

Table 1. Descriptive statistics.

Statistics	RVt	NPI	ANP	SNP	LPI	NPI2
Observations	383	383	383	383	383	383
Mean	2.699	0.674	−0.783	−0.109	0.02	0.117
Std. dev	0.982	2.388	4.579	5.266	0.055	0.322
Skewness	0.619	4.619	−8.512	−5.364	4.344	2.385
Kurtosis	0.778	25.302	84.597	51.531	27.963	3.708
Jarque-Bera	33.270***	11,291.273***	115,804.672***	43,078.301***	13,336.992***	572.226***
Q(5)	486.600**	98.615 *	72.812	87.343***	22.949***	168.659***
Q (22)	955.607***	115.23	77.157	92.702	31.160***	201.661***
ADF	−8.822 * **	−12.147 * **	−12.261 * *	−11.911 * **	−15.494*	−11.312 * **

Note: Descriptive statistical analysis shows the distribution of samples and the situation of each test index.

Our employed data are from Global Financial Data between 1st January, 1989 and 31st January, 2020. As depicted in Table 1, the statistical properties of the variables involved are described in detail ,in which ***, **and * denote rejections of the null hypothesis at the 1%, 5% and 10% levels, respectively.

The realization volatility of the S&P 500 Index is signed as RVt. Five oil shocks are measured by five measurements mentioned before. Q(n) represents the test that Ljung and Box (1978) proposed as the Ljung-Box statistic (Choi et al., 2018). In Table 1, it can be found that the kurtosis of variables above show high feature. The Ljung-Box test shows there are serial auto-correlations of 22nd order among the RV, LPI, and NPI2. Moreover, it can be obviously seen from the ADF test that sample data are stationary since no sign of unit root is observed at the 1 % significance level.

4. Empirical analysis

4.1. Regression by Gradient Descent

Using the RV of given data, we then put it into the model established as follow. By the equation we obtained before, which is,

$$Ln(\bar{R}V_t) = \beta_0' + \beta_1' Ln(RV_{t-1}) + \beta_2' NPI_t + \beta_3' ANP_t + \beta_3' SNP_t + \beta_4' LPI_t + \beta_5' NPI2_t + \epsilon_t' \tag{32}$$

The loss function of it is,

$$E_{\hat{\omega}} = \frac{1}{2t} \sum_{i=1}^{t} (ln(\bar{R}V_t) - y_t)^2 \tag{33}$$

in which y_t represents the true value of $ln(RV_t)$ while $ln(\bar{R}V_t)$ denotes the prediction value calculated in equation 32.

In solving for the minimum of the loss function, gradient descent can be conducted to solve iteratively to minimize the loss function and then the values of parameter can be refreshingly converged to the final result. After the previous model, the individual β coefficients can be obtained after 100000 cycles of stabilization.

The result is (0.6966, 0.7273, 0.0024, −0.0005, 0.0018, 0.0153, 0.0420), from which we could conclude that the first, third, fourth, fifth shock has positive effect to RVt while the second shock has negative effect to RVt. Since the regression coefficient of the previous period on the current period in this regression equation is as high as 0.727308431, which makes it the most main factor to prediction. To reduce the complexity of the model, the following forecasts use time series data to predict RV_t rather than forecast other index. The input items are RV_{t-1}, NPI,ANP,SNP,LPI,$NPI2$, and the output of them is RVt, corresponding to our prediction work.

4.2. Finding the Period by Fourier Transform

To predict the RVt, at first we look at the RVt changing by time.It is easy to see through the curves that there is a clear cyclical nature to volatility changes, but the random wandering behavior makes forecasting harder than regular indicator movements, which is caused by the many endogenous effects of oil shocks.

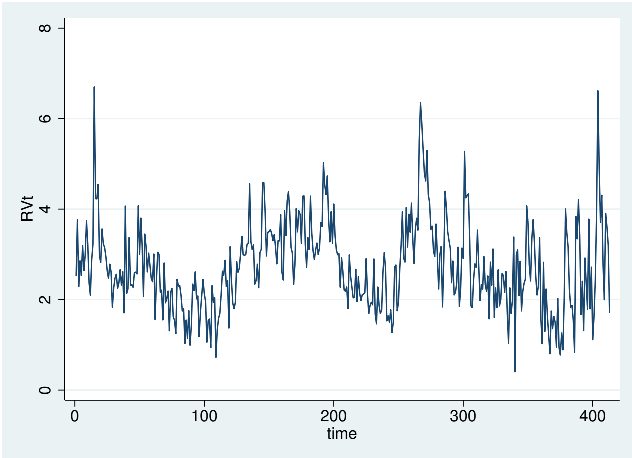

Figure 1. The trend of RVt

Note: This chart shows the trend of return volatility. It can be found that although there is a random walk phenomenon, it still has a certain periodicity.

Combining the previous equation for the calculation, we obtain its period as 68.7 months.As mentioned, studies usually take the relation between oil shock and the volatility stock as a liner relation while it actually isn't, which results in neglecting the slope heterogeneity of them according to specific occasions. This may be one of the reasons we don't find the same way on the nature of their relationship since it might lead to several biased findings in such relationship. The results will be improved if their periods are obtained by fourier analysis and combined with oil shocks to perform LSTM forecasting.

4.3. Prediction by LSTM

In the model setup, we use a three-layer structure, with 512 neurons in the first layer and 128 neurons in the second layer, and then use sigmoid to finally map the data between (0,1) after summing the 128 data in the third layer, and compared with the labeled values to calculate the loss after optimizing the coefficients and then entered the next round of training.We set a total of 1500 epochs to train the

LSTM model and observe the loss at the end. If the loss is too high, we add a dropout layer to retrain the model for that data. The final prediction results and images can be obtained.

Then, the prediction for the later 72 months is combined with the previously derived later period.

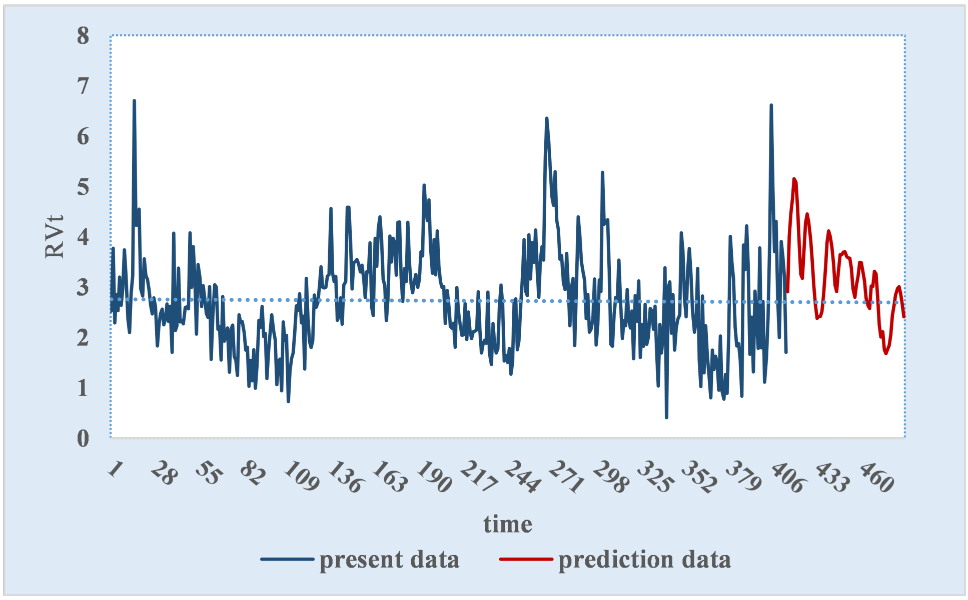

Figure 2. The predicted results.

Note: The blue interval represents data from 1st January, 1989 to 31st January, 2020 and the red interval indicates the prediction results.

4.4. Forecast Analysis

It is obvious to find that our proposed model is better than the current model in all aspects than the test set outside the sampling range by all metrics as shown in Table 2. Besides the six variables we mentioned before, there are 5 other generally adopted ways to make prediction, which are mean, median, trimmed mean, and discount mean squared which contains two ways, namely the DMSPE1 and DMSPE2 (Gokmenoglu et al., 2021). The following findings are presented. First, regarding the results of forecasts, the $R^2_{OOS}(\%)$ values of AR-LPI are higher than 0 while others have a $R^2_{OOS}(\%)$ value lower than 0, which means such models behave worse than the benchmark model in forecasting. Second, the $R^2_{OOS}(\%)$ of Fourier-LSTM is 19.984, which means fourier transform based LSTM increases the correctness by 19.9841% more than the benchmark, showing the fourier transform way contributes to forecasting to some extent.

The sudden jump in financial returns is partly due to structural breaks which is because of the investors' multiple expectations about the future of the stock as well. The excellent performance of the LSTM can also be attributed to the treatment of noise in fourier transform that allows the model to effectively identify cyclically varying anomalies and reject them, which make it avoid noise and overfitting problems.

Table 2. Forecasting performance.

AR or LSTM models	$R^2_{OOS}(\%)$	p-value	CER
RV			2.68
NPI	−0.050	0.748	2.672
ANP	−7.780	0.931	2.675
SNP	−0.004	0.299	2.678
LPI	0.018	0.396	2.689
NPI2	−0.101	0.822	2.68
MEAN	−0.459	0.907	2.68
MEDIAN	−0.047	0.845	2.679
TMC	−0.098	0.904	2.679
DMSPE1	−0.449	0.903	2.68
DMSPE2	−0.456	0.904	2.68
Fourier-LSTM	19.984	0.014	2.89

Note: This chart shows the performance of different prediction methods, and obtains the advantages of prediction methods through comparison.

Table 3. Robustness test.

AR or LSTM models	QLIKE(value)	MSE(value)
RV	34.7%	24.6%
NPI	26.1%	24.6%
ANP	26.1%	17.6%
SNP	26.1%	24.6%
LPI	34.7%	24.6%
NPI2	34.7%	24.6%
MEAN	34.7%	18.7%
MEDIAN	34.7%	24.6%
TMC	34.7%	24.6%
DMSPE1	34.7%	22.0%
DMSPE2	34.7%	20.2%
Fourier-LSTM	57.2%	25.1%

Note: This chart shows the robustness test of different prediction methods.

4.5. Robustness Test

The QLIKE and MSE are adopted as an index to evaluate the robustness. Refering to Hansen et al. (2011)'s conclusion, the model confidence set (MCS) is used to measure how the performance the model will act in test set. Additionally, the stationary bootstrap method is applied to compute the p-value of model confidence set test. Table 3 shows all the results.

In the research of predicting international crude oil shock by models, in order to enhance the interpretability of model prediction results, most of the research scholars analyze and predict the future international crude oil price movements by studying the influencing factors related to international crude oil prices and using statistical methods such as correlation coefficients and multiple regression models, while the above-mentioned due to the different degrees of influence of different influencing factors on international crude oil prices in different periods. In particular, it is difficult to identify the impact of short-term international crude oil shock on stock market returns due to the combined effect of multiple influencing factors.

However, our experimental results effectively confirm that the prediction method solves the above problems. Table 3 shows all the results. Some findings are listed. First, only the based on QLIKE is the MCS p-values greater than 0.25 in five forecasts, all of which are less than 0.25 when computing the MSE value of them. Moreover, the Fourier transform-LSTM' MCS p-values are greater than 0.25 in each occasion, indicating the fourier transform is efficient to improve the performance of predicting the volatility of stock market.

5. Conclusions

This paper analyses the impact of various oil shocks on the stock volatility prediction by using Fourier transform based Long Short-Term Memory(LSTM). The oil shocks are decomposed into five components in a comprehensive manner. Based on a daily dataset involving the volatility of the S&P 500 index and WTI oil futures' price, we find that various oil shocks perform differential impacts on stock volatility dynamics by using gradient descent. The stock volatility is featured by an evident autoregressive feature that the current stock volatility it is found to be significantly driven by its past values. Moreover, by considering the role of five oil shocks on the stock volatility dynamics, our results show that the LSTM method containing fourier transform can contribute to the performance of forecasting which is valid in economic and test meanings. Additional analyses reassure robustness of our findings.

Our paper sheds new lights on clear interpretation regarding the role of oil shocks in stock market dynamics and prediction of the latter, both of which help market participants make sensible decisions to investment risk reduction. Moreover, our obtained results further help policymakers and investors anticipate the market volatility by using a superior forecasting method, formulate reasonable risk management when encountering unexpected oil shocks, reduce systemic risk, and thus enhance effectiveness of financial market interventions. In addition, as a research extension, in the future research agenda, other potential stock volatility drivers on the stock market dynamics such as macroeconomic factors and policy shifts will be further considered in the stock volatility prediction. In addition, future agenda for research extension might be considerations of the role of other potential drivers in the field of macroeconomics and policy for stock volatility.

Acknowledgement

This research was supported by the Natural Science Fund of Hunan Province (2022JJ40647).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

Aastveit KA (2014) Oil price shocks in a data-rich environment. *Energ Econ* 45: 268–279. https://doi.org/10.1016/j.eneco.2014.07.006

Ahmed AD, Huo R (2021) Volatility transmissions across international oil market, commodity futures and stock markets: Empirical evidence from china. *Energ Econ* 93: 104741. https://doi.org/10.1016/j.eneco.2020.104741

An Y, Sun M, Gao C, et al. (2018) Analysis of the impact of crude oil price fluctuations on china's stock market in different periods—based on time series network model. *Physica A* 492: 1016–1031. https://doi.org/10.1016/j.physa.2017.11.032

Barsky RB, Kilian L (2001) Do we really know that oil caused the great stagflation? a monetary alternative. *NBER Macroecon Annu* 16: 137–183. https://doi.org/10.3386/w8389

Barsky RB, Kilian L (2004) Oil and the macroeconomy since the 1970s. *J Econ Perspect* 18: 115–134. https://doi.org/10.3386/w10855

Charfi S, Mselmi F (2022) Modeling exchange rate volatility: application of GARCH models with a Normal Tempered Stable distribution *Quant Financ Econ* 6: 206–222. https://doi.org/10.3934/QFE.2022009

Chen MW, Zhu J (2007) Volatility clustering within industries: An empirical investigation. *Am J Bus* 22: 33–44. https://doi.org/10.1108/19355181200700008

Chen Q, Zhou S (2002) A study on the volatility of sse index returns. *Res Quant Econ Technol Econ* 19: 101–110.

Chen Y (2019) Shareholding structure and governance of inefficient investment behavior of state-owned enterprises–empirical evidence based on data of state-owned listed companies analysis. *Econ Manage Res* 5: 49–54.

Cheng A, Chang Q (2016) Research on the dynamic relationship between oil prices and share prices of oil companies in china. *South China Univ Technol J (Social Science Edition)* 18: 1–11.

Chiou JS, Lee YH (2009) Jump dynamics and volatility: Oil and the stock markets. *Energy* 34: 788–796. https://doi.org/10.1016/j.energy.2009.02.011

Choi S, Furceri D, Loungani P, et al. (2018) Oil prices and inflation dynamics: Evidence from advanced and developing economies. *J Int Money Financ* 82: 71–96. https://doi.org/10.2139/ssrn.3036910

Christiansen C, Schmeling M, Schrimpf A (2012) A comprehensive look at financial volatility prediction by economic variables. *J Appl Econ* 27: 956–977. https://doi.org/10.2139/ssrn.2016994

Clark TE, West KD (2007) Approximately normal tests for equal predictive accuracy in nested models. *J Econometrics* 138: 291–311. https://doi.org/10.1016/j.jeconom.2006.05.023

Cunado J, de Gracia FP (2014) Oil price shocks and stock market returns: Evidence for some european countries. *Energ Econ* 42: 365–377. https://doi.org/10.1016/j.eneco.2013.10.017

Dickey DA, Fuller WA (1979) Distribution of the estimators for autoregressive time series with a unit root. *J Am stat Assoc* 74: 427–431. https://doi.org/10.2307/2286348

Duan K, Ren X, Shi Y, et al. (2021) The marginal impacts of energy prices on carbon price variations: evidence from a quantile-on-quantile approach. *Energ Econ* 95: 105131. https://doi.org/10.1016/j.eneco.2021.105131

Duan Y, Wang H (2016) A study on the spillover correlation between emerging market stock prices and crude oil price volatility. *Stat DecI Making* 03: 165–168.

Engle RF, Ghysels E, Sohn B (2013) Stock market volatility and macroeconomic fundamentals. *Rev Econ Stat* 95: 776–797. https://doi.org/10.1162/REST_a_00300

Fama EF (1970) Efficient capital markets: A review of theory and empirical work. *J Financ* 25: 383–417. https://doi.org/10.2307/2325486

Georgellis Y (1994) The oil price shocks and the unit root hypothesis: the uk experience. *Appl Econ* 26: 827–830. https://doi.org/10.1080/00036849400000098

Gokmenoglu K, Eren BM, Hesami S (2021) Exchange rates and stock markets in emerging economies: new evidence using the Quantile-on-Quantile approach. *Quant Financ Econ* 5: 94–110. https://doi.org/10.3934/QFE.2021005

Hansen PR, Lunde A, Nason JM (2011) The model confidence set. *Econometrica* 79: 453–497. https://doi.org/10.2139/ssrn.522382

Haugom E, Langeland H, Molnár P, et al. (2014) Forecasting volatility of the us oil market. *J Bank & Financ* 47: 1–14. https://doi.org/10.2139/ssrn.2691391

He W (2012) Study on the "spillover effect" of international crude oil prices on chinese stock market and its transmission mechanism. *Int J Financ & Econ* 35: 61–68.

He W, Wang M (2013) A comparative study on the impact of domestic and international crude oil prices on chinese stock market. *J Yan'an Univ (Social Science Edition)* 35: 61–68. https://doi.org/10.3969/j.issn.1004-9975.2013.04.013

He Z, Chen J, Zhou F, et al. (2020) Oil price uncertainty and the risk-return relation in stock markets: evidence from oil-importing and oil-exporting countries. *Int J Financ & Econ* 35: 61–68. https://doi.org/10.1002/ijfe.2206

Hu L, Ying W (2017) Study on volatility of stock market: Empirical analysis based on arma-tgarch-m model. *J Beijing Univ Aeronaut Astronaut (Social Sciences Edition)* 30: 5.

Kilian L, Vigfusson RJ (2013) Do oil prices help forecast us real gdp? the role of nonlinearities and asymmetries. *J Bus & Econ Stat* 31: 78–93. https://doi.org/10.1080/07350015.2012.740436

Liow KH, Song J, Zhou X (2021) Volatility connectedness and market dependence across major financial markets in China economy. *Quant Financ Econ* 5: 397–420. https://doi.org/10.3934/QFE.2021018

LIU H, WANG Y (2017) Study on volatility of stock market: Empirical analysis based on arma-tgarch-m model. *J Beijing Univ Aeronaut Astronaut (Social Sciences Edition)* 30: 56.

Lu X, Ma F, Wang J, et al. (2021) Oil shocks and stock market volatility: New evidence. *Energ Econ* 103: 10556. https://doi.org/10.1016/j.eneco.2021.105567

Maheu JM, Song Y, Yang Q (2020) Oil price shocks and economic growth: The volatility link. *Int J Forecast* 36: 570–587. https://doi.org/10.1016/j.ijforecast.2019.07.008

Paye BS (2012) 'déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *J Financ Econ* 106: 527–546. https://doi.org/10.2139/ssrn.783986

Qiao K, Sun Y, Wang S (2019) Market inefficiencies associated with pricing oil stocks during shocks. *Energ Econ* 81: 661–671. https://doi.org/10.2139/ssrn.3285838

Rapach DE, Strauss JK, Zhou G (2010) Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Rev Financ Stud* 23: 821–862. https://doi.org/10.2139/ssrn.1257858

Ren X, Dou Y, Dong K, et al. (2022a) Information spillover and market connectedness: multi-scale quantile-on-quantile analysis of the crude oil and carbon markets. *Appl Econ* 54: 4465–4485, 2022a. https://doi.org/10.1080/00036846.2022.2030855

Ren X, Duan K, Tao L, et al. (2022b) Carbon prices forecasting in quantiles. *Energ Econ* 108: 105862, 2022b. https://doi.org/10.1016/j.eneco.2022.105862

Ren X, Tong Z, Sun X, et al. (2022) Dynamic impacts of energy consumption on economic growth in china: Evidence from a non-parametric panel data model. *Energ Econ* 107: 105855, 2022c. https://doi.org/10.1016/j.eneco.2022.105855

Shahrestani P, Rafei M (2020) The impact of oil price shocks on tehran stock exchange returns: Application of the markov switching vector autoregressive models. *Resour Policy* 65: 101579. https://doi.org/10.1016/j.resourpol.2020.101579

Sukharev OS (2020) Structural analysis of income and risk dynamics in models of economic growth. *Quant Financ Econ* 4: 1–18. https://doi.org/10.3934/QFE.2020001

Sun M, Yang T (2012) Analysis of the impact of oil prices on chinese stock market. *J Shenzhen Univ (Humanities and Social Sciences Edition)* 29: 113–116.

Wei C (2003) Energy, the stock market, and the putty-clay investment model. *Am Econ Rev* 93: 311–323. https://doi.org/10.1257/000282803321455313

Wen F, Tong X, Ren X (2022) Gold or bitcoin, which is the safe haven during the covid-19 pandemic? *Int Rev Financ Anal* 81: 102121. https://doi.org/10.1016/j.irfa.2022.102121

Xiao J, Hu C, Ouyang G, et al. (2019) Impacts of oil implied volatility shocks on stock implied volatility in china: Empirical evidence from a quantile regression approach. *Energ Econ* 80: 297–309. https://doi.org/10.1016/j.eneco.2019.01.016

Xu W, Ma F, Chen W, et al. (2019) Asymmetric volatility spillovers between oil and stock markets: Evidence from china and the united states. *Energ Econ* 80: 310–320. https://doi.org/10.1016/j.eneco.2019.01.014

Zhang D (2017) Oil shocks and stock markets revisited: measuring connectedness from a global perspective. *Energ Econ* 62: 323–333. https://doi.org/10.1016/j.eneco.2017.01.009

Zhang Y, Ma F, Wang Y (2019) Forecasting crude oil prices with a large set of predictors: Can lasso select powerful predictors? *J Empir Financ* 54: 97–117. https://doi.org/10.1016/j.jempfin.2019.08.007

Zhao L, Zhang X, Wang S, et al. (2016) The effects of oil price shocks on output and inflation in china. *Energ Econ* 53: 101–110. https://doi.org/10.1016/j.eneco.2014.11.017

AIMS Press