Resources Policy 83 (2023) 103614

Contents lists available at ScienceDirect

Resources Policy

journal homepage: www.elsevier.com/locate/resourpol

Using econometric and machine learning models to forecast crude oil prices: Insights from economic history

Zilin Xu [a], Muhammad Mohsin [b], Kaleem Ullah [c], Xiaoyu Ma [d,*]

[a] School of History, Anhui Normal University, Wuhu, 241000, China
[b] School of Finance and Economics, Jiangsu University, Zhenjiang, 212013, PR China
[c] National University of Sciences and Technology (NUST), Islamabad, Pakistan
[d] School of Finance and Trade, Wenzhou Business College, 325035, Wenzhou, China

ARTICLE INFO

Keywords:
Crude oil
Machine learning
Price prediction
Artificial neural networks
Economic history

ABSTRACT

The volatility of the crude oil market and its effects on the global economy increased the concerns of individual investors, states/governments, and corporations. Forecasting the price of crude oil is difficult owing to its complicated, nonlinear, and chaotic nature in economic history. Multiple variables influence crude oil prices, such as the economic history, economic cycle, international relations, and geopolitics. Predicting the price of crude oil is a complex but valuable endeavor. Crude oil price forecasting is done using historical data (time series method) or dependent variables/factors (regression method) using traditional econometric or machine learning models. In this study, we use both methods (regression and time series) to examine the prediction performance of both models (econometric and machine learning models) for daily WTI crude oil prices covering the period December 18, 2011, through December 31, 2018. We present a performance analysis of conventional econometric models (ARIMA, GARCH, and OLS), Artificial Neural Network (ANN) regression models, and ANN Time Series models to compare their results to find out the best-performing method (time series or regression) and the best model (econometric or machine learning model). Based on our study results, we propose a novel Artificial Neural Network model to improve the prediction performance of existing models by adjusting the bias and weights of ANN hidden layers. We used historical prices of 14 different variables, including gold, silver, S&P500, USD Index price, and US-EU conversion rates for regression models, whereas historical time series data of WTI crude oil for time series models. Analysis of the results reveals that the performance of our proposed model remained better than all tested models. The comparative results of existing models show that the overall performance of Neural Networks remained better than econometric models. Our results have substantial implications for governments, businesses, and investors, and for the sustainable growth of economies that rely on energy.

1. Introduction

Crude oil is the unrefined, thick liquid that is pumped out of oil reserves located hundreds of meters below the earth's surface. It is fundamental to every economy and cannot be replaced. Crude oil is classified into many categories based on quality and region of production. Since it is used in so many different ways, its prices significantly influence the global economy and security. Since crude oil accounts for a sizable percentage of exports for certain countries, a rapid change in the price of crude oil can have far-reaching financial consequences, with crashes in crude oil prices leading to slower economic activity and booms in crude oil prices initiating substantial inflation as well (Ullah et al., 2020).

Furthermore, since the 1970s, oil prices have grown much more unstable. Because of oil's rising importance in the economy, precise predictions of its future price are crucial for policy makers to make sound economic choices. In addition, governments can better prepare for unexpected spikes in oil prices with the help of precise predictions. According to (Xiuzhen et al., 2022), the oil market is volatile. While supply and demand are the primary drivers of crude oil prices, other variables such as the stock market, economic activity, political situation, and other external factors also have a role (Wu et al., 2022). Therefore, predicting the future price of oil is crucial but also highly challenging owing to the market's inherent volatility and uncertainty. The capacity

* Corresponding author.
E-mail addresses: 591370172@qq.com (Z. Xu), m.mohsin3801@yahoo.com (M. Mohsin), kaleem_7198@hotmail.com (K. Ullah), mxy@wzbc.edu.cn (X. Ma).

https://doi.org/10.1016/j.resourpol.2023.103614
Received 26 February 2023; Received in revised form 16 April 2023; Accepted 18 April 2023
Available online 12 May 2023

to correctly estimate the future price of crude oil has emerged as one of the most pressing research questions in the world of forecasting (Pan et al., 2023).

When it comes to the design of programs to stabilize the economy and financial plans, the capability to estimate crude oil prices is vital for nations that are both oil producers and oil consumers (Fang et al., 2022). On the other hand, due to the nonlinear character of the oil market, it is not easy to precisely forecast the prices of the oil market. Forecasting errors for the price of crude oil are primarily caused by the complexity of the supply and demand structure and the presence of many unexpected factors that disrupt the market's equilibrium. Both exogenous factors, such as the status of the global economy, and endogenous components of the oil market, such as consumption, inventory, and supply of oil, each have a unique role in influencing the trajectory of the price of crude oil. Exogenous factors include conflicts, war, economic development, and political instabilities, which affect the prices of oil as well as disturb demand and supply ratios. For example, the price of crude oil in 2020 has been unpredictable because several factors contribute to its price. The oil price war between Russia and Saudi Arabia and the pandemic produced by coronavirus infection in 2019 (COVID-19) are examples of these factors. There has been a decline in demand for oil worldwide due to the global economic downturn brought by the COVID-19 pandemic (Liu et al., 2023).

The United States is the second-largest consumer and the fourth-largest exporter of crude oil. Before 2005, the United States was among the world's leading net energy importers. Imports met about a third of the United States' energy needs. The United States has become a net overall energy exporter, with exports increasing steadily since 2005 and peaking in 2019. The U.S. economy affects the dynamics of the global energy market since it is one of the world's largest consumers and suppliers of energy. Oil prices worldwide are susceptible to the changes in the economy of the United States. Supply and demand pressures in the crude oil industry are shaped by these triggers, which are primarily influenced by macroeconomic and financial variables, and therefore impact investment behavior(Li and Umair, 2023). Such behavior-driving variables influence the quantity and cost of WTI crude oil. However, the fluctuations of the WTI crude oil prices are determined by elements unique to the US economy, such as the cost of refining crude oil and the volume of oil output. The ups and downs of oil prices are greatly influenced by political events, crises, and conflicts on a global scale, as well as by the policies of OPEC. After considering this, it is essential to establish a robust model for predicting future crude oil prices. It is difficult to make predictions based on past prices due to non-linearity, uncertainty, and dynamics in these prices (Iram et al., 2020).

An overview of the literature on economic history and forecasting crude oil prices indicates that most of the existing literature relied on ARIMA and ANNs to make the prediction(Iqbal et al., 2019). Moreover, generally in the literature, only one type of model (time series models or regression models) and a single method (econometrics or machine learning) were used to forecast crude oil prices. To the best of our knowledge, none of those works has thoroughly compared the prediction performance of different models (time series and regression models) and methods (econometrics and machine learning) on a single data set to determine the optimal model and the best method. It indicates a research gap concerning comparing the performance of time series and regression methods using both econometrics and machine learning models. Moreover, there is also room for improvement in the detection performance of existing predictors. With this background, the current study presents the following:

- We propose a novel ANN model to improve the existing prediction results (prediction of crude oil future prices) of ANN regression model by redefining the biases and weights of hidden layers. Moreover, the performance of regression neural networks remained better than the time series network for predicting WTI crude oil prices due to the non-linearity of time series data.
- We use a data set comprising historical values of 15 different commodities (including WTI crude oil) and implemented ANNs (regression and time series) to forecast the price of crude oil.
- We use traditional econometric models to predict crude oil prices based on historical time series values and past values of 14 other commodities.
- We compare the performance of econometric models, ANN regression models, and ANN time series models to find out the best model and method for prediction of the future price of crude oil.

The rest of the paper is organized as under: section 2 describes the literature review, section 3 presents the analysis of data and a brief overview of the methodology adopted in this paper, section 4 analyzes the results of implemented models in detail and presents a comparison of results and section 5 concludes the paper.

2. Literature review

Standard methods for forecasting future energy prices may be broken down into three broad categories: econometric models, models based on artificial intelligence, and hybrid models. Some examples of econometric models are the Autoregressive Integrated Moving Average (ARIMA) (Sebestyén and Abonyi, 2021), the Vector Autoregressive model(Mohsin et al., 2020a), and the Generalized Auto-regressive Conditional Heteroskedasticity (GARCH) (Shah et al., 2019)(Xia et al., 2020). Generally, econometric models perform poorly in nonlinear settings, despite their success with linear and stationary time series data (Adebayo et al., 2021). For analyzing complicated time series, AI techniques are preferred over econometric models due to their ability to accurately reflect the nonlinear and non-stationary features. One of the most common uses of artificial intelligence is energy price forecasting, and two of the most used methods are support vector machine (SVM) and support vector regression (SVR) models(D. D. Zhang et al., 2021), (Ikram et al., 2019). The Generalized Auto-Regressive Conditional Heteroskedasticity (GARCH) model was one of the first to be used in this domain due to its ability to reflect emotional volatility across time (Mohsin et al., 2020b)(Iqbal et al., 2022).

To accurately depict sparse data samples exhibiting nonlinear behavior, some writers have turned to the SVM model for price prediction. The SVM model is more adaptable and convenient due to multiple reasons (Mohsin et al., 2018a). Although some have praised the wavelet methodology for accurately forecasting oil prices (Pesaran, 2014), a significant drawback is that it is susceptible to the sample size used for prediction. Recent studies have favored using hybrid approaches for price forecasting (Mohsin et al., 2018b). It has been argued that it is feasible to develop a single hybrid framework that uses the benefits of both the soft-computing method and the econometric technique by merging them or vice versa (Pedroni, 1999) (Kao, 1999). Time series models have also been widely used in the literature, but time series for oil price estimation has several inherent problems. They are not constant across time but rather exhibit cyclicity and volatility. Because of this, it isn't easy to make reliable predictions about the price of energy in the future (L. Chang et al., 2022c). Another technique that may be used to forecast the future of oil prices is the neural network (NNET) approach (Huang et al., 2022). In recent years, several studies have shown the efficacy of ANNs for the prediction of prices (Mele and Magazzino, 2020),(Naz et al., 2021). There has been a lot of focus on using deep learning to tackle the problem of estimating future energy expenditures in the last several years (Gengenbach et al., 2010). Some of the issues linked to NNET approach include over-fitting, local minima, and a need for more generalizability(Batool et al., 2022). This has led to the use of hybrid models for predicting oil prices. Better predicting results may be achieved via hybrid models, which combine the best features of many existing models(Liu et al., 2022). Phase space reconstruction (PSR) is the cornerstone of artificial intelligence (AI) methods; it involves mapping time series to a vector space in which the dimension and current value have a functional link between initial values (Dilanchiev and Taktakishvili, 2021). Decomposition-ensemble forecasting as a paradigm for estimating energy expenses in the future has gained popularity in recent years. In this method, complex time series data is broken down into a collection of more specific datasets. Then the predictions made from each of these datasets are combined to form a master forecast (L. Chang et al., 2022e). Wavelet decomposition (Luo et al., 2023), empirical mode decomposition (EMD), compressed sensing (Hillis et al., 2023), and sparse representation are some of the most well-known methods of decomposition (Sharma et al., 2022). EM decomposition (EMD) is proposed by (Zhou and Li, 2022) to separate a raw signal into a set of intrinsic mode functions: complete, nearly orthogonal sub-signals. This operation occurs alone in the temporal domain. Instead of competing for signal decomposition methods, EMD's main benefit is that its essential functionalities can be extracted directly from the signal itself (L. Chang et al., 2022d). In addition, EMD is easy to use and provides an accurate estimation of even minor frequency shifts. As used by the EMD, Multiscale decomposition can potentially simplify models drastically. The application of EMD in the construction of ensemble learning models is widespread in the field of energy price forecasting. EMD-based neural network ensemble learning for prediction of crude oil has been proposed (Wang et al., 2022). The findings show that the technique provides superior performance compared to using a single ARIMA model, a single feedforward neural network (FNN) model, or an EMD-based ARIMA model. A decomposition-ensemble model based on ensemble EMD (EEMD) was created using

Fig. 1. Graphical representation of all variables from 2012 to 2018.

data-characteristic-driven reconstruction and AI techniques to forecast the price of crude oil(L. Chang et al., 2022b). Foreseeing future energy costs using a decomposition-ensemble learning technique on the lines of authors in (Bull, 2001) is a fast and efficient method that draws inspiration from randomness. West Texas Intermediate crude oil daily, weekly, and monthly price forecasts are generated using the GARCH and PSO-LSSVM model developed by (Chang et al., 2023a) and based on the estimating equations of multiple dimensions (EEMD). Combining the ARIMA and PSO-LSSVM decomposition ensemble methods with the EEMD (Bei and Wang, 2023), provides a better forecasting model than econometric models. A unique approach for predicting the price of crude oil across a wide variety of periods was presented by (L. Chang et al., 2022c). EMD alongwith other methods are used in the model for prediction. Using ARIMA created a more precise forecasting model to delink the high volatility and daily seasonality in power prices in Australia. Using Akaike's information criteria (Umair, 2022), introduces an innovative decomposition-ensemble model to improve the precision of crude oil price prediction according to the economic history. In a recent article, describe a new EEMD-based network learning strategy for crude oil price forecasting.

3. Data and methodology

3.1. Data

This research collects data from the UCI machine learning repository from December 18, 2011 to December 31, 2018. The data comprises a total of 1718 rows and 81 columns. Gold Price, Standard and Poor's (S&P) 500 index, WTI Crude oil price, Dow Jones Index, US Bond rates, Euro-USD exchange rates, prices of precious metals (Silver, Platinum, Palladium, and Rhodium), prices of US Dollar Index, Eldorado Gold Corporation, and Gold Miners ETF are included in this data set. Each element in the dataset comprises several values, such as the day's opening value, the day's maximum value, the day's lowest value, and the day's closing price, among others. During preprocessing, we choose the opening value of all commodities to reduce the dimensionality of the data for better analysis. Figure-1 gives a graphical representation of all variables before the normalization of the data.

We use WTI Crude oil price as our response variable for both regression and time series models whereas other 14 variable are used as predictors (see Table 1). Multiple institutions, including US Department of Energy utilize the price of WTI to approximate the value of worldwide crude oil. Changes in the WTI crude oil market prices affect the pricing of crude oil on other marketplaces (Wang et al., 2023). We initially normalized the raw data using MATLAB preprocessing to increase the accuracy of our predictions and to avoid the neurons of the neural network from becoming saturated during the computation (Agbonifo, 2021). The data has been divided in training data (70%), cross-validation (15%), and testing (15%). Table 2 shows the statistics of dataset, whereas Fig. 2 represents all variables on same scale.

A thorough analysis of the input dataset has been undertaken to gain valuable insight; however, all details are not included in this study to focus on the implementation part. Detailed analysis in the form of plots and tables can be forwarded on request. Analysis using the following features/plots for each variable has been undertaken:

- Frequency distribution plot
- Scatter plot with least square fit
- Estimated density
- Frequency spectrum plot
- Time spectrum plot
- Auto-Correlation and Partial Auto- Correlation plots

3.2. Methodology

We present a methodology for forecasting WTI crude oil prices using econometric and machine learning techniques. We propose a novel ANN model to improve the prediction results of the existing ANN regression model by redefining the biases and weights. For analysis of results, we compare traditional econometric models such as ARIMA, GARCH, and OLS with machine learning models such as ANNs (ANN time series and ANN regression). In ANN regression models, we use single-layer, bi-layer, and tri-layer ANNs. An HP Pavilion G Series laptop with 8 GB of RAM, 500 GB of hard disc space, a 64-bit operating system, and a 3.00 GHz Intel Core i5 CPU was utilized to get these results. MATLAB R2022b and Gnu Regression Econometrics and Time-series Library are used to analyze econometrics and ANN models.

Table 1
Acronyms and their meaning.

Term	Meaning
AI	Artificial Intelligence
ARIMA	Autoregressive Integrated Moving Average
GARCH	Generalized Auto-regressive Conditional Heteroskedasticity
OLS	Ordinary Least Square
ANN	Artificial neural networks
S&P500	Standard and Poor's (S&P) 500 index
WTI	West Texas Intermediate
OPEC	Organization of the Petroleum Exporting Countries
SVR	Support Vector Regression
SVM	Support Vector Machine
NNET	Neural Network
PSR	Phase Space Reconstruction
EMD	Empirical Mode Decomposition
EEMD	Ensemble EMD
FNN	Feed Forward Neural Network
PSO-LSSVM	Particle swarm optimization least squares support vector machine
GA	Generalized Algorithm
MSE	Mean Square Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Square Error
ACF	Auto-Correlation Function
PACF	Partial-Auto correlation function

3.2.1. ARIMA Model

The ARIMA model is widely employed because of its ability to forecast trends in time series reliably (Geng, 2021), (Y. Y. Zhang et al., 2021). The ARIMA model is often represented as (p,d,q), where p = AR order, q = MA order, and d = differencing order. The equation for this model can be written as:

$$\varphi(B)\nabla^d z_t = \theta(B)\alpha_t \text{or} z_t = \sum_{i=0}^{p} \varphi_{ii} z_{t-i} + \alpha_t - \sum_{k=1}^{q} \theta_i \alpha_{t-k} \tag{1}$$

where

$$\varphi(B) = 1 - \varphi_1 B – \varphi_2 B^2 - \ldots - \varphi_p B^p \tag{2}$$

is the regressive equation of order p and

$$\theta(B) = 1 - \theta_1 B – \theta_2 B^2 - \ldots.. - \theta_q B^q \tag{3}$$

is the equation of order q, whereas B is the shift operator and can be described as:

$$B^p z_t = z_{t-p} \tag{4}$$

where$\varphi_1, \varphi_2 \ldots, \varphi_p$ $\theta_1, \theta_2, \ldots, \theta_q$ are the unknown coefficients to be predicted and

$$\nabla^d = (1 - B)^d \tag{5}$$

backward shift operator can be written as under:

$$\nabla z_t = z_t – z_{t-1} \text{with} \nabla^d = \nabla \nabla^{d-1} \tag{6}$$

Table 2
Statistics of input dataset.

Variable Name	Mean	Median	Minimum	Maximum	Std. Dev.	Skewness	5% Perc.	95% Perc.	Missing obs.
Gold	127.32	121.92	100.92	173.20	17.527	1.1646	107.02	164.75	0
SP500	204.49	205.46	122.06	293.09	43.832	0.087813	135.47	278.37	0
Dow Jones Index	18161.	17601.	11769.	26833.	3889.8	0.53231	12806.	25345.	0
Eldordo Gold Corporation	28.277	22.800	2.7700	80.200	20.326	0.88739	4.6500	71.100	0
EURUSD Exchange rate	1.2085	1.1841	1.0390	1.3933	0.10058	0.19779	1.0664	1.3673	0
Brent Crudeoil Futures	77.505	70.115	27.880	126.22	27.401	0.17232	42.130	115.41	0
Crude Oil WTI USD	70.275	64.840	27.340	110.34	23.480	0.18648	39.817	105.24	0
Silver Futures	43309.	40528.	33146.	65400.	7550.4	1.0442	34758.	58596.	0
US Bond Rate	2.2631	2.2590	1.3660	3.2370	0.43398	0.11421	1.5889	2.9701	0
Platinum Price	1184.4	1098.2	765.30	1737.8	273.98	0.32770	833.29	1634.4	0
Palladium Price	766.36	748.00	458.60	1196.0	148.08	0.57711	552.40	1048.9	0
Rhodium Prices	1130.4	1100.0	0.00000	2600.0	570.01	0.46962	0.00000	2380.0	0
US dollar Index Price	89.805	92.905	78.220	103.35	7.5208	−0.15043	79.620	100.23	0
Gold Miners ETF	26.747	23.115	12.700	57.520	10.621	1.3535	14.220	51.356	0
Oil ETF USO	22.113	16.450	7.8200	41.600	11.431	0.30823	9.6295	38.682	0

Fig. 2. Graphical representation of all variables on the same scale.

3.2.2. *GARCH model*

Whenever working with predictions, the issue is not always whether a model is a good match for our data; additionally, we desire to assess the model's accuracy and the validity of its predictions. Examining the variance of our error terms is one approach for evaluating accuracy. Calculating the standard deviation of our error terms is one way to assess the reliability of our findings and identify problem areas that need further attention. Before Engle's release of the ARCH model in 1982, econometricians often used the idea of a "rolling standard deviation" to make sense of the inconsistencies between various models. With this method, the weight of each observation was normalized to the variance in the total number of readings. Using this method, we can account for the differences found in the models.

$$\sigma_{u_{t+1}} = \frac{1}{n} \sum_{i=0}^{n} \sigma u_{t-i} \tag{7}$$

ARCH may be seen as an extension of this formulation; instead of giving equal weight to each value, the weights are viewed as estimation parameters. This is the more practical approach due to the following reasons:

- Assuming equal weights appears wrong, given our presumption that more recent data are more likely to be important
- Limiting our weights to some fixed number of observations is not ideal.

The ARCH formulation views the weights not as fixed points in a range but as variables that must be evaluated. This paradigm provides valuable insight through which to view ARCH. If more recent data are more likely to be necessary, then 1) assuming equal weights appears unreasonable, and 2) restricting our weights totally to some finite number of observations, this is a more accurate picture of the situation. Given our opinion that more recent data are more likely to be more significant, this is a more truthful depiction of the situation than 1) assuming equal weights. One possible formulation of this is the m-order generalized ARCH model:

$$u_t^2 = c + \left(\sum_{i=1}^{m} \alpha_i u_{t-1}^2 \right) + w_t \tag{8}$$

Where w_t is white-noise-based random (Bag et al., 2020). Like the generalization of ARMA to ARIMA, GARCH is an extension of ARCH that proposes the long-run average variance is the most significant predictor of future variation (Lee, 2020).

3.2.3. Ordinary Least Squares (OLS)

Ordinary Least Squares (OLS) is a widely used statistical method in quantitative analysis. Whenever feasible, economists look for (pseudo) linear correlations between many variables to make the interaction between them more transparent so that it is easy to see how a shift in one variable affects the others. This model can effectively predict futures and retrospective explanations due to linkages between the variables. Finding a function that precisely fits every single point in any given data set is feasible. Still, this kind of function is useless since it doesn't provide anything in the way of explaining or predicting the data. If an equation is a complicated polynomial involving trigonometric functions, it will be challenging to come up with a reasonable connection between two variables. So, economists and scientists often try to establish straightforward linear correlations. Suppose we have n parameters $\{A_{1,i}, A_{2,i} \cdots, A_{n-1,i}, B_i\}$ where we designate each A_j to be an "independent" variable and B_i is defined to be the "dependent" variable; formula for OLS can be written as:

$$B_i = \beta_0 + \beta_1 A_{1,i} + \cdots + \beta_{n-1} - A_{n-1,i} + u_i \tag{9}$$

where B_i is constant and u_i is defined as the disturbance term.

3.2.4. Neural network

A Neural network, as shown in Fig. 3, is formed up using input, hidden, and output layers of neurons. There is no specific rule about the number of hidden layers present to replicate a function correctly; nonetheless, theoretical work such as (Golosnoy et al., 2019) argues that

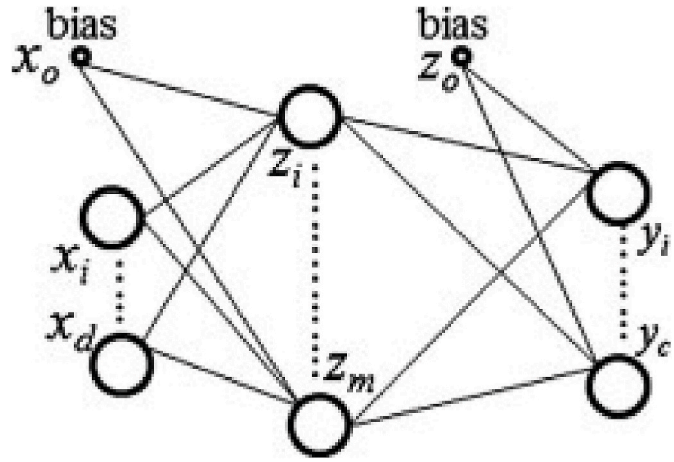

Fig. 3. Structure of neural network.

just one hidden layer is essential. In this particular investigation, the inputs and outputs of neurons stand in for the independent and dependent variables, respectively.

The equation of input neuron (d) and output neurons (c) can be written as:

$$a_j = \sum_{i=1}^{d} w_{ji} x_i + w_{jo} (10)$$

where w_{ji} and w_{jo} represents the weights of the input neuron and bias of the hidden neuron j and

$$a_j = \sum_{i=o}^{d} w_{ji}^{(1)} x_i \tag{11}$$

The activation function $g(x)$ in hidden layers can be used to transform Eq. (11):

$$z_j = g(a_j) \tag{12}$$

output of neuron k, can be given by:

$$a_k = \sum_{j=0}^{M} w_{kj} + w_{ko} \tag{13}$$

After the bias is absorbed:

$$a_k = \sum_{j=0}^{M} w_{kj}^{(2)} z_j \tag{14}$$

After transformation of Eq. (14) with nonlinear activation function:

$$y_k = \tilde{g}(a_k) \tag{16}$$

To demonstrate that the activation function at the output neuron is different from that of the hidden neurons, we used the symbol g (•) for its activation. Combining Eq. (12) through (16) will result in the full function seen in Fig. 3:

$$y_k = \tilde{g}\left(\sum_{j=0}^{M} w_{kj}^{(2)} g\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i \right) \right) \tag{17}$$

According to reasoning in (Marmolejo-Saucedo, 2020), sigmoid activation should be employed in the hidden layer instead of another nonlinear activation function because it will confine the output values of the network to a certain range. The sigmoid is also used because of its adaptability in differentiating.

3.2.5. Proposed novel generalized improved hybrid neural network (GINN)

Within this research scope, we propose using the Generalized Algorithm Improved Neural Network (GINN) approach for concurrently optimizing the parameters of neural network (NN) architecture (i.e hidden layers weights and biases). The general algorithm is used to iteratively evolve neural nets several times to lower the fitness functions. Before GA searches begin, starting random values are assigned to the connecting weights, bias, and neurons in the hidden layer. Because there are 14 predictors in our data, we use 14 neurons in hidden layers. The search parameters for the GA are saved in chromosomes, which are themselves constructed from thirty-two-bit strings. Genetic algorithm approaches are used to search for encoding chromosomes (f) to lower the fitness function. Because the significance of f varies according to the characteristics of the problem at hand, each application has to be tuned to perform optimally for a particular f. Within the scope of our application, one of the GA's key focuses is optimizing the neural network (NN) structure. Since it reflects a model's overall capacity for generalization, the WTI prediction accuracy metric is used to evaluate its performance in achieving its goals. Therefore, the f parameter considers the accuracy level that can be achieved while making predictions using the test dataset. As a result of the fact that the mean square error (MSE) is the benchmark against which crude oil price forecasting models are evaluated, we have decided to use it as our f value. In addition, it is suggested in the literature that MSE is more beneficial than other statistical indices when assessing the effectiveness of multiple models on the same dataset. This is because they take into account all of the variables involved. On the subject of the value of f, a conclusion is drawn using

$$f(x,y) = \text{MSE} = \frac{\sum_{j=1}^{N}(x(j) - y(j))^2}{N} \tag{18}$$

Where $x(j)$ is the value of the first observation in the dataset, $y(j)$ is the predicted value, and N is the total number of predictions generated by the model. The f is spelled with f(x, y). If the MSE is near zero, then the forecast will likely be accurate. In reality, zero-point predictions are uncommon. The suggested hybrid GINN search approach seeks to identify the optimal discretization parameters that yield the lowest possible prediction performance, much as the GA seeks to decrease f to achieve the ideal parameters. The initial chromosomes are subject to crossover and mutation until the threshold is met. The mutation and crossover rates and the total number of chromosomes in the population are all derived from (Wu and Xu, 2020). According to (Klein et al., 2018), this is the most accurate way to predict critical demographic variables, including population size, crossover rate, and mutation frequency. The experiments aim to determine the nature of the mutations, the crossovers, and the selection process at work. The genetic algorithm (GA) searches are programmed to end if the best minimal MSE is not improved upon after five generations.

3.3. Result evaluation parameters

Different models' results are discussed using standards such as R-squared, RMSE, MSE, and MAE. A brief explanation of these terms is given below:

We denote the predicted value with X_i, and the actual value *with* Y_i. The regression model estimates the X_i element for the corresponding $Y_i$element from our dataset. The mean of the true values can be written as:

$$\overline{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i (19)$$

mean total sum of squares can be defined as:

$$\text{MST} = \frac{1}{m}\sum_{i=1}^{m}(Y_i - \overline{Y})^2 \tag{20}$$

R-squared For a given set of independent variables, the proportion of their effect on the dependent variable is the R-squared value.

$$\text{R}^2 = 1 - \frac{\sum_{i=1}^{m}(X_i - Y_i)^2}{\sum_{i=1}^{m}(\overline{Y} - Y_i)^2} \tag{21}$$

(worst value $= -\infty$; best value $= +1$).

Mean Square Error (MSE) MSE is a tool that may be used to find abnormalities. It is evident that the squaring element of the function will exacerbate the error if the model finally yields a single bad forecast. Since the relationship between R^2 and MSE is monotonic (a negative monotonic link), ranking regression models by R^2 is likely to provide comparable results (albeit in reverse order).

$$\text{MSE} = \frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2 (22)$$

(worst value $= +\infty$; best value $= 0$).

Root mean square error (RMSE) Square root of MSE is RMSE and can be written as:

$$\text{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(X_i - Y_i)^2} \tag{23}$$

(best value $= 0$; worst value $= +\infty$).

Mean Absolute Error (MAE) It measures the median magnitude of mistakes over a set of forecasts without considering the trend of those errors. The mean absolute difference is the weighted average of the individual anomalies between the forecast and the actual observation across the entire test sample.

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}|X_i - X_i| \tag{24}$$

(best value $= 0$; worst value $= +\infty$).

4. Results and discussion

4.1. Results of ARIMA model

We implement the ARIMA model using MATLAB R 2022b and Gnu Regression Econometrics and Time-series Library (GRETL). Value of the coefficient and p-values are given in Table 3. By analyzing the values, we reject the null hypothesis for Dow Jones Index, SP500, Gold, Silver Futures, Platinum Price, and Oil ETF USO which means these variables have a significance effect on response variable whereas Eldorado Gold Corporation, EUR-USD Exchange rate US Bond Rate, Palladium price, US dollar Index Price, Gold Miners ETF and Rhodium Prices do not have significant impact on value of response variable (WTI Crude Oil Price).

Moreover, the coefficient value of Rhodium Prices is the highest, whereas the coefficient value of Silver Futures is the lowest among all predictors indicating the maximum change in response variable (WTI Crude Oil) value by changing the value of Rhodium Prices and minimum change in the value of response variable by changing the value of Silver Futures.

Fig. 4 shows the actual plot of crude oil price and the plot of the predicted price in different colors. Both the lines (actual vs predicted) are closer to each other from 2012 to 2016. From 2016 to 2019, prediction performance of the model degraded considerably as depicted in Fig. 4. This indicates that the model predicted values from 2012 to 2016 more accurately than 2016–2019.

The value of MAE in Table 4 is 5.1911, indicating a modest difference between the model's fitted and observed values. Due to the model's poor performance between 2016 and 2019, the MAE value reached 5.1911. Fig. 5 shows how well the model predicts various response values. The predicted response of a perfect regression model is identical to the actual response thus, all the points fall on the diagonal line. Predicted Points on

Table 3
ARIMA model using observations 2011–2018 (T = 1718) Dependent variable WTI Crude Oil.

Variables	Coefficient	Std. Error	Z	p-value	
Const	18.1997	23.4805	0.7751	0.4383	
phi_1	0.999175	0.000760051	1315.	<0.0001	***
theta_1	−0.557641	0.0218534	−25.52	<0.0001	***
Gold	−0.131414	0.0258975	−5.074	<0.0001	***
SP500	−0.215245	0.0213850	−10.07	<0.0001	***
DowJonesIndex	0.00289881	0.000235459	12.31	<0.0001	***
EldoradoGoldCorporation	−0.0323457	0.0240203	−1.347	0.1781	
EURUSDExchangerate	3.17691	8.59458	0.3696	0.7117	
SilverFutures	0.000166780	3.62812e-05	4.597	<0.0001	***
USBondRate	−0.294886	0.446927	−0.6598	0.5094	
PlatinumPrice	0.00523757	0.00169592	3.088	0.0020	***
Palladiumprice	0.00156985	0.00166627	0.9421	0.3461	
RhodiumPrices	−9.52797e-05	5.64604e-05	−1.688	0.0915	*
USdollarIndexPrice	−0.168200	0.135395	−1.242	0.2141	
GoldMinersETF	0.0137093	0.0586335	0.2338	0.8151	
OilETFUSO	2.52570	0.0538180	46.93	<0.0001	***

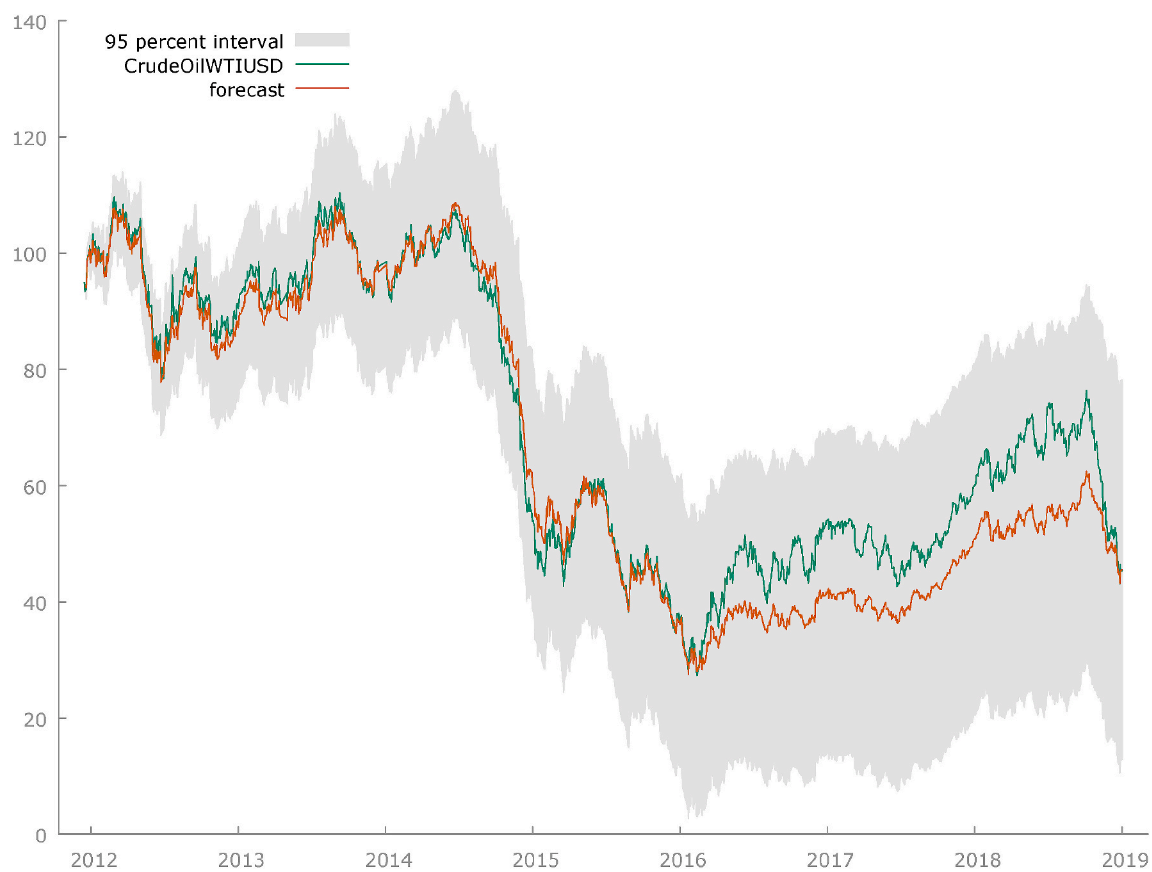

Fig. 4. Actual values of response variable vs predicted values.

Table 4
Prediction performance parameters of ARIMA Model.

Result parameter	Value	Result parameter	value
dependent variable mean	70.27540	S.D. dependent variable	23.48005
R^2	0.998792	Adjusted R^2	0.998782
Log	−2093.282	Akaike criterion	4220.564
Schwarz criterion	4313.196	Hannan-Quinn	4254.839
Mean Error	4.0901	RMSE	6.7931
MAE	5.1911	MPE	7.2519
MAPE	8.9014	Theil's U2	5.7439
Proportion - bias	0.47182	Proportion- Regression	0.26202
Proportion - disturbance	0.26616		

the left or right of the plot, which are furthest from the mean, exert the most significant force on the fitted line, attempting to drag it towards a point away from the diagonal line. Possible outliers are points that are far from the line vertically. Both sorts of points might affect the fit negatively. MAPE = 8.9014 represents a lower average absolute difference (in percentage) between fitted model values and observed values. RMSE = 6.7931 shows residual (distance between anticipated and actual values) deviation that explains the residuals' dispersion around the best-fit line. Using Euclidean distance, it presents the distance between the predicted values and the true values. Fig. 6 shows the residuals produced from the model. The residuals plot illustrates the deviation between the expected and actual responses. Generally, symmetrical distribution of residuals about 0 indicates a good model. If the

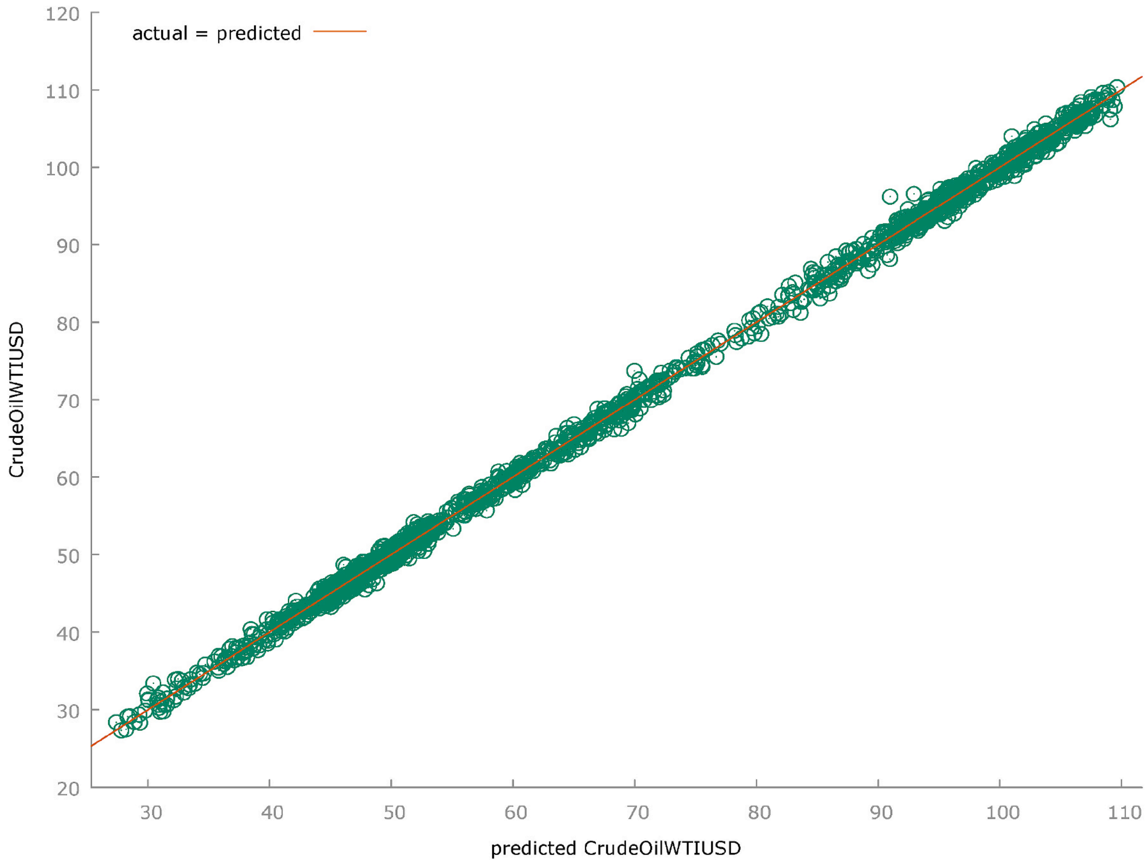

Fig. 5. Predicted fitted values vs actual values.

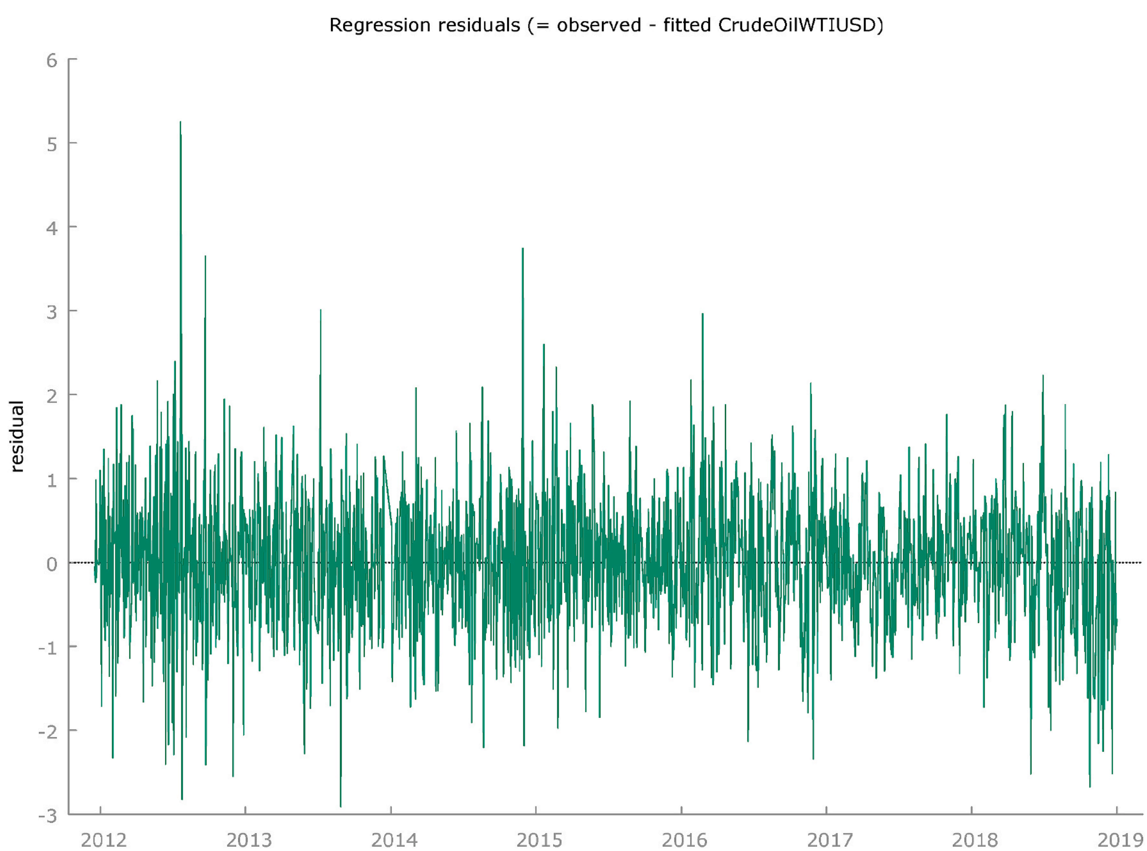

Fig. 6. Regression residuals from ARIMA model.

residuals exhibit a distinct pattern, it means the model needs improvement. The following may be concluded from Fig. 6 as there is no predictable pattern in the residuals:

- Residuals are symmetrically distributed around 0
- The amount of residuals does not vary considerably from left to right in the plot.
- Outliers are uncommon in the plot.

The autocorrelation function (ACF) evaluates the correlation between observations in a time series given a collection of delays. By calculating the autocorrelation function, we can test for statistically significant correlations between delays to assess the data's randomness and stationarity. Our data show no obvious trends or seasonal patterns, as seen by the plot in Appendix-1. Each bar in an ACF plot reflects the magnitude and direction of the association. Statistical significance is shown by a bar that exceeds the dotted line. Plot in Appendix-1 demonstrates that autocorrelations are close to zero for all delays. This occurrence is known as white noise. The absence of a substantial lag in the ACF indicates that the data is random and there is no particular pattern in the data. We may also infer from the graph that the data are stationary, meaning there is no trend, constant variance, autocorrelation pattern, or seasonal pattern in the time series. Keeping in view the scope of this research and to keep the analysis short, the following tables/plots are given in Appendix-1:

- Last 50 values of Actual, Fitted, and Residuals
- Normality test of residuals
- Coefficient covariance matrix

4.2. Results of the GARCH model

We implement the GARCH model using MATLAB R 2022b and Gnu Regression Econometrics and Time-series Library (GRETL). Table 5 shows the t-ratio and p-value of the GARCH model. By analyzing values in Table 5, we reject the null hypothesis for all predictors except Eldorado Gold Co and Rhodium Prices, which means all predictors significantly contribute to the response value (crude oil price) except Eldorado Gold Co and Rhodium Prices. Moreover, the coefficient value of the US Bond Rate is the highest whereas the coefficient value of Rhodium Prices is the lowest among all predictors indicating the maximum change in response value (WTI crude oil price) by changing the value of the US Bond Rate and minimum change in the value of response variable (WTI crude oil price) by changing value of Rhodium Prices.

Fig. 7 shows the plot for actual price and predicted price of WTI crude oil. Both the lines (actual vs predicted) are close to each other; however, they are not overlapping (ideal condition) each other. A relatively close predicted line with the actual line indicates better prediction results of the model; however, not ideal results.

Table 6 shows the performance parameters of the GARCH model. The value of MAE = 4.5027 indicates a relatively small difference between the fitted and observed values of the model. The same has also been shown in Fig. 8, which indicates all predicted values lie fairly close to the diagonal line. It may be noted that all the points fall on the diagonal line for a perfect regression model(L. Chang et al., 2022a)(Chang et al., 2023b). However, if the predicted Points lie on the left or right of the plot, they attempt to drag the prediction line towards a point away from the diagonal line. Possible outliers are another reason for the movement of prediction points away from the diagonal line. Both points might affect the fit negatively. The value of MAPE = 7.0454 indicates a minor average absolute difference (in percentage) between the fitted values of the model and the observed values. RMSE value of 5.8877 marks the deviation of the residuals (distance between predicted and actual values) and explains the spread of residuals around the best fit line. Keeping in view the scope of this research, and to keep analysis short, following tables/plots are given in Appendix-1:

- Last 50 values of Actual, Fitted and Residuals
- Normality test of residuals
- Coefficient covariance matrix

4.3. Results of Ordinary Least Squares (OLS)

We implement the OLS model using MATLAB R2022b and Gnu Regression Econometrics and Time-series Library (GRETL). Table 7 shows the values of the t-ratio and p-value. By analyzing values in Table 7, we reject the null hypothesis for all predictors except Gold prices, which means all predictors significantly contribute to predicting the response variable's future value on the response value (crude oil price) except gold prices. Moreover, the coefficient value of the EUR-USD Exchange rate is the highest, whereas the coefficient value of silver future is the lowest among all predictors indicating the maximum change in response variable value (WTI crude oil price) by changing the value of the EUR-USD Exchange rate and minimum change in the value of response variable by changing the value of silver future.

Fig. 9 shows the plot of actual price and predicted price of crude oil. A close predicted line with the actual price line indicates good prediction results of the model. Table 8 shows R^2 value of 0.993 which indicates that variance in the crude oil prices is predictable from the independent variables. Value of MAE is 1.526 which indicates small difference between the fitted values and observed values of the model. Same has also been shown in Fig. 10 which indicates all predicted values lie close to diagonal line. Value of MAPE = 2.448 indicates a smaller average absolute difference (in percentage) between fitted values of the model and the observed values. RMSE value of = 1.9 indicates a smaller deviation of the residuals (distance between predicted and actual values) which is also shown in Fig. 11 to prove that the spread of residuals around best fit line is tight without much outliers.

The leverage/influence graph in Fig. 12 demonstrates that all data

Table 5
GARCH model using observations 2011–2018 (T = 1718) Dependent variable WTI Crude Oil.

Variables	Coefficient	Std. Error	Z	p-value	
Const	200.534	27.2876	7.349	<0.0001	***
GoldPrice	−0.782616	0.0316201	−24.75	<0.0001	***
SP500	−0.225597	0.0349519	−6.455	<0.0001	***
DowJonesIndex	0.00203696	0.000439042	4.640	<0.0001	***
EldoradoGoldCō	**−0.0428818**	**0.0336366**	**−1.275**	**0.2024**	
EURUSDExchange~	−22.4143	10.8721	−2.062	0.0392	**
SilverFutures	0.000428565	4.34910e-05	9.854	<0.0001	***
USBondRate	19.6940	0.601947	32.72	<0.0001	***
PlatinumPrice	0.0440027	0.00250199	17.59	<0.0001	***
Palladiumprice	0.0119095	0.00243349	4.894	<0.0001	***
RhodiumPrices	**0.000135990**	**0.000162849**	**0.8351**	**0.4037**	
USdollarIndexP~	−1.61596	0.166108	−9.728	<0.0001	***
GoldMinersETF	1.01206	0.0417168	24.26	<0.0001	***

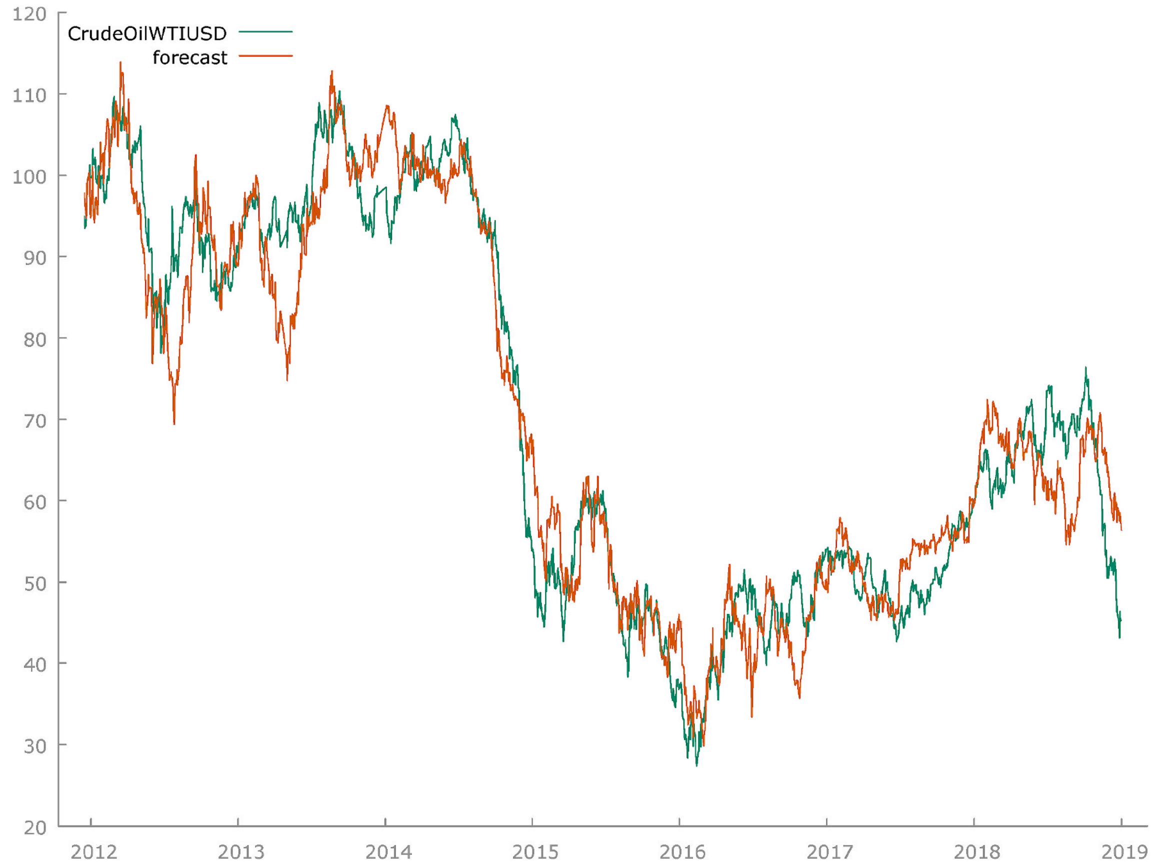

Fig. 7. Actual and predicted values of the response variable (WTI crude oil) obtained from the GARCH model.

Table 6
Prediction performance parameters of the GARCH Model.

Result Parameter	Value	Result Parameter	Value
Mean Error	0.3516	RMSE	5.8877
MAE	4.5027	MPE	−0.43429
MAPE	7.0454	Theil's U2	4.4787
Proportion - Bias	0.0035662	Proportion - regression	0.0011362
Proportion - disturbance	0.9953		

points conform to the overall trend of the remaining data; there are no outliers or extreme values. The graph indicates that there are no points of significant leverage. Overall, none of the data points have any bearing on the position of the best-fitting line. The following tables/plots are given in Appendix-1:

- Last 50 values of Actual, Fitted and Residuals
- Normality test of residuals
- Coefficient covariance matrix

4.4. *Results of Artificial Neural Networks and GINN*

ANNs and GINN are implemented using MATLAB R2022b. Pseudo code of proposed GINN regression model is given as under:

Table 9 shows the validation and testing results of Neural Networks, whereas Fig. 13 shows the flow diagram of the proposed GINN. Validation results from Table 9 indicate R-squared value of all ANNs is 1 (ideal predictions) except time series ANN. It may be noted that the R-squared value of time series ANN is also close to 1 (0.9972). Values of R-squared indicate that variance in the crude oil prices is predictable from the independent variables. MAE value of single-layer ANN is the lowest (0.83695) amongst layered and time series ANN whereas it has been further reduced by GINN (0.9148). The MAE value of time series ANN is the highest (1.3654), implying that the proposed GINN model has the smallest difference between the fitted and observed values. In contrast, time series ANN has the highest difference between fitted and observed values. RMSE value of two-layered ANN is the lowest (0.8737) amongst layered and time series ANN which is further reduced by GINN (0.8737), whereas the RMSE value of time series ANN is the highest (1.2261).

Comparison of RMSE values indicates a smaller deviation of the residuals (distance between predicted and actual values) for GINN and a relatively greater distance between residuals of time series ANNs. Fig. 14 shows the plot of true response and predicted response against single layered and double layer ANNs, whereas Fig. 15 represents the plot of true response and predicted response of tri-layered and GINN models. Keeping in mind the fact that the predicted response of a perfect regression model is identical to the actual response and in ideal conditions, all points shall fall on the diagonal line. Figs. 14 and 15 indicate that in the case of layered ANNs (single layered, bi-layered and tri-layerd ANNs), some of the predicted points lie on the left or right of the mean line due to which they dragged the predicted line/points away from the diagonal line. However, the plot of the proposed GINN in Fig. 15 indicates a better fit than layered ANNs, as all predicted points seem to lie close to the diagonal line with fewer outliers.

After training and validation, the model was tested on the dataset to check the prediction accuracy of the model. If we compare the test results of Neural networks, the R-squared values of all the models remained the same as those obtained during validation. The test RMSE value of bi-layer ANN is the lowest (0.9148) amongst layered and time series ANNs, which is further reduced (0.8737) by GINN, whereas the RMSE of time series ANN is 1.2261. RMSE values comparison reveals a smaller deviation of the residuals (distance between predicted and actual values) for GINN and a relatively greater distance between residuals of time series ANNs during the testing of the model. MAE (test)

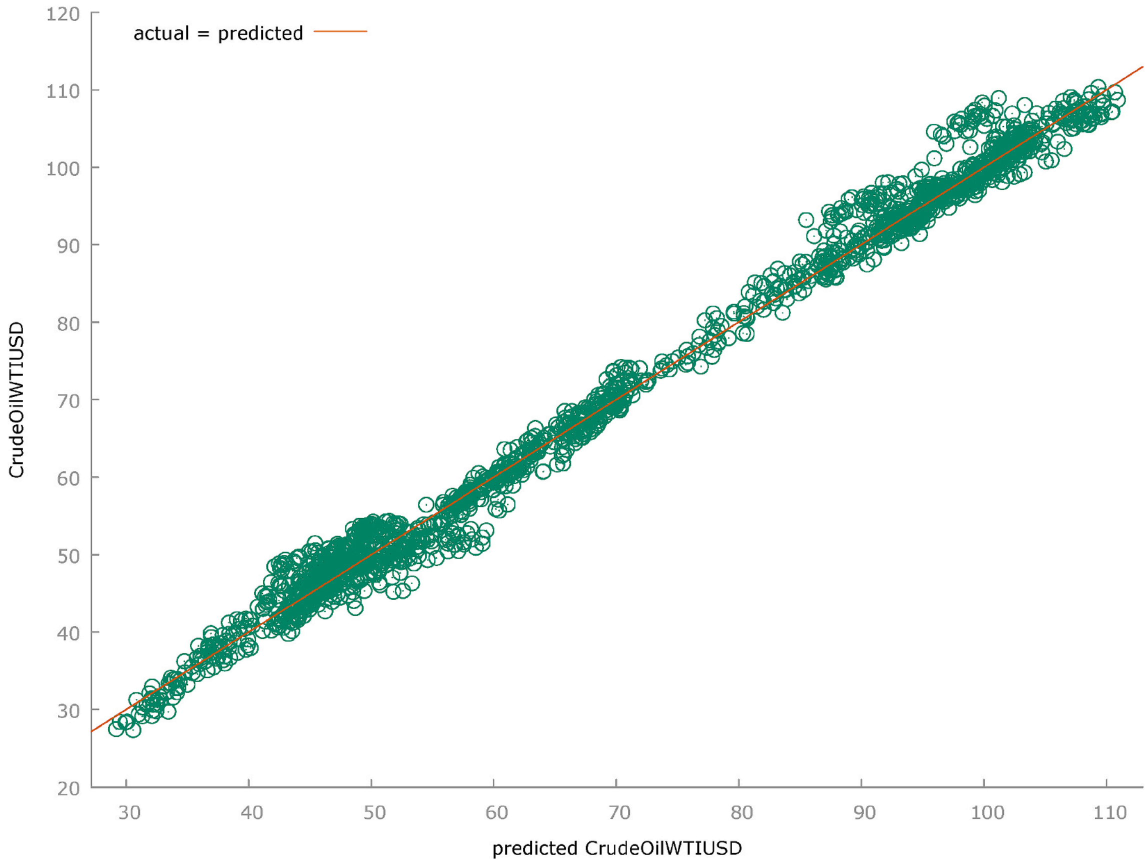

Fig. 8. Actual values and predicted values from GARCH model.

Table 7
OLS model using observations 2011–2018 (T = 1718) Dependent variable WTI Crude Oil.

Variable	Coefficient	Std. Error	t-ratio	p-value	
const	−188.969	13.7821	−13.71	<0.0001	***
GoldPrice	−0.0408529	0.0216662	−1.886	0.0595	*
SP500	−0.0797892	0.0153908	−5.184	<0.0001	***
DowJonesIndex	0.00299866	0.000204024	14.70	<0.0001	***
EldoradoGoldCorporation	−0.288474	0.0145522	−19.82	<0.0001	***
EURUSDExchangerate	63.5561	5.16733	12.30	<0.0001	***
BrentCrudeoilFutures	0.246070	0.0157419	15.63	<0.0001	***
SilverFutures	0.000386292	2.83563e-05	13.62	<0.0001	***
USBondRate	−1.23534	0.333013	−3.710	0.0002	***
PlatinumPrice	−0.00565379	0.00105860	−5.341	<0.0001	***
Palladiumprice	−0.0193553	0.000964055	−20.08	<0.0001	***
RhodiumPrices	−0.00108667	0.000129833	−8.370	<0.0001	***
USdollarIndexPrice	0.997922	0.0833503	11.97	<0.0001	***
GoldMinersETF	0.398390	0.0277669	14.35	<0.0001	***
OilETFUSO	2.13120	0.0491641	43.35	<0.0001	***

value of bi-layered ANNs remained the lowest (0.97599) among layered and time series ANN which is further reduced by GINN (0.6634). MAE value of time series ANN remained the highest. MAE value analysis indicates that GINN models have the smallest difference between the fitted values and observed values, whereas time series ANN has the highest difference between fitted values and observed values(L. Chang et al., 2022f). Fig. 16 shows the true response and predicted response of single-layered ANNs against residuals, whereas Fig. 17 shows the true response and predicted response of single layered ANNs against residuals. Residuals in Fig. 17 show a more symmetrical distribution of about 0 which is an indication of a superior model. Fig. 18 indicates the response plot of GINN, which offers the line for predicted prices almost overlaps the line for true prices, indicating a good prediction model.

4.5. *Comparison/analysis of results of all implemented models*

Table 10 shows the comparison of the results of all models discussed in this paper. If we compare the RMSE value of econometric functions, the OLS model outperformed ARIMA and GARCH models. Overall, we can see that the RMSE value of GINN is the lowest, whereas the RMSE value of the ARIMA model is the highest amongst all models indicating the smallest deviation of the residuals (distance between predicted and actual values) for GINN and the highest deviation of the residuals for ARIMA model. MAE values comparison shows the lowest value for the GINN model (0.6634) and the highest value for the ARIMA model (5.1911). From the comparison of results, we conclude that the prediction performance of the proposed neural network GINN remained better than econometric and existing neural networks models.

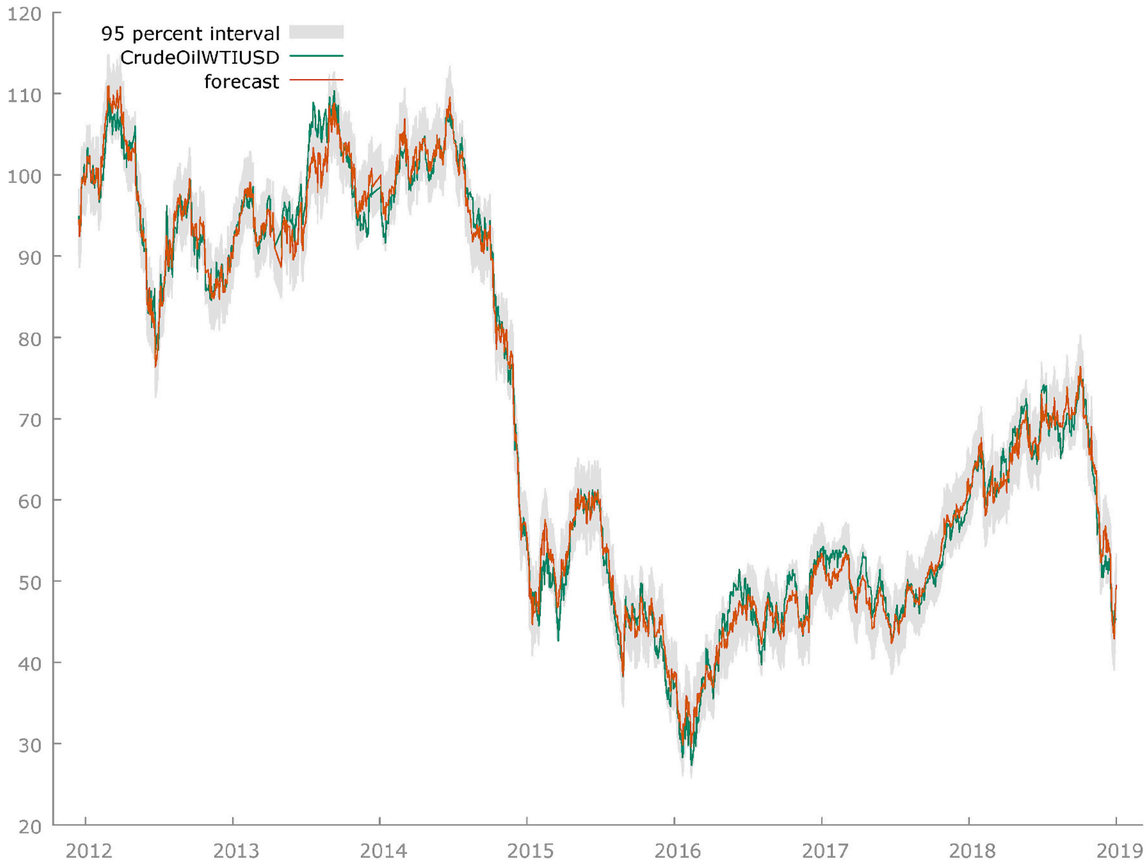

Fig. 9. Actual and predicted values of response variable (WTI crude oil) obtained from OLS model.

Table 8
Prediction performance parameters of OLS Model.

Result Parameter	Value	Result Parameter	Value
Dependent variable mean	70.27540	S.D. dependent variable	23.48005
R^2	0.993394	Adjusted R^2	0.993340
F(14, 1703)	18292.45	P-value (F)	0.000000
Log	−3547.511	Akaike criterion	7125.022
Schwarz criterion	7206.756	Hannan-Quinn	7155.264
Mean Error	−3.7564e-013	Root Mean Squared Error	1.9078
MAE	1.5267	Mean Percentage Error	−0.13079
MAPE	2.4487	Theil's U2	1.5378
Proportion - bias	3.1958e-026	Proportion- regression	7.3009e-024
Proportion - disturbance	1		

5. Conclusion and policy implications

5.1. Conclusion

In the global economy, crude oil is becoming more and more significant. To forecast WTI crude oil prices, we proposed a novel Generalized Algorithm Improved Neural network model from economic history. We suggested optimizing the weights, bias, and topology of the existing neural network. We also present the performance analysis of conventional econometric models (ARIMA, GARCH, and OLS), Artificial Neural Network (single-layer, bi-layer, and tri-layer ANNs) regression models, and ANN Time Series models and compare their results with our proposed model to find out best-performing method (time series or regression) and the best model (econometric or machine learning model). We used historical prices of 14 different variables, including gold, silver, S&P500, USD Index price, and US-EU conversion rates for regression models, whereas historical time series data of WTI crude oil for time series models. Results are compared using RMSE, R-Squared, and MAE values. Analysis of the results reveals that the performance of our proposed model remained better than all tested models. The comparative results of existing models show that the overall performance of Neural Networks remained better than econometric models. Moreover, the performance of regression neural networks remained better than the time series network for predicting WTI crude oil prices due to the non-linearity of time series data by viewing economic history. The model suggested in this research has the potential to increase the forecast accuracy of WTI crude oil prices significantly and is appropriate for prediction and interpretation tasks in other disciplines.

5.2. Policy implications

The results of this study will have a wide range of policy implications for export countries, financiers, decision-makers in politics, and market regulators. It will be helpful in the decision-making process at different management levels in both the public and private sectors since the price of crude oil plays a significant role in choices about development and industrial production by using evidences from economic history. Making better judgements is implied by an accurate understanding of the future behavior of crude oil price changes. Future work may include the selection of more macro-economic and socio-economic variables to enhance the performance of the predictions relevant to the crude oil market using various macroeconomics and financial indicators.

Author statement

Zilin Xu, Muhammad Mohsin, kaleem Ullah, Xiaoyu Ma: In the process of writing the article, the first author conceived in the idea of the

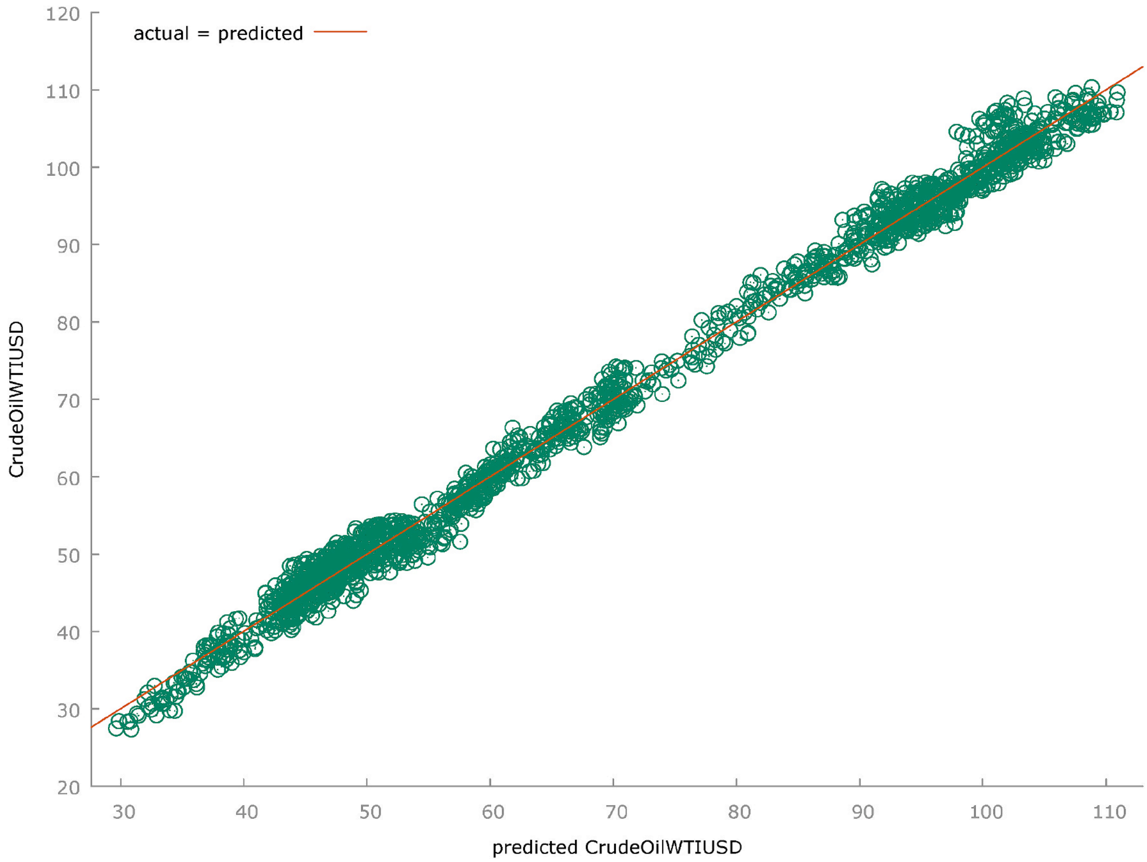

Fig. 10. Actual values and predicted values from OLS model.

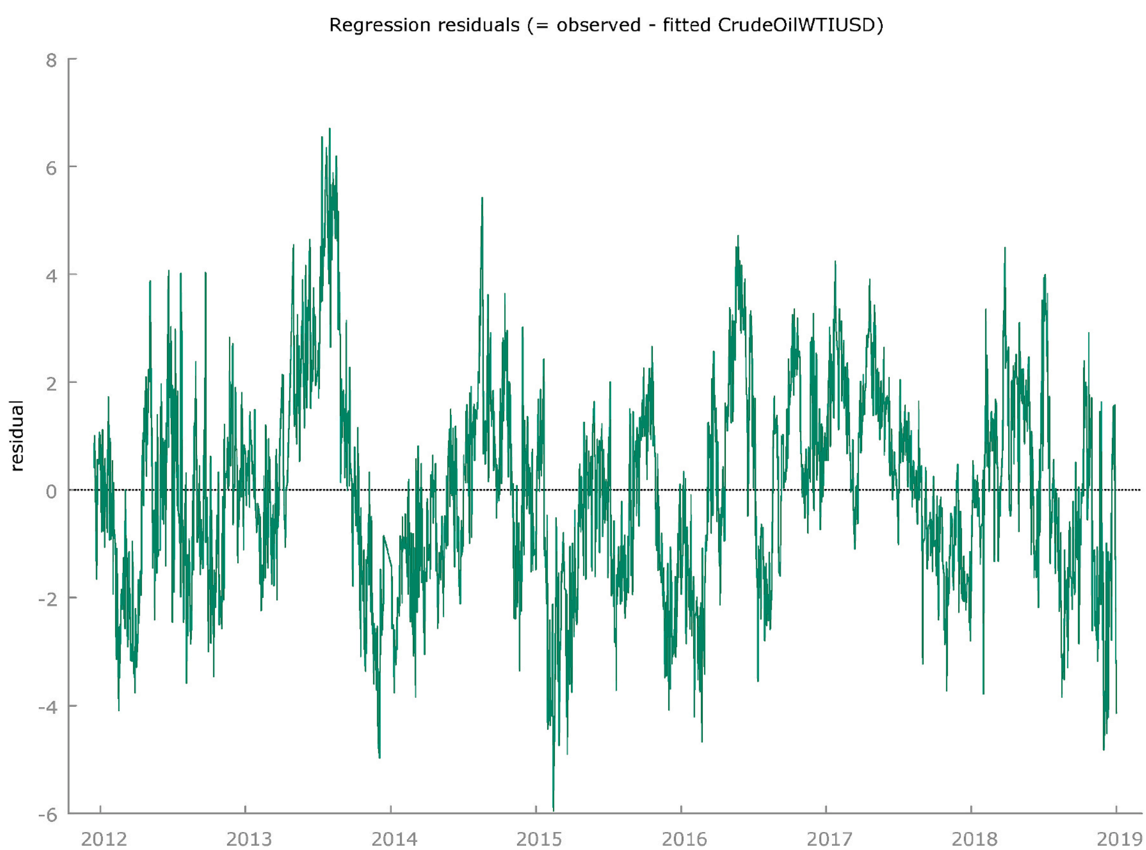

Fig. 11. Regression residuals from OLS model.

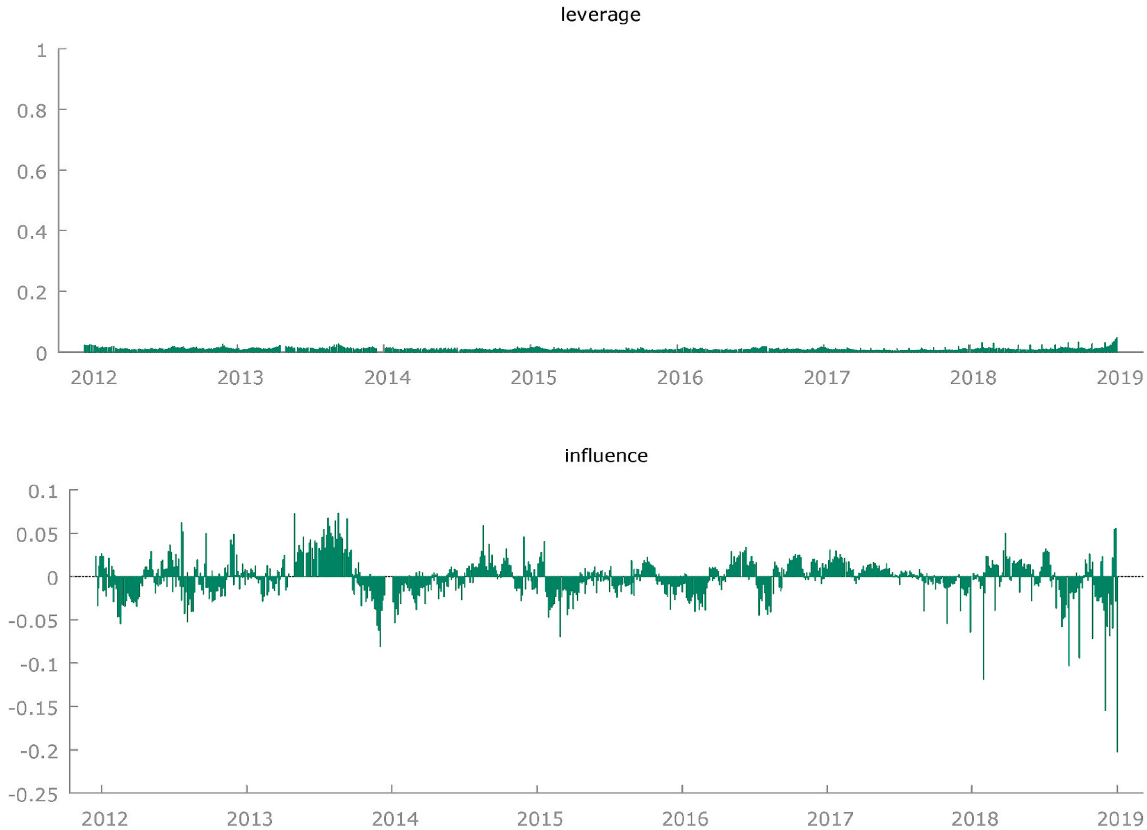

Fig. 12. Regression residuals from OLS model.

```
Load training data file in MATLAB Directory/ work space
Input: 70 percent of dataset (training Data;
Define predictors (from training data)with this function: predictors = inputTable(:,
predictorNames);
Define response variable with this function: response = inputTable.response variable
define whether predictor is Categorical or not?
Train a regression model using regressionNeuralNetwork = fitrnet(...
Define layer size
Define activation function
Define iteration limit
Define weights
Define biases
Create the result structures with predict function
Add additional fields to the result structure as require
Carryout cross-validation by using:
partitionedModel = crossval(trainedModel.RegressionNeuralNetwork, 'KFold', 14);
Calculate validation predictions by using:
validationPredictions = kfoldPredict(partitionedModel);
Calculate validation RMSE by:
validationRMSE = sqrt(kfoldLoss(partitionedModel, 'LossFun', 'mse'));
```

article, implementation, and the collation and analysis of data, and the; **Zilin Xu, Muhammad Mohsin:** carried out the implementation of relevant methods, project support, data collection and proofreading, **kaleem Ullah, Xiaoyu Ma**: Contributed in implementation, Reviewing, Editing, Monitoring.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All of the authors consented to publish this manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 9
Prediction performance parameters of ANNs.

Parameter	Single layer ANN	Bi-Layer ANN	Tri-Layer ANN	GINN	Time Series ANN
Validation					
RMSE (Validation)	1.0974	1.084	1.1635	1.0017	1.2592
R-squared (Validation)	1.00	1.00	1.00	1.00	0.9972
MSE (Validation)	1.1651	1.1673	1.3538	1.0034	1.5855
MAE (Validation)	0.83695	0.8426	0.87744	0.7682	1.3654
Testing (Test)					
RMSE (Test)	0.98792	0.9148	0.98834	0.8737	1.2261
R-squared (Test)	1.00	1.00	1.00	1.00	0.9972
MSE (Test)	0.97599	0.83628	0.97683	0.7634	1.5033
MAE (Test)	0.78712	0.70368	0.76243	0.6634	1.2322
Prediction Speed (Obs/Sec)	21000	73000	120000	145000	–
Training Time (Sec)	4.8862	8.2011	20.22	3.234	–

Data availability

Supplementary data has been included

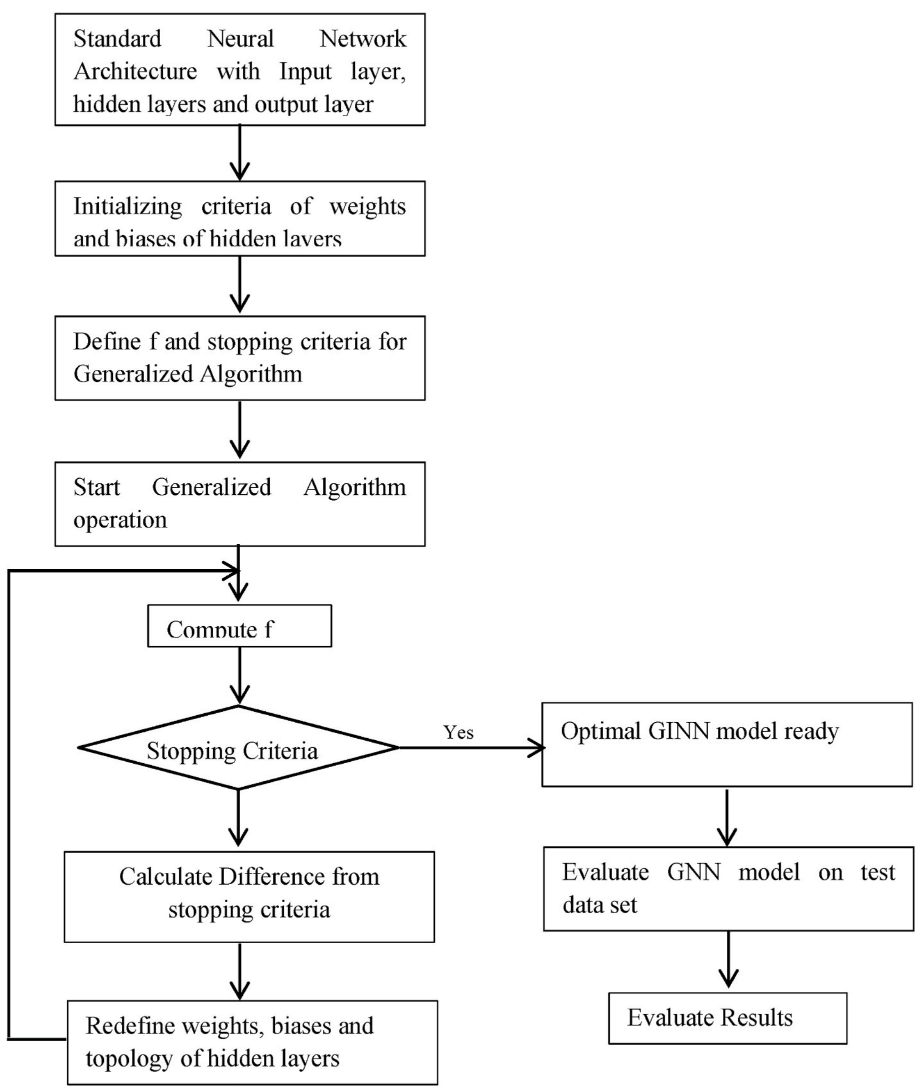

Fig. 13. Work flow of proposed GINN model.

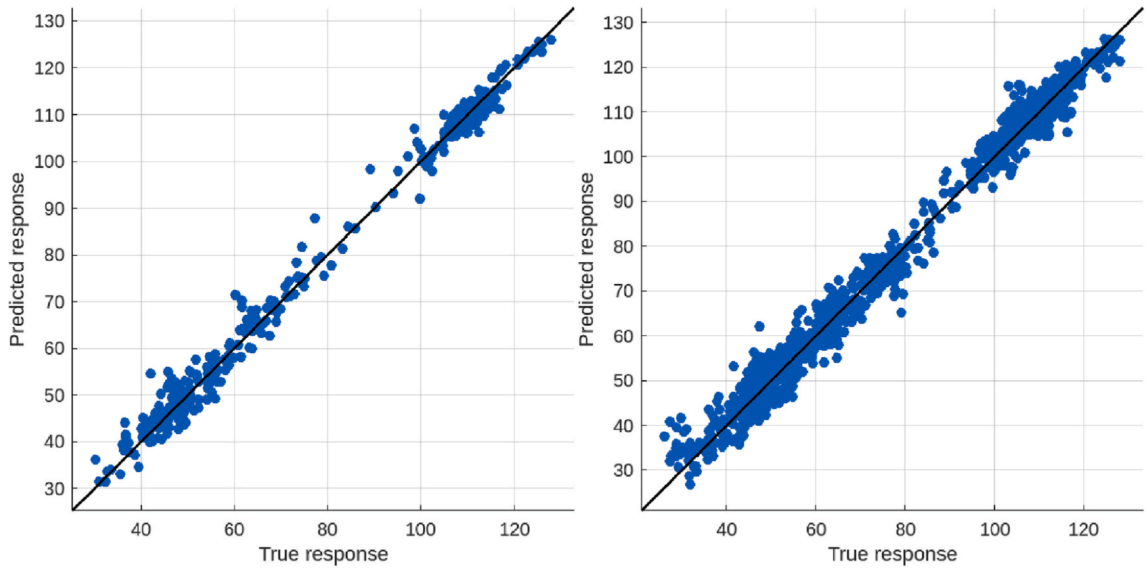

Fig. 14. True response and prediction response of test results of single layer and bilayer ANNs.

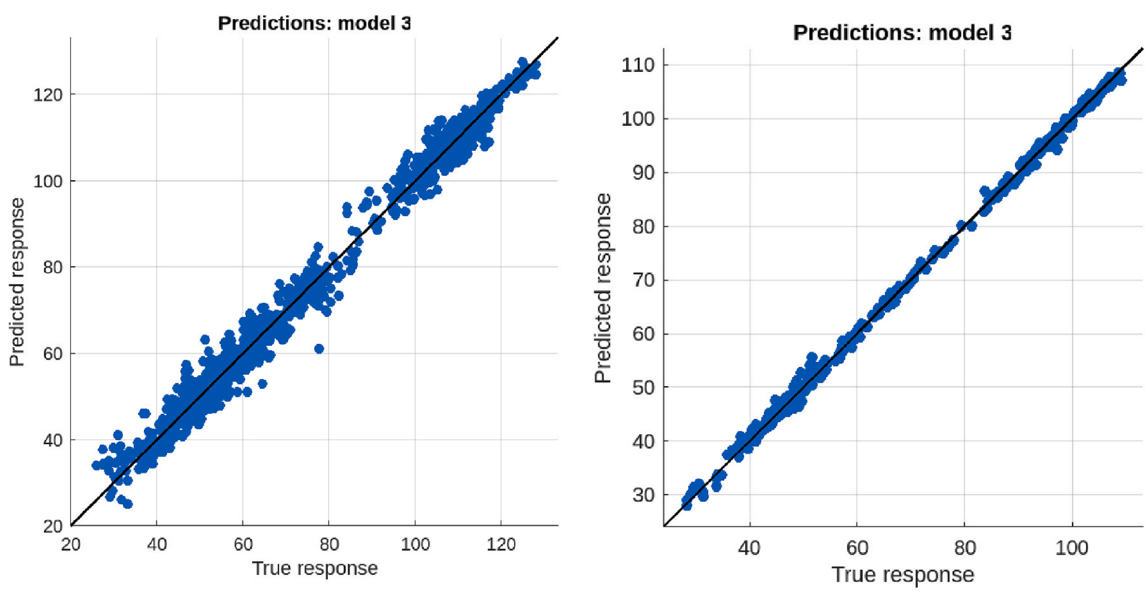

Fig. 15. True response and prediction response of test results of tri layer ANNs and GINN.

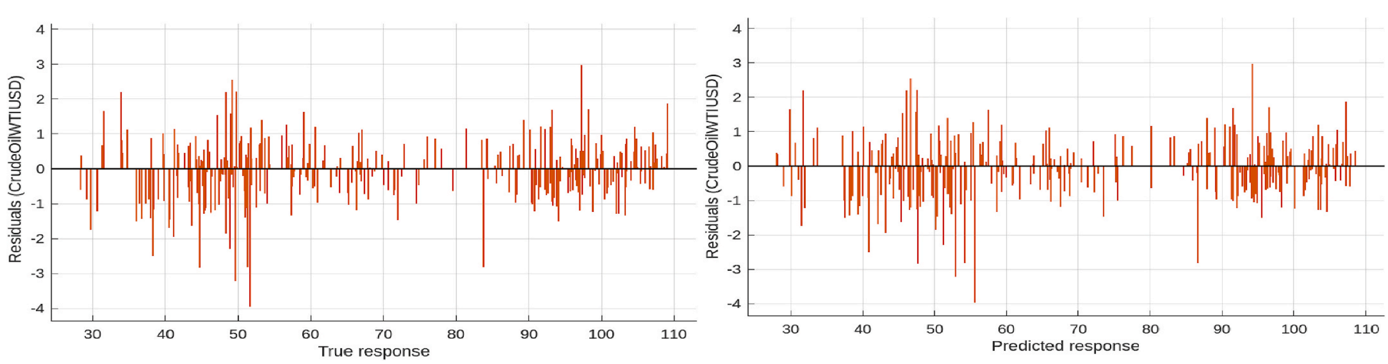

Fig. 16. True response and predicted response against residuals for single layered ANNs.

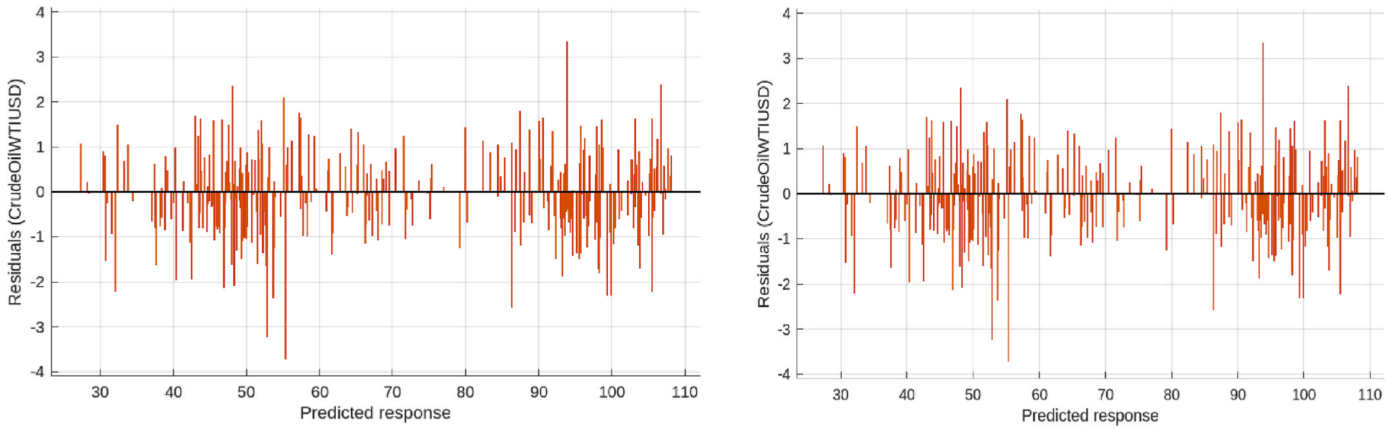

Fig. 17. True response and predicted response against residuals for GINN.

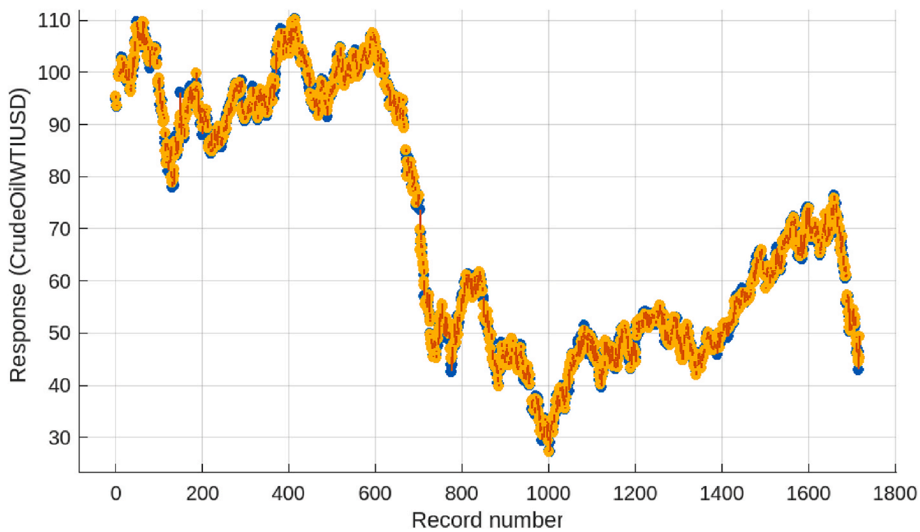

Fig. 18. Response plot of GINN indicating true line (blue color) and predicted line (yellow colour).

Table 10
Comparison of results of econometric vs ANNs.

Model Type	RMSE	R-Squared	MAE
Neural Network Single layer	0.98792	1.00	0.78712
Neural Network Bi layer	0.9148	1.00	0.70368
Neural Network Tri layer	0.98834	1.00	0.76243
GINN	0.8737	1.00	0.6634
Time series ANN	1.2261	0.9972	1.2322
ARIMA	6.7931	0.998792	5.1911
GARCH	5.8877	0.99734	4.5027
OLS	1.9078	0.993394	1.5267

Appendix-1

ADDITIONAL DATA OF ECONOMETRIC MODELS

1. Auto-Correlation plot of ARIMA Model is given as under:

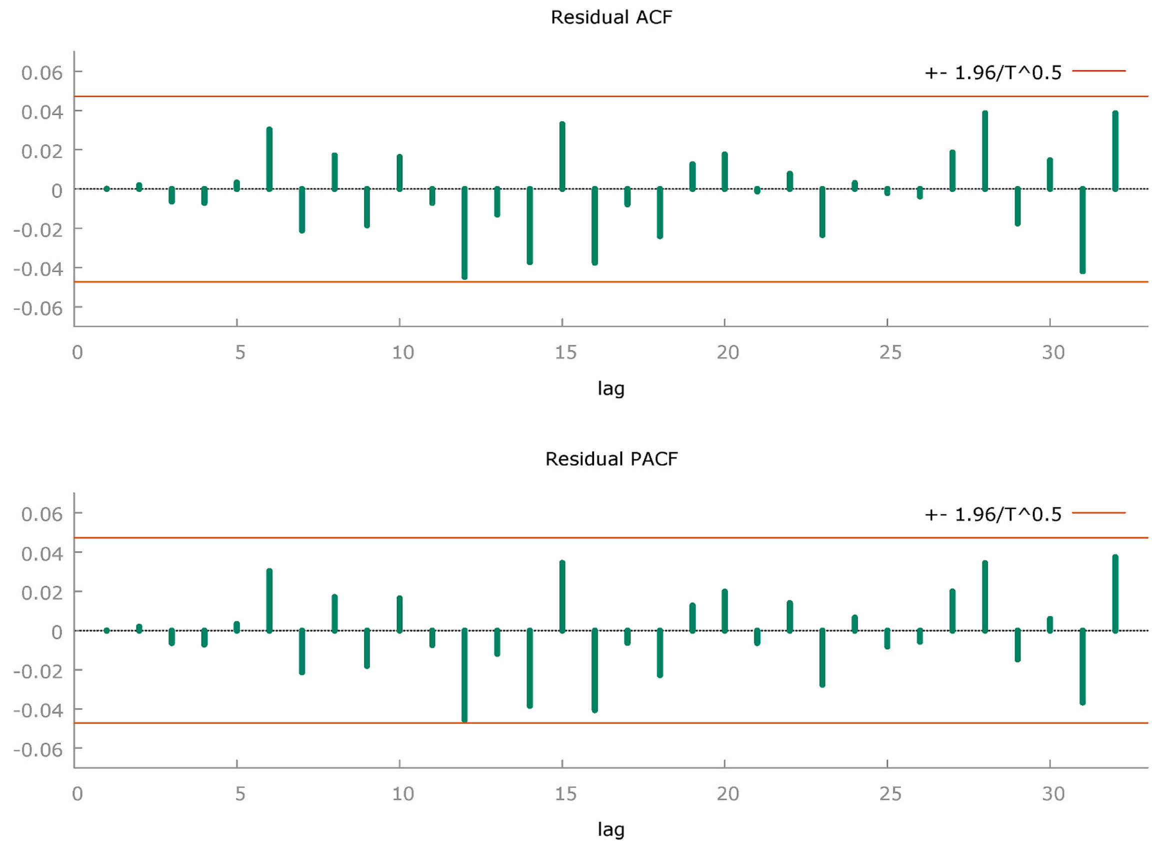

Fig. 19. Autocorrelation plot of ARIMA model

Table. 11
Last 50 values of Actual, Fitted and Residuals of econometric Models

		ARIMA		GARCH		OLS	
Date	WTI Crude Oil Actual Value	fitted	residual	fitted	residual	fitted	residual
2018-09-04	70.6000	69.8459	−0.565906	70.6246	−0.0246489	68.9714	0.308645
2018-09-05	69.8400	69.1685	−0.548542	69.8538	−0.0137889	68.9704	−0.350413
2018-09-06	69.0100	69.1825	0.547519	69.9453	−0.935270	69.5174	0.212626
2018-09-07	68.3300	69.7340	−0.504030	68.9165	−0.586456	70.1124	−0.882410
2018-09-10	68.3900	69.4264	0.0736387	69.7032	−1.31321	70.6151	−1.11514
2018-09-11	67.8400	69.1687	0.621279	69.2850	−1.44500	69.5433	0.246708
2018-09-12	70.1100	68.7519	−1.45194	70.7809	−0.670922	69.0941	−1.79415
2018-09-13	70.4200	68.1236	−0.883636	70.3961	0.0239187	68.9018	−1.66176
2018-09-14	69.0600	67.3643	0.825687	70.0946	−1.03461	68.5664	−0.376355
2018-09-17	68.8600	67.8440	−0.113971	70.5161	−1.65607	68.8245	−1.09451
2018-09-18	68.6200	66.2043	−1.11435	70.2517	−1.63173	68.0769	−2.98689
2018-09-19	69.8800	66.9844	−1.40439	70.4176	−0.537632	68.1328	−2.55280
2018-09-20	71.4100	66.2149	−0.354935	71.0256	0.384382	68.0976	−2.23764
2018-09-21	70.5600	66.7751	−0.465074	71.1483	−0.588344	68.5719	−2.26186
2018-09-24	71.9200	67.2021	−0.312104	72.3919	−0.471883	69.6191	−2.72912
2018-09-25	72.4700	66.9388	1.88120	72.5135	−0.0434620	69.0982	−0.278207
2018-09-26	72.2500	67.9387	0.731300	72.3192	−0.0692148	69.7632	−1.09316
2018-09-27	72.4400	68.6060	0.764015	71.9122	0.527826	69.9757	−0.605708
2018-09-28	72.4200	69.4779	0.182121	72.5349	−0.114879	70.3339	−0.673924
2018-10-01	73.7300	69.6028	−0.382810	73.5749	0.155070	70.8897	−1.66965
2018-10-03	75.4100	69.8623	0.447666	75.1910	0.219023	70.9128	−0.602764
2018-10-04	76.4100	70.1698	0.500236	75.8798	0.530187	73.4587	−2.78874
2018-10-05	74.9500	70.3127	0.287340	74.6461	0.303929	70.4663	0.133725
2018-10-08	74.0300	69.6606	0.179372	73.6055	0.424543	69.4829	0.357143
2018-10-09	74.4400	69.8464	−0.836415	74.0210	0.418974	69.5046	−0.494614

(continued on next page)

Table. 11 (*continued*)

		ARIMA		GARCH		OLS	
Date	WTI Crude Oil Actual Value	fitted	residual	fitted	residual	fitted	residual
2018-10-10	74.9000	69.1853	−0.855276	73.6186	1.28135	69.0888	−0.758807
2018-10-11	72.7800	69.0533	−0.663349	70.7687	2.01133	69.3841	−0.994107
2018-10-12	71.1900	68.3753	−0.535263	70.6380	0.551979	69.4929	−1.65289
2018-10-15	72.4200	69.3827	0.727285	70.9387	1.48135	70.7428	−0.632842
2018-10-16	71.9300	69.2451	1.17492	70.8967	1.03326	70.0867	0.333295
2018-10-17	72.3300	69.7906	−0.730638	70.9907	1.33928	70.0085	−0.948473
2018-10-18	69.9400	69.8092	−0.949162	69.8230	0.117025	70.3113	−1.45126
2018-10-19	68.8500	69.7332	−1.11316	70.0531	−1.20309	70.4661	−1.84609
2018-10-22	69.3100	69.7981	0.0819011	69.7808	−0.470802	70.8400	−0.959988
2018-10-23	69.4800	70.8172	0.592842	68.2525	1.22751	71.1909	0.219148
2018-10-24	66.2200	71.8828	−1.32278	67.9010	−1.68097	72.1198	−1.55976
2018-10-25	66.0800	72.0068	−0.0867666	67.0386	−0.958585	73.3367	−1.41670
2018-10-26	66.8200	71.8640	0.605994	67.0863	−0.266253	73.1261	−0.656092
2018-10-29	67.6000	71.8305	0.419480	67.5513	0.0487426	72.9824	−0.732379
2018-10-30	66.5800	72.1188	0.321205	65.8967	0.683345	72.6754	−0.235389
2018-10-31	66.3600	72.2231	0.196868	67.5986	−1.23859	75.7710	−3.35096
2018-11-01	64.9200	72.8738	0.856243	66.6880	−1.76804	74.4804	−0.750375
2018-11-02	63.5200	74.4325	0.977457	65.8394	−2.31945	76.3562	−0.946231
2018-11-05	62.9100	75.5314	0.878580	65.5022	−2.59225	76.3218	0.0881843
2018-11-06	62.7800	75.1238	−0.173785	65.3172	−2.53716	75.2245	−0.274522
2018-11-07	61.7600	74.8973	−0.867263	65.3270	−3.56697	74.3477	−0.317725
2018-11-08	61.6300	75.0390	−0.599049	64.9175	−3.28748	74.9803	−0.540318
2018-11-09	60.6900	74.8620	0.0380204	63.8395	−3.14946	74.4229	0.477136
2018-11-12	60.7400	72.7305	0.0495322	64.0352	−3.29516	70.8498	1.93020
2018-11-13	58.9200	72.3847	−1.19467	61.4084	−2.48836	70.5343	0.655720
2018-11-14	55.6900	71.7750	0.645035	60.1835	−4.49348	70.7871	1.63294
2018-11-15	55.8900	71.7568	0.173228	59.9620	−4.07199	70.7789	1.15113
2018-11-16	56.4800	72.0579	0.272102	60.9866	−4.50657	70.9593	1.37065
2018-11-19	57.1000	70.7847	−0.844685	60.0656	−2.96557	69.7197	0.220325
2018-11-20	57.1700	70.4656	−1.61561	58.3837	−1.21369	69.7699	−0.919899
2018-11-21	53.3300	69.9361	−0.626084	57.8952	−4.56521	69.6712	−0.361221
2018-11-23	53.8800	68.7669	0.713051	55.2815	−1.40155	66.8280	2.65199
2018-11-26	50.4500	68.8905	−2.67047	55.9089	−5.45891	66.6093	−0.389302
2018-11-27	51.4700	67.5086	−1.42860	55.8778	−4.40779	65.9454	0.134595
2018-11-28	51.9200	66.6916	0.128444	55.7898	−3.86979	66.1301	0.689868
2018-11-29	50.4800	66.7008	0.899186	56.8734	−6.39342	66.4261	1.17387
2018-11-30	51.3000	66.3431	0.236916	56.5604	−5.26044	65.0156	1.56441
2018-12-03	52.2200	67.0609	−0.700880	58.9238	−6.70377	69.3214	−2.96138
2018-12-04	53.1400	66.2607	−1.34074	59.3683	−6.22831	65.5241	−0.604142
2018-12-06	52.5800	65.1110	−1.59099	56.9801	−4.40008	65.2269	−1.70692
2018-12-07	51.4200	64.5370	−1.62700	59.0659	−7.64588	64.7949	−1.88494
2018-12-10	52.3600	63.8803	−1.10029	57.0037	−4.64372	64.5263	−1.74632
2018-12-11	50.9000	63.1093	−1.34930	58.6211	−7.72106	65.1373	−3.37728
2018-12-12	51.9900	62.3485	−0.718514	58.1859	−6.19588	64.5829	−2.95289
2018-12-13	51.3200	61.2338	−0.543818	57.4498	−6.12981	63.8286	−3.13860
2018-12-14	52.7800	61.1940	−0.454040	57.7629	−4.98292	63.7915	−3.05155
2018-12-17	51.3000	58.8269	0.0930831	56.7694	−5.46943	60.4226	−1.50259
2018-12-18	49.3100	57.8454	−2.15538	54.1934	−4.88343	59.5117	−3.82174
2018-12-19	46.3100	56.9633	−1.07330	52.9138	−6.60376	58.9582	−3.06821
2018-12-20	47.0700	56.7385	−0.258481	51.3782	−4.30819	59.3514	−2.87142
2018-12-21	46.0800	56.2060	0.894035	50.2227	−4.14270	58.7654	−1.66541
2018-12-24	45.3400	55.9792	1.19080	49.0184	−3.67843	56.1864	0.983633
2018-12-26	43.0900	55.0906	−1.76061	48.0803	−4.99033	55.7523	−2.42234
2018-12-27	46.4100	53.1445	0.735513	50.2089	−3.79886	52.8759	1.00412
2018-12-28	45.2300	52.6973	−2.24731	51.3942	−6.16422	53.6735	−3.22351
2018-12-31	45.3400	52.3814	−0.911426	52.2899	−6.94988	53.6648	−2.19477

Note: * denotes a residual in excess of 2.5 standard errors.

2. Normality Test of Residuals

The estimated experimental error known as residuals is calculated by deducting the observed responses from the expected responses. Following the estimation of all the unknown model parameters from the experimental data, the predicted response is determined using the selected model. All statistical modelling, including DOEs, must include examining residuals as a crucial component. We can determine if our model selection and assumptions are sound by carefully examining residuals. Residuals may be regarded an aspects of variance that the fitted model does not account for. Given that this is an error, the same basic presumptions that we make for errors in general apply to the group of residuals: one expects them to be (nearly) normal and (roughly) independently distributed with a mean of 0 and some constant variance. Running a linear model requires the residuals to be normal. Therefore, if your residuals are normal, your assumption is true, then model inference (confidence intervals, model predictions) should be true as well.

2.1. Normality test of residuals for ARIMA model

Table. 12
Normality test of residuals for ARIMA model

interval	midpt	frequency	Rel	cum.
< −2.7626	−2.9083	2	0.12%	0.12%
−2.7626 - 2.4713	−2.6169	4	0.23%	0.35%
−2.4713 - 2.1799	−2.3256	9	0.52%	0.87%
−2.1799 - 1.8886	−2.0342	9	0.52%	1.40%
−1.8886 - 1.5972	−1.7429	19	1.11%	2.50%
−1.5972 - 1.3059	−1.4515	40	2.33%	4.83%
−1.3059 - 1.0145	−1.1602	77	4.48%	9.32% *
−1.0145 - 0.72317	−0.86884	147	8.56%	17.88% ***
−0.72317 - 0.43182	0.57749	186	10.83%	28.71% ***
−0.43182 - 0.14047-	0.28614	230	13.40%	42.11% ****
−0.14047 - 0.15088	0.0052056	258	15.03%	57.13% *****
0.15088–0.44223	0.29656	246	14.33%	71.46% *****
0.44223–0.73358	0.58791	206	12.00%	83.46% ****
0.73358–1.0249	0.87926	133	7.75%	91.21% **
1.0249–1.3163	1.1706	74	4.31%	95.52% *
1.3163–1.6076	1.4620	37	2.15%	97.67%
1.6076–1.8990	1.7533	21	1.22%	98.89%
1.8990–2.1903	2.0447	10	0.58%	99.48%
2.1903–2.4817	2.3360	3	0.17%	99.65%
2.4817–2.7730	2.6274	1	0.06%	99.71%
2.7730–3.0644	2.9187	2	0.12%	99.83%
3.0644–3.3557	3.2101	0	0.00%	99.83%
3.3557–3.6471	3.5014	0	0.00%	99.83%
3.6471–3.9384	3.7928	2	0.12%	99.94%
3.9384–4.2298	4.0841	0	0.00%	99.94%
4.2298–4.5211	4.3755	0	0.00%	99.94%
4.5211–4.8125	4.6668	0	0.00%	99.94%
4.8125–5.1038	4.9582	0	0.00%	99.94%
≥ 5.1038	5.2495	1	0.06%	100.00%

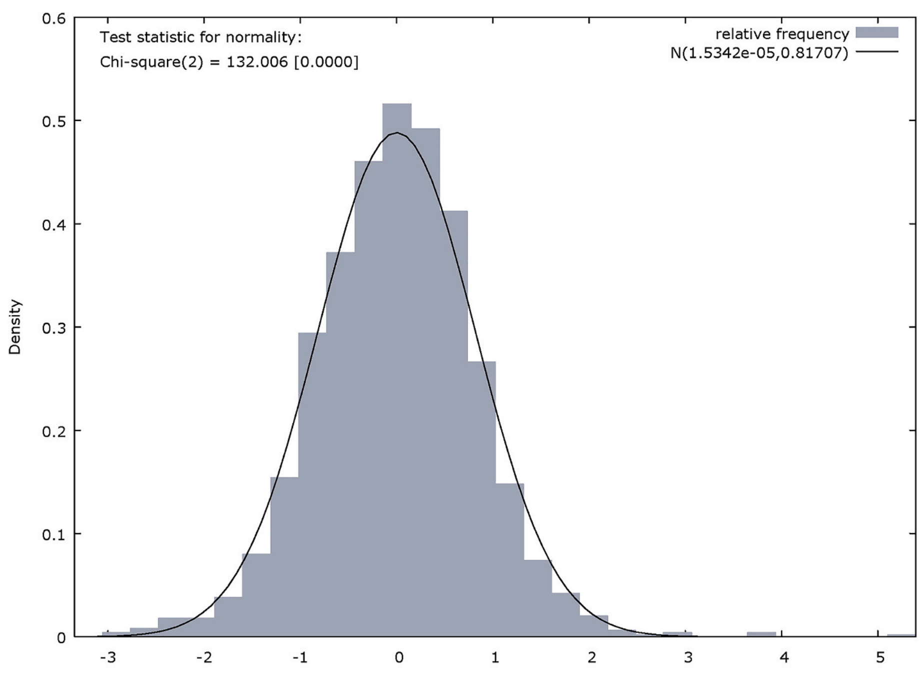

Fig. 20. Normality test of residuals - ARIMA model

2.2. NORMALITY TEST OF RESIDUALS FOR GARCH MODEL. Frequency distribution for residual, obs 1-1718
Number of bins = 29, mean = 0.254607, sd = 2.17651.

Table. 13
Normality test of residuals for ARIMA model

interval	midpt	frequency	rel.	cum.
< −7.4323	−7.7211	2	0.12%	0.12%
−7.4323–−6.8547	−7.1435	1	0.06%	0.17%
−6.8547–−6.2772	−6.5660	3	0.17%	0.35%
−6.2772–−5.6996	−5.9884	6	0.35%	0.70%
−5.6996–−5.1221	−5.4109	3	0.17%	0.87%
−5.1221–−4.5446	−4.8333	9	0.52%	1.40%
−4.5446–−3.9670	−4.2558	12	0.70%	2.10%
−3.9670–−3.3895	−3.6782	19	1.11%	3.20%
−3.3895–−2.8119	−3.1007	30	1.75%	4.95%
−2.8119–−2.2344	−2.5231	59	3.43%	8.38% *
−2.2344–−1.6568	−1.9456	111	6.46%	14.84% **
−1.6568–−1.0793	−1.3680	152	8.85%	23.69% ***
−1.0793–−0.50172	−0.79050	222	12.92%	36.61% ****
−0.50172 - 0.075822	−0.21295	248	14.44%	51.05% *****
0.075822–0.65337	0.36460	215	12.51%	63.56% ****
0.65337–1.2309	0.94214	195	11.35%	74.91% ****
1.2309–1.8085	1.5197	110	6.40%	81.32% **
1.8085–2.3860	2.0972	84	4.89%	86.20% *
2.3860–2.9636	2.6748	73	4.25%	90.45% *
2.9636–3.5411	3.2523	35	2.04%	92.49%
3.5411–4.1186	3.8299	42	2.44%	94.94%
4.1186–4.6962	4.4074	12	0.70%	95.63%
4.6962–5.2737	4.9850	21	1.22%	96.86%
5.2737–5.8513	5.5625	18	1.05%	97.90%
5.8513–6.4288	6.1401	12	0.70%	98.60%
6.4288–7.0064	6.7176	11	0.64%	99.24%
7.0064–7.5839	7.2952	6	0.35%	99.59%
7.5839–8.1615	7.8727	5	0.29%	99.88%
≥ 8.1615	8.4502	2	0.12%	100.00%

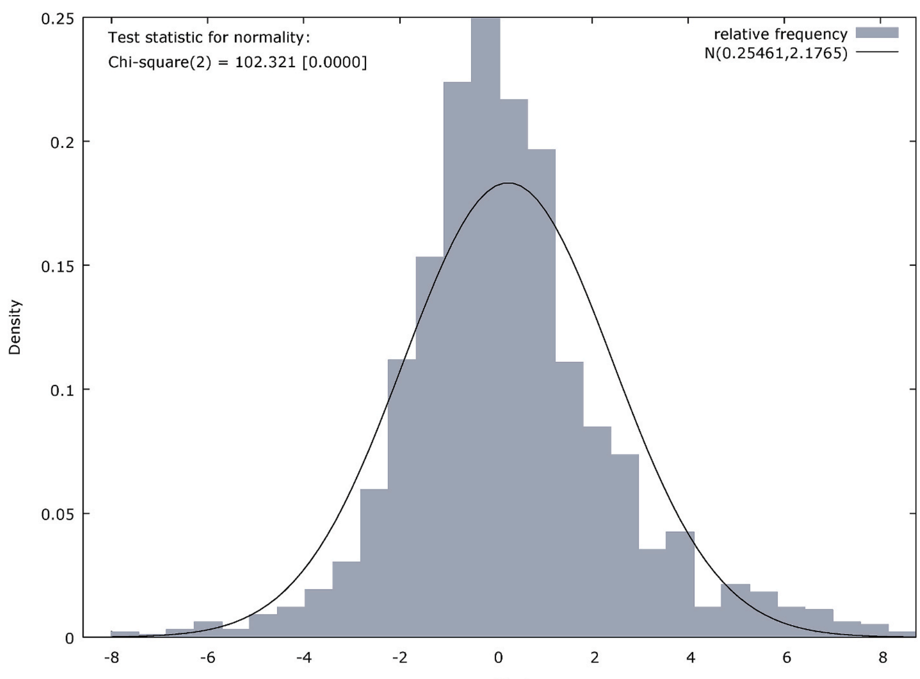

Fig. 21. Normality test of residuals - GARCH model

2.3. Normality test of residuals for OLS model

Table. 14
Normality test of residuals for ARIMA model

interval	midpt	frequency	rel.	cum.
< −5.9280	−6.1608	1	0.06%	0.06%
−5.9280–−5.4624	−5.6952	2	0.12%	0.17%
−5.4624–−4.9969	−5.2297	4	0.23%	0.41%
−4.9969–−4.5314	−4.7641	9	0.52%	0.93%

(*continued on next page*)

Table. 14 (*continued*)

interval	midpt	frequency	rel.	cum.
−4.5314–−4.0658	−4.2986	12	0.70%	1.63%
−4.0658–−3.6003	−3.8331	25	1.46%	3.08%
−3.6003–−3.1347	−3.3675	45	2.62%	5.70%
−3.1347–−2.6692	−2.9020	62	3.61%	9.31% *
−2.6692–−2.2037	−2.4364	79	4.60%	13.91% *
−2.2037–−1.7381	−1.9709	88	5.12%	19.03% *
−1.7381–−1.2726	−1.5053	135	7.86%	26.89% **
−1.2726–−0.80704	−1.0398	143	8.32%	35.22% **
−0.80704–−0.34150	−0.57427	154	8.96%	44.18% ***
−0.34150 - 0.12404	−0.10873	153	8.91%	53.08% ***
0.12404–0.58958	0.35681	165	9.60%	62.69% ***
0.58958–1.0551	0.82235	145	8.44%	71.13% ***
1.0551–1.5207	1.2879	115	6.69%	77.82% **
1.5207–1.9862	1.7534	109	6.34%	84.17% **
1.9862–2.4517	2.2190	79	4.60%	88.77% *
2.4517–2.9173	2.6845	63	3.67%	92.43% *
2.9173–3.3828	3.1501	41	2.39%	94.82%
3.3828–3.8484	3.6156	36	2.10%	96.92%
3.8484–4.3139	4.0811	17	0.99%	97.90%
4.3139–4.7794	4.5467	16	0.93%	98.84%
4.7794–5.2450	5.0122	7	0.41%	99.24%
5.2450–5.7105	5.4778	6	0.35%	99.59%
5.7105–6.1761	5.9433	4	0.23%	99.83%
6.1761–6.6416	6.4088	1	0.06%	99.88%
≥ 6.6416	6.8744	2	0.12%	100.00%

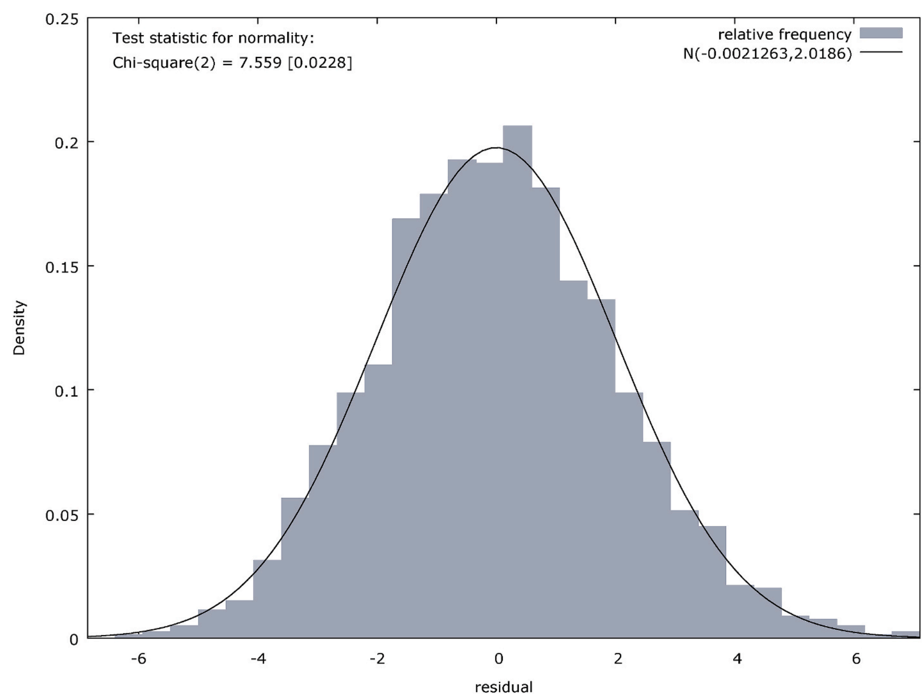

Fig. 22. Normality test of residuals - OLS model

3. Covariance matrix of regression coefficients

3.1. Coefficient covariance matrix of ARIMA Model

Table. 15
Coefficient covariance matrix of ARIMA Model

const	phi_1	theta_1	GoldPrice	SP500	
9.68490e-005	7.62311e-006	−1.98659e-005	−1.26449e-005	−9.48426e-006	const
	0.00198576	−0.00158938	3.08245e-005	−2.20698e-005	phi_1
		0.00201367	6.27755e-007	−1.50321e-005	theta_1
			0.000668241	4.42441e-005	GoldPrice
				0.000452949	SP500
DowJonesIndex	EldoradoGoldCorporation	EURUSDExchangerate	BrentCrudeoilFutures	SilverFutures	

(*continued on next page*)

Table. 15 (*continued*)

const	phi_1	theta_1	GoldPrice	SP500	
−1.38729e-008	8.44640e-006	−0.00177639	5.07970e-006	−6.63155e-010	const
6.37114e-008	4.06650e-006	−0.00404965	−2.16364e-005	4.87654e-010	phi_1
6.37661e-007	−3.08875e-005	0.00783935	−0.000170957	−6.01988e-008	theta_1
1.00860e-007	3.61505e-005	−0.00519782	−2.30908e-005	−1.97107e-007	GoldPrice
−4.22580e-006	−1.59059e-005	0.00455998	−1.32339e-005	−6.82174e-009	SP500
5.45022e-008	1.44616e-007	−6.63790e-005	−1.21946e-007	3.51506e-010	DowJonesIndex
	0.000595122	−0.00569068	2.23219e-005	−1.89782e-008	EldoradoGoldCorporation
		60.6664	0.00986736	8.97239e-006	EURUSDExchangerate
			0.000348095	4.40937e-009	BrentCrudeoilFutures
				1.32787e-009	SilverFutures

USBondRate	PlatinumPrice	Palladiumprice	RhodiumPrices	USdollarIndexPrice	
−5.38793e-005	6.82575e-007	−1.20378e-006	−1.07185e-009	−3.10845e-005	const
−0.000682005	8.16068e-008	−4.12335e-006	−8.35903e-008	−6.00315e-006	phi_1
0.00138381	7.62192e-006	8.51186e-006	5.61556e-008	−7.06717e-005	theta_1
0.00209720	−4.20322e-006	−6.81597e-007	7.00360e-009	−1.59065e-006	GoldPrice
0.000120063	1.26744e-006	−5.91290e-008	2.33509e-008	0.000106461	SP500
−1.20624e-005	−3.10607e-008	−3.74476e-008	−8.00168e-011	−1.55369e-006	DowJonesIndex
−0.000539045	−8.57216e-007	7.24449e-007	3.27321e-008	−0.000108276	EldoradoGoldCorporation
−0.519161	−0.000158097	−5.76175e-005	−6.60096e-006	0.891393	EURUSDExchangerate
−0.000659952	−1.00289e-006	−2.06314e-006	3.19940e-008	0.000186174	BrentCrudeoilFutures
−9.52635e-007	−2.09083e-008	−2.82180e-009	−4.09803e-011	1.60332e-007	SilverFutures
0.182137	3.58841e-005	4.27244e-006	−8.41041e-007	−0.00942988	USBondRate
	2.90608e-006	−9.02261e-007	−2.12544e-009	1.29596e-005	PlatinumPrice
		2.83423e-006	1.15231e-009	−5.65082e-007	Palladiumprice
			3.14834e-009	−1.14021e-007	RhodiumPrices
				0.0151788	USdollarIndexPrice

GoldMinersETF	OilETFUSO	
1.47699e-005	4.40880e-005	const
−8.70123e-005	0.000389301	phi_1
1.39758e-005	−0.000645906	theta_1
−0.000907004	−8.53805e-005	GoldPrice
−7.37516e-005	−0.000142340	SP500
−5.60726e-008	7.26127e-007	DowJonesIndex
−0.000877844	−7.23825e-005	EldoradoGoldCorporation
0.0209523	−0.000104572	EURUSDExchangerate
1.64841e-005	−0.000497706	BrentCrudeoilFutures
−6.38154e-008	1.17633e-007	SilverFutures
−0.000399816	−0.00407033	USBondRate
−5.58207e-006	−1.19634e-005	PlatinumPrice
−2.67219e-006	−3.73922e-006	Palladiumprice
−4.13470e-008	−7.50602e-008	RhodiumPrices
0.000408126	0.000177818	USdollarIndexPrice
0.00349339	4.02921e-005	GoldMinersETF
	0.00371443	OilETFUSO

3.2. *Coefficient covariance matrix of GARCH Model*

Table. 16
Coefficient covariance matrix of GARCH Model

const	GoldPrice	SP500	DowJonesIndex	EldoradoGoldCõ	
257.535	−0.131292	0.115626	−0.00156712	−0.0353186	const
	0.000716003	7.40162e-005	−1.15005e-006	−6.97921e-005	GoldPrice
		0.000250812	−3.01286e-006	−2.40123e-005	SP500
			4.16589e-008	6.04119e-007	DowJonesIndex
				0.000171461	EldoradoGoldCõ

EURUSDExchange∼	BrentCrudeoilF∼	SilverFutures	USBondRate	PlatinumPrice	
−86.3769	−0.0248033	2.38688e-005	3.78744	0.00360113	const
0.0197606	−0.000118022	−2.49935e-007	0.00374201	−1.00216e-005	GoldPrice
−0.0408795	1.28604e-006	−8.63743e-008	0.00323044	−2.59048e-006	SP500
0.000548897	−9.94967e-008	1.46153e-009	−5.15852e-005	3.56512e-008	DowJonesIndex
0.0202547	6.06500e-005	4.18298e-008	−0.00179777	−4.72695e-006	EldoradoGoldCõ
31.3199	0.0161060	−4.95143e-006	−1.55433	−0.00160725	EURUSDExchange∼
	0.000174387	−1.43113e-007	−0.00176274	3.27325e-006	BrentCrudeoilF∼
		6.58356e-010	−2.55825e-006	−6.75693e-009	SilverFutures
			0.171993	6.10995e-005	USBondRate
				1.15609e-006	PlatinumPrice

Palladiumprice	RhodiumPrices	USdollarIndexP∼	GoldMinersETF	OilETFUSO	
0.00747064	4.00194e-005	−1.62811	0.154778	−0.292254	const

(continued on next page)

Table. 16 (*continued*)

const	GoldPrice	SP500	DowJonesIndex	EldoradoGoldCõ	
5.29671e-007	3.42777e-007	0.000569621	−0.000593250	0.000513251	GoldPrice
5.70340e-006	1.31435e-007	−0.000837248	7.29443e-005	−0.000160423	SP500
−1.37079e-007	−2.98235e-009	1.09560e-005	−9.74379e-007	2.88701e-006	DowJonesIndex
−3.38077e-006	−2.64777e-007	0.000230593	−0.000117536	−0.000122850	EldoradoGoldCõ
−0.00257876	−2.08310e-005	0.552533	−0.0346014	0.0716399	EURUSDExchange~
−1.99453e-006	−1.49977e-007	0.000247471	7.90543e-005	−0.000430433	BrentCrudeoilF~
2.39396e-009	−1.09223e-010	−1.20242e-007	−1.59682e-008	4.31130e-007	SilverFutures
0.000206549	5.64036e-006	−0.0263425	4.85438e-005	−0.00338637	USBondRate
−2.68146e-007	5.92476e-009	−1.77858e-005	1.48478e-005	−1.92496e-005	PlatinumPrice
1.42081e-006	9.22627e-009	−4.58955e-005	3.25317e-006	−8.23217e-006	Palladiumprice
	6.74533e-009	−4.26343e-007	−1.45111e-007	1.78748e-008	RhodiumPrices
		0.0104847	−0.000773961	0.00164494	USdollarIndexP~
			0.000895204	−0.000509102	GoldMinersETF
				0.00186914	OilETFUSO
	alpha(0)	alpha(1)	alpha(2)	beta(1)	
	0.0411641	0.0458184	−0.0332630	−0.0512312	const
	−0.000171520	−0.000129887	0.000178960	4.11972e-005	GoldPrice
	−5.27638e-005	2.10376e-005	−9.37701e-005	8.18338e-005	SP500
	1.24161e-006	−1.06686e-008	1.68784e-006	−1.91331e-006	DowJonesIndex
	6.58701e-005	6.64477e-005	−5.24792e-005	−3.70857e-005	EldoradoGoldCõ
	−0.00944229	−0.0127245	−0.00463727	0.0265055	EURUSDExchange~
	3.94941e-005	−4.70328e-005	2.29052e-005	6.23659e-006	BrentCrudeoilF~
	8.99081e-008	1.92596e-007	−5.87915e-008	−1.51739e-007	SilverFutures
	−0.00145313	0.000278139	−0.000875049	0.000848182	USBondRate
	5.84470e-006	7.07290e-007	9.23863e-006	−1.14552e-005	PlatinumPrice
	−1.07060e-005	2.39420e-007	−1.56503e-005	1.73591e-005	Palladiumprice
	−1.32170e-007	1.88633e-007	2.34778e-008	−1.39575e-007	RhodiumPrices
	−0.000232352	−0.000301312	0.000169090	0.000356137	USdollarIndexP~
	1.90592e-005	3.48119e-005	−0.000254377	0.000173108	GoldMinersETF
	−0.000144830	−1.60081e-005	8.99940e-005	2.26978e-005	OilETFUSO
	0.00324878	0.000414734	0.00288489	−0.00436987	alpha(0)
		0.00285796	−0.00150238	−0.000943579	alpha(1)
			0.00603641	−0.00534244	alpha(2)
				0.00747205	beta(1)

3.3. Coefficient covariance matrix of OLS Model

Table. 17
Coefficient covariance matrix of OLS Model

GoldPrice	SP500	DowJonesIndex	EldoradoGoldCorporation	EURUSDExchangerate	
0.000461054	0.000158900	−1.75935e-006	−4.59885e-005	−0.0226038	GoldPrice
	0.000259338	−3.18897e-006	−3.62340e-005	−0.00573983	SP500
		4.51822e-008	6.02156e-007	1.91590e-005	DowJonesIndex
			0.000217498	0.00666437	EldoradoGoldCorporation
				2.56586	EURUSDExchangerate
BrentCrudeoilFutures	SilverFutures	USBondRate	PlatinumPrice	Palladiumprice	
−8.28585e-005	−1.73864e-007	0.00321080	−6.39170e-006	6.12237e-007	GoldPrice
4.84151e-005	−1.51621e-007	0.00185780	−3.66077e-006	3.91656e-006	SP500
−1.15827e-006	2.37325e-009	−2.65372e-005	5.62000e-008	−1.00355e-007	DowJonesIndex
7.17215e-005	−1.58018e-008	−0.000566236	−2.42664e-006	2.63690e-007	EldoradoGoldCorporation
0.00989251	5.94919e-008	−0.191174	−0.000253063	0.000297677	EURUSDExchangerate
0.000274255	−2.44276e-007	−0.00108255	1.39803e-006	3.57014e-006	BrentCrudeoilFutures
	8.91026e-010	−2.74705e-006	−7.86099e-009	−8.62180e-010	SilverFutures
		0.104764	5.41973e-005	−2.80523e-006	USBondRate
			1.12056e-006	−4.20526e-007	PlatinumPrice
				9.71782e-007	Palladiumprice
	RhodiumPrices	USdollarIndexPrice	GoldMinersETF	OilETFUSO	
	−7.85576e-007	−0.000159804	−0.000343899	0.000343615	GoldPrice
	−1.37840e-008	−7.07856e-005	2.05981e-005	−7.33509e-005	SP500
	−1.35257e-009	3.14547e-007	−6.66566e-007	2.88895e-006	DowJonesIndex
	−3.07248e-007	−1.21867e-005	−0.000191570	−0.000281711	EldoradoGoldCorporation
	4.65871e-005	0.00659485	0.0153893	−0.0409853	EURUSDExchangerate
	4.27383e-009	0.000114521	2.97925e-005	−0.000723714	BrentCrudeoilFutures
	5.61243e-010	−5.41749e-008	−1.79508e-007	7.22604e-007	SilverFutures
	−3.88410e-006	−0.00167994	0.000303601	−0.000400574	USBondRate
	3.69802e-008	9.27381e-007	8.92891e-006	−1.55215e-005	PlatinumPrice
	−2.07685e-008	3.65323e-006	5.18544e-007	−1.00252e-005	Palladiumprice
	1.85799e-008	2.06516e-007	5.16954e-007	−7.82992e-007	RhodiumPrices
		0.000176947	0.000150836	−0.000239071	USdollarIndexPrice
			0.000821151	−0.000171440	GoldMinersETF
				0.00248684	OilETFUSO

References

Adebayo, T.S., Kirikkaleli, D., Adeshola, I., Oluwajana, D., Akinsola, G.D., Osemeahon, O.S., 2021. Coal consumption and environmental sustainability in South Africa: the role of financial development and globalization. Int. J. Renew. Energy Dev. 10, 527–536. https://doi.org/10.14710/ijred.2021.34982.

Agbonifo, P.E., 2021. Renewable energy development: opportunities and barriers within the context of global energy politics. Int. J. Energy Econ. Pol. 11 https://doi.org/10.32479/ijeep.10773.

Bag, S., Wood, L.C., Mangla, S.K., Luthra, S., 2020. Procurement 4.0 and its implications on business process performance in a circular economy. Resour. Conserv. Recycl. 152, 104502 https://doi.org/10.1016/j.resconrec.2019.104502.

Batool, K., Zhao, Z.-Y., Atif, F., Dilanchiev, A., 2022. Nexus between energy poverty and technological innovations: a pathway for addressing energy sustainability. Front. Environ. Sci. 10, 888080 https://doi.org/10.3389/fenvs.2022.888080.

Bei, J., Wang, C., 2023. Renewable energy resources and sustainable development goals: evidence based on green finance, clean energy and environmentally friendly investment. Res. Pol. 80, 103194 https://doi.org/10.1016/j.resourpol.2022.103194.

Bull, S.R., 2001. Renewable energy today and tomorrow. Proc. IEEE 89, 1216–1226. https://doi.org/10.1109/5.940290.

Chang, L., Chen, K., Saydaliev, H.B., Faridi, M.Z., 2022a. Asymmetric impact of pandemics-related uncertainty on CO2 emissions: evidence from top-10 polluted countries. Stoch. Environ. Res. Risk Assess. 36, 4103–4117. https://doi.org/10.1007/s00477-022-02248-5.

Chang, L., Lu, Q., Ali, S., Mohsin, M., 2022b. How does hydropower energy asymmetrically affect environmental quality? Evidence from quantile-based econometric estimation. Sustain. Energy Technol. Assessments 53, 102564. https://doi.org/10.1016/j.seta.2022.102564.

Chang, L., Moldir, M., Zhang, Y., Nazar, R., 2023a. Asymmetric impact of green bonds on energy efficiency: fresh evidence from quantile estimation. Util. Pol. 80, 101474 https://doi.org/10.1016/j.jup.2022.101474.

Chang, L., Qian, C., Dilanchiev, A., 2022c. Nexus between financial development and renewable energy: empirical evidence from nonlinear autoregression distributed lag. Renew. Energy 193, 475–483. https://doi.org/10.1016/j.renene.2022.04.160.

Chang, L., Saydaliev, H.B., Meo, M.S., Mohsin, M., 2022d. How renewable energy matter for environmental sustainability: evidence from top-10 wind energy consumer countries of European Union. Sustain. Energy, Grids Networks 31, 100716. https://doi.org/10.1016/j.segan.2022.100716.

Chang, L., Shi, F., Taghizadeh-Hesary, F., Saydaliev, H.B., 2023b. Information and communication technologies development and the resource curse. Res. Pol. 80, 103123 https://doi.org/10.1016/j.resourpol.2022.103123.

Chang, L., Taghizadeh-Hesary, F., Chen, H., Mohsin, M., 2022e. Do green bonds have environmental benefits? Energy Econ. 115, 106356 https://doi.org/10.1016/j.eneco.2022.106356.

Chang, L., Taghizadeh-Hesary, F., Saydaliev, H.B., 2022f. How do ICT and renewable energy impact sustainable development? Renew. Energy 199, 123–131. https://doi.org/10.1016/j.renene.2022.08.082.

Dilanchiev, A., Taktakishvili, T., 2021. Currency depreciation nexus country's export: evidence from Georgia. Univers. J. Account. Financ. 9, 1116–1124. https://doi.org/10.13189/ujaf.2021.090521.

Fang, W., Liu, Z., Surya Putra, A.R., 2022. Role of research and development in green economic growth through renewable energy development: empirical evidence from South Asia. Renew. Energy 194, 1142–1152. https://doi.org/10.1016/j.renene.2022.04.125.

Geng, Q., 2021. The belt and road initiative and its implications for global renewable energy development. Curr. Sustain. Energy Rep. https://doi.org/10.1007/s40518-020-00172-2.

Gengenbach, C., Palm, F.C., Urbain, J.P., 2010. Panel unit root tests in the presence of cross-sectional dependencies: comparison and implications for modelling. Econom. Rev. 29, 111–145. https://doi.org/10.1080/07474930903382125.

Golosnoy, A.S., Provotorov, V.V., Sergeev, S.M., Raikhelgauz, L.B., Kravets, O.J., 2019. Software engineering math for network applications. J. Phys. Conf. Ser. 1399, 044047 https://doi.org/10.1088/1742-6596/1399/4/044047.

Hillis, A., Germain, J., Whitfield, M., Halsall, D., Mcveigh, J., Abbasi, Y., Claire, M., Hout, V., 2023. Emerging Trends in Drugs , Addictions , and Health Internet sourcing and unsafe use of controlled drugs (opioids , sedatives and GABA drugs) in the UK : an in depth case study of consumer dynamics during COVID-19. Emerg. Trends Drugs, Addict. Health 3, 100049. https://doi.org/10.1016/j.etdah.2022.100049.

Huang, W., Chau, K.Y., Kit, I.Y., Nureen, N., Irfan, M., Dilanchiev, A., 2022. Relating sustainable business development practices and information management in promoting digital green innovation: evidence from China. Front. Psychol. 13, 930138 https://doi.org/10.3389/fpsyg.2022.930138.

Ikram, M., Mahmoudi, A., Shah, S.Z.A., Mohsin, M., 2019. Forecasting number of ISO 14001 certifications of selected countries: application of even GM (1,1), DGM, and NDGM models. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-019-04534-2.

Iqbal, N., Tufail, M.S., Mohsin, M., Sandhu, M.A., 2022. Assessing social and financial efficiency: the evidence from microfinance institutions in Pakistan. Pakistan J. Soc. Sci. 39, 149–161.

Iqbal, W., Yumei, H., Abbas, Q., Hafeez, M., Mohsin, M., Fatima, A., Jamali, M.A., Jamali, M., Siyal, A., Sohail, N., 2019. Assessment of wind energy potential for the production of renewable hydrogen in Sindh Province of Pakistan. Processes 7. https://doi.org/10.3390/pr7040196.

Iram, R., Zhang, J., Erdogan, S., Abbas, Q., Mohsin, M., 2020. Economics of energy and environmental efficiency: evidence from OECD countries. Environ. Sci. Pollut. Res. https://doi.org/10.1007/s11356-019-07020-x.

Kao, C., 1999. Spurious regression and residual-based tests for cointegration in panel data. J. Econom. 90, 1–44. https://doi.org/10.1016/S0304-4076(98)00023-2.

Klein, T., Pham Thu, H., Walther, T., 2018. Bitcoin is not the New Gold – a comparison of volatility, correlation, and portfolio performance. Int. Rev. Financ. Anal. 59, 105–116. https://doi.org/10.1016/j.irfa.2018.07.010.

Lee, W., 2020. Unravelling consumer responses to omni-channel approach. J. Theor. Appl. Electron. Commer. Res. 15, 37–49. https://doi.org/10.4067/S0718-18762020000300104.

Li, C., Umair, M., 2023. Does green finance development goals affects renewable energy in China. Renew. Energy 203, 898–905. https://doi.org/10.1016/j.renene.2022.12.066.

Liu, F., Umair, M., Gao, J., 2023. Assessing oil price volatility co-movement with stock market volatility through quantile regression approach. Res. Pol. 81, 103375 https://doi.org/10.1016/j.resourpol.2023.103375.

Liu, Y., Dilanchiev, A., Xu, K., Hajiyeva, A.M., 2022. Financing SMEs and business development as new post Covid-19 economic recovery determinants. Econ. Anal. Pol. 76, 554–567. https://doi.org/10.1016/j.eap.2022.09.006.

Luo, J., Wan, L., Zhang, Q., Cui, B., Li, C., Jiang, Y., Jiang, M., Wang, K.K., 2023. Constructing a drug release model by central composite design to investigate the interaction between drugs and temperature-sensitive controlled release nanoparticles. Eur. J. Pharm. Biopharm. 183, 24–32. https://doi.org/10.1016/j.ejpb.2022.12.009.

Marmolejo-Saucedo, J.A., 2020. Design and development of digital twins: a case study in supply chains. Mobile Network. Appl. 25, 2141–2160. https://doi.org/10.1007/s11036-020-01557-9.

Mele, M., Magazzino, C., 2020. A Machine Learning analysis of the relationship among iron and steel industries, air pollution, and economic growth in China. J. Clean. Prod. 277, 123293 https://doi.org/10.1016/j.jclepro.2020.123293.

Mohsin, M., Nurunnabi, M., Zhang, J., Sun, H., Iqbal, N., Iram, R., Abbas, Q., 2020a. The evaluation of efficiency and value addition of IFRS endorsement towards earnings timeliness disclosure. Int. J. Financ. Econ. https://doi.org/10.1002/ijfe.1878.

Mohsin, M., Rasheed, A.K., Saidur, R., 2018a. Economic viability and production capacity of wind generated renewable hydrogen. Int. J. Hydrogen Energy 43, 2621–2630. https://doi.org/10.1016/J.IJHYDENE.2017.12.113.

Mohsin, M., Zaidi, U., Abbas, Q., Rao, H., Iqbal, N., Chaudhry, S., 2020b. Relationship between multi-factor pricing and equity price fragility: evidence from Pakistan. Int. J. Sci. Technol. Res. 8.

Mohsin, M., Zhou, P., Iqbal, N., Shah, S.A.A., 2018b. Assessing oil supply security of South Asia. Energy 155, 438–447. https://doi.org/10.1016/J.ENERGY.2018.04.116.

Naz, L., Patel, K.K., Dilanchiev, A., 2021. Are socioeconomic status and type of residence critical risk factors of under-five mortality in Pakistan? Evidence from nationally representative survey. Clin. Epidemiol. Glob. Heal. 10, 100670 https://doi.org/10.1016/j.cegh.2020.11.003.

Pan, W., Cao, H., Liu, Y., 2023. Green" innovation, privacy regulation and environmental policy. Renew. Energy 203, 245–254. https://doi.org/10.1016/j.renene.2022.12.025.

Pedroni, P., 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. Oxf. Bull. Econ. Stat. 61, 653–670. https://doi.org/10.1111/1468-0084.61.s1.14.

Pesaran, M.H., 2014. J. Appl. Econom. 21, 1–21. https://doi.org/10.1002/jae.

Sebestyén, V., Abonyi, J., 2021. Data-driven comparative analysis of national adaptation pathways for Sustainable Development Goals. J. Clean. Prod. 319, 128657 https://doi.org/10.1016/j.jclepro.2021.128657.

Shah, S.A.A., Zhou, P., Walasai, G.D., Mohsin, M., 2019. Energy security and environmental sustainability index of South Asian countries: a composite index approach. Ecol. Indicat. 106, 105507 https://doi.org/10.1016/j.ecolind.2019.105507.

Sharma, G.D., Verma, M., Shahbaz, M., Gupta, M., Chopra, R., 2022. Transitioning green finance from theory to practice for renewable energy development. Renew. Energy 195, 554–565. https://doi.org/10.1016/j.renene.2022.06.041.

Ullah, K., Rashid, I., Afzal, H., Iqbal, M.M.W., Bangash, Y.A., Abbas, H., 2020. SS7 vulnerabilities - a survey and implementation of machine learning vs rule based filtering for detection of SS7 network attacks. IEEE Commun. Surv. Tutorials 22. https://doi.org/10.1109/COMST.2020.2971757.

Umair, M., 2022. Muhammad Umair 45689 5847.

Wang, J., Cui, M., Chang, L., 2023. Evaluating economic recovery by measuring the COVID-19 spillover impact on business practices: evidence from Asian markets intermediaries. Econ. Change Restruct. https://doi.org/10.1007/s10644-023-09482-z.

Wang, S., Sun, L., Iqbal, S., 2022. Green financing role on renewable energy dependence and energy transition in E7 economies. Renew. Energy 200, 1561–1572. https://doi.org/10.1016/j.renene.2022.10.067.

Wu, L., Xu, L., 2020. The role of venture capital in SME loans in China. Res. Int. Bus. Finance 51. https://doi.org/10.1016/j.ribaf.2019.101081.

Wu, Q., Yan, D., Umair, M., 2022. Assessing the role of competitive intelligence and practices of dynamic capabilities in business accommodation of SMEs. Econ. Anal. Pol. https://doi.org/10.1016/j.eap.2022.11.024.

Xia, Z., Abbas, Q., Mohsin, M., Song, G., 2020. Trilemma among energy, economic and environmental efficiency: can dilemma of EEE address simultaneously in era of COP 21? J. Environ. Manag. https://doi.org/10.1016/j.jenvman.2020.111322.

Xiuzhen, X., Zheng, W., Umair, M., 2022. Testing the fluctuations of oil resource price volatility: a hurdle for economic recovery. Res. Pol. 79, 102982 https://doi.org/10.1016/j.resourpol.2022.102982.

Zhang, D., Mohsin, M., Rasheed, A.K., Chang, Y., Taghizadeh-Hesary, F., 2021. Public spending and green economic growth in BRI region: mediating role of green finance. Energy Pol. 153 https://doi.org/10.1016/j.enpol.2021.112256.

Zhang, Y., Prayag, G., Song, H., 2021. Attribution theory and negative emotions in tourism experiences. Tourism Manag. Perspect. https://doi.org/10.1016/j.tmp.2021.100904.

Zhou, M., Li, X., 2022. Influence of green finance and renewable energy resources over the sustainable development goal of clean energy in China. Res. Pol. 78, 102816 https://doi.org/10.1016/j.resourpol.2022.102816.