Comparing the Performances of Artificial Neural Networks Models Based on Autoregressive Fractionally Integrated Moving Average Models

Remal Shaher Al-Gounmeein and Mohd Tahir Ismail

Abstract—**The autoregressive fractional integrated moving average (ARFIMA) has become one of the popular linear models in time series modeling and forecasting in the past decades. Recent research in modeling and forecasting with artificial neural networks (ANN) suggests that these networks are a promising alternative to the traditional linear and nonlinear methods. ARFIMA models and ANNs are often compared with mixed conclusions in terms of superiority in forecasting performance. In this research, a hybrid methodology that combines both ARFIMA and multilayer perceptron (MLP) models is proposed to take advantage of the unique strength of the ARFIMA and MLP models in linear and nonlinear modeling, which is the primary objective of this study. This research uses the monthly Brent crude oil price series for the period of January 1979-July 2019. As for our other goal, the researchers' previous works are also extended by examining the linear and nonlinear methods for the dataset simultaneously and comparing individual models with the hybrid models. The best model is determined by comparing 19 individual and hybrid models in terms of forecasting accuracy based on the root mean squared error and Ljung–Box test. Empirical results with real datasets indicate that the ARFIMA (1,0.3589648,0)–MLP (1,2,1) hybrid model outperforms the separately used models and the other hybrid models, and the Akaike information criterion value is not the smallest for this model.**

Index Terms—**Autoregressive Fractionally Integrated Moving Average, Multilayer perceptron, Modeling and Forecasting, Hybrid Model.**

I. Introduction

TIMEseries forecasting is a vital field in developing and extending a model and describing the primary relationship of a dataset to study its future trend. Although modeling is a useful and important approach when the general pattern of the data sequence is unknown, it cannot describe the current and future patterns of the data. Numerous efforts in the past decades attempted to develop and improve time series forecasting models in various fields and describe data through illustrative, satisfactory, and accurate mathematical rules by describing the current and future patterns of the model. The individual model is a common forecasting method and a good approach used in many previous studies. Examples include the autoregressive integral moving average (ARIMA) model, autoregressive fractional integral moving average (ARFIMA) model, and artificial neural network (ANN) model. ARIMA (p, d, q) models are a popular class of models for time series data by assuming the differencing parameter (d) as an integer value. In the event that this model is extended, assuming the differential parameter (d) has a fractional value between $-0.5 < d < 0.5$, this kind of model with long memory behavior can be classified as an ARFIMA (p, d, q) model. ARFIMA models are linear time series models, but they are unsuitable for time series containing nonlinear structures.

ANN models mimic how the human brain works and rely on the principle of parallel processing in parallel layers. [1] studied in detail ANN models by conducting a literature review to forecast a time series. A similar research has recently been conducted by [2]. ANN models have been used for modeling and forecasting in several fields, both for linear and nonlinear time series; by contrast, ARFIMA models deal with linear series only, as mentioned above. ANNs have been applied to a wide range of disciplines, such as system identification and control, decision making, pattern recognition, medical diagnosis, finance, data mining, and visualization, among others [3]. ANN models can model any time series regardless of the structure of the series, and they are known to yield good forecasting results. Another approach in obtaining accurate forecasts is to use hybridization methods (i.e., based on more than one model) as a means of knowing future data forecasts and overcoming the disadvantages of individual models, such as those that deal with non-normal residuals. These methods can help solve problems of having both linear and nonlinear structures. Accordingly, these models are known as hybrid models.

In this study, forecasting the prices of Brent crude oil is regarded as an important task. Such prices are employed in future economic planning, especially in the context of significant challenges faced by the global economy, including the current volatilities caused by COVID-19. Brent crude oil prices are also important parameters in achieving economic development for all countries, whether for export or import programs, hence the equal importance of choosing the correct type of data in this research. This study aims to obtain the best accuracy for forecasting by modeling both linear and nonlinear patterns based on a long memory for a time series dataset by constructing individual and hybrid models. In view of determining the appropriate model, an additional procedure is performed, that is, to compare the results obtained for the individual models and the proposed hybrid models. Our study also aims to present an advanced methodology for ANN as an alternative to the generalized autoregressive conditional heteroscedastic (GARCH) and ARCH models when dealing with the hybridization approaches. In this manner, the problems and defects of individual models, such as those dealing with non-normal residuals, can be addressed.

Manuscript received August 24, 2020; revised January 12, 2021.

R. Al-Gounmeein, PhD. student, School of Mathematical Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia and a full-time lecturer, Department of Mathematics, Al-Hussein Bin Talal University, Ma'an, Jordan. e-mail: ralgounmeein@yahoo.com.

M. Ismail, Associate professor, School of Mathematical Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia e-mail:m.tahir@usm.my.

If highly accurate forecasts can be obtained, then the best plans can also be developed as a means of ensuring the best price for Brent crude oil through investor and decision-maker channels.

The remainder of this paper is organized as follows. Section 2 reviews the previous literature related to this study. Section 3 provides the actual time series data used in this study and a brief description of long memory tests and estimation and the model specifications of the ARFIMA and ANN models. In addition, the proposed hybrid method is discussed with the tests, criteria, and accuracy measures used to choose the best model. The results obtained from the study, the discussion, and future work are presented in Section 4. Section 5 provides the conclusions.

II. Literature review

The literature lists a number of studies that use different methods of modeling and forecasting, ranging from simple methods (individual models) to complex ones (hybrid models), dealing with different components of a time series. This section provides a brief presentation for some of the models in sequential order. According to [4], the ARFIMA model was created by Granjer and Joyeux in 1980 to capture the long memory behavior of a time series dataset. [5] examined numerous ARFIMA models. The ARFIMA (1,0.47,2) model was appropriate for the West Texas Intermediate series, while the ARFIMA (2,0.09,0) model was suitable for the Brent series, for the period of May 15, 1987 to December 20, 2013. [6] used the ARIMA, ARFIMA, and error correction models for the modeling and forecasting of monthly prices of wholesale mustard of Sri Ganganagar District in Rajasthan. The ARIMA (1,1,1) model was appropriate and even performed better compared with other models based on the mean absolute percentage error (MAPE) value of 6.60 percent. [7] examined different individual models (ARFIMA (1,0.474,2), ARFIMA (1,0.876,2), ARFIMA (5,0.786,1), ARFIMA (5,0.722,1), ARFIMA (1,0.413,2), and ARIMA (5,1,4)) to introduce an appropriate model for modeling and forecasting the total value of traded securities of the Arab Republic of Egypt. The ARFIMA (1,0.413,2) model outperformed and was even more accurate for forecasting compared with other models based on the root mean squared error (RMSE), mean absolute error (MAE), and MAPE values. [8] forecasted the sales volume of motorcycles in Indonesia by comparing the ARFIMA model with the singular spectrum analysis (SSA) model for the period of January 2005 to December 2016. The comparison results showed that the SSA model was superior to the ARFIMA model based on the MAPE value of 13.57 percent. [9] proposed ARIMA and ARFIMA to model the data of domestic air passengers in India for the period of January 2012 to December 2018. The forecast accuracy of the ARFIMA (1,-0.347,1) model was better than that of the ARIMA (1,1,1) model based on the RMSE, MAE, and MAPE values. [10] found that the ARFIMA (1,1.05716,[3]) model outperforms the ARFIMA (1,1.05716,0), ARFIMA ([3],1.05716,0), ARFIMA (0,1.05716,[3]), ARFIMA ([3],1.05716,1), ARFIMA ([3],1.05716,[3]), and ARFIMA ([3],1.05716,[1,3]) models when using the gold price data of Indonesia. Many other authors have become interested recently in obtaining and estimating the ARFIMA models as a means of choosing the best predictive model. [11] proposed a new class of long memory models with a flexible time-varying fractional parameter. The resulting model is based on the theory of generalized autoregressive (AR) score models and allows the long memory parameter to vary dynamically over time. Their results are promising for both simulated and real-time series. [12] showed that the horizon dependence of the cluster entropy is related to long-range positive correlations in financial markets. Their result was obtained by applying the moving average (MA) cluster entropy approach to long-range correlated stochastic processes, such as ARFIMA and fractional Brownian motion. Their proposed approach could also capture detailed horizon dependence over relatively short horizons (i.e., 1-12 months), thus highlighting its relevance in defining risk analysis indices. [13] found that ARFIMA models can achieve better forecasting performance compared with short memory alternatives for all long memory generating mechanisms and forecast horizons. This ability is achieved whenever the long memory of the processes has a high degree, regardless of the generated mechanism. The obtained results are particularly useful for climate econometrics, financial econometrics, and other models that deal with forecasts at different horizons. [14] assessed the performance of various time series models for electroencephalography (EEG) data by using the Akaike information criterion (AIC) as the metric. Their results confirmed that EEG signals could exhibit long-range dependencies, and the ARFIMA models are better suited for capturing temporal correlations compared with the conventional ARMA models.

The literature on neural networks is enormous, and their application has spread over many scientific areas [15]. [16] applied the method of the multilayer perceptron (MLP) model and a causal method based on the ANN model. The components of a decomposed time series were used as the input variables to deal with the demand variability of a seasonal time series by using ANN. The data used in the study were the airline passenger monthly dataset covering the period of January 1949 to December 1960. The ANN model yielded good accuracy regardless of whether the real-time series was decomposed. [17] applied the feed-forward neural network (FFNN) model procedure to model and forecast Indonesian financial data by using a set of monthly data (76 observations in total) for the period covering January 2009 to April 2015. The FFNN model performed better than the ARIMA $(1,0,0)(1,0,0)_{11}$ and ARIMAX models based on the MSE value. [18] compared the seasonal ARIMA (SARIMA) and MLP models that were used for US' quarterly energy consumption dataset covering the period of January 1973 to June 2015. The neural network model was slightly superior to the SARIMA $(1,0,1)(0,1,1)_4$ model based on the MAE and MAPE values. [19] used various ANN models with different learning algorithms, activations functions, and performance measures to model and forecast the Turkish Lira/US Dollar (TRY/USD) exchange rate. The ANN models were built on the weekly dataset covering the period of January 2010 to April 2016. The variable learning rate backpropagation (BP) learning algorithm with the tan–sigmoid activation function attained the best performance for the TRY/USD exchange rate forecasting. A comparative study of the BP algorithm with two types of networks, namely, the general regression neural network (GRNN) and the GARCH model, was conducted by [20] to study the effectiveness and suitability of ANN in handling the non-homogeneous variance of a financial series. The application part was applied to the Egyptian exchange market to study the local currency exchange rate volatility for the period between January 1, 2009 and June 4, 2013. The BP network was more accurate than the other two models in representing the financial series based on the RMSE and MAPE values. [21] used MLP with a feed-forwarded BP algorithm and the sigmoid activation function to forecast air passengers traveling by domestic flights in India. The MLP network with three hidden neurons and one hidden layer achieved a minimum error and attained the lowest RMSE and MAE values compared with the other two networks, namely, the MLP network with one hidden neuron and the MLP network with two hidden neurons. [22] tested MLP neural network models with different combinations of transfer functions and a net input function and different numbers of neurons and layers to forecast solar power. The evaluation showed that the error decreases with the increase in the number of layers and number of neurons based on the obtained MSE values.

The most important finding in the reviews presented above is the obtainment of mixed results when selecting the appropriate model for modeling and forecasting. The finding is apparent when either the ARFIMA or ANN model was compared with other models, indicating that no single model can appropriately deal with the linear and nonlinear characteristics of a time series. As a means of confirming this observation, some of the previous studies dealing with the hybridization approach, particularly the models combined with ARFIMA and ANN to model and forecast any type of time series, are highlighted in the succeeding discussions.

Several techniques can be used to build a nonlinear combination model (NCM), including the ARFIMA model with the support vector machine model and the BP model. [23] used the Renminbi exchange rate against the US dollar (RMB/USD) and the Euro (RMB/EUR) as experimental examples for NCM performance evaluation. The basic idea of the proposed model was to ensure effective prediction by combining the different models' advantages. The results showed that the forecasting performance of the NCM was suitable and even better compared with the single models and linear combination models. [24] studied the sales forecasts of a global furniture retailer operating in Turkey by using state-space models, ARIMA and ARFIMA models, neural networks, adaptive network-based fuzzy inference system, and their combined versions. The empirical results showed that most of the combined forecasts could achieve statistically significant increases, and their forecasting accuracies are better compared with the individual models. [25] proposed a novel hybrid forecasting model by combining ARIMA and ANN. The MA component and the annual seasonal index were incorporated into the analysis of Thailand's cassava export (native starch, modified starch, and sago). Their proposed model attained the lowest error in contrast to those of the ARIMA, ANN, and ARIMA–ANN models based on prediction accuracy measures, namely, MSE, MAE, and MAPE. [26] compared a hybrid Pakistan Stock Exchange (PSX) forecasting model with the ARIMA model, long-short term memory (LSTM) model, and generalized regression radial basis neural network (i.e., a GRNN). The experimental results showed that the proposed AFRIMA–LSTM hybrid model not only could minimize the volatility problem but also could overcome the overfitting problem of neural networks. The obtained forecasting performance also indicates the effectiveness of the proposed hybrid model relative to the other models based on the RMSE, MSE, and MAPE values. [27] introduced a class of ARFIMA–GARCH models with level shift (LS)-type intervention that can capture three major time-series features: long-range dependence, volatility, and LS.

Numerous past and ongoing studies have used time series analysis for crude oil price data, such as the works of [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40]. Their results indicate that crude oil prices change significantly over time. Consequently, our current study focuses on constructing a time series model to forecast monthly Brent crude oil prices by using ARFIMA, ANN, hybridization models. These models are also compared in this study.

III. Materials and methods

We elaborate in this section the methods and the data used to apply the methods by adopting actual time series data. The first method deals with a long memory and adopts several tests and estimations. The second method deals with ANN models, specifically MLP. The third method corresponds to the proposed hybrid method and is discussed along with the applied tests, criteria, and accuracy measures.

A. Dataset

The dataset used in this study are the monthly Brent crude oil prices (all prices are per barrel in USA $), which are obtainable from www.indexmundi.com/commodities/?commodity=crude-oil-bren, totaling 487 observations. The dataset covers the period of January 1979 to July 2019. The data from January 1979 to July 2018 are used as the training dataset (475 observations), while the rest of the data starting from August 2018 are used as the testing dataset (12 observations). The datasets are processed because of the significant challenges presently faced by the global economy, e.g., perceptible volatilities of Brent crude oil prices due to COVID-19. The R-software (version 3.5.3) is used to perform all statistical analyses.

B. Long memory tests and estimations

Long memory, a phenomenon that can be observed in a time series, manifests when the distance between two points is increased [41], and it greatly impacts the financial field [42]. This long memory feature can be identified when the autocorrelation function (ACF) decays more slowly than the exponential decay [41]. Several statistical methods can be applied to check the existence of the long memory feature [43]. The methods include the R/S Hurst, Higuchi, and aggregated variance methods. The approaches used to test and estimate long memory parameters are the Hurst exponent and Geweke and Porter–Hudak (GPH) methods (see [43], [44] and [45] for details). The smoothed periodogram (Sperio) and fractionally differenced (Fracdiff) are used as functions in the R-software for fractional difference (d) value estimation. [46] proposed the use of the Sperio function to estimate the fractional difference (d) in the ARFIMA (p, d, q) model. [47] explained the importance of the Fracdiff operator, which uses the regression estimation method to estimate the fractional difference (d) for the ARFIMA (p, d, q) model [48]. The fractional difference factor (d) is defined by a binomial series as follows:

$$\begin{aligned} \nabla^d &= (1-B)^d \\ &= \sum_{k=0}^{\infty} \binom{d}{k}(-B)^k \\ &= 1 - dB - \tfrac{1}{2}d(1-d)B^2 \\ &\quad -\tfrac{1}{6}d(1-d)(2-d)B^3 - \cdots \end{aligned} \tag{1}$$

(See [46]-[48] for details.)

Long memory characteristics are generated by a non-stationary structural break [49]. The structural breaks of a time series therefore, should be tested because they determine whether a long memory is present or imaginary, as pointed out by [50], [51], and [49]. Chow introduced the single break test in [52] and since then has been modified as the Quandt likelihood ratio (QLR) test. The QLR test is performed to determine the break between (t_0 and t_1), also called the supremum F-statistic [5], which is given by

$$\sup F = \max\{F(t_0), F(t_0+1), \cdots, F(t_1)\} \tag{2}$$

where the Sup F-statistic is the largest among the given values. If the P-value of the F-statistic is < 0.05, then the test rejects the null hypothesis (H_0, i.e., no structural break).

C. ARFIMA models

The general formula of the ARFIMA (p, d, q) model is similar to that of the ARIMA (p, d, q) model, except for the difference value (d), which is given as follows:

$$\phi_p(B)(1-B)^d x_t = \theta_q(B)\epsilon_t \qquad \text{for } 0 < d < 0.5 \tag{3}$$

where the value (d) is a non-integer and a non-seasonal difference order, $\{x_t\}$ is a dependent variable at time t, (ϵ_t) is a white noise process, and $\phi_p(B)$ and $\theta_q(B)$ represent the AR for the order (p) components and the MA for the order (q) components with backward shift operators (B), respectively (see [44], [53] for details). The factor (d) is defined by a binomial series, as depicted in Equation (1).

D. ANN models

ANN models are a common topic in modern data analysis, and they can be classified as a semi-parametric method [15]. This class of models can learn complex tasks, such as recognition, decision making, or predictions [54], and can deal with nonlinear data [55]. Moreover, ANNs can solve a wide range of problems in several areas of artificial intelligence and machine learning [56]. The power of ANN models emanates from the parallel processing of information in the data, and no prior assumptions are required when building the model; thus, the network model can be simply determined on the basis of the data's properties [55].

The ANN structure includes the number of layers and the total number of nodes in each layer, but the determination of these layers is accomplished through experimentation because a theoretical basis is lacking [15], [55]-[57]. ANNs include one input layer where external information is received, an output layer where the problem is solved, and one or more hidden layer/s that separate the input layer from the output layer, each of them containing one node or more to connect each layer to the next top layer. The activation (transfer) function determines the relationship between the inputs and the outputs of a node and a network, in which the relationship is represented by the sigmoid (logistic), hyperbolic tangent (tanh), sine or cosine, or linear function [1]. Among them, the logistic function is the most popular option in the literature and thus is also used in our study. The relationship between the output (y_t) and the inputs ($y_{t-1}, y_{t-2}, \cdots, y_{t-p}$) has the following mathematical formula [55], [24]:

$$\begin{aligned} y_t &= \omega_0 + \sum_{j=1}^{q} \omega_j f\left(\omega_{oj} + \sum_{i=1}^{p} \omega_{ij} y_{t-i}\right) + \epsilon_t \\ &\text{with } \epsilon_t^{iid} \sim N(0, \sigma^2) \end{aligned} \tag{4}$$

where $(j = 1, 2, \ldots, q)$, $(i = 1, 2, \ldots, p), \omega_0$ and ω_{0j} are the biases on the nodes, ω_j and ω_{ij} are the connection weights between the layers of the model, $f(\cdot)$ is the activation function of the hidden layer, (p) is the number of input nodes, (q) is the number of hidden nodes, ϵ_t is a white noise, and iid is independently identically distributed. Furthermore, the logistic function used in this study has the following mathematical formula [55]-[56]:

$$f(x) = (1 + e^{-x})^{-1} \tag{5}$$

Hence, the ANN model represented in Equation (4) basically performs a nonlinear functional mapping, from the past values ($y_{t-1}, y_{t-2}, \cdots, y_{t-p}$) to the future value ($y_t$), which is expressed as follows:

$$y_t = g(y_{t-1}, y_{t-2}, \cdots, y_{t-p}, \omega) + \epsilon_t \tag{6}$$

where ω is a vector consisting of all parameters, and $g(\cdot)$ is a function specified by the network architecture (i.e., network structure and connection weights). In other words, the ANN model is equivalent to a general nonlinear AR model. When nonlinear activation functions, such as the logistic function, are used at the output nodes, the desired output values must be transformed to the range of the actual outputs of the network, which is typically represented by (0,1) or (-1,1). Data normalization is often performed prior the training process (see [1] for details), hence use the same approach in this study.

An ANN model with a single hidden layer feedforward network is the most widely and preferred model when building a network model for modeling and forecasting time series data [1]. Feedforward multilayer neural networks (FFMNNs), also known as MLPs, with a single hidden layer are commonly used in the area of this study (Figure 1).

The MLP model is expressed as follows:

$$Y_t = \alpha + \sum_{j=1}^{n} \alpha_j f\left(\beta_j + \sum_{i=1}^{m} \beta_{ij} Y_{t-i}\right) + \epsilon_t \tag{7}$$

where (α) is a vector for the weights between the (n) hidden nodes and the output node, (β) is a vector for the

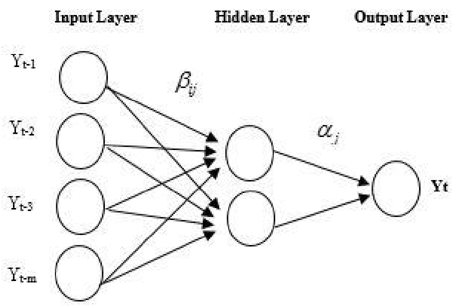

Fig. 1. MLP neural network [58]

weights between the (m) input nodes and the hidden node, (j) denotes the number of nodes in (i) depth of the network, $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$, and $\alpha, \beta \in [0, 1]$. MLP neural networks are used to solve a variety of problems, especially in modeling and forecasting, because of their ability to perform arbitrary input-output charting [1].

E. Hybrid ARFIMA–ANN models

Both ARFIMA and ANN models have achieved successes in their own linear or nonlinear domains when used for time series data. However, none of them is a comprehensive model that is appropriate for all modeling and forecasting circumstances. In other words, ARFIMA is an insufficient model for complex nonlinear problems, especially when studying the trend of series. Similarly, ANN models have shown mixed results when modeling linear problems, as depicted by previous studies. Consequently, the hybrid models have been proposed to model simultaneously a time series with both linear and nonlinear properties.

The motivation for including the hybrid technique in this study emanates from the defects that appear in the modeling of the aforementioned individual models. However, a difficulty encountered in proposing hybrid techniques is determining how single models should be hybridized together, especially when dealing with linear and nonlinear data (i.e., ARFIMA and ANN models), to obtain accurate forecasts. The resolution process is illustrated here by modifying Zhang's [55] hybrid approach. That is,

$$y_t = L_t + N_t \tag{8}$$

where y_t denotes the original time series dataset for the period (t), L_t is the linear component, and N_t is the nonlinear component. The linear component is estimated by the ARFIMA model; then, the residuals obtained from the ARFIMA model are estimated by the ANN model, which represents the nonlinear component of the series (see [55], [59] for details). In this manner, forecasting values of the time series dataset can be obtained as follows:

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \tag{9}$$

where $\hat{Y}_t$ denotes the forecasting values of the time series dataset for the period (t), $\hat{L}_t$ is the forecast values for the period (t) from the estimated relationship (3), and $\hat{N}_t$ is the forecast values for the period (t) from the estimated relationship ((4) or (7)). Subsequently, the forecasting obtained from the two models are summed separately to obtain the final predicted value of the proposed hybridization model.

The hybrid methods' applications in the previous literature show that combining different methods can effectively and efficiently improve forecasts [59]; this aspect is also discussed in our study. The proposed hybrid method is expected to exploit and combine the strengths and advantages of both ARFIMA and ANN models.

F. Stationary test

The augmented Dickey–Fuller (ADF) test and the Phillips–Perron (PP) test are used to check the stationary feature of the dataset [60]-[61].

G. Ljung–Box test

The Ljung–Box test is an essential step in examining the correlation between residuals in the model (see [62] for details).

H. Information criteria and accuracy measures

The fit model selection of the above mentioned models is based on a set of criteria, including the AIC (Equation (10)) [53], [63]. The AIC is given by the following formula:

$$AIC = -2\ln(l) + 2k \tag{10}$$

where (l) is a maximum likelihood for the model, and (k) is the total number of parameters in Equation (3). The model with the lowest AIC value is treated as the best model. Furthermore, the RMSE (Equation (11)) is used as an accuracy measure to evaluate the performance of the model [62].

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2} \tag{11}$$

where (Y_t) is the actual value, and $(\hat{Y}_t)$ is the forecasted value.

IV. Results and discussion

The monthly crude oil price for the Brent series is shown in (Figure 2). The $\{x_t\}$ in the plot denotes the price, and (t) represents the time in all months. The series depicts a stable price followed by a gradual increase and a decrease, and so on. The descriptive statistics of the Brent price is as follows: a mean of 42.95, a median of 30.20, and a positive skewness of 1.177466. All the structural breaks are visible in the series, i.e., the breaks in 1986, 1999, 2005, and 2013.

The preliminary result of the structural break by using the QLR test shows that the null hypothesis is rejected due to the extensive Sup-F statistic (1190) and the extremely small P-value ($< 2.2e-16$), which is less than 0.05 for the Brent series. The ACF of the Brent price (Figure 3) has a slow decay, which is a typical behavior of long memory processes. The long memory is confirmed, as shown in Table I, according to the statistical methods.

Table I shows that all (H) values are higher than 0.5, which firmly concludes the existence of the long memory

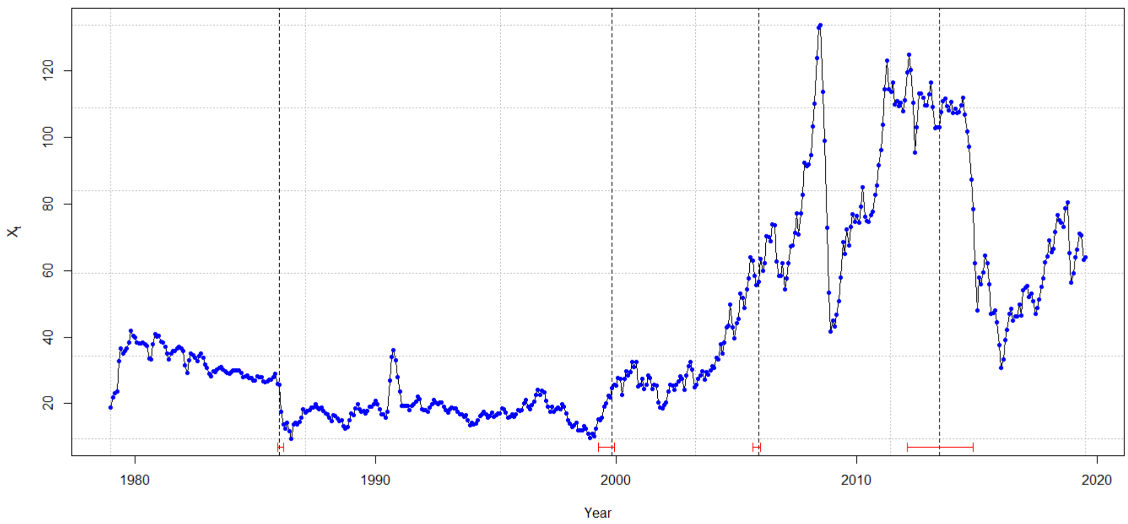

Fig. 2. Monthly Brent crude oil prices (\$/bbl) with all breaks and their confidence intervals

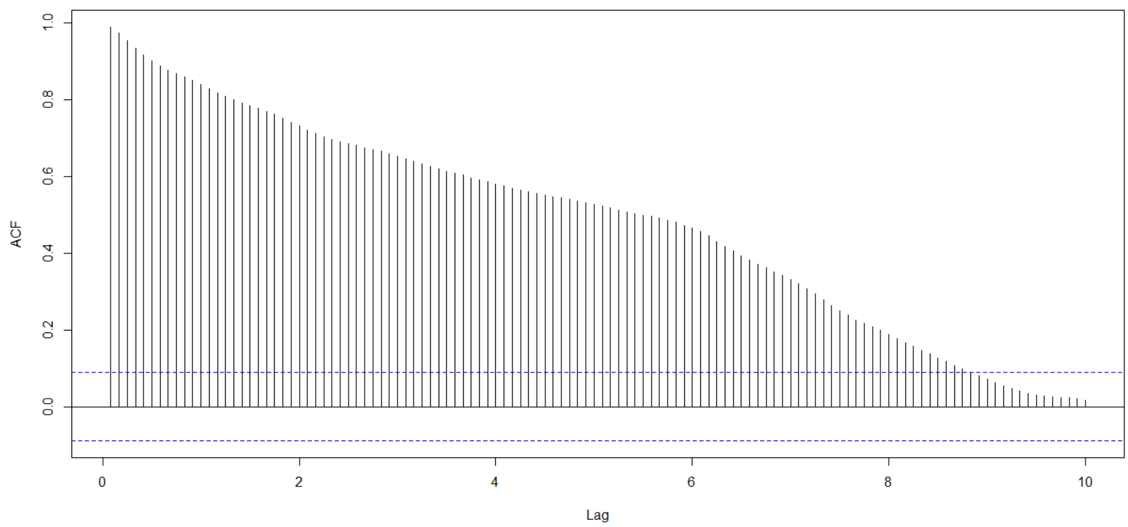

Fig. 3. ACF of the $\{x_t\}$ series

TABLE I
LONG MEMORY TESTS FOR THE $\{x_t\}$ TIME SERIES

R/S Hurst Method	Aggregated Variance Method	Higuchi Method
H = 0.8531864	H = 0.7910981	H = 0.9578515

characteristic of the Brent price. This time series is not normal (P-value of $< 2.2e - 16$) based on the Jarque–Bera test and the coefficient of skewness mentioned earlier. Moreover, the $\{x_t\}$ transformation must be accomplished. Here, $\{Y_t\}$ represents the growth rate of $\{x_t\}$, as shown in the following formula:

$$Y_t = \log(x_t) \tag{12}$$

Figure 4 shows the ACF and partial autocorrelation function (PACF) values of the $\{Y_t\}$ series. No forms of white noise can be observed.

The fractional difference (d) of the $\{Y_t\}$ series is estimated using different methods and functions (Table II). The (d) value is 0.3589648 in the R/S Hurst analysis; 0.4984955 is the Sperio estimate, and 0.4994726 is the result of the Fracdiff estimate. GPH estimation is excluded because its value is > 0.5.

After the fractional difference (d_i) is computed using Equation (12), the series is transformed as follows:

$$Z_{t(di)} = diff(Y_t) = Y_t \triangledown^{di} \tag{13}$$

where $d_i = d_1, d_2$ and d_3. The stationary test results of the $Z_{t(d_i)}$ series are illustrated in Table III and (Figure 5). The P-values of the ADF and PP tests (Table III) reveals that the series has become stationary after computing for the fractional difference, which is also confirmed by (Figure 5).

In concordance with Equation (10), a qualifying model is

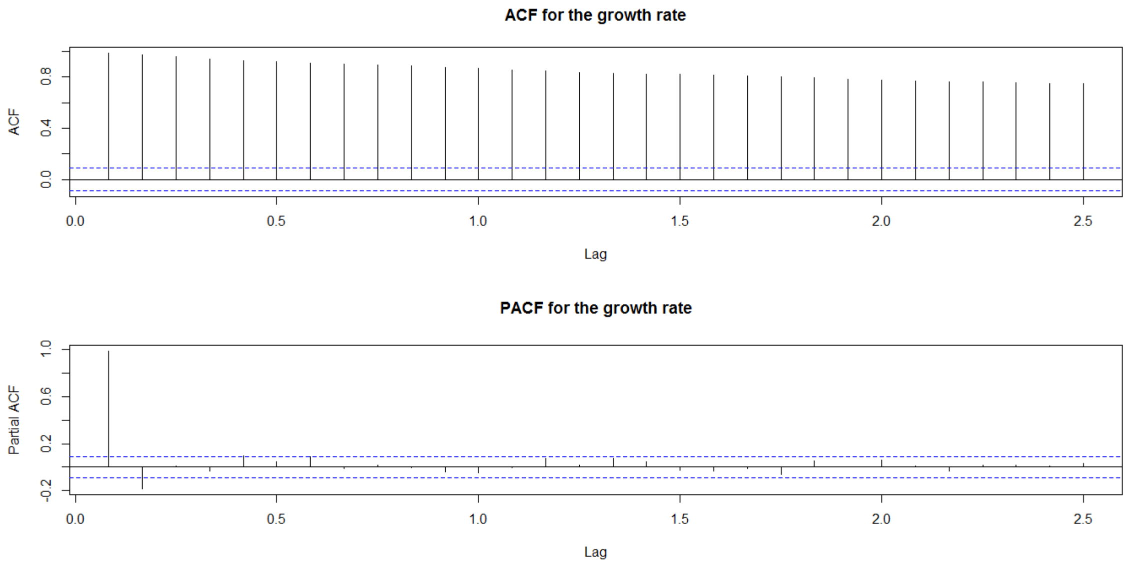

Fig. 4. ACF and PACF plots of the $\{Y_t\}$ series

TABLE II
LONG MEMORY ESTIMATION FOR $\{Y_t\}$ SERIES

Method / Function	d	State
R/S Hurst (d = H-0.5)	$d_1 = 0.3589648$	$0 < d_1 < 0.5$
Sperio (bandw.exp = 0.3, beta = 0.74)	$d_2 = 0.4984955$	$0 < d_2 < 0.5$
Fractionally Differenced (Fracdiff)	$d_3 = 0.4994726$	$0 < d_3 < 0.5$
Geweke and Porter–Hudak (GPH)	$d_4 = 0.7676326$	$0.5 < d_4$

TABLE III
STATIONARY TEST FOR THE $Z_{t(d_i)}$ SERIES

Method/Function	Test	Value	P-value	State
R/S Hurst	ADF Test for $Z_{t(d_1)}$ series	-4.1727	0.01	Stationary
	PP Test for $Z_{t(d_1)}$ series	-82.923	0.01	Stationary
Sperio	ADF Test for $Z_{t(d_2)}$ series	-5.1927	0.01	Stationary
	PP Test for $Z_{t(d_2)}$ series	-151.34	0.01	Stationary
Fracdiff	ADF Test for $Z_{t(d_3)}$ series	-5.2001	0.01	Stationary
	PP Test for $Z_{t(d_3)}$ series	-151.89	0.01	Stationary

a model with the lowest AIC value. Thus, for the said dataset (Table IV), the best-selected models for the training period are ARFIMA $(1, 0.3589648, 0)$, ARFIMA $(2, 0.3589648, 1)$, and ARFIMA (2,0.3589648,2) based on the AIC values of $-962.91, -966.25$, and -966.07, respectively. Besides, the three models belong to the Hurst estimate, which has the lowest value for the fractional difference (d) estimate.

Ljung–Box, Jarque–Bera, and Shapiro–Wilk normality tests are then performed to assess the existence of a nonlinear pattern in the residuals extracted from the selected ARFIMA models. The resulting P-values of these tests (Table V) reject the null hypothesis of iid, suggesting that nonlinear structures exist in the dataset. In other words, the three models do not have the property of the unit root for the residuals. Besides, the residuals of the models are not normally distributed. Subsequently, the obtained residuals from the selected three ARFIMA models are shaped using the neural network models.

As a result of the preceding discussions, the models should be able to deal with the problem of having both linear and nonlinear structures for the selected type of time series; these models are ANN models. Therefore, the three models mentioned above may be taken and hybridized with the multilayer ANN models (i.e., MLP) depending on the results of the second phase, in which the residuals obtained in the first phase (ARFIMA modeling) are analyzed using the MLP model. The result belongs to a hybrid method called the ARFIMA–MLP model. Furthermore, in the hybridization method involving forecasting steps, the predictions can be summed after obtaining separately the forecasted values form the ARFIMA model and MLP model.

In view of finding the best MLP architecture, a set of one to five neurons is tested with three network layers. Each model in the training set is trained more than 52 times for the network's steps. Moreover, the error and AIC of the set is randomized to compare their performance. Table VI provides a summary of the information related to the selected network architectures and models. Before the data can be entered into the network, they must be processed via the normalization step, as mentioned earlier. Residuals belonging to the interval of $[0, 1]$ are suitable for the sigmoid function used in the hidden layer.

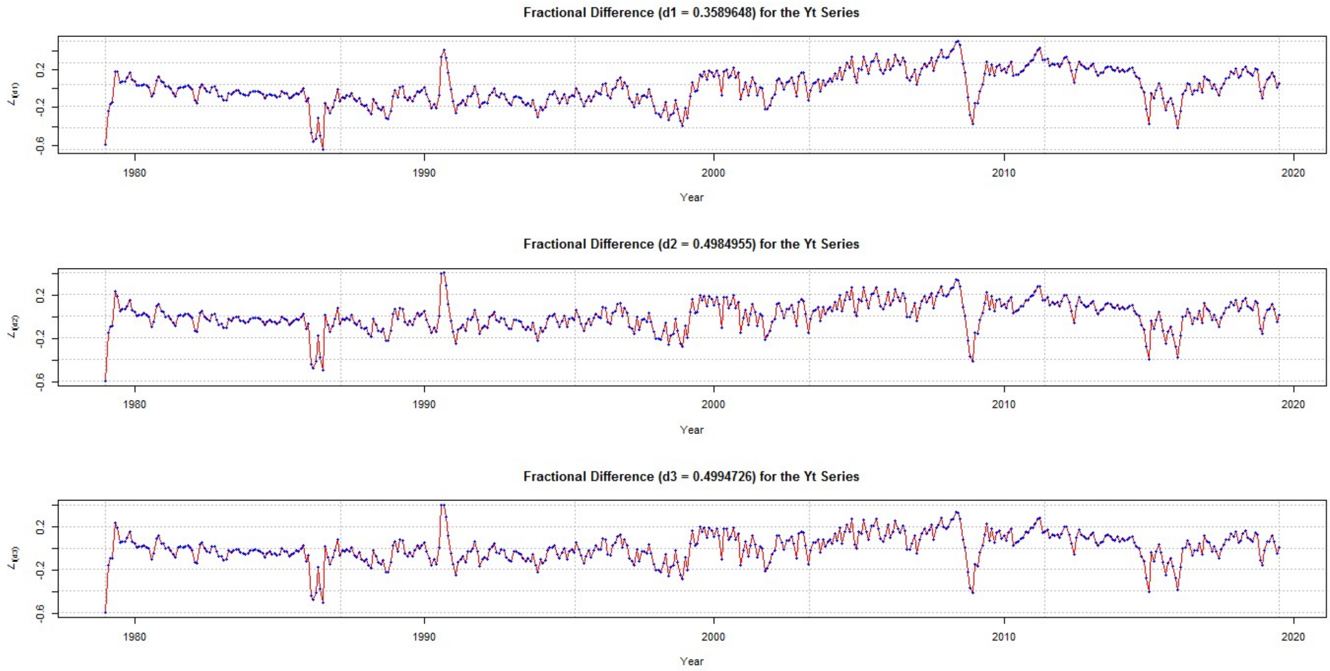

Fig. 5. Time series plots for $Z_{t(d_i)}$ after calculating the fractional difference (d_i)

TABLE IV
AIC VALUES OF THE ARFIMA MODELS

Procedure/d	Model	AIC
Hurst $d_1 = 0.3589648$	$(1, d_1, 0)$	-962.91
	$(0, d_1, 1)$	-639.85
	$(1, d_1, 1)$	-961.07
	$(2, d_1, 0)$	-961.04
	$(0, d_1, 2)$	-769.74
	$(2, d_1, 1)$	-966.25
	$(1, d_1, 2)$	-962.70
	$(2, d_1, 2)$	-966.07
Sperio $d_2 = 0.4984955$	$(1, d_2, 0)$	-956.67
	$(0, d_2, 1)$	-818.60
	$(1, d_2, 1)$	-954.80
	$(2, d_2, 0)$	-954.77
	$(0, d_2, 2)$	-876.59
	$(2, d_2, 1)$	-959.16
	$(1, d_2, 2)$	-955.83
	$(2, d_2, 2)$	-957.37
Fractionally Differenced (Fracdiff) $d_3 = 0.4994726$	$(1, d_3, 0)$	-956.64
	$(0, d_3, 1)$	-819.47
	$(1, d_3, 1)$	-954.77
	$(2, d_3, 0)$	-954.74
	$(0, d_3, 2)$	-877.09
	$(2, d_3, 1)$	-959.11
	$(1, d_3, 2)$	-955.79
	$(2, d_3, 2)$	-957.31

Table VI summarizes the error measure (i.e., RMSE) depicting the performance of all individual and hybrid models related to the test set. The empirical analysis confirms that the performances of all models (RMSEs) are within 0.1 and therefore, close to the real values of the series. These results indicate that all models are likely to perform well in the forecasting phase. The ARFIMA (2,0.3589648,2), ARFIMA (1,0.3589648,0)–MLP (1,2,1), and MLP (1,3,1) models have the smallest values for this measurement. The best ARFIMA individual model is the ARFIMA (2,0.3589648,2) model. The individual model for the MLP network is the MLP (1,3,1) model, which consists of the following three layers: the input layer with one neuron representing the values of the real dataset; a hidden layer with three neurons; and the output layer with one neuron representing the current Brent price values. In Zhang's hybrid ARFIMA–ANN model, the appropriate ARFIMA models were first set, then the neural networks (i.e., MLPs) were trained using the residual values of those ARFIMA models. In consonance with the above findings, the best hybrid model is the ARFIMA (1,0.3589648,0)–MLP (1,2,1) model.

The Ljung–Box test of the residuals for the individual models shows that the residuals are not white noise and not independent [64], which differ for the hybrid model (Table VII).

The experiential results using the dataset for the Brent crude oil prices are shown in (Figure 6). The hybrid ARFIMA (1,0.3589648,0)–MLP (1,2,1) model significantly outperformed the individual models, namely, ARFIMA (2,0.3589648,2) and MLP (1,3,1), in the Ljung–Box test. Thus, the proposed hybridization method is an effective forecasting technique in obtaining an accurate hybrid model. Another important finding is the inability of the individual models to model accurately and forecast effectively the dataset used in this study. The numbers of nodes in the hidden layers are an important approach for a neural network to forecast a time series based on the work of [65]. The performance of the individual MLP models in the test dataset are better than the individual ARFIMA models in terms of forecasting accuracy based on the RMSE value. As mentioned in [66] as the ANN method is considered to be more efficient in forecasting results with less errors. The

TABLE V
STATISTICAL TESTS OF THE RESIDUALS

Model/Test	Ljung–Box Test			Jarque–Bera Test	Shapiro–Wilk Normality Test
	Lag (12)	Lag (24)	Lag (36)		
ARFIMA (1,0.3589648,0)	0.02570	0.02195	0.04968	$< 2.2e^{-16}$	$1.585e^{-8}$
ARFIMA (2,0.3589648,1)	0.08763	0.06623	0.15880		$1.919e^{-9}$
ARFIMA (2,0.3589648,2)	0.12860	0.11040	0.22870		$1.686e^{-9}$

TABLE VI
SUMMARY OF MODELS AND NETWORK ARCHITECTURE

Model	Training Set			Test Set
	AIC	Error Network	Network Steps	RMSE
ARFIMA (1,0.3589648,0)	-962.91	–	–	0.08981462
ARFIMA (2,0.3589648,1)	-966.25	–	–	0.08886256
ARFIMA (2,0.3589648,2)	-966.07	–	–	0.08800826
ARFIMA (1,0.3589648,0)-MLP (1,2,1)	19.661926144	2.878314000	76	0.1128193
ARFIMA (1,0.3589648,0)-MLP (1,3,1)	25.661604851	2.830802425	75	0.1134101
ARFIMA (1,0.3589648,0)-MLP (1,4,1)	31.662736293	2.831368146	52	0.1131051
ARFIMA (1,0.3589648,0)-MLP (1,5,1)	37.663901979	2.831950990	66	0.1135345
ARFIMA (2,0.3589648,1)-MLP (1,2,1)	19.697854137	2.848927068	69	0.1197690
ARFIMA (2,0.3589648,1)-MLP (1,3,1)	25.697898662	2.848949331	83	0.1196893
ARFIMA (2,0.3589648,1)-MLP (1,4,1)	31.697602383	2.848801192	58	0.1190429
ARFIMA (2,0.3589648,1)-MLP (1,5,1)	37.705028386	2.852514193	66	0.1196126
ARFIMA (2,0.3589648,2)-MLP (1,2,1)	19.756627322	2.878313661	76	0.1159453
ARFIMA (2,0.3589648,2)-MLP (1,3,1)	25.756985628	2.878492814	60	0.1157175
ARFIMA (2,0.3589648,2)-MLP (1,4,1)	31.756654775	2.878327388	57	0.1159331
ARFIMA (2,0.3589648,2)-MLP (1,5,1)	37.761009731	2.880504866	62	0.1156413
MLP (1,2,1)	18.80420	2.402098	34096	0.05630954
MLP (1,3,1)	23.02168	1.510841	47484	0.05603607
MLP (1,4,1)	30.59673	2.298363	9478	0.05615925
MLP (1,5,1)	35.09408	1.547039	58352	0.05656235

TABLE VII
LJUNG–BOX TEST STATISTIC FOR THE RESIDUALS

Model	Ljung–Box test	State
ARFIMA (2,0.3589648,2)	0.02449	P-value$<$0.05
ARFIMA (1,0.3589648,0)-MLP (1,2,1)	0.07631	P-value$>$ 0.05
MLP (1,3,1)	0.02526	P-value $<$ 0.05

results are clearly better when the MLP (1,2,1) model is used as part of the hybrid model. Therefore, the MLP model is preferred in the proposed hybrid method. However, the best model fails to attain the smallest AIC value. Thus, we can fully utilize the strength of each model by combining the flexibility of ARFIMA and the power of ANN in linear and nonlinear modeling for complex problems facing researchers. This combination method can be an effective way to improve forecasting performance in future studies. The empirical flexible result of this method enables addressing complex issues in any dataset in terms of predictive nonlinear behavior. Thus, the proposed predictive model has excellent test accuracy with respect to the residuals. The result also justifies the role of the test sample, which is not to over-increase the training steps of the network. The obtained results also indicate that increasing the examined neurons in the hidden layer in the training set do not have a significant effect, and the values show unstable volatility in increase and decrease in terms of error network and network steps. Thus, the predicted values of these models may be unstable. The obtained results are interesting in the sense that the method has difficulty achieving precision in the modeling of Brent crude oil prices. These findings highlight the relevance of the proposed hybrid methodology.

The change in the Brent oil prices are likely to show significant changes in the future movement of the US dollar exchange rate over time. Increases in Brent oil prices are linked to dollar appreciation in the long run. Subsequently, the changes can affect the liquidity of the global market, international trade, and economic activity in all countries, as mentioned by [67]. Thus, for a future research direction, a comparison of the models and the hybridization with other types of neural networks, such as recurrent neural networks or feedback neural networks (e.g., Elman neural network), maybe pursued to improve forecasting accuracy. Especially, through the use of new activation functions to analyze the complex-valued neural networks [68]. In addition, the effect of an increased number of hidden layers on network accuracy may be studied.

V. CONCLUSION

The real datasets of a long memory time series are analyzed using ARFIMA models, which are based on linear structures but are not adequate for nonlinear structures.

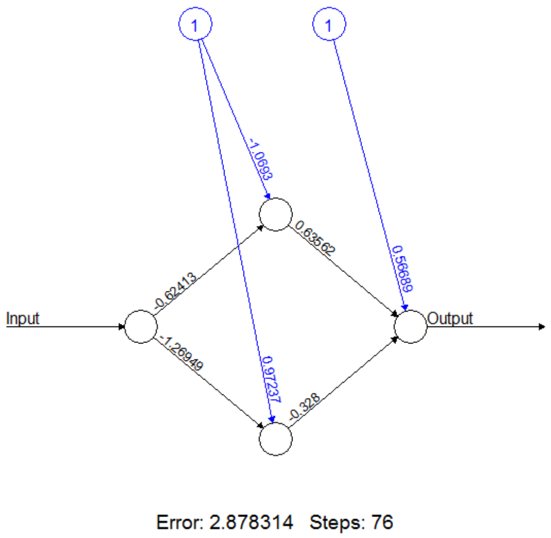

Fig. 6. Hybrid ARFIMA (1,0.3589648,0)–MLP (1,2,1) model

Hybrid methods that combine linear and nonlinear models may effectively model and improve forecasting performance. Motivated by this idea on hybridization, this study estimates and evaluates the ARFIMA and MLP individual models by using the real-life dataset of Brent crude oil prices and proposes a hybrid ARFIMA–MLP model to increase the forecasting accuracy. The experimental results prove that the ARFIMA (1,0.3589648,0)–MLP (1,2,1) hybrid model can outperform the other models as depicted by the Ljung–Box test. Furthermore, the obtained results verify that a hybrid model that combines ARFIMA and MLP can increase forecasting accuracy. Good forecasts can be used to adjust the supply and demand of future Brent prices.

REFERENCES

[1] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," *International journal of forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[2] M. N. Aziz, "A review on artificial neural networks and its' applicability," *Bangladesh Journal of Multidisciplinary Scientific Research*, vol. 2, no. 1, pp. 48–51, 2020.

[3] T. C. T. Chen, C. L. Liu, and H. D. Lin, "Advanced artificial neural networks," *Algorithms*, vol. 11, no. 7, pp. 1–2, 2018.

[4] H. Mostafaei and L. Sakhabakhsh, "Using sarfima model to study and predict the iran's oil supply," *International Journal of Energy Economics and Policy*, vol. 2, no. 1, pp. 41–49, 2012.

[5] S. A. Jibrin *et al.*, "Arfima modelling and investigation of structural break(s) in west texas intermediate and brent series," *CBN Journal of Applied Statistics*, vol. 6, no. 2, pp. 59–79, 2015.

[6] R. K. Bannor and M. Melkamu, "Arfima, arima and ecm models forecasting of wholesale price of mustard in sri ganganagar district of rajasthan of india," *Journal of Business Management & Social Sciences Research*, vol. 5, no. 1, pp. 1–13, 2016.

[7] R. A. H. Mohamed, "Using arfima models in forecasting the total value of traded securities on the arab republic of egypt," *IJRRAS*, vol. 27, no. 1, pp. 26–34, 2016.

[8] Y. O. Sitohang and G. Darmawan, "The accuracy comparison between arfima and singular spectrum analysis for forecasting the sales volume of motorcycle in indonesia," in *AIP Conference Proceedings*, vol. 1868, no. 1, 2017, pp. 1–8.

[9] M. Dingari *et al.*, "Time series analysis for long memory process of air traffic using arfima," *International Journal of Scientific & Technology Research*, vol. 8, no. 10, pp. 395–400, 2019.

[10] D. Safitri *et al.*, "Gold price modeling in indonesia using arfima method," *Journal of Physics, Conference Series*, vol. 1217, pp. 1–11, 2019.

[11] L. Bisaglia and M. Grigoletto, "A new time-varying model for forecasting long-memory series," *Statistical Methods & Applications*, 2020.

[12] P. Murialdo *et al.*, "Long-range dependence in financial markets: a moving average cluster entropy approach," *Entropy*, vol. 22, no. 6, pp. 1–19, 2020.

[13] J. E. Vera-Valdés, "On long memory origins and forecast horizons," *Journal of Forecasting*, vol. 39, no. 5, pp. 811–826, 2020.

[14] A. Tokhmpash *et al.*, "Fractional order modeling of brain signals," *Advances in Intelligent Systems and Computing*, vol. 1201, pp. 9–15, 2021.

[15] R. S. Tsay, *Analysis of Financial Time Series*, 3rd ed. John Wiley & Sons, Inc., 2010.

[16] S. Benkachcha *et al.*, "Seasonal time series forecasting models based on artificial neural network," *International Journal of Computer Applications*, vol. 116, no. 20, pp. 9–14, 2015.

[17] Sukono *et al.*, "Indonesian financial data modeling and forecasting by using econometrics time series and neural network," *Global Journal of Pure and Applied Mathematics*, vol. 12, no. 4, pp. 3745–3757, 2016.

[18] A. Camara *et al.*, "Energy consumption forecasting using seasonal arima with artificial neural networks models," *International Journal of Business and Management*, vol. 11, no. 5, pp. 231–243, 2016.

[19] C. Bal and S. Demir, "Forecasting try/usd exchange rate with various artificial neural network models," *TEM Journal*, vol. 6, no. 1, pp. 11–16, 2017.

[20] M. A. H. Ashour *et al.*, "Effectiveness of artificial neural networks in solving financial time series," *International Journal of Engineering & Technology*, vol. 7, no. 4.11, pp. 99–105, 2018.

[21] M. Dingari *et al.*, "Air traffic forecasting using artificial neural networks," *International Journal of Scientific & Technology Research*, vol. 8, no. 10, pp. 556–559, 2019.

[22] Isha *et al.*, "Effects of activation function and input function of ann for solar power forecasting," *Advances in Data and Information Sciences*, vol. 94, pp. 329–342, 2020.

[23] C. Xie, Z. Mao, and G. J. Wang, "Forecasting rmb exchange rate based on a nonlinear combination model of arfima, svm, and bpnn," *Mathematical Problems in Engineering*, vol. 2015, pp. 1–10, 2015.

[24] S. Aras *et al.*, "Comparative study on retail sales forecasting between single and combination methods," *Journal of Business Economics and Management*, vol. 18, no. 5, pp. 803–832, 2017.

[25] W. Pannakkong *et al.*, "A novel hybrid autoregressive integrated moving average and artificial neural network model for cassava export forecasting," *International Journal of Computational Intelligence Systems*, vol. 12, no. 2, pp. 1047–1061, 2019.

[26] A. H. Bukhari *et al.*, "Fractional neuro-sequential arfima-lstm for financial market forecasting," *IEEE Access*, vol. 8, pp. 71 326–71 338, 2020.

[27] L. Dhliwayo *et al.*, "Autoregressive fractionally integrated moving average-generalized autoregressive conditional heteroskedasticity model with level shift intervention," *Open Journal of Statistics*, vol. 10, no. 2, pp. 341–362, 2020.

[28] M. Hamdi and C. Aloui, "Forecasting crude oil price using artificial neural networks: A literature survey," *Economics Bulletin*, vol. 35, no. 2, pp. 1339–1359, 2015.

[29] K. S. Mohammed *et al.*, "Can oil prices forecast the algerian exchange rate? ," *International Journal of Engineering Research and Management*, vol. 2, no. 4, pp. 78–82, 2015.

[30] Z. Ftiti *et al.*, "Relationship between crude oil prices and economic growth in selected opec countries," *The Journal of Applied Business Research*, vol. 32, no. 1, pp. 11–22, 2016.

[31] J. Xiao *et al.*, "A hybrid transfer learning model for crude oil price forecasting," *Statistics and Its Interface*, vol. 10, no. 1, pp. 119–130, 2017.

[32] D. Ambach and O. Ambach, "Forecasting the oil price with a periodic regression arfima-garch process." IEEE Second International Conference on Data Stream Mining & Processing (DSMP), 2018, pp. 212–217.

[33] N. Gupta and S. Nigam, "Crude oil price prediction using artificial neural network," *Procedia Computer Science*, vol. 170, pp. 642–647, 2020.

[34] G. B. Adesi *et al.*, "Wti crude oil option implied var and cvar: An empirical application," *Journal of Forecasting*, vol. 38, no. 6, pp. 552–563, 2019.

[35] L. Han, Q. Lv, and L. Yin, "The effect of oil returns on the stock markets network," *Physica A: Statistical Mechanics and its Applications*, vol. 533, pp. 1–17, 2019.

[36] A. Nyangarika *et al.*, "Oil price factors: Forecasting on the base of modified auto-regressive integrated moving average model," *Interna-*

tional Journal of Energy Economics and Policy*, vol. 9, no. 1, pp. 149–159, 2019.

[37] M. Youssef and K. Mokni, “Do crude oil prices drive the relationship between stock markets of oil-importing and oil-exporting countries?” *Economies*, vol. 7, no. 3, pp. 1–22, 2019.

[38] H. Abdollahi and S. B. Ebrahimi, “A new hybrid model for forecasting brent crude oil price,” *Energy*, vol. 200, pp. 1–13, 2020.

[39] M. A. Ijomah and P. Enewari, “Modelling volatility transmission between crude oil price and the nigerian exchange rate using multivariate garch models,” *International Journal of Innovative Mathematics, Statistics & Energy Policies*, vol. 8, no. 2, pp. 1–12, 2020.

[40] Y. Zhang and S. Hamori, “Forecasting crude oil market crashes using machine learning technologies,” *Energies*, vol. 13, no. 10, pp. 1–14, 2020.

[41] N. M. Noh, A. Bahar, and Z. M. Zainuddin, “Forecasting model for crude oil price with structural break,” *Malaysian Journal of Fundamental and Applied Sciences*, pp. 421–424, 2017.

[42] G. Bhardwaj and N. R. Swanson, “An empirical investigation of the usefulness of arfima models for predicting macroeconomic and financial time series,” *Journal of Econometrics*, vol. 131, no. 1-2, pp. 539–578, 2006.

[43] M. Boutahar *et al.*, “Estimation methods of the long memory parameter: Monte carlo analysis and application,” *Journal of Applied Statistics*, vol. 34, no. 3, pp. 261–301, 2007.

[44] J. Beran, *Statistics for Long-Memory Processes*, 1st ed. Chapman and Hall, 1994.

[45] S. Telbany and M. Sous, “Using arfima models in forecasting indicator of the food and agriculture organization,” *IUGJEBS*, vol. 24, no. 1, pp. 168–187, 2016.

[46] V. A. Reisen, “Estimation of the fractional difference parameter in the arima(p,d,q) model using the smoothed periodogram,” *Journal of Time Series Analysis*, vol. 15, no. 3, pp. 335–350, 1994.

[47] J. R. M. Hosking, “Fractional differencing,” *Biometrika*, vol. 86, no. 1, pp. 165–176, 1981.

[48] T. O. Olatayo and A. F. Adedotun, “On the test and estimation of fractional parameter in arfima model: Bootstrap approach,” *Applied Mathematical Sciences*, vol. 8, no. 96, pp. 4783–4792, 2014.

[49] A. Ohanissian *et al.*, “True or spurious long memory? a new test,” *Journal of Business & Economic Statistics*, vol. 26, no. 2, pp. 161–175, 2008.

[50] F. X. Diebold and A. Inoue, “Long memory and regime switching,” *Journal of Econometrics*, vol. 105, no. 1, pp. 131–159, 2001.

[51] C. W. J. Granger and N. Hyung, “Occasional structural breaks and long memory with an application to the s&p 500 absolute stock returns,” *Journal of Empirical Finance*, vol. 11, no. 3, pp. 399–421, 2004.

[52] G. C. Chow, “Tests of equality between sets of coefficients in two linear regressions,” *Econometrica: Journal of the Econometric Society*, vol. 28, no. 3, pp. 591–605, 1960.

[53] G. E. P. Box *et al.*, *Time series analysis: forecasting and control*, 4th ed. Wiley & Sons, Inc., 2008.

[54] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*, 2nd ed. Cambridge University Press, 2003.

[55] G. P. Zhang, “Time series forecasting using a hybrid arima and neural network model,” *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[56] G. Ciaburro and B. Venkateswaran, *Neural Networks with R*. Packt Publishing Ltd., 2017.

[57] K. Naveena *et al.*, “Hybrid arima-ann modelling for forecasting the price of robusta coffee in india,” *Int. J. Curr. Microbiol. App. Sci.*, vol. 6, no. 7, pp. 1721–1726, 2017.

[58] M. Khairalla *et al.*, “Hybrid forecasting scheme for financial time-series data using neural network and statistical methods,” *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 9, pp. 319–327, 2017.

[59] C. H. Aladag *et al.*, “Improvement in forecasting accuracy using the hybrid model of arfima and feed forward neural network,” *American Journal of Intelligent Systems*, vol. 2, no. 2, pp. 12–17, 2012.

[60] D. A. Dickey and W. A. Fuller, “Distribution of the estimators for autoregressive time series with a unit root,” *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 427–431, 1979.

[61] P. C. B. Phillips and P. Perron, “Testing for a unit root in time series regression,” *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.

[62] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, 2nd ed. Wiley & Sons, Inc., 2015.

[63] J. D. Cryer and K. S. Chan, *Time Series Analysis: with Application in R*, 2nd ed. Springer, 2008.

[64] K. B. Tadesse and M. O. Dinka, “Application of sarima model to forecasting monthly flows in waterval river, south africa,” *Journal of Water and Land Development*, no. 35, pp. 229–236, 2017.

[65] Z. Berradi *et al.*, “Effect of architecture in recurrent neural network applied on the prediction of stock price,” *IAENG International Journal of Computer Science*, vol. 47, no. 3, pp. 436–441, 2020.

[66] W. Wongsinlatam and S. Buchitchon, “Criminal cases forecasting model using a new intelligent hybrid artificial neural network with cuckoo search algorithm,” *IAENG International Journal of Computer Science*, vol. 47, no. 3, pp. 481–490, 2020.

[67] R. S. Al-Gounmeein and M. T. Ismail, “Forecasting the exchange rate of the jordanian dinar versus the us dollar using a box-jenkins seasonal arima model,” *International Journal of Mathematics and Computer Science*, vol. 15, no. 1, pp. 27–40, 2020.

[68] Z. Pu and X. Wang, “Finite-time convergent complex-valued neural networks for computing square root of complex matrices,” *IAENG International Journal of Applied Mathematics*, vol. 50, no. 3, pp. 512–518, 2020.