




Research Article

Forecasting Stock Market Volatility Using Hybrid of Adaptive Network of Fuzzy Inference System and Wavelet Functions

Abdullah H. Alenezy,^{1,2} Mohd Tahir Ismail ,² S. Al Wadi,³ Muhammad Tahir ,⁴ Nawaf N. Hamadne ,⁵ Jamil J. Jaber,³ and Waqar A. Khan ⁶

¹Department of Mathematics, College of Science, University of Ha'il, Hail, Saudi Arabia

²School of Mathematical Science, Universiti Sains Malaysia, Penang, Malaysia

³Department of Risk Management and Insurance, Faculty of Business, The University of Jordan, Amman, Jordan

⁴College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia

⁵Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Riyadh 11673, Saudi Arabia

⁶Department of Mechanical Engineering, College of Engineering, Prince Mohammad Bin Fahd University, Al Khobar 31952, Saudi Arabia

Correspondence should be addressed to Muhammad Tahir; m.tahir@seu.edu.sa

Received 5 March 2021; Accepted 16 August 2021; Published 27 August 2021

Academic Editor: Basil K. Papadopoulos

Copyright © 2021 Abdullah H. Alenezy et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study aims to model and enhance the forecasting accuracy of Saudi Arabia stock exchange (Tadawul) data patterns using the daily stock price indices data with 2026 observations from October 2011 to December 2019. This study employs a nonlinear spectral model of maximum overlapping discrete wavelet transform (MODWT) with five mathematical functions, namely, Haar, Daubechies (Db), Least Square (LA-8), Best localization (BL14), and Coiflet (C6) in conjunction with adaptive network-based fuzzy inference system (ANFIS). We have selected oil price (Loil) and repo rate (Repo) as input values according to correlation, the Engle and Granger Causality test, and multiple regressions. The input variables in this study have been collected from Saudi Authority for Statistics and Saudi Central Bank. The output variable is obtained from Tadawul. The performance of the proposed model (MODWT-LA8-ANFIS) is evaluated in terms of mean error (ME), root mean square error (RMSE), and mean absolute percentage error (MAPE). Also, we have compared the MODWT-LA8-ANFIS model with traditional models, which are autoregressive integrated moving average (ARIMA) model and ANFIS model. The obtained results show that the performance of MODWT-LA8-ANFIS is better than that of the traditional models. Therefore, the proposed forecasting model is capable of decomposing in the stock markets.

1. Introduction

The stock price movements are evaluated by volatility in the stock exchange market. Volatility explains the action of stock exchange market. It reflects that if the price of a stock fluctuates a lot over time (high volatility) or if a stock price fluctuates slowly over time (low volatility). Volatility is measured as the standard deviation of stock price [1]. Stock market volatility is a metric that measures riskiness of stocks and is relevant to both market policy makers and practitioners, mainly in emerging

markets [2]. Indeed, an effective quantitative approach is needed to model the volatility of stock market in order to protect against unexpected price changes. Previous studies have shown that volatility in the stock market is time-varying; thus, the movements in volatility are nonrandom. Therefore, a number of time-varying volatility models have been developed by financial econometricians and other practitioners [3–8].

Artificial neural networks (ANNs) are widely used to support applications across a various business and scientific disciplines [9–12]. There are many articles that

employed ANN for the prediction of stock market. For example, the authors [13] also predicted price fluctuations using the Haar wavelet and Takagi–Sugeno–Kang (TSK) fuzzy rule-based system. The TSK fuzzy rule-based method is used to forecast stock prices using a number of technical indices. The model has successfully predicted stock price fluctuations in Taiwan Stock Exchange market with an accuracy of up to 99.1% according to simulation results [14]. The authors proposed a forecasting fusion model that combined wavelet as a data preparation tool, fuzzy logic, and neural network. The proposed model was trained on dataset that covers duration from 2005 to 2010. The results indicate that this hybrid model achieves better forecasting accuracy than either of the models used separately. Similarly, the authors in [15] presented fuzzy wavelet neural network (FWNN) for the prediction of stock prices. The daily stock prices for the last three years have been used as dataset of 1000 samples where 950 samples were used for training and the remaining 50 samples were used for testing. The simulation results demonstrated that the proposed FWNN system with differential evaluation (DE) training has achieved better performance compared with other models.

The ANFIS uses both fuzzy logic and ANN [16, 17]. The forward and the backward processes comprise the ANFIS learning algorithm. The forward process goes via the five layers as given in [17–19]. The ANFIS models are used successfully in many fields such as engineering, computer science, and chemistry. Moreover, many models are used successfully for predictions when combining with MODWT. Note that, MODWT is a mathematical model based on five functions, namely, Haar, Db, LA-8, BL14, and C6 [20]. The literature reveals that a number of research works have been published that used ANFIS. The authors in [21] presented an ANFIS approach for long-term prediction of electricity consumption. They introduced ANFIS and AR models to forecast the prediction of long-term natural electrical demand for some European countries. The authors in [22] applied the ANFIS model to the Yamadu Hydrological Station annual runoff forecast in China. The results show that the ANFIS model has better forecasting efficiency than the ANN model on the basis of relative percentage errors. The authors in [23] used ANFIS to forecast the future sales of an online shop. The sample size was 80 day's sale of 200 products. The results show that the ANFIS model can partially improve the accuracy in time series prediction. The literature also reveals that few articles have focused on the wavelet with fuzzy logic. In [24], the weekly data have been taken from January 2012 to November 2014. They used the fuzzy wavelet model to forecast the exchange rate of IDR of USD. The fuzzy wavelet model is the combination of the fuzzy Mamdani model and Discrete Wavelet Transforms (DWTs). The authors in [25] used a fuzzy wavelet neural control scheme for the micro-electro-mechanical system (MEMS). A novel time series forecasting model based on fuzzy cognitive maps and empirical wavelet transformation is proposed in [26]. The performance of wavelet neural network (WNN) and ANFIS models was compared using small datasets by [27].

The stock market volatility has been affected by macroeconomic variables such as inflation rate, unemployment, interest rate, gross domestic product, and oil prices. The Repo plays an important role in macroeconomy. The Repo is the monetary policy interest rate as it is used by the central bank to lend money to the banks for short term. The impact of Repo on stock market is studied by [28]. Furthermore, the oil price refers to the close price of a barrel of crude oil. The effect of oil prices on stock markets is studied by [29]. Indeed, the financial covariates (Repo and Loil) are investigated as input variables in our study.

According to the literature review, no one has concentrated on MODWT for modeling and enhancing the prediction accuracy in Tadawul over the last decade. In terms of the study's objectives, a variety of comparative studies have been conducted on the usage of various MODWT functions individually as well as in combination with other MODWT models over the last ten years in various methodologies. However, potential room exists for further investigation about comparative applications of all MODWT functions, which include Haar, Db, LA-8, BL14, and C6 in combination with fitting ANFIS model in terms of single particular context or financial market. In this connection, this study undertakes this work in relation to Tadawul since some researchers in the literature have only used one feature of MODWT. The current research aims at using MODWT functions to analyze fluctuations in Tadawul. The index refers to the average performance of firms listed in Saudi exchange market. Additionally, the causes of stock market volatility and modeling of variance behavior are also specified to represent the accuracy of expectations and the percentage of possible risks. Furthermore, by combining MODWT functions with the ANFIS model and using the statistical criterion such as MSE, RMSE, MAE, and MAPE, the forecasting accuracy is enhanced and the new forecasting model is proposed.

This study is organized as follows. Materials and methods are explained in Section 2. The research design and methodology are discussed in Section 3. The empirical results are analyzed in Section 4. The conclusions are drawn in Section 5.

2. Materials and Methods

2.1. Dataset. The dataset for closing prices is obtained from Saudi Arabia stock market (Tadawul), Saudi Authority for Statistics, and Saudi Central Bank [30, 31]. The day-to-day closing prices were gathered from August 2011 to December 2019. The size of observations is 2026 [20, 32]. Table 1 shows the descriptive statistic of dataset.

LSCS refers to the logarithm of standard deviation for closing stock prices, which can be expressed as $\sigma = \sqrt{(\sum_{i=1}^n (x_i - \bar{x})^2)/(n - 1)}$, where x is the closing stock price.

LSCS has a mean of 6.75 and a standard deviation of 0.6923. LSCS has a minimum value of 3.83 and a maximum value of 7.22. It should be noted that Repo has a mean of 0.70 and a standard deviation of 0.28 whereas Repo has a

TABLE 1: Descriptive statistic of Saudi stock market dataset.

	N	Mean	Std. deviation	Minimum	Maximum
LSCS	2026	6.7489	0.6923	3.8294	7.2194
Repo	2026	0.6963	0.2796	0.1271	4.5490
Loil	2026	4.2992	0.3535	3.3279	4.8380

minimum value of 0.13 and a maximum value of 4.55. Loil has a mean of 4.30 and a standard deviation of 0.35 whereas Loil has a minimum value of 3.33 and a maximum value of 4.84.

2.2. Wavelet Transform Formula. Wavelet transform (WT) is a mathematical formulation for transforming the original time series data into a time-scale domain. WT is an appealing option for nonstationary data, especially stock exchange market data because of its inherent nature. WT can be categorized as continuous wavelet transforms (CWTs), DWT, and MODWT. Note that, these transforms demonstrate similar behavior in general. The key difference among DWT and MODWT is that DWT can be applied to a certain number of observations (the size of samples should be 2 raised to the power J) whereas MODWT can handle data of any size. In this study, we focus on MODWT due to its flexible behavior [32, 33].

Theoretically, WT is an extension of Fourier transform (FT) [34] that is the output on sine and cosine functions. WT should meet the following criterion [20]:

$$C\varphi = \int_0^{\infty} \varphi(f) V \frac{\varphi(t)}{f} df < \infty, \quad (1)$$

where $\varphi(f)$ is a function of frequency f and known as the FT. WT is used in a variety of applications including signal

processing and image analysis. It was developed to solve the FT issue, essentially when treating with time, space, and frequency. As shown in equations (2a) and (2b), the father wavelet represents the low-fluctuate (smooth data) components whereas the mother wavelet represents the high-fluctuate (detailed data) components, with $j = 1, 2, 3, \dots, J$ in the J -level wavelet decomposition:

$$\phi_{j,k} = 2^{(-j/2)} \phi\left(t - \frac{2^j k}{2^j}\right), \quad (2a)$$

$$\varphi_{j,k} = 2^{(-j/2)} \varphi\left(t - \frac{2^j k}{2^j}\right), \quad (2b)$$

where J defines the maximum scale supported by the number of data points and the two forms of wavelets, father and mother wavelets, and meets the following criteria as expressed by the following equations:

$$\int \phi(t) dt = 1, \quad (3a)$$

$$\int \varphi(t) dt = 0. \quad (3b)$$

The general mathematical model is presented in the following equation:

$$\begin{aligned} S_{j,k} &= \int \phi_{j,k} f(t) dt, \\ d_{j,k} &= \int \varphi_{j,k} f(t) dt. \end{aligned} \quad (4)$$

In more detail, as expressed by the following equation,

$$F(t) = \sum S_{j,k} \phi_{j,k}(t) + \sum d_{j,k} \varphi_{j,k}(t) + \sum d_{j-1,k} \varphi_{j-1,k}(t) + \dots + \sum d_{1,k} \varphi_{1,k}(t), \quad (5)$$

$$S_j(t) = \sum S_{j,k} \phi_{j,k}(t), \quad (6a)$$

$$D_j(t) = \sum d_{j,k} \varphi_{j,k}(t). \quad (6b)$$

In equations (6a) and (6b) and, where $S_j(t)$ and $D_j(t)$ are the smooth and detailed coefficients, respectively, the WT is used to measure the approximation coefficient. The detailed coefficients are used to measure the significant fluctuations of the original data, while the smooth coefficients contain the most significant features of the original data. In general, Haar, Db, LA-8, BL14, and C6 are common transform functions in WT [32]. The following are some of the key characteristics of these functions. Except for the Haar model, the WT functions are arbitrarily regular. WT functions, with the exception of the Haar model, do not have explicit expression. WT functions are applied to real numbers. WT functions support an arbitrary number of zero moments,

orthogonal, compact, bio-orthogonal analysis, orthogonal analysis, continuous/discrete transformation, fast algorithm, and exact reconstruction.

2.3. ARIMA Model. The autoregressive moving average (ARMA) model is considered one of the most important mathematical models, which is widely used in time series analysis. A moving average (MA) and an autoregressive (AR) models are combined in the ARMA model. A time series e_t denotes a white noise (WN) process, and Y_t denotes Gaussian process iff for all t , e_t is iid $N(0, \sigma^2)$. A time series Y_t , given by equation (7), follows the ARMA (p, q) model of [20, 35, 36]:

$$Y_t = \mu + \phi_1 Y_{t-1} + \phi_2 Y_{t-2}, \dots, + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2}, \dots, - \theta_q e_{t-q}, \quad (7)$$

where q and p are used as nonnegative integers, p is defined as the order of autoregressive part (AR), q is used as an order of the first (MA) part, and e_t is defined as the white noise (WN) process. (ARIMA(p, d, q)) is an extension of the ordinary ARMA model. (ARIMA(p, d, q)) is given by [20].

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B)e_t, \quad (8)$$

where p , d , and q are the orders of auto-regression, integration (differencing), and moving average, respectively. When $d = 0$, the ARIMA model is reduced to the ordinary ARMA model.

2.4. ANSIF Model. ANFIS utilizes both fuzzy logic and ANN [37] where the ANN learning algorithm is used for training. Its operations consist of forward and backward steps that collectively comprise the ANFIS learning algorithm. The forward step consists of five layers. The fuzzy inference system under consideration is supposed to have two inputs (x, y) and one output (z) to simplify the explanations. Note that the input x represents the variable oil price, the input y represents the variable Repo, and the output z represents the logarithm of standard deviation for closing stock prices (LSCS). A standard rule base of fuzzy if-then rules for a first order of the Sugeno fuzzy model can be expressed as follows: if x is A_1 and y is B_1 , then $f_1 = p_1 x + q_1 y + r_1$, where p, r , and q are denoted as linear output parameters. Figure 1 depicts the ANFIS architecture, which has two inputs and one output.

Layer 1. Every node i in this layer is a square node with a node function as given by the following equations:

$$O_{1,i} = \mu A_i(x), \quad \text{for } i = 1, 2, 3, \quad (9a)$$

$$O_{1,i} = \mu B_{i-3}(y), \quad \text{for } i = 4, 5, 6, \quad (9b)$$

where x and y are denoted as inputs to node i and A_i and B_i are defined as linguistic labels for the inputs. In other words, $O_{1,i}$ is the membership function of A_i and B_i . Typically, $\mu A_i(x)$ and $\mu B_i(y)$ are selected to be bell-shaped with maximum value of 1 and minimum value of 0, such as $\mu A_i(x), \mu B_{i-3}(y) = \exp\left(\frac{-(x-c_i)}{a_i}\right)^2$, where the set of parameter is a_i and c_i . These parameters are referred to as premise parameters in this layer. Indeed, using the Gaussian function as the shape of the membership function, the fuzzification process transforms crisp values into linguistic values.

Layer 2. Each node in this layer is a circle node labeled \prod that multiplies the incoming signals and sends out the product as expressed by the following equation:

$$O_{2,i} = w_i = \mu A_i(x) \cdot \mu B_{i-3}(y), \quad i = 1, 2, 3, \dots, 9. \quad (10)$$

Each node output describes the firing strength of a rule. In this layer, the t -norm operator (the AND operator) is used by the inference stage.

Layer 3. Each node in this layer is a circle node called N . The i_{th} node measures the ratio of the i_{th} rule firing strength to the sum of all rules firing strengths as given by the following equation:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2 + \dots + w_9}, \quad i = 1, 2, 3, \dots, 9. \quad (11)$$

In short, the ratio of the strengths of the rules is calculated in this layer.

Layer 4. Each node i in this layer is a square node with a node function as expressed by the following equation:

$$O_{4,i} = \bar{w}_i \cdot f_i = w_i \cdot (p_i x + q_i y + r_i), \quad i = 1, 2, 3, \dots, 9, \quad (12)$$

where w_i is the output of layer 3 and p_i, q_i , and r_i are the parameter set. Parameters in this layer are referred to as consequent parameters. In short, the parameters for the consequent parts are measured in this layer.

Layer 5. A circle node called \sum is the single node in this layer that calculates the overall output as the summation of all incoming signals as given by the following equation:

$$O_{5,i} = \text{overall output} = \sum_i \bar{w}_i \cdot f_i = \frac{\sum_i w_i \cdot f_i}{\sum_i w_i}. \quad (13)$$

The backward step is a database estimation method consisting of the membership function parameters in the antecedent part and the linear equation coefficients in the consequent part. Since the Gaussian function is used as the membership function in this process, two parameters, namely, mean and variance of this function are optimized. The least squares method is used to perform the parameter learning in this step.

2.5. Performance Measures. We use a number of accuracy criteria including the mean absolute percentage error (MAPE), the mean absolute error (MAE), the mean error (ME), and the root mean squared error (RMSE) [38, 39] as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right|, \quad (14)$$

$$\text{MPE} = \frac{1}{n} \sum_{t=1}^n \frac{X_t - F_t}{X_t}, \quad (15)$$

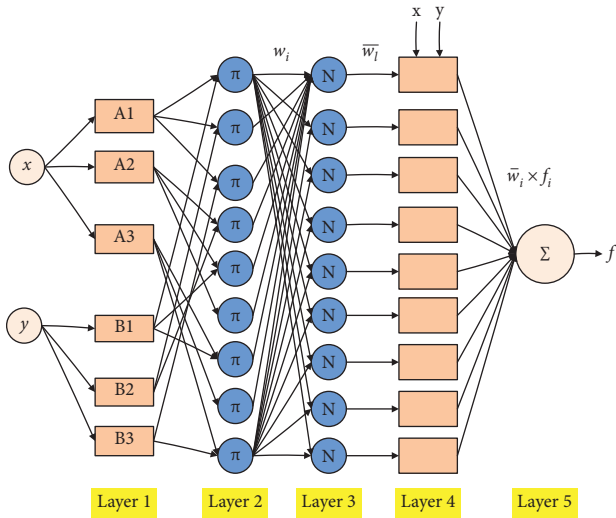


FIGURE 1: ANFIS architecture of two inputs and one output with four rules.

$$MAE = \frac{1}{n} \sum_{t=1}^n |X_t - F_t|, \quad (16)$$

$$ME = \frac{1}{n} \sum_{t=1}^n (X_t - F_t), \quad (17)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{t=1}^n (\text{actual value} - \text{predicted value})^2}{n}}. \quad (18)$$

The MAPE criterion, also referred to as mean absolute percentage deviation (MAPD), is a criterion of prediction accuracy for the forecasting method in statistics. It always expresses accuracy as a percentage and is determined by equation (14), where X_t represents the actual value, F_t represents the forecasted value, and n represents the sample size. In this equation, the absolute value is added for each forecasted point in time and divided by the number of fitted points. In addition, the MPE is defined by equation (15), the MAE is given by equation (16), and ME is expressed by equation (17). The root mean square deviation (RMSD), also known as the RMSE, is a widely used criterion of the estimator differences. It estimates the mean error produced by the model in predicting the outcome for an observation. It is determined by equation (18), where N denotes the number of observations.

3. Research Design and Methodology

The aim of this study is to propose a new model to forecast the closed price data from the Tadawul stock market, which covers the period from 2011 to 2019. The proposed model coupled the ANFIS model with MODWT-LA8. In this connection, we have employed five MODWT functions, namely, Haar, Db, LA-8, BL14, and C6. Note that, the statistical tests are used to evaluate the accuracy of the

models. Moreover, the original data were transformed into a time-scale domain using MODWT. The different phases of the MODWT forecasting mechanism are depicted in Figure 2. It should be noted that the wavelet process is used repeatedly while the data pattern was fluctuating. The objective of preprocessing is to reduce the statistical error criteria such as RMSE in the data before and after transformation. In this way, the noise in the original data can be eliminated. Essentially, the adaptive noise in the training pattern can help to minimize the risk of overfitting in the training process. Thus, we used MODWT twice for the preprocessing of the training data in this study. Further, MODWT converts the data into two sets, namely, detail series and approximation series. Since the financial data fluctuates significantly, we have employed these two series due to the fact that they show good behavior on such data. This helps in anticipating the transformed data more precisely. The MODWT's filtering effect is responsible for these two series' positive behavior.

In order to propose our new model, we designed the following methodology. Firstly, we have collected the closed price data from Tadawal. Secondly, the closed price data have been treated using logarithm standard deviation to find LSCS. Thirdly, The LSCS data have been decomposed using MODWT function that divides the LSCS data into two partitions, namely, details coefficient (high fluctuated data) and approximation coefficient (low fluctuated data). We have employed five MODWT functions, namely, Haar, Db, LA-8, C6, and BL14. The approximation coefficient for each function consists of the main features of the data and is used as output in the forecasting model. Fourthly, the approximation coefficient (LSCS) for each function is used with input variables (Repo and Loil) inside ANFIS to our proposed model MODWT-ANFIS. Finally, the best MODWT-ANFIS model has been compared with other functions of MODWT-ANFIS and also with traditional models (ARIMA and ANFIS models).

In order to make a fair comparison, we first applied 80% of the original data and the converted data to the proposed model and then selected the best performing model, which is further used with other suggested models for the remaining 20% data. This confirms the outstanding performance of our proposed model.

3.1. Endogeneity Issues. In this section, we discuss to select suitable variables by removing multicollinearity, causality test, and multiple regression analysis.

3.1.1. Correlation. In this section, we carefully picked independent variables from various other variables, which are removed depending on the statistical test. Firstly, we removed variables as a result of multicollinearity among independent variables as shown in Table 2. The absence of perfect multicollinearity, an exact (nonstochastic) linear relationship between two or more independent variables, is generally referred to as no multicollinearity. We removed some variables from input variables based on their strong

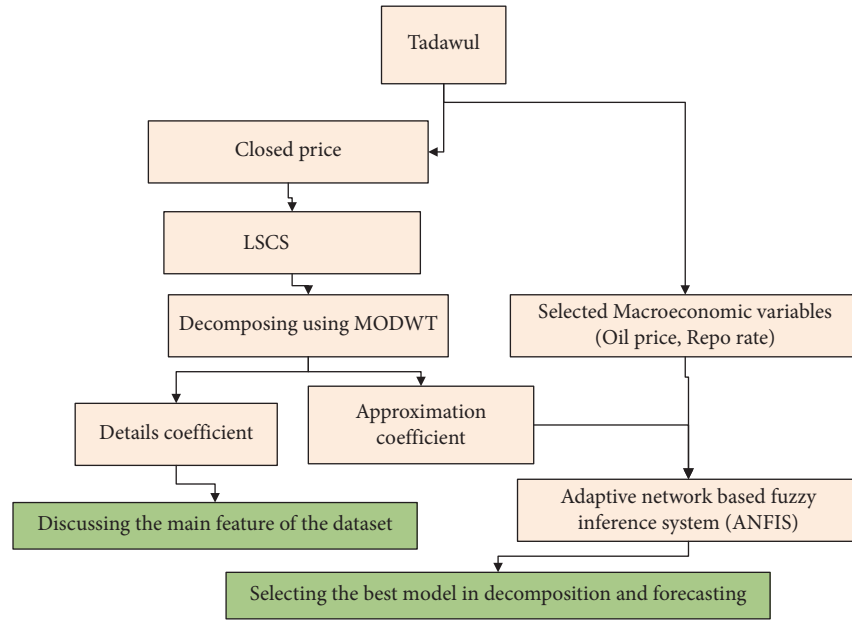


FIGURE 2: The flowchart of the MODWT with ANFIS.

correlation with other input variables. The correlation between the input and output variables is shown in Table 2.

3.1.2. Engle and Granger Causality Test. Engle and Granger’s test uses co-integration to represent causal relationships. Based on static regression, it creates residuals (errors). An augmented Dickey–Fuller test or another similar test uses residuals to detect unit roots. The residuals will be almost stationary if the time series is co-integrated [30, 31, 39, 40]. $d(Y_t) = c + \sum_{i=1}^n \beta_n^* d(Y_{t-i}) + \sum_{i=0}^n \alpha_n^* d(X_{t-i} + \phi ECT_{t-1} + \varepsilon_t)$, where Y is the output variable, X is the input variable set, ECT is the word for error correction whereas β , α , and ϕ are the parameters. If ϕ is negative or greater than 1.96, the null hypothesis for the Engle Granger test (H_0 : there is no co-integration) is rejected. In more details, the rule for hypothesis testing says that if the p value is less than or equal to the critical value, then we reject the null hypothesis. Table 3 explains the Engle and Granger test for output and input variables and shows that all values of “ p value” are less than 0.05. Accordingly, the null hypothesis rejected the input variables, and we conclude that there is sufficient evidence to support the claim: there is co-integration with output variables at a significant level of 5%. This result almost implies that the output variable is influenced by the input variables.

3.1.3. The Results of Multiple Regressions. An extension of simple linear regression is multiple regression that is used when we want the value of a variable to be predicted based on the value of two or more other variables. In this study, the variable that needs to be predicted is the dependent variable (LSCS) whereas Repo and Loil are the independent variables that are used to predict LSCS. The multiple regression analysis is shown in Table 4. At 5% significant level, the Repo rate and Loil are significant. In addition, R -square and

TABLE 2: The correlations between the input and the output variables.

	LSCS	Repo	Loil
LSCS	1	-0.149	-0.673
Repo		1	0.327
Loil			1

TABLE 3: Engle and granger causality test for output and input variables.

Variables	Estimate	Std. error	t-stat	p value	Causality detection
Repo	-0.0060	0.0014	-4.2010	2.77E-05	Causality
Loil	-0.0061	0.0013	-4.7100	2.65E-06	Causality
All	-0.0070	0.0016	-4.2440	2.30E-05	Causality

TABLE 4: The multiple regression.

	Estimate	Std. error	t value	p value
Intercept	12.427	0.147	84.410	$<2E - 16$
Repo	0.264	0.047	5.683	$1.54E - 08$
Loil	-1.368	0.036	-38.424	$<2E - 16$
R -square	0.4552	Adjusted R -square	0.4546	
F -statistic	760.3	p value	$<2.2E - 16$	

adjusted R -square are approximately 46%, which implies that the independent variables can explain about 46% of the output variable. F -statistic at 1 percent signifies that the linear regression model is better suited to the results.

There is negative relationship between oil prices and volatility risk ($\beta = -1.368$), which measures the standard deviation of closing stock prices. This indicates that the

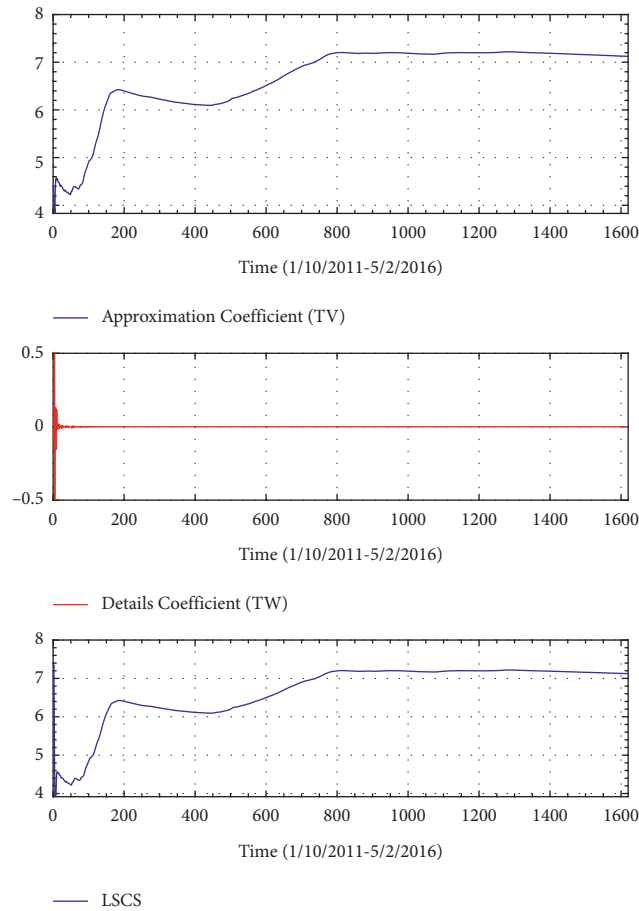


FIGURE 3: Decomposing the data using MODWT based on LA-8 function.

increase in oil prices will reduce the volatility risk in stock market. On the other hand, the Repo rate has positive relationship with volatility risk ($\beta = 0.264$). This implies that the increase in Repo will increase the volatility risk.

4. Results and Discussion

The current study investigates the closing price data from Tadawul. It has been selected for a variety of reasons. In terms of financial market volatility, emerging markets have a deep characteristic. Due to lack of information, stochastic trading, unprofessional financial analysis, and the Saudi sector have experienced considerable volatility. Furthermore, the investors from outside the Gulf Cooperation Council (GCC) are not permitted to invest in Saudi stocks. Tadawul is the largest exchange market in the Middle East. As seen in Figure 3, the volatility data are decomposed using MODWT with the LA-8 function.

Firstly, the closed price data have been treated using logarithm standard deviation to find LSCS. Secondly, The LSCS data have been decomposed using MODWT via R-statistic software. The MODWT mechanism divides the LSCS data into two parts, namely, details coefficient (high fluctuated data) and approximation coefficient (low fluctuated data). The approximation coefficient for each function consists of the main features of data and is used as

output in the forecasting model. We have employed five MODWT functions, namely, Haar, Daubechies (Db), least Asymmetric (LA-8), Coiflet (C6), and the best-localized function (BL14). As a result the best function with MODWT is LA-8 (see Table 5). The best result of MODWT with the LA-8 is described in Figure 3.

MODWT-based decomposition is an effective approach for revealing variations, magnitudes, and phases of the data. The WT will determine the levels of decomposition using the formula, according to the WT mechanism $X = TV1 + TW1$ where the original signal is referred to as X . The next component shows the plot of the transformed data approximation coefficients at one level of approximation (TV1). The TW1 reflects the level of detail, whereby TW1 is the plot of the first level of the coefficients of detail, so the fluctuation can be explained by this level. Note that, 80% of the data are represented by 1620 samples out of 2026 total samples that are given along x-axis of 3.

Tadawul has numerous fluctuations from 2011 to 2019. It was recorded that the general index of stock exchange market dropped to 6417.7 points in 2011 whereas it grew up to 8535 points in 2013. Market management modified the trading process from SAXESS to X-Stream INET, and an interactive multiuser system (IFSAH) was developed to improve the market's efficiency and effectiveness [41]. The fluctuation of stock prices is one of the issues that confronted

TABLE 5: The WT function of output variable for 80% dataset.

WT function		ME	MAE	MAPE
Haar	ARIMA(0, 1, 1) with drift	0.0003965	0.0041307	0.0889295
Db4	ARIMA(0, 1, 0)	0.0005757	0.0047082	0.0954798
LA-8	ARIMA(1, 1, 0) with drift	0.0000053	0.0032142	0.0644968
BL14	ARIMA(1, 1, 0) with drift	0.0000096	0.0032940	0.0655779
C6	ARIMA(1, 1, 0) with drift	0.0000330	0.0033147	0.0660498

TABLE 6: The WT function of output variable for 20% dataset.

Models		ME	RMSE	MAE	MPE
ANFIS	—	3.1721832	3.2151644	3.1782445	82.8745374
ANFIS + Haar	ARIMA(0, 1, 1) with drift	3.1615475	3.1830709	3.1650890	81.4528974
ANFIS + Db4	ARIMA(0, 1, 0)	3.2816144	3.2816653	3.2816144	86.2793642
ANFIS + LA-8	ARIMA(1, 1, 0) with drift	3.1677598	3.1678125	3.1677598	80.8652744
ANFIS + BL14	ARIMA(1, 1, 0) with drift	3.2463071	3.2463585	3.2463071	84.5660539
ANFIS + C6	ARIMA(1, 1, 0) with drift	3.2132334	3.2132853	3.2132334	82.9894738
ANFIS + ARIMA direct	ARIMA(3, 2, 0)	3.3292731	3.3513009	3.3317951	89.6907631

various economies around the world. Tadawul has been influenced by both domestic and international economies. Consequently, financial crises in other countries are passed to domestic market. As a result, Tadawul was hit hard by the global financial crisis in 2008 [42].

4.1. The Result of MODWT Function. Table 5 provides the results of the suggested models using the first 80% of the dataset. The original LSCS data are provided by MODWT models. Based on the comparison, MODWT (LA-8) is identified as the best model since it has a minimum value of 0.0000053, 0.0032142, and 0.0644968, respectively, for ME, MAE, and MAPE. The MODWT (LA-8) model used LSCS as output variable whereas employed Repo and Loil as input variables to construct the ANFIS model.

4.2. The Result of Forecasting WT Models with ANFIS. In order to validate our findings, the forecasting is conducted using the remaining 20% of the transformed and original data with the same proposed models. The best model is ARIMA-MODWT (LA-8) with ANFIS since it has the lowest ME, RMSE, MAE, and MPE-fit as shown in Table 6. Similar to the training phase, LSCS is used as output variable whereas Repo and Loil are used as input variables by MODWT to construct ANFIS and ANFIS + MODWT models.

5. Conclusion

In this study, we have proposed a new model (MODWT-LA8-ANFIS) successfully. The model is used to forecast the closing price in stock market. We have selected oil price and Repo rate as input values based on correlation, the Engle and Granger Causality test, and multiple regressions. We found that there is weak correlation between the input variables ($r=0.327$). On the other hand, the correlation between oil price and output variable (closing price) is strong ($r=-0.673$). Moreover, the input variables have causality with output variable based on the Engle and Granger

Causality test. In order to check the significant effect, the multiple regression test is used. As a result, the input variables are significant at level 5%. The output variable is collected from Tadawul from October 2011 to December 2019 with 2026 observations. The input variables in this study have been collected from Saudi Authority for Statistics and Saudi Central Bank. The MODWT mechanism splits variables into details coefficient and approximation coefficient. MODWT has five functions, namely, Haar, Daubechies (Db), least Square (LA-8), Best localization (BL14), and Coiflet (C6). Therefore, the output variable is split into details coefficient (high fluctuated data) and approximation coefficient (which consists of the main features of data). The approximation coefficient data (MODWT-LA8) are used with input variables to build our model MODWT-LA8-ANFIS. The MODWT-LA8-ANFIS has been evaluated using statistical tests, namely, mean error (ME), root mean squared error (RMSE), and mean absolute percentage error (MAPE). The MODWT-LA8-ANFIS model has been compared with traditional models (ARIMA and ANFIS models). The MODWT-LA8-ANFIS is more accurate than the traditional models. Therefore, the new proposed forecasting model can be generalized to forecast in other international stock markets. Furthermore, this model is sufficiently powerful to optimize business processes for economic development of a country.

Data Availability

The data are publicly available online that are referred to in the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] J. Hull, *Risk Management and Financial Institutions*, John Wiley & Sons, Hoboken, NJ, USA, 2012.

- [2] Y. Seetharam, "Investigating the low-risk anomaly in South Africa," *Review of Behavioral Finance*, vol. 79, 2021.
- [3] T. G. Andersen, T. Bollerslev, P. Christoffersen, and F. X. Diebold, "Practical volatility and correlation modeling for financial market risk management," in *The Risks of Financial Institutions*, pp. 513–548, University of Chicago Press, Chicago, IL, USA, 2007.
- [4] J. Chen and D. N. Politis, "Time-varying NoVaS versus GARCH: point prediction, volatility estimation and prediction intervals," *Journal of Time Series Econometrics*, vol. 1, 2020.
- [5] A. Khera and M. P. Yadav, "Predicting the volatility in stock return of emerging economy: an empirical approach," *Theoretical & Applied Economics*, vol. 27, no. 4, 2020.
- [6] R. Matkovskyy, A. Jalan, and M. Dowling, "Effects of economic policy uncertainty shocks on the interdependence between Bitcoin and traditional financial markets," *The Quarterly Review of Economics and Finance*, vol. 77, pp. 150–155, 2020.
- [7] S. Mozumder, M. H. Kabir, M. Dempsey, and T. Choudhry, "Risk management under time varying volatility and Pareto-stable distributions," *Applied Economics Letters*, vol. 27, no. 3, pp. 161–167, 2020.
- [8] C.-Y. Tan, Y.-B. Koh, K.-H. Ng, and K.-H. Ng, "Dynamic volatility modelling of bitcoin using time-varying transition probability markov-switching GARCH model," *The North American Journal of Economics and Finance*, vol. 39, Article ID 101377, 2021.
- [9] K. A. Rashedi, M. T. Ismail, N. N. Hamadneh, S. Wadi, J. J. Jaber, and M. Tahir, "Application of radial basis function neural network coupling particle swarm optimization algorithm to classification of Saudi Arabia stock returns," *Journal of Mathematics*, vol. 2021, Article ID 5593705, 8 pages, 2021.
- [10] N. N. Hamadneh, M. Tahir, and W. A. Khan, "Using artificial neural network with prey predator algorithm for prediction of the COVID-19: the case of Brazil and Mexico," *Mathematics*, vol. 9, no. 2, p. 180, 2021.
- [11] N. N. Hamadneh, "Dead sea water levels analysis using artificial neural networks and firefly algorithm," *International Journal of Swarm Intelligence Research*, vol. 11, no. 3, pp. 19–29, 2020.
- [12] W. S. Khan, N. N. Hamadneh, and W. A. Khan, "Prediction of thermal conductivity of polyvinylpyrrolidone (PVP) electrospun nanocomposite fibers using artificial neural network and prey-predator algorithm," *PLoS One*, vol. 12, no. 9, Article ID e0183920, 2017.
- [13] P.-C. Chang and C.-Y. Fan, "A hybrid system integrating a wavelet and TSK fuzzy rules for stock price forecasting," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 6, pp. 802–815, 2008.
- [14] N. Homayouni and A. Amiri, "Stock price prediction using a fusion model of wavelet, fuzzy logic and ANN," in *Proceedings of the International Conference on E-Business, Management and Economics*, vol. 25, pp. 277–281, Dubai, UAE, December 2011.
- [15] R. H. Abiyev and V. H. Abiyev, "Differential evaluation learning of fuzzy wavelet neural networks for stock price prediction," *Journal of Information and Computing Science*, vol. 7, no. 2, pp. 121–130, 2012.
- [16] J.-S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.
- [17] H. Alawad, M. An, and S. Kaewunruen, "Utilizing an adaptive neuro-fuzzy inference system (ANFIS) for overcrowding level risk assessment in railway stations," *Applied Sciences*, vol. 10, no. 15, p. 5156, 2020.
- [18] M. A. Boyacioglu and D. Avci, "An adaptive network-based fuzzy inference system (ANFIS) for the prediction of stock market return: the case of the Istanbul stock exchange," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7908–7912, 2010.
- [19] L. S. Riza, C. N. Bergmeir, F. Herrera Triguero, J. M. Benítez Sánchez et al., "frbs: fuzzy rule-based systems for classification and regression in R," *American Statistical Association*, vol. 19, 2015.
- [20] J. J. Jaber, N. Ismail, S. Ramli, S. Al Wadi, and D. Boughaci, "Assessment of credit losses based on ARIMA-wavelet method," *Journal of Theoretical and Applied Information Technology*, vol. 98, no. 9, pp. 1379–1392, 2020.
- [21] V. Nadimi, A. Azadeh, P. Pazhoheshfar, and M. Saberi, "An adaptive-network-based fuzzy inference system for long-term electric consumption forecasting (2008- 2015): a case study of the group of seven (G7) industrialized nations: USA, Canada, Germany, United Kingdom, Japan, France and Italy," in *Proceedings of the 2010 Fourth UKSim European Symposium on Computer Modeling and Simulation*, pp. 301–305, IEEE, Pisa, Italy, November 2010.
- [22] W. Wang and L. Qiu, "Prediction of annual runoff using adaptive network based fuzzy inference system," in, vol. 3, pp. 1324–1327, in *Proceedings of the 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 3, , IEEE, Yantai, China, August 2010.
- [23] Y. Liang, Y. Jia, J. Li et al., "Online shop daily sale prediction using adaptive network-based fuzzy inference system," in *Proceedings of the 2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–66, IEEE, Suzhou, China, October 2019.
- [24] T. W. Septiarini, A. M. Abadi, A. Maman Abadi, and M. Rifki Taufik, "Application of wavelet fuzzy model to forecast the exchange rate IDR of USD," *International Journal of Modeling and Optimization*, vol. 6, no. 1, pp. 66–70, 2016.
- [25] X. Shao, H. Si, and W. Zhang, "Fuzzy wavelet neural control with improved prescribed performance for MEMS gyroscope subject to input quantization," *Fuzzy Sets and Systems*, vol. 411, pp. 136–154, 2020.
- [26] R. Gao, L. Du, and K. F. Yuen, "Robust empirical wavelet fuzzy cognitive map for time series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 96, Article ID 103978, 2020.
- [27] R. Tabaraki and M. Khodabakhshi, "Performance comparison of wavelet neural network and adaptive neuro-fuzzy inference system with small data sets," *Journal of Molecular Graphics and Modelling*, vol. 100, Article ID 107698, 2020.
- [28] A. Alhodiry, H. Rjoub, and A. Samour, "Impact of oil prices, the US interest rates on Turkey's real estate market. New evidence from combined co-integration and bootstrap ARDL tests," *PLoS One*, vol. 16, no. 1, Article ID e0242672, 2021.
- [29] S. Murali, S. Thiyagarajan, and N. Gopal, "Oil prices and stock market interplay in Dubai," *International Journal of Management Practice*, vol. 14, no. 1, pp. 107–127, 2021.
- [30] Saudi Authority for Statistics, "Saudi authority for statistics," 2021, <https://www.stats.gov.sa/en>.
- [31] Saudi Central Bank, <https://www.sama.gov.sa/en-us/pages/default.aspx>, 2021.
- [32] R. Gençay, F. Selçuk, and B. J. Whitcher, *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*, Elsevier, Amsterdam, Netherlands, 2001.

- [33] P. Zhang and H. Wang, "Fuzzy wavelet neural networks for city electric energy consumption forecasting," *Energy Procedia*, vol. 17, pp. 1332–1338, 2012.
- [34] I. H. Adil and A. U. R. Irshad, "A modified approach for detection of outliers," *Pakistan Journal of Statistics and Operation Research*, vol. 11, no. 1, pp. 91–102, 2015.
- [35] V. D. Sharma, "Board of director characteristics, institutional ownership, and fraud: evidence from Australia," *Auditing: A Journal of Practice & Theory*, vol. 23, no. 2, pp. 105–117, 2004.
- [36] A. M. A. Tapanjeh and A. R. Al Tarawneh, "Applicability of forensic accounting to reduce fraud and its effects on financial statement of Jordanian shareholding companies from the perspective of judiciary and certified public accountant," *International Journal of Financial Research*, vol. 11, no. 2, 2020.
- [37] H. Harandizadeh and D. J. Armaghani, "Prediction of air-overpressure induced by blasting using an ANFIS-PNN model optimized by GA," *Applied Soft Computing*, vol. 99, Article ID 106904, 2021.
- [38] N. N. Hamadneh, W. A. Khan, W. Ashraf, S. H. Atawneh, I. Khan, and B. N. Hamadneh, "Artificial neural networks for prediction of COVID-19 in Saudi Arabia," *Computers, Materials & Continua*, vol. 66, no. 3, 2021.
- [39] N. Hamadneh, *Logic programming in radial basis function neural networks*, PhD thesis, Universiti Sains Malaysia, George Town, Malaysia, 2013.
- [40] R. F. Engle and C. W. J. Granger, "Co-integration and error correction: representation, estimation, and testing," *Journal of the Econometric Society*, vol. 55, no. 2, pp. 251–276, 1987.
- [41] T. Alotaibi, A. Nazir, R. Alroobaea et al., "Saudi Arabia stock market prediction using neural network," *International Journal on Computer Science and Engineering*, vol. 9, no. 2, pp. 62–70, 2018.
- [42] A. R. Shaik and A. M. Syed, "Intraday return volatility in Saudi Stock Market: an evidence from Tadawul all share index," *Management Science Letters*, vol. 9, no. 7, pp. 1131–1140, 2019.