IEEE Access
Multidisciplinary ⁞ Rapid Review ⁞ Open Access Journal

Received 8 March 2023, accepted 3 May 2023, date of publication 11 May 2023, date of current version 24 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3275534

RESEARCH ARTICLE

Intelligent Hybrid ARIMA-NARNET Time Series Model to Forecast Coconut Price

ABDULLAH [1], RICHARD MANU NANA YAW SARPONG-STREETOR [2], RAJALINGAM SOKKALINGAM [2], MAHMOD OTHMAN [2], ABDUS SAMAD AZAD [2], GUNAWAN SYAHRANTAU [3], YUSRIWARTI [4], AND ZAINAL ARIFIN [5]

[1]Department of Information System, Faculty of Engineering and Computer Science, Universitas Islam Indragiri, Tembilahan Hulu 29213, Indonesia
[2]Department of Foundation and Applied Science, Universiti Teknologi PETRONAS, Perak 32610, Malaysia
[3]Department of Agribusiness, Faculty of Agriculture, Universitas Islam Indragiri, Tembilahan Hulu 29213, Indonesia
[4]Department of Accounting, Faculty of Economics and Business, Universitas Islam Indragiri, Tembilahan 29212, Indonesia
[5]Department of Management, Faculty of Economics and Business, Universitas Islam Indragiri, Tembilahan 29212, Indonesia

Corresponding author: Abdullah (abdullah@unisi.ac.id)

This work was supported by International Collaborative Research Grant -015MEO-220-223, between Universitas Islam Indragiri Indonesia and Universiti Teknologi PETRONAS Malaysia. The National Collaborative Research Fund (015MC0-032) at Universiti Teknologi PETRONAS Malaysia also supported the work.

ABSTRACT The global demand for coconut and coconut-based products has increased rapidly over the past decades. Coconut price continues to fluctuate; thus, it is not easy to make predictions. Good price modelling is important to accurately predict the future coconut price. Several studies have been conducted to predict the price of coconut using various models. One of the most important and widely used models in time series forecasting is the autoregressive integrated moving average (ARIMA). However, price fluctuations is considered a problem with uncertain behaviour. The existing ARIMA time series model is unsuitable for solving this problem, because of the nonlinear series. Artificial neural networks (ANN) have been an effective method in solving nonlinear data pattern problems in the last two decades. The non-linear autoregressive neural network (NARNET) gives good forecast, most especially when series are non-linear. Therefore ARIMA- NARNET is considered a universal approach to forecasting the coconut price. The aim of the study is to establish a linear and nonlinear model in time series to forecast coconut prices. The ability of a hybrid approach that combines ARIMA and NARNET(ANN) models is investigated. Based on the experimental study, the experimental results show that the proposed method ARIMA- NARNET, is better at forecasting the price of coconut, an agriculture commodity, than both the ARIMA model and NARNET models. The expected benefit of the proposed forecasting model is it can help farmers, exporters, and the government to maximize profits in the future.

INDEX TERMS ARIMA- NARNET, intelligent hybrid, coconut price, forecasting, time series.

I. INTRODUCTION

Indonesia produced 17.13 million tons of coconut in 2019. Based on the World Atlas report, coconut production in Indonesia is the highest in the world. Referring to the data from the Indonesian Central Statistics Agency (BPS), coconut exports from Indonesia reached 1.53 million tons or US$ 819.26 million as of the third quarter of 2020. The countries which are the destinations for Indonesia's coconut exports includes fthe United States, Netherlands, South Korea, China, Japan, Singapore, Philippines and Malaysia [1]. Therefore, the coconut price forecast has a fundamental importance in the trading strategy of Indonesia. A good forecasting model is critically important to predicting the future price of coconut accurately, thus proper planning could be made by the farmers, exporters, and the government to maximize future profit. The forecast of the coconut price in time series is considered one of the most challenging because of fluctuation issues of coconut price. Fluctuations in agricultural prices affect the supply and demand of commodities and have a significant impact on consumers and farmers [2]. Fluctuations in coconut price lead to uncertainty of income

The associate editor coordinating the review of this manuscript and approving it for publication was Xianzhi Wang.

for the farmer, making it difficult for the government to put in place policies and stabilize supply and demand.

The Autoregressive integrated moving average (ARIMA) model has been one of the vital and widely used methods in time series forecasting [3], [4], [5]. The popularity of the ARIMA model is due to its statistical properties as well as the use of the well-known Box-Jenkins methodology in the model-building process [6]. This model assumes the time series under study is generated from a linear process. Several methods have been used to model and predict coconut prices including the ARIMA model [7], [8], [9]. Results showed the potential of the ARIMA model accurately predict coconut price data. However, ARIMA time series models are generated from linear processes and therefore may be unsuitable for most practical problems that are nonlinear. Prices of industrial agriculture are largely influenced by eventualities, and seasonality, consequently prices are nonlinear and difficult to predict [10]. The fluctuation of coconut prices is considered an uncertain behaviors and changes over time. Some factors that influence the fluctuations of coconut price are company pricing, falling market demand for coconut, and declining quality, and quantity of coconut products. Therefore, it is a challenge to propose an appropriate approach to forecast coconut prices.

Recently, time series data can be modelled using artificial neural networks (ANN). The main advantage of a neural network is its flexible functional form and universal functional approximator. ANN is effective in solving nonlinear data pattern problems. Many non-linear problems are relevant today, including forecasting stock markets with uncertain behaviour and changing over time. There are several studies where neural networks are used to address agricultural commodity price forecasting [11], [12], [13], [14]. The results conclude that the ANN is a better model for forecasting agriculture commodity prices than the ARIMA model [15]. However, the use of a single ANN model could not be complementary in capturing patterns to obtain an optimal prediction. The literature review demonstrates that the ANN model is suitable for nonlinear time series data and the ARIMA model is suitable for linear time series data. In this paper, a hybrid model of coconut price prediction is proposed. The motivation behind this hybrid model is the fact that coconut price fluctuation is complex. The hybrid methodology combines both ARIMA and ANN models to take advantage of ARIMA and ANN models in linear and nonlinear modelling. The ability of a hybrid approach combining ARIMA and ANN models for coconut price forecasting is investigated. The use of hybrid models could be complementary in capturing patterns of coconut price data and could improve forecasting accuracy. Quite different from the ARIMA model and ANN model, this proposed hybrid model combines the ARIMA, and ANN to more accurate predictions for coconut prices. The coconut price prediction model proposed in this study will help farmers, exporters, and the government to maximize profits in the future.

This paper is organized as follows: Section II discusses the methods used in the research work. Section III presents the results of ARIMA-NARNET modelling process for coconut price prediction. NARNET is a version of ANN. Section IV summarizes the present study and draws some conclusions.

II. RESEARCH METHODS

The research methods section involves acquiring the data, process' used to develop the models, forecasting process and lastly the evaluation of the forecasts.

A. DATA

The data for the study is the monthly price of coconut obtained from the Department of Industry and Trade, Indragiri Hilir Regency, Indonesia. The data sample is the average coconut price per month starting from January 2014 to February 2022. Therefore, there are 115 coconut price data points in rupiah currency which have been collected for 8 years and 3 months. The historical data of the coconut prices per month fluctuated over time. This is shown in Figure 1. The data can be found online at: https://bit.ly/3M7TvcM

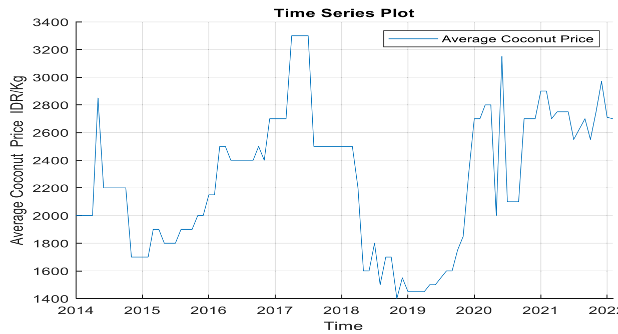

FIGURE 1. Average coconut price in rupees per kilogram from 2014-2022.

B. ARIMA MODEL

$ARIMA(p, d, q)$ model is stochastic in nature and has been used in diverse fields for prediction studies [16], [17], [18]. ARIMA was first put into use in time series for modelling and forecasting by Box Jenkins in 1970 [19]. The model is made up of three parts; autoregressive (AR), integrated (I), and moving average (MA). ARIMA model is developed in three steps: (1) Model Identification, (2) Parameter Estimation and (3) Diagnostic Checking

1) MODEL IDENTIFICATION

Model identification involves finding the order of the ARIMA model using the sample autocorrelation function (SACF) and sample partial autocorrelation function (SPACF) charts [20]. SACF of a time series is the correlation of its past's values with its future values. Given that data points of the time series with first $N-1$ observation is $X_t : t = 2, 3, \ldots, N$, where

$t = 1, 2, \ldots, N-1$, the relationship between X_t and X_{t+1} is defined as equation (1) and equation (2);

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - X_1)(x_{t+1} - X_2)}{\left[\sum_{t=1}^{N-1} (x_t - X_1)^2\right]\left[\sum_{t=1}^{N-1} (x_t - X_1)^2\right]} \quad (1)$$

$$r_1 = \frac{\sum_{t=1}^{N-1} (x_t - X)(x_{t+1} - X)}{\sum_{t=1}^{N-1} (x_t - X)^2} \quad (2)$$

X_1 is the first $N-1$ observation's mean. When N is substantially large, variations among sub-period means X_1 and X_2 are neglected and r_1 is calculated by equation (3):

$$r_k = \frac{\sum_{t=1}^{N-1} (x_t - X)(x_{t+k} - X)}{\sum_{t=1}^{N-1} (x_t - X)^2}. \quad (3)$$

SACF is used to identify the moving average order of a stationary time series. SPACF is the correlation between lag values with other shorter lags of the group at various lags k, where $k = 1, 2, 3 \ldots$. SPACF at varied lags k is defined by equation (4);

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-1}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \quad (4)$$

$r_{k,j} = r_{k-1,j} - r_{kk} r_{k-1,k-j}, j = 1, 2, \ldots, k-1$. SPACF is used to identify the autoregressive order of a stationary time series. SPACF of an $AR(p)$ process at lag $p+1$ and beyond is zero.

2) PARAMETER ESTIMATION

The Box–Jenkins model of order $ARIMA(p, d, q)$ is given by equation (5).

$$\phi(1+L)^p (1+L)^d y_t = c + \theta(1+L)^q \varepsilon_t \quad (5)$$

The variable y_t, the future value at time step t, is taken to be a linear function of several past observations $y_{t-n}, 1, 2, \ldots, n < t$ and random errors, ε_t as demonstrated by equation (5). p is the autoregressive order; q is the moving average order and d represents the differencing order of the coconut price time series. L is the Lag operator. ϕ and θ are the coefficients of regressions for the autoregressions and moving averages [21], [22].

3) DIAGNOSTIC CHECKING

The residual (white noise) of models is assessed using the correlogram (SACF and SPACF), Ljung–Box Q tests [23] and Durbin–Watson test [24] to test the sufficiency of the models.

C. ANN MODEL

The artificial neural network has the potential to represent complex, nonlinear relationships [25], [26], [27], [28], [29], [30], [31], [32], [33], [34]. The evolution of ANN has given rise to the multilayer perceptron (deep learning), which is effective at modelling and predicting complex, nonlinear relationships in time series. It is made up of an input layer, hidden layer; and an output layer. The hidden layer is a network of three layers connected by open-chain linkages as shown in Figure 2.

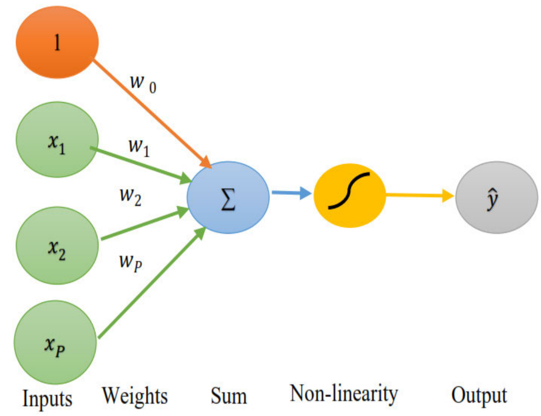

FIGURE 2. The perceptron forward propagation.

$w_{i,j}$ and w_j where $i = 0, 1, 2, \ldots, P$ $j = 1, 2, \ldots, Q$ and are the model parameters. Also referred to as connection weights, the model parameters have P as the number of input nodes and Q as the number of hidden nodes.

After the hidden layers, the sigmoid function, equation (6), among others is employed as an activation

$$Sig(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

function to introduce nonlinearity to the output of the neural network. The nonlinearity allows the network to arbitrarily approximate complex functions as the perceptron is a linear combination of the weights and the input vector.

The network is trained after the activation function is applied. Training is done through Optimization (backpropagation) of the activated perceptron. This allows some of the activated perceptron to drop out as the weight approaches zero (regularization). The perceptron is trained in mini-batches to allow the central processing unit (CPU) or graphical processing unit (GPU) to process the network in a fast, accurate estimation of gradients, smooth convergence of gradients and also allow large learning rates [35]. The remaining perceptron acts as an input node again, and weights are added to form a new network. The backpropagation is done again. The training process continues until there is one perceptron node left. Equation (7) is the mathematical equation between inputs $(y_{t-1}, \ldots, y_{t-p})$ and output (y_t).

$$y_t = w_0 + \sum_{j=1}^{Q} w_g g\left(w_{0j} + \sum_{i=1}^{p} w_{i,j} y_{t-i}\right) + e_t \quad (7)$$

$$y_t = f(y_{t-i}, \ldots, y_{t-P}, W) + e_t. \quad (8)$$

The ANN model, equation (7) maps the input data to the forecast values, y_t. The connection weight, W is a vector containing all parameters [31]. Equation (7) implies one output node emerges as the step-ahead forecast. It shows that the network is robust and can model any function when the number of neurons of the hidden nodes (Q) are high enough [36]. An out-of-sample forecasting can be effectively done using a primary network layout with a modest number of hidden nodes (Q) [37]. The parameter Q is influenced by the input data and therefore there is no alternative process for determining it. The selection of the number of input vectors, P, and its dimensionality is critical to ANN modelling [37]. The autocorrelation, a nonlinear framework of the time series is defined by P. It is one of the most vital parameters to estimate in an ANN model. Known hypothesis have not been able to assist in P selection. Research is mostly done to identify appropriate Q and P. In the implementation phase, we select the NARNET, a shallow learning model in MATLAB to model the residual of the ARIMA model.

D. INTELLIGENT HYBRID ARIMA-NARNET MODEL

The hybrid modelling process has the linear (ARIMA), and nonlinear (NARNET) components of the model defined respectively as $\hat{L}_t$ and $\hat{J}_t$ [38]. The Intelligent hybrid model $\hat{y}_t$ is estimated using the equation (9).

$$\hat{y}_t = \hat{L}_t + \hat{J}_t \tag{9}$$

$\hat{J}_t$ is the NARNET model trained from the residual of the ARIMA model $\hat{J}_t$ at time t.

E. FORECAST EVALUATION

The forecasts are evaluated using the Multiple Forecast Comparison method (MDM) and if possible, the Diebold Mariano (DM) test. The multiple forecast comparison method investigates whether three or more forecasts for example, the Hybrid ARIMA-NARNET, ARIMA, and NARNET perform equally in terms of specific loss B () functions such as mean absolute error (MAE) and mean squared error (MSE). The Diebold Mariano test is used to compare whether two forecasts performed equally.

The hypothesis for the Multiple forecasts' comparison test of Equal predictive ability (EPA) is;

$$H_o\colon \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{1,t}\right)\right] = \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{2,t}\right)\right], \ldots, = \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{k+1,t}\right)\right].$$

The alternative hypothesis is;

$$H_1\colon \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{1,t}\right)\right] \neq E\left[\mathrm{B}\left(\mathrm{e}_{2,t}\right)\right], \ldots, \neq \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{k+1,t}\right)\right].$$

A test of significance level is conducted by rejecting the null hypothesis of EPA when;

$$SorS_c > X^2_{k,1-\alpha}.$$

Her, $X^2_{k,1-\alpha}$ is the quantile of $X^2_{k,}$ distribution. Rejection of the null hypothesis using S *or* Sc implies that one or more of the alternative models stands out in terms of predictive ability [39]. The hypothesis for the Diebold Marino test of EPA is

$$H_o\colon \mathrm{B}\left[\mathrm{L}\left(\mathrm{e}_{1,t}\right)\right] = \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{2,t}\right)\right],$$

implies that the observed differences between the performance of two forecasts are not significant, while the alternative hypothesis,

$$H_1\colon \mathrm{E} \neq \mathrm{E}\left[\mathrm{B}\left(\mathrm{e}_{2,t}\right)\right]$$

implies that the observed differences between the performance of two forecasts are significant. The DM test has a normal distribution [40]. The assumption for the test is that the models are not nested. Alternative models are invariant to any permutation (reordering) [39], [40], [41], [42], [43], [44].

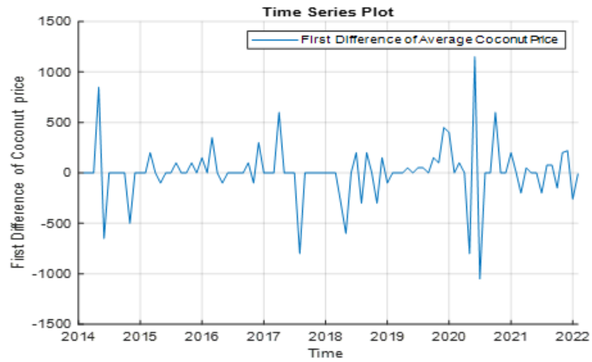

FIGURE 3. First difference plot of average coconut price.

III. RESULTS AND ANALYSIS

In this section, The Intelligent ARIMA-NARNET model is developed and its forecasting power is assessed. A 12-month forecast of coconut price is made using the ARIMA, NARNET and ARIMA-NARNET Hybrid models.

A. COCONUT PRICE ARIMA MODELING

The ARIMA modelling is presented in this section.

1) ARIMA MODEL IDENTIFICATION

The entire data obtained is used to train the ARIMA model. The data as seen in Figure 1 is not stationary. A condition necessary to train the ARIMA model is that the data is stationary. The ARIMA model is anticipated by identifying a stationary time series at the first difference, $d = 1$. This is shown in Figure 3.

The SACF and the SPACF are plotted from the stationary time series. The ARIMA p, q parameters were identified using SACF and SPACF plots which are shown in Figure 4 and Figure 5 respectively. Observing the SPACF, the autocorrelations spike at lag 1, and die off sharply for the other lags, hence the p is estimated to be 1. The identified tentative model for the coconut price data is an $ARIMA(1, 1, 0)$ with equation (10).

$$(1 - \emptyset_1 L)(1 - L)\, y_t = \epsilon_t \tag{10}$$

Equation (10) is expanded to give equation (11).

$$y_t = y_{(t-1)}(1 + \emptyset_1) - \emptyset_1 y_{(t-2)} + \epsilon_t \tag{11}$$

2) ARIMA MODEL PARAMETER ESTIMATION

The model parameters are estimated using the MATLAB Econometric Modeler [45]. Tentative models are assessed and compared, using the AICs and BICs; for instance, the ARIMA models with and without the constant terms were compared and the models trained under the Gaussian, and t

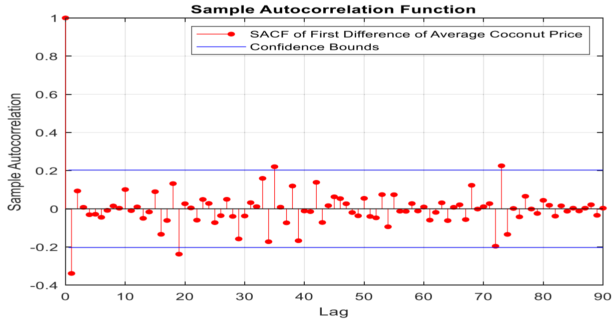

FIGURE 4. Sample autocorrelation function for the first difference of coconut price.

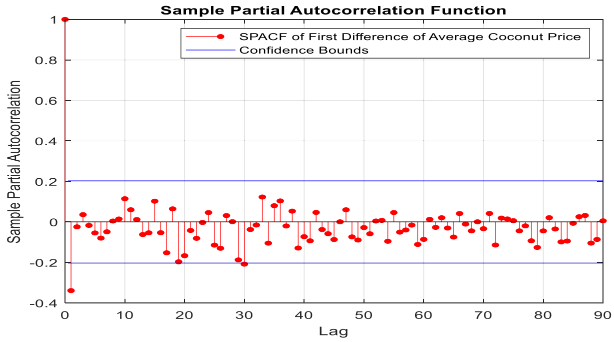

FIGURE 5. Sample partial autocorrelation function for the first difference of coconut price.

TABLE 1. Parameter estimates for ARIMA (1,1,0) model.

Parameter	Value	Standard Error	***t*** Statistic	***P***-Value
Constant	-			
AR **(1)**	-0.33438	0.060473	-5.5294	3.214e-08
Variance	71167.9454	6807.4798	10.4544	1.3992e-25

TABLE 2. Performance values of ARIMA(1, 1, 0) model.

Model	AIC	BIC
ARIMA (1,1,0)	1374.1748	1379.3035

distributions are compared. The results obtained by following the iterative procedure of ARIMA model estimation are given in Table 1 and Table 2.

The estimation is done with the Gaussian probability distribution and the constant term omitted to optimize the model. The parameters in Table 1 are substituted into the model, equation (11) which gives equation (12).

$$y_t = 0.66562y_{(t-1)} + 0.33438y_{(t-2)} + \epsilon_t \tag{12}$$

The MATLAB code for equation (12) is Appendix I. Appendix I is then applied in Appendix II to carry out the ARIMA forecast. Figure 6 and Figure 7 are the ARIMA Model Fit Plot and Residual Plot of the Average Coconut Price respectively.

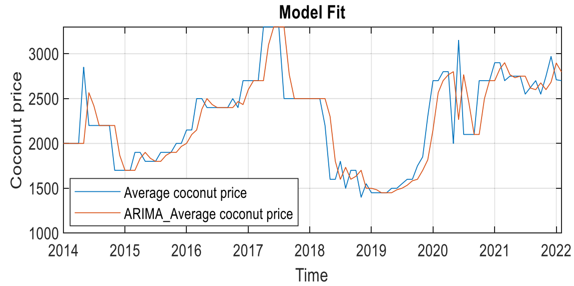

FIGURE 6. Plot of ARIMA(1,1,0) model.

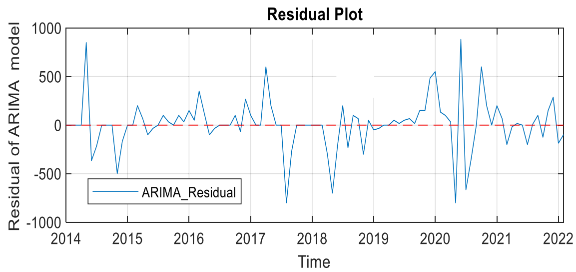

FIGURE 7. Plot of residual of ARIMA (1,1,0) model.

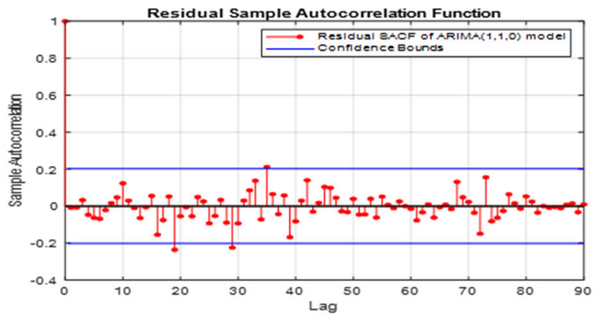

FIGURE 8. Residual sample autocorrelation function.

3) DIAGNOSTIC CHECKING

The ARIMA Model Fit Plot and Residual Plot model is assessed using the Residual Sample Autocorrelation Plot as shown in Figure 8. There are spikes in the Residual SACF which indicates autocorrelation in the residual data, thus the ARIMA model is still not sufficient for the coconut price data hence the need to model the residual data. The Nonlinear Autoregressive Neural network (NARNET) is used to model the residual.

B. RESIDUAL NARNET MODELING

The NARNET modelling process involves three steps (1) setting the input parameters for the NARNET training, (2) training the Network and (3) deploying the Neural network. The NARENT modelling process is discussed below.

1) SETTING INPUT PARAMETERS FOR NARNET TRAINING

The training process is achieved using the Neural Net Time Series application which is part of the Machine learning and Deep learning Applications cluster in MATLAB [46]. The residual time series is used as the only input for the NARNET and requires a continuous feed of forecasted data to allow the network to continue working. The input of the neural network is the residual of the ARIMA model. The residual is retrieved using the code provided in Appendix III. The input data is 98 months data points; short by 1 month because of the first differencing at the ARIMA modelling stage. The delay time step is set at 2 months. 70% of the data is used for training the neural network, 15% is used to test the trained network and the other 15% is used to validate the Network. The neural network architecture is set as per Figure 9 for a single horizon forecast and Figure 10 for multiple horizons forecast.

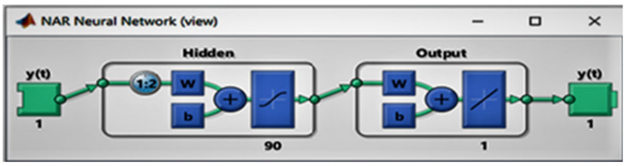

FIGURE 9. Closed loop NARNET architecture.

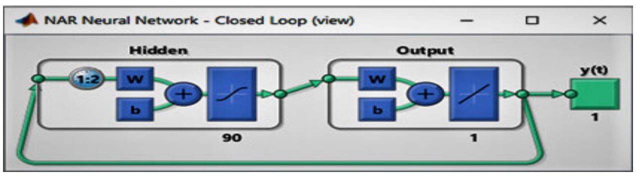

FIGURE 10. Open loop NARNET architecture.

There are 90 hidden layers and one output layer with an output node. The 90 hidden layers are optimal and were arrived at through continuous testing process.

The NARNET is first initialized using random weights at the start of the training process. A Levenberg-Marquardt Back Propagation (LMBP), an iterative algorithm is chosen to train the NARNET model. The LMBP algorithm locates the minimum of a function expressed as the sum of the squares of nonlinear functions through an iterative process. The training cycle 'epoch', is set automatically by converging at the minimum point of the function. The least MSE is used in the NARNET training to identify the best number of layers and associated neurons in each hidden layer [39].

2) TRAINING RESULT

After the NARNET architecture has been set, in the workflow, Appendix IV, or the training application window, the training command is executed by clicking the train button and waiting up until it is done. The training outputs of the neural network have several parameters which are necessary for the neural network to be trained optimally.

The Progress box in Figure 11, the Trained output, shows the error performance of the network which is initialized at 4.06 × 10+7 MSE and stopped at 2.53 × 10+4 MSE. The training performance window in Figure 12 shows that overfitting and underfitting are avoided by training the network such that training, testing, and validation performance graphs are parallel. The R, an indication of the linear relationship between the outputs and targets, which measures the goodness of fit, of the neural network model is above 71% for the training set. Figure 13 shows the model summary where the testing and validation sets are above 42%. Here there is a little compromise on the R for the validation and testing R. The training is repeated until R above 50% is achieved for the training and the testing and validation sets. The fitted model for the residual is presented in Figure 14. Figure 15 is the errors associated with the neural network model. The sufficiency of the neural network is assessed using the autocorrelation of errors. There is no autocorrelation in the errors (Erro1) as shown in Figure 16.

The spikes do not die sharply beyond the first lag for the non-zero correlations from the neural network errors. It is the same case in the autocorrelation correlation between Input1 and Error1 (Target1 – Output1) as shown in Figure 17.

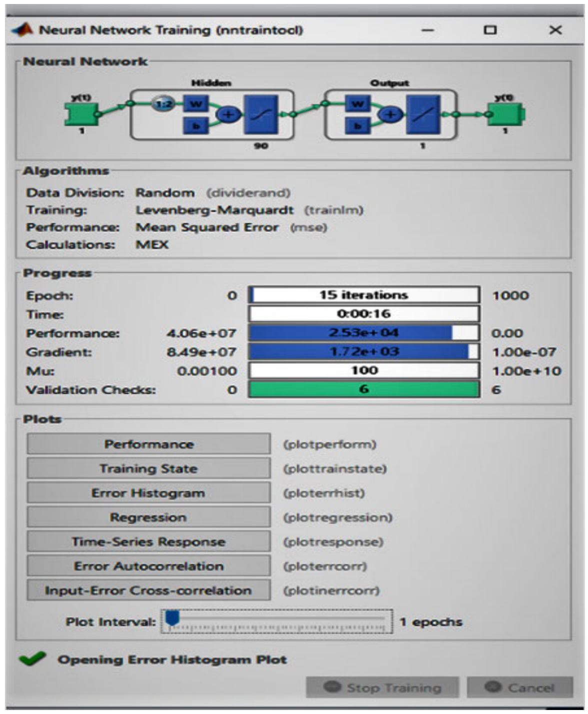

FIGURE 11. Trained output.

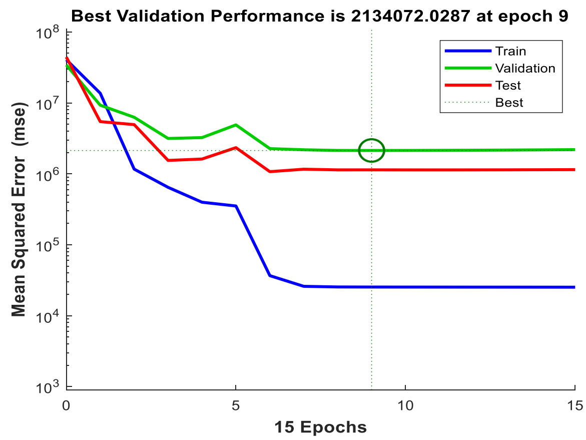

FIGURE 12. Training performance.

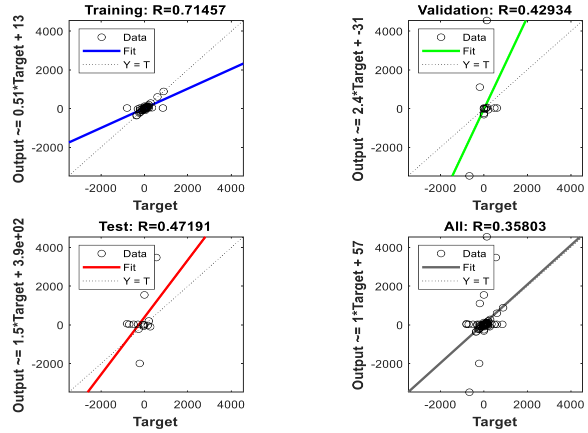

FIGURE 13. **Model summary.**

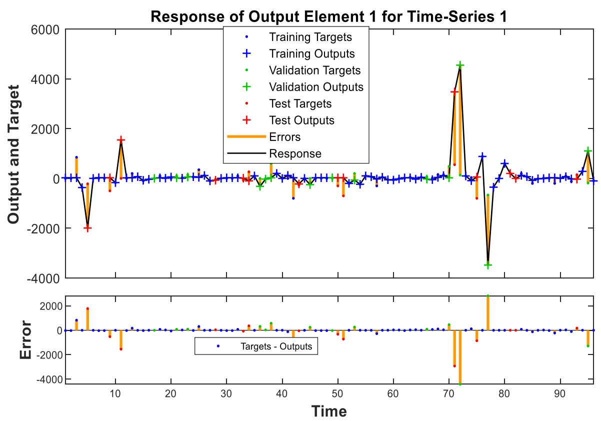

FIGURE 14. **Response of output element for residual time-series.**

There is no evidence of a correlation between errors (Erro1), and input (ARIMA residual). The inference is that the NARNET model is now sufficient to model the residual component of the coconut price.

3) DEPLOYING THE NARNET

The NARNET model is deployed as a function with the input Arguments stored in the trained network structure in the MATLAB workspace, Appendix V is the function. The advantage of deploying a trained network in such way is to avoid the network behaving as a stochastic model but as a deterministic function.

C. INTELLIGENT HYBRID ARIMA-NARNET MODELING

The hybrid model is deduced from equation (9). The ARIMA model can be expressed mathematically as equation (12), which is also in MATLAB code as Appendix II, but the NARNET model as shown in Appendix IV cannot easily be expressed in a single mathematical equation, instead, the hybrid model is expressed in the code form as presented in Appendix VI.

D. A 12-MONTH FORECAST USING THE MODELS

In Figure 18, the forecast of coconut price is plotted, for three different models, ARIMA, NARNET and Intelligent

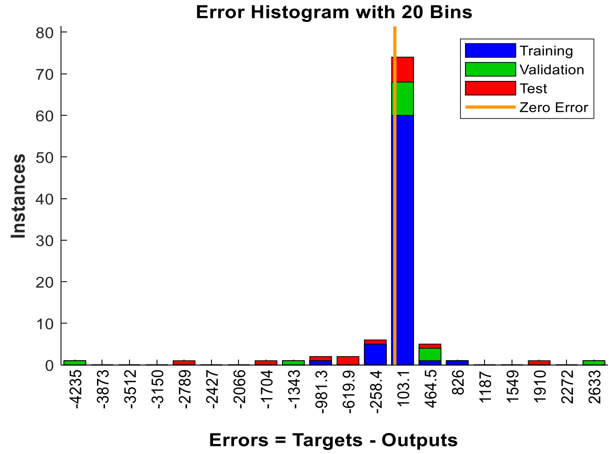

FIGURE 15. **Errors of the fitted residual neural network.**

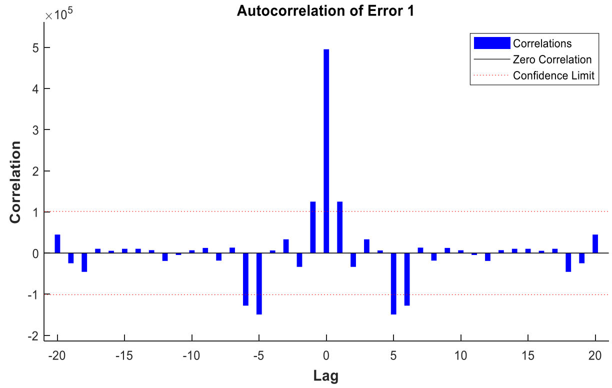

FIGURE 16. **Autocorrelation of Error1.**

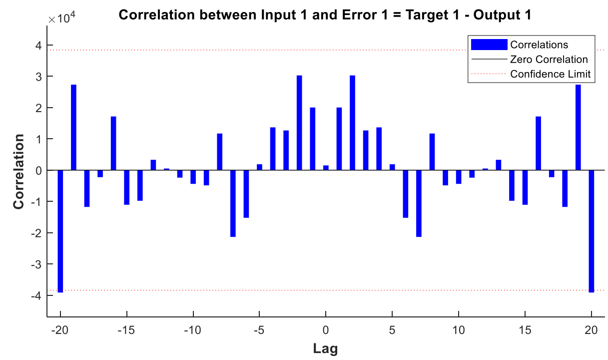

FIGURE 17. **Autocorrelation correlation between Input1 and Error1 (Target1 – Output1).**

Hybrid ARIMA-NARNET. Other features of the plot are: the observed price time series and the 95% confidence bound of the Intelligent Hybrid ARIMA-NARNET model. The are some missing prices from March 2023–August 2023 from the observed prices time series; however these do not have any impact on the forecast as the data used for the model span from January 2014-February 2023. The previous month January 2023 is used to estimate the missing month's prices.

E. FORECAST EVALUATION

The MDM test showed a test statistic of S or Sc at infinity for the first forecast horizon and NAN for the other forecast

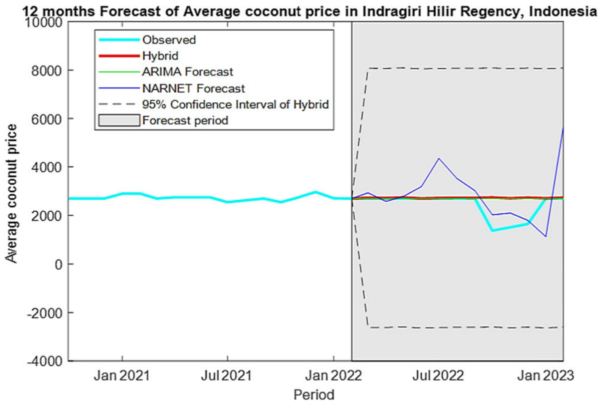

FIGURE 18. Forecasted monthly coconut price.

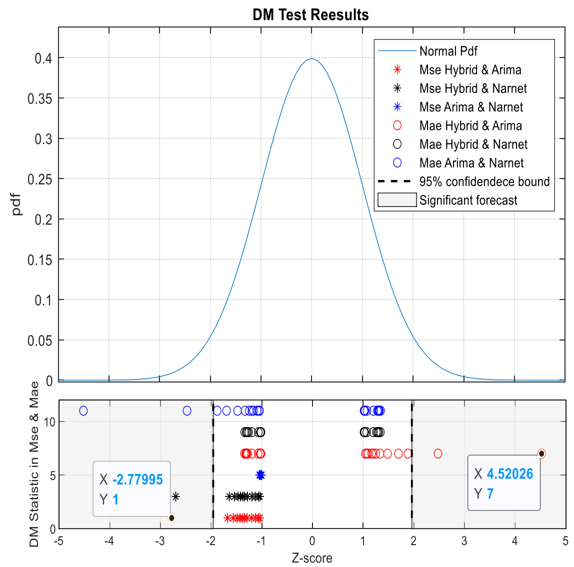

FIGURE 19. The DM test result.

horizons (2nd to 12th). The NAN signifies that the test was not successful, this may be due to the models having nested properties and this results in the singular matrixes in calculating the S and Sc Statistics. On the other hand, the infinity on the chi-square scale signifies that there is at least one of the models with superior predictive ability, concerning the other models in the first forecast horizon. The DM test is resorted to identifying the model with the superior predictive ability. Here the assumption was that the models are 4th-order polynomials as can be seen in Figure 1. Per the nature of the DM statistic in this particular test, it may produce equivalent statistics at both ends of the normal distribution curve as can be seen in Figure 19. The DM test statistics generally reduce as the forecast horizon increases. Rejection of the null hypothesis using DM statistic implies that one or more of the alternative models have superior predictive ability. The models are characterized in a 95% confidence interval bound which is equivalent to the test statistic 1.96. or below. It is expected that the alternative model has superior predictive ability, if the DM statistic >1.96. as shown in Figure 19. The DM test result shows, the hybrid ARIMA-NARNET and ARIMA forecast comparison for the first month/horizon has a superior predictive ability. The ARIMA and NARNET comparisons are not considered as their results are inconsistent for both loss functions. Comparatively hybrid ARIMA–NARNET is better than ARIMA from the forecast graph (Figure 18). The hybrid ARIMA–NARNET blends some nonlinear features which are captured by the NARNET with the ARIMA.

IV. CONCLUSION

The results conclude that the Hybrid ARIMA-NARNET model is better for forecasting agriculture commodity prices than both the ARIMA and NARNET models. This is because a single ARIMA model cannot capture all patterns for an optimal forecast, it captures mostly the linear patterns. Per the analysis above, the NARNET model is ideal for nonlinear time series. In this paper, a hybrid model of coconut price prediction is proposed. The forecast evaluation indicates that the hybrid ARIMA-NARNET model is the best at forecasting coconut prices as it has the strongest predictive ability. Hybrid models can complementarily capture patterns of coconut price data and improve forecasting accuracy. The proposed hybrid forecasting model blends linear and nonlinear model features. The coconut price forecast model suggested in this study will help farmers, exporters, and the government to maximize profits in the future.

APPENDIX A
SUPPORTING INFORMATION

Supplementary codes associated with this article can be found online at: https://bit.ly/3XopI2B

ACKNOWLEDGMENT

The authors are grateful to the department of industry and trade, Indragiri Hilir Regency, Indonesia for providing coconut price data. The authors are grateful to the International Collaborative Research Grant -015MEO-220-223, between Universitas Islam Indragiri Indonesia and Universiti Teknologi PETRONAS Malaysia. The authors are also grateful to the National Collaborative Research Fund (015MC0-032) at Universiti Teknologi PETRONAS Malaysia for supporting the Research.

REFERENCES

[1] J. C. Alouw and S. Wulandari, "Present status and outlook of coconut development in Indonesia," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 418, no. 1, Jan. 2020, Art. no. 012035, doi: 10.1088/1755-1315/418/1/012035.

[2] Y. H. Gu, D. Jin, H. Yin, R. Zheng, X. Piao, and S. J. Yoo, "Forecasting agricultural commodity prices using dual input attention LSTM," *Agriculture*, vol. 12, no. 2, pp. 1–18, 2022, doi: 10.3390/agriculture12020256.

[3] S. Khan and H. Alghulaiakh, "ARIMA model for accurate time series stocks forecasting," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 524–528, 2020, doi: 10.14569/IJACSA.2020.0110765.

[4] C. N. Babu and B. E. Reddy, "Performance comparison of four new ARIMA-ANN prediction models on internet traffic data," *J. Telecommun. Inf. Technol.*, vol. 2015, no. 1, pp. 67–75, 2015. [Online]. Available: https://core.ac.uk/download/pdf/235207347.pdf

[5] U. M. Butt, S. Letchmunan, F. H. Hassan, M. Ali, A. Baqir, T. W. Koh, and H. H. R. Sherazi, ''Spatio-temporal crime predictions by leveraging artificial intelligence for citizens security in smart cities,'' *IEEE Access*, vol. 9, pp. 47516–47529, 2021, doi: 10.1109/ACCESS.2021.3068306.
[6] D. Aryal, ''Time-series analysis with a hybrid Box-Jenkins ARIMA and neural network model,'' *J. Harbin Inst. Technol.*, vol. 11, no. 4, pp. 413–421, 2004.
[7] T. Prasert and V. Rungreunganun, ''Thai coconut price forecasting using ARIMA model,'' *Int. J. Ind. Eng. Res. Develop.*, vol. 12, no. 1, pp. 950–961, Feb. 2021, doi: 10.34218/ijierd.12.1.2021.001.
[8] B. D. P. Rangoda, L. M. Abeywickrama, and M. T. N. Fernando, ''An analysis of different forecasting models for prices of coconut products in Sri Lanka,'' in *Proc. 3rd Academic Sessions*, 2004, pp. 8–15. [Online]. Available: http://ir.lib.ruh.ac.lk/bitstream/handle/iruor/1018/AP-6387-8.pdf?sequence=1&isAllowed=y
[9] K. Khan and D. Singh, ''Stock price forecasting of Maruti Suzuki using auto regressive integrated moving average (ARIMA) model,'' *Rev. Prof. Manage.*, vol. 20, no. 1, pp. 2164–2169, Jul. 2022. [Online]. Available: https://www.phytojournal.com/archives/2019/vol8issue3/PartAC/8-3-5-424.pdf
[10] M. G. D. Abeysekara and K. Waidyarathne, ''The coconut industry: A review of price forecasting modelling in major coconut producing countries,'' *Cord*, vol. 36, pp. 17–26, Jan. 2020, doi: 10.37833/cord.v36i.422.
[11] G. K. Jha and K. Sinha, ''Agricultural price forecasting using neural network model: An innovative information delivery system,'' *Agricult. Econ. Res. Rev.*, vol. 26, no. 26, pp. 229–239, 2013, doi: 10.22004/ag.econ.162150.
[12] A. K. Mahto, M. A. Alam, R. Biswas, J. Ahmed, and S. I. Alam, ''Short-term forecasting of agriculture commodities in context of Indian market for sustainable agriculture by using the artificial neural network,'' *J. Food Quality*, vol. 2021, Jan. 2021, Art. no. 9939906, doi: 10.1155/2021/9939906.
[13] W. Anggraeni, F. Mahananto, M. A. Rofiq, K. B. Andri, Z. Zaini, and A. P. Subriadi, ''Agricultural strategic commodity price forecasting using artificial neural network,'' in *Proc. Int. Seminar Res. Inf. Technol. Intell. Syst. (ISRITI)*, Nov. 2018, pp. 347–352, doi: 10.1109/ISRITI.2018.8864442.
[14] K. P. Waidyarathne, T. H. Chandrathilake, and W. S. Wickramarachchi, ''Application of artificial neural network to predict copra conversion factor,'' *Neural Comput. Appl.*, vol. 34, no. 10, pp. 7909–7918, May 2022, doi: 10.1007/s00521-022-06893-3.
[15] V. S. Pandey and A. Bajpai, ''Predictive efficiency of ARIMA and ANN models: A case analysis of nifty fifty in Indian stock market,'' *Int. J. Appl. Eng. Res.*, vol. 14, no. 2, pp. 232–244, 2019.
[16] B. Choubin and A. Malekian, ''Combined gamma and M-test-based ANN and ARIMA models for groundwater fluctuation forecasting in semiarid regions,'' *Environ. Earth Sci.*, vol. 76, no. 15, pp. 1–10, Aug. 2017, doi: 10.1007/s12665-017-6870-8.
[17] A. Hernandez-Matamoros, H. Fujita, T. Hayashi, and H. Perez-Meana, ''Forecasting of COVID19 per regions using ARIMA models and polynomial functions,'' *Appl. Soft Comput.*, vol. 96, Nov. 2020, Art. no. 106610, doi: 10.1016/j.asoc.2020.106610.
[18] Y. Song and J. Cao, ''An ARIMA-based study of bibliometric index prediction,'' *Aslib J. Inf. Manage.*, vol. 74, no. 1, pp. 94–109, Jan. 2022, doi: 10.1108/AJIM-03-2021-0072.
[19] G. E. P. Box and D. A. Pierce, ''Distribution of residual autocorrelations in autoregressive-integrated moving average time series models,'' *J. Amer. Stat. Assoc.*, vol. 65, no. 332, pp. 1509–1526, Dec. 1970, doi: 10.1080/01621459.1970.10481180.
[20] A. Chuang and W. W. S. Wei, ''Time series analysis: Univariate and multivariate methods,'' *Technometrics*, vol. 33, no. 1, p. 108, 1991, doi: 10.2307/1269015.
[21] R. T. O'Connell, A. B. Koehler, B. L. Bowerman, R. T. O'Connell, and A. B. Koehler, *Forecasting, Time Series, and Regression: An Applied Approach*, vol. 4. San Francisco, CA, USA: Thomson Brooks/Cole, 2005.
[22] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. Hoboken, NJ, USA: Wiley, 2015.
[23] G. M. Ljung and G. E. P. Box, ''On a measure of lack of fit in time series models,'' *Biometrika*, vol. 65, no. 2, pp. 297–303, Aug. 1978, doi: 10.1093/biomet/65.2.297.
[24] J. Durbin and G. S. Watson, ''Testing for serial correlation in least squares regression. III,'' *Biometrika*, vol. 58, no. 1, pp. 1–19, 1971, doi: 10.2307/2334313.
[25] H. X. Li and X. L. Da, ''A neural network representation of linear programming,'' *Eur. J. Oper. Res.*, vol. 124, no. 2, pp. 224–234, 2000, doi: 10.1016/S0377-2217(99)00376-8.
[26] H. X. Li and L. X. Li, ''Representing diverse mathematical problems using neural networks in hybrid intelligent systems Hong,'' *Expert Syst. Appl.*, vol. 16, no. 4, pp. 271–281, 1999, doi: 10.1111/1468-0394.00118.
[27] P. Wang, L. Xu, S.-M. Zhou, Z. Fan, Y. Li, and S. Feng, ''A novel Bayesian learning method for information aggregation in modular neural networks,'' *Expert Syst. Appl.*, vol. 37, no. 2, pp. 1071–1074, Mar. 2010, doi: 10.1016/j.eswa.2009.06.104.
[28] L. Li, R.-L. Ge, S.-M. Zhou, and R. Valerdi, ''Guest editorial integrated healthcare information systems,'' *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 515–517, Jul. 2012, doi: 10.1109/TITB.2012.2198317.
[29] S. M. Zhou and L. Da Xu, ''A new type of recurrent fuzzy neural network for modeling dynamic systems,'' *Knowl.-Based Syst.*, vol. 14, nos. 5–6, pp. 243–251, 2001, doi: 10.1016/S0950-7051(01)00102-2.
[30] Y. H. Yin, Y. J. Fan, and L. D. Xu, ''EMG and EPP-integrated human–machine interface between the paralyzed and rehabilitation exoskeleton,'' *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 4, pp. 542–549, Jul. 2012, doi: 10.1109/TITB.2011.2178034.
[31] R. Rathipriya, A. A. Abdul Rahman, S. Dhamodharavadhani, A. Meero, and G. Yoganandan, ''Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model,'' *Neural Comput. Appl.*, vol. 35, no. 2, pp. 1945–1957, Jan. 2023, doi: 10.1007/s00521-022-07889-9.
[32] S. Ai, A. Chakravorty, and C. Rong, ''Household power demand prediction using evolutionary ensemble neural network pool with multiple network structures,'' *Sensors*, vol. 19, no. 3, p. 721, Feb. 2019, doi: 10.3390/s19030721.
[33] W. W. Y. Ng, S. Xu, T. Wang, S. Zhang, and C. Nugent, ''Radial basis function neural network with localized stochastic-sensitive autoencoder for home-based activity recognition,'' *Sensors*, vol. 20, no. 5, p. 1479, Mar. 2020, doi: 10.3390/s20051479.
[34] Z. Chao and H.-J. Kim, ''Removal of computed tomography ring artifacts via radial basis function artificial neural networks,'' *Phys. Med. Biol.*, vol. 64, no. 23, Dec. 2019, Art. no. 235015, doi: 10.1088/1361-6560/ab5035.
[35] A. Amini and A. Amini. (2023). *MIT Introduction to Deep Learning*. [Online]. Available: http://introtodeeplearning.com
[36] G. Zhang, B. E. Patuwo, and M. Y. Hu, ''Forecasting with artificial neural networks: The state of the art,'' *Int. J. Forecasting*, vol. 14, no. 1, pp. 35–62, 1998, doi: 10.1016/S0169-2070(97)00044-7.
[37] N. Morgan and H. Bourlad, ''Generalization and parameter estimation in feedforward nets: Some experiments,'' in *Proc. Adv. Neural Inf. Process. Syst.*, 1990, pp. 630–637. [Online]. Available: https://proceedings.neurips.cc/paper/1989/file/63923f49e5241343aa7acb6a06a751e7-Paper.pdf
[38] G. P. Zhang, ''Time series forecasting using a hybrid ARIMA and neural network model,'' *Neurocomputing*, vol. 50, pp. 159–175, Jan. 2003, doi: 10.1016/S0925-2312(01)00702-0.
[39] R. S. Mariano and D. Preve, ''Statistical tests for multiple forecast comparison,'' *J. Econometrics*, vol. 169, no. 1, pp. 123–130, Jul. 2012, doi: 10.1016/j.jeconom.2012.01.014.
[40] H. Chen, Q. Wan, and Y. Wang, ''Refined Diebold–Mariano test methods for the evaluation of wind power forecasting models,'' *Energies*, vol. 7, no. 7, pp. 4185–4198, Jul. 2014, doi: 10.3390/en7074185.
[41] F. X. Diebold and R. S. Mariano, ''Comparing predictive accuracy,'' *J. Bus. Econ. Statist.*, vol. 13, no. 3, pp. 253–263, Jul. 1995, doi: 10.1080/07350015.1995.10524599.
[42] D. Harvey, S. Leybourne, and P. Newbold, ''Testing the equality of prediction mean squared errors,'' *Int. J. Forecasting*, vol. 13, no. 2, pp. 281–291, Jun. 1997, doi: 10.1016/S0169-2070(96)00719-4.
[43] F. X. Diebold, ''Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold–Mariano tests,'' *J. Bus. Econ. Statist.*, vol. 33, no. 1, p. 1, Jan. 2015, doi: 10.1080/07350015.2014.983236.
[44] F. Busetti and J. Marcucci, ''Comparing forecast accuracy: A Monte Carlo investigation,'' *Int. J. Forecasting*, vol. 29, no. 1, pp. 13–27, Jan. 2013, doi: 10.1016/J.IJFORECAST.2012.04.011.
[45] J. Manjon. (2020). *Econometric Modeler*. MathWorks Inc. [Online]. Available: https://ch.mathworks.com/help/econ/econometric-modeler-overview.html#mw_b77d8b24-a3e7-4574-934f-58ee1a195fea
[46] Mathworks. (2023). *Modeling and Prediction With NARX and Time-Delay Networks*. Accessed: Jan. 28, 2023. [Online]. Available: https://ch.mathworks.com/help/deeplearning/modeling-and-prediction-with-narx-and-time-delay-networks.html

ABDULLAH received the bachelor's and master's degrees in computer science from Gadjah Mada University, Yogyakarta, Indonesia, and the Ph.D. degree in computer science from Universiti Utara Malaysia (UUM), in 2015. He is currently pursuing the bachelor's degree with the Information System Program, Engineering and Computer Science Faculty, Universitas Islam Indragiri, Indonesia. He has been with Universitas Islam Indragiri, since December 2020. He has authored and coauthored research work in several national and international conferences, and some national and international journals. His research interests include data mining, optimization algorithm, multimedia database, and pattern recognition and classification.

RICHARD MANU NANA YAW SARPONG-STREETOR received the Bachelor of Science degree in mathematics from KNUST, Ghana, and the Master of Science degree in applied sciences (mathematics) from Universiti Teknologi PETRONAS, Malaysia, where he is currently pursuing the Ph.D. degree in applied sciences (mathematics) with the Department of Fundamental and Applied Sciences. He was also a Student Accountant with ACCA Global, Ghana. His competencies include mathematical and statistical modeling and finance management. He is familiar with related data science software, such as MATLAB and Python. He has coauthored several published international journals and conference papers.

RAJALINGAM SOKKALINGAM received the M.Sc. degree in industrial technology from UKM and the Ph.D. degree in mathematics from UMS.

He has total of 25 years working experience with 12 years in manufacturing industries in engineering field and another 13 years in education industries as an academician. He has taught from the range of first-year bachelor's degree in mathematics and statistics up to the master's level in risk, project and cost management. He possesses good experience in computer laboratory sessions and very familiar with some related mathematics and statistical software, such as Maple, SPSS, and expert design. Previously, he was with the Department of Electrical and Electronic Engineering, Curtin University, Malaysia, conducting teaching and research in mathematics and statistics. He is currently a Senior Lecturer with the Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS. He has managed to present his research work in several national and international conferences, and published some results at national and international journals.

MAHMOD OTHMAN received the Ph.D. degree in mathematics from Universiti Utara Malaysia, in 2005. He has a total of 23 years of working experience in education industries as an academician, 21 years with Universiti Teknologi MARA and another six years with Universiti Teknologi PETRONAS, Malaysia. He has authored and coauthored research work in several national and international conferences, and some national and international journals. His research interests include fuzzy mathematics, artificial intelligent, optimization, and decision making. He is a Professional Technologists registered with the Malaysia Board of Technologists (MBOT). He is a member of the International Association of Survey Statisticians, Malaysian Mathematical Sciences Society/Ahli Erastian Sains dan Matematik (PERSAMA), and Management Science/Operation Research Society Malaysia, and a HRDF Certified Trainer.

ABDUS SAMAD AZAD received the B.Sc. degree in computer science engineering from International Islamic University Chittagong, and the Master of Science (M.Sc.) degree from the Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, Seri Iskandar, Perak, Malaysia, where he is currently pursuing the Ph.D. degree. He has published a book chapter, three conference papers, one technical article, and a review article in the metaheuristic algorithms and machine learning field. He is also working in the domain of machine learning and has expertise in the research area.

GUNAWAN SYAHRANTAU received the master's degree in agribusiness management from Universitas Islam Riau, Indonesia, in 2012. He is currently a Lecturer with the Agribusiness Study Program, Faculty of Agriculture, Universitas Islam Indragiri, Indonesia. He has been active with Indonesian Agricultural Economic Association Organization, Komda Riau, for the last five years. He has authored and coauthored in several scientific journals. His latest research is "The Role of Farmer Readiness in the Sustainable Palm Oil Industry."

YUSRIWARTI received the bachelor's degree in accounting science from Universitas Bung Hatta, Indonesia, and the master's degree in management science from Universitas Riau, Indonesia, in 2011. She is currently a Lecturer with the Accounting Study Program, Faculty of Economics and Business, Universitas Islam Indragiri, Indonesia. She has also been active with Koppas Kasuma Cooperative for 24 years. She has also conducted several researches in the field of accounting and has authored and coauthored in several scientific journals. Her latest research is "Analysis of Factors Affecting Implementation of Entity Financial Accounting Standards Without Public Accountability (SAK ETAP) in Middle Small Micro Businesses (UMKM) District in Indragiri District Region."

ZAINAL ARIFIN received the bachelor's degree from the Management Study Program, and the master's degree from the Economics Studies Program. He is currently pursuing the Ph.D. degree with the Economics Study Program, Universitas Jambi, Indonesia. He is also a Lecturer with the Faculty of Economics and Business, Universitas Islam Indragiri, Indonesia, where he is also the Dean of the Faculty of Economics and Business. Previously, he was the Secretary of the Islamic Economics Study Program and the Chairperson of the Research and Community Service Institute, Universitas Islam Indragiri.