IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 2, June 2024, pp. 1858~1867
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i2.pp1858-1867

Hybrid approach for vegetable price forecasting in electronic commerce platform

Kar Yan Choong[1], Suhizaz Sudin[1,2], Rafikha Aliana A. Raof[1,2], Rhui Jaan Ong[1]
[1]Faculty of Electronic Engineering and Technology, Universiti Malaysia Perlis, Kampus Pauh Putra, Malaysia
[2]Sports Engineering Research Center, Universiti Malaysia Perlis, Kampus Pauh Putra, Malaysia

Article Info

Article history:

Received Jul 10, 2023
Revised Oct 10, 2023
Accepted Jan 6, 2024

Keywords:

Genetic algorithm
Hybrid model
Neural network
Time series forecasting
Wavelet transform

ABSTRACT

The significance of the agriculture sector in Malaysia is often overlooked, and there is a notable deficiency in the advancement of digitalization within the country's agricultural domain. The integration of a price forecasting model in the platform enables the relevant parties, including farmers, to make informed decisions and plan their crop selection based on projected future prices. In this research, the authors proposed the hybrid approach with the combination of a linear model and a non-linear model in doing the vegetable price forecasting model. The hybrid model combining seasonal autoregressive integrated moving average (SARIMA)-discrete wavelet transform (DWT)-genetic algorithm neural network (GANN), referred to SARIMA-DWT-GANN, was used to forecast monthly vegetable prices in Malaysia. The historical vegetable price data is collected from the federal agricultural marketing authority Malaysia and split into training/test sets for modeling. The performance of the models is evaluated on the accuracy metrics including mean absolute error (MAE), mean absolute percentage error, and root mean square error (RMSE). The forecasted results using the proposed hybrid model are compared to those using the single SARIMA model. In conclusion, the hybrid SARIMA-DWT-GANN model is superior to the individual model, which obtained the smaller MAE and RMSE, and got the forecast accuracy of at least 95%.

Corresponding Author:

Suhizaz Sudin
Faculty of Electronic Engineering and Technology, Universiti Malaysia Perlis
Pauh Putra Campus, 02600 Arau, Perlis, Malaysia
Email: suhizaz@unimap.edu.my

1. INTRODUCTION

Forecasts are estimates of the timing and magnitude of the occurrence of future events while forecasting is the process of determining what will occur in the future by analyzing past and present events [1]. Essentially, it is a tool for making decisions that help the organization's management to deal with the impact of future uncertainty by analyzing historical data and trends. Forecasting is important for decision-making and future planning as it allows businesses to plan their future steps and build budgets that will ideally account for any unforeseen circumstances.

There are two main types of forecasting methods, which are qualitative forecasting and quantitative forecasting [2]. Qualitative forecasting methods, also known as judgmental methods, are subjective in nature and rely on emotions, intuitions, judgments, personal experiences, and opinions of the forecaster. Examples of qualitative forecasting methods include executive opinion, market research, and the Delphi method. On the other hand, quantitative forecasting methods are based on mathematical or quantitative models, eliminating human bias. These methods utilize historical demand data and statistical computations to predict the future.

Time series and causal models are examples of quantitative forecasting methods. In summary, quantitative forecasting emphasizes the analysis of historical data and statistical calculations, while qualitative forecasting relies on subjective judgments and insights [3].

Time series forecasting methods are used when a temporal or time component is associated with the forecast and the corresponding data is called time series. Time series can be defined as a series of observations taken over time. Time series forecasting is the process of analyzing the time series with the aid of statistics and modeling to create forecasts and intelligent strategic decisions, which study and analyze historical data to predict the future. There are three categories of time series forecasting methods, including statistical models, artificial intelligence (AI) models, and hybrid models [4].

Statistical models are mathematical representations of the underlying relationships between variables in a time series, and they are utilized to forecast future values. There are several examples of statistical models for time series forecasting such as Box-Jenkins basic models. On the other hand, AI models are self-correcting and powerful if compared with traditional statistical forecasting methods. For example, artificial neural networks (ANN) and support vector machines (SVM). Instead of using single forecasting models, the hybrid or combination models can perform better with higher forecast accuracy by combining the advantages of different single models [5], [6].

In Malaysia, fresh produce is commonly obtained through wholesale markets and hypermarkets/supermarkets [7], [8]. Agricultural challenges include production-consumption imbalance, oil palm bias, aging/unskilled farmers, low wages, unfair prices, reliance on imports, supply chain constraints, and market ignorance [9]. There are numerous flaws in the current agricultural marketing system, and it is difficult to address them immediately; therefore, new techniques and trends or even new concepts must be adopted to eliminate these flaws. The authors propose the use of agricultural E-commerce or E-agribusiness to eliminate middlemen and increase farmers' income and Malaysia's food supply, which is also the federal agricultural marketing authority (FAMA) chairman's goal [10].

The Agrobazaar, which is an agricultural E-commerce platform developed by FAMA Malaysia with the funding support of Malaysia's government, is observed as lacking a price forecasting model. The author found that the importance of a price forecasting model in the platform would be helpful for the relevant parties for their decision-making, supply chain management, and even the crop recommendation system [11]. With the idea of developing the agricultural E-commerce platform which supposedly included the price forecasting model, the authors believe that it could meet the requirements of the National Agrofood Policy 2021-2025 (NAP 2.0), which aims to make the agro-food industry more efficient, and the national E-commerce strategic roadmap [12], which aims to speed up the growth of E-commerce in Malaysia.

The organization of the paper continued by section 2 to introduce the related works of price forecasting in the agriculture sector. Section 3 explains the time series forecasting models including Zhang's hybrid model and the proposed hybrid model in this research. Section 4 presents the forecasting results and evaluation of the proposed hybrid model while section 5 concludes the paper.

2. RELATED WORKS

Authors engage in reading and performing literature reviews in journals relevant to their chosen research topic. The purpose is to acquire comprehensive knowledge pertaining to the research field and aid researchers in identifying research problems, thereby providing support for the research process. In related works, Reddy [13] suggested a seasonal autoregressive integrated moving average (SARIMA) model anticipates tomato prices during harvest, which showed that prices fluctuated widely, indicating inadequate tomato econometric model forecasting capacity. Mutwiri [14] used the SARIMA model to analyze tomato price changes in Kenya. He found that the $SARIMA(2,1,1)(1,0,1)_{12}$ model best predicts tomato prices in Nairobi Country, Kenya.

Besides, Vibas and Raqueño [15] analyzed the retail price movement of various fruit and vegetable commodities using three-time series analytic models: Autoregressive integrated moving average (ARIMA), SARIMA, and ARIMAx. For banana and mango monthly retail prices, *ARIMAX(5, 2, 2, x=mango)* and *ARIMAX(2, 2, 1, x=banana)* were the best models. The research results show that the *ARIMAX(3, 2, 1, x=pechay)*, $SARIMA(1\ ,1\ ,1)(1,\ 1,\ 1)_{12}$, and $SARIMA(2,\ 1,\ 1)(2,\ 1,\ 1)_{12}$ were the best models for estimating cabbage prices monthly.

Rathnayake *et al.* [16] used ARIMA and SARIMA models to analyze and forecast wholesale green chili and tomato prices. The best models to predict green chili and tomato wholesale prices were $SARIMA(1,\ 1,\ 1)(2,\ 1,\ 2)_{12}$ and *ARIMA(2, 1, 2)*. The Granger-causation test and vector autoregressive (VAR) technique showed a unidirectional causality between wholesale green chili and tomato prices. Moreover, Sabu and Kumar [17] used three models, which are SARIMA, Holt-Winter's seasonal method, and long short-term memory (LSTM) neural networks in forecasting the monthly prices of arecanut in Kerala. They conclude that the LSTM neural network model was found to be the best model that fits that data, in which LSTM has the smallest root mean square error (RMSE) value compared to the other two models.

Paul *et al.* [18] carried out an empirical comparison of the predictive accuracies of different models (machine learning models) with the usual stochastic model for forecasting wholesale price of Brinjal in seventeen major markets of Odisha, India by means of model confidence set (MCS), and other accuracy measures such as mean error (ME), RMSE, mean absolute error (MAE), and mean absolute percentage error (MAPE). The best performer among the machine learning techniques is the general regression neural network (GRNN). As the literature review mentioned above, many researchers implemented price forecasting models not only with the use of traditional statistical models, but also machine learning models. In this research, the authors aim for the utilization of both the statistical models and machine learning models in implementing a price forecasting model in an agricultural E-commerce platform, which would fulfill the policies of Malaysia's government such as NAP 2.0.

3. MODELS

This research mainly aims to implement the hybrid model with combinations of traditional statistical forecasting methods with ANN to forecast future vegetable prices and then it could be displayed on an agricultural E-commerce platform. Specifically, the hybrid model is according to the theory of Zhang's hybrid model and utilizes with combination of SARIMA and ANN with discrete wavelet transform (DWT) in handling the possible overfitting events of ANN.

3.1. Zhang's hybrid model

In the realm of time series forecasting, Zhang introduced a notable approach by proposing an ARIMA-ANN hybrid model [13]. This innovative technique operates under the assumption that the time series data, denoted as y_t, comprises a combination of both linear (L_t) and non-linear (N_t) components depicted in (1). By synergizing the strengths of ARIMA and ANN, this hybrid model demonstrates enhanced predictive capabilities, capitalizing on the complementary nature of these methodologies. Such integration allows for a more comprehensive and accurate representation of the underlying patterns within the time series data.

$$y_t = L_t + N_t \tag{1}$$

First, ARIMA predicts linearly from time series data. From Zhang's model, it presumes that the linear component residuals have only non-linear relationships (N_t). The residuals e_t of the ARIMA models at time t is calculated with the formula as (2):

$$e_t = y_t - \bar{L}_t \tag{2}$$

while the $\bar{L}_t$ denotes the linear forecast of ARIMA.

In summary, ARIMA predicts linear components while ANN predicts nonlinear components in this strategy. Combining these models improves predicting performance. The experiments result in the Wolf's sunspot data, Canadian lynx data, and British pound/US dollar exchange rate data indicates that this method outperforms ARIMA and ANN methods in forecasting. The Zhang's hybrid model can also be known as the additive hybrid method. To forecast the monthly retail and wholesale price of tomato, onion, and potato (TOP) in India, Purohit *et al.* [19] use statistical models, machine learning models, and also hybrid models and compare the results with the accuracy metrics including MAE, symmetric mean absolute percentage error (SMAPE), and RMSE. They claim that the hybrid methods provide better results than the individual models in predicting crop prices.

Sanjeev and Bhardwaj [20] also proposed the hybrid approach with ARIMA and ANN in forecasting the price of Yamunanagar and Panipat sugarcane. The results indicate the superiority of the hybrid ARIMA-ANN model over the single ARIMA models in terms of RMSE and MAPE. Similarly, Chi and Chi [21] claim that using the hybrid method with the combination of the ARIMA and non-linear autoregressive neural network (NARNN) is good at modeling linear and non-linear problems for time series forecasting as the hybrid method has both linear and non-linear capabilities.

In the comparison of several neural network models in doing forecasts, Subhasree and Priya [22] concluded that the genetic algorithm based neural network (GANN) has the highest forecast accuracy compared to that of back propagation neural network (BPNN) and radial basic function (RBF). Another research was conducted by Inthachot and his team to predict the Thai stock price index trend using the combination of genetic algorithm and ANN [23]. They highlighted the outstanding performance of the hybrid approach compared to the single models.

Furthermore, multiple research studies have been conducted to demonstrate the effectiveness of the GANN approach in improving the accuracy of vegetable price forecasting, showcasing its potential for enhancing pricing strategies in the E-commerce platform [24], [25]. Therefore, this research referred the Zhang's hybrid model with the combination of SARIMA and GANN, but with some modifications in handling both of their limitations. The SARIMA model is used to support the seasonality of the time series instead of ARIMA.

3.2. Proposed hybrid model

The overall concept of the proposed hybrid model with the combination of SARIMA and GANN is shown in Figure 1. The time series data used in this research is the 12 years of monthly vegetable price data in Malaysia between the years 2010 to 2021 provided by FAMA Malaysia, which is a total of 144 observations. The first 75% of the time series, which is 108 observations that is from 2010 to 2018, is used as the training set while the remaining 25% (2019 to 2021) which consists of 36 observations is the test set. Firstly, the training set is fitted into the SARIMA modeling to forecast its linear part.

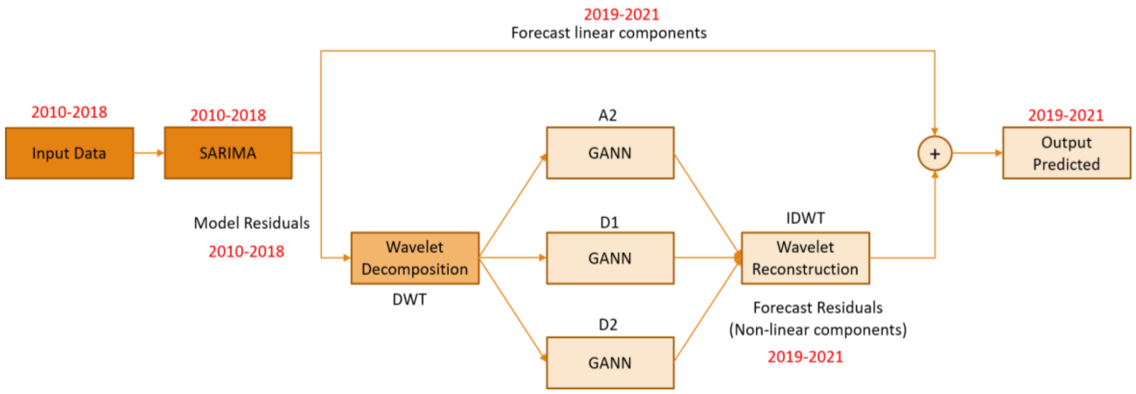


Figure 1. Conceptual diagram of the proposed hybrid method

The SARIMA model is an extension of the ARIMA model that incorporates three additional seasonal factors as shown in Figure 2. These factors include hyper-parameters like those in the non-seasonal components, but they involve seasonal period backshifts. In the SARIMA model, the parameters *p*, *d*, *q* represents the orders of autoregressive *AR(p)*, integration *(d)*, and moving average *MA(q)* respectively, for the non-seasonal component. On the other hand, the parameters *P*, *D*, *Q* represent the corresponding *AR*, *I*, *MA* orders for the seasonal period. The parameter m defines the number of time steps within a single seasonal period.

$$SARIMA\underbrace{(p,d,q)}_{non-seasonal}\underbrace{(P,D,Q)_m}_{seasonal}$$

Figure 2. The notation of SARIMA model

The process of building a suitable SARIMA model includes model identification, parameter estimation, diagnostic checking, and forecasting. During the model identification, the augmented Dickey-Fuller (ADF) test is used to check for trend stationary. The data transformation such as differences needs to be taken for the non-stationary data to become stationary. Using the autocorrelation function (ACF) and partial autocorrelation function (PACF), the data stationary is further investigated and then the number of candidate parameters could be shortlisted.

After shortlisting the target model coefficients, the grid search is conducted to determine the optimal values of the parameter of SARIMA model. The time series is used to train the model coefficients and then fit into the selected model using Akaike information criterion (AIC) approach. The best SARIMA model with the smallest AIC would be chosen. The quality of the best-fitted SARIMA model is evaluated using the Ljung-Box test and through the plots of residuals for model diagnostic checking. The residuals of the best SARIMA model should be uncorrelated and normally distributed. If the residuals are not well-behaved, then the step of parameter estimation needs to be repeated.

After that, the best SARIMA model with the uncorrelated and normally distributed residuals is ready to make the linear forecast. Next step, the DWT is utilized to decompose the SARIMA model residuals into their details and approximations for wavelet decomposition level 2 using the Haar wavelet. The DWT is done to provide better inputs to the GANN to prevent the possible overfitting events that might occur in the neural network. The resulting coefficients included the A_2, D_1, and D_2.

The GANN steps are shown in Figure 3 which consists of 10 basic steps including the initialization of population, fitness evaluation, stopping criterion, selection, crossover, mutation, replacement, and looping. The recurrent neural network with five neurons in three respective hidden layers is developed for modeling the resulting coefficients from DWT in combination with the genetic algorithm. After the adoption of GANN, the forecasted A_2, D_1, and D_2 would be reconstructed to get the final non-linear forecast through inverse discrete wavelet transform (IDWT).

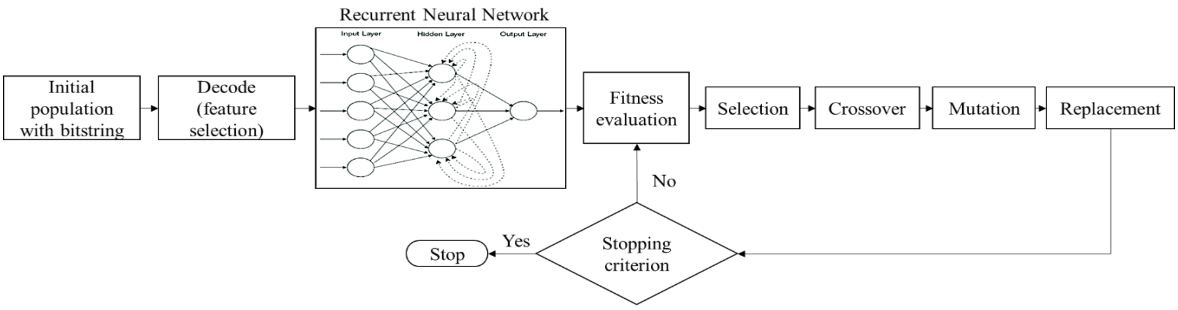


Figure 3. Steps of GANN

Ultimately, the linear forecast produced by the SARIMA model and the non-linear forecast generated by the DWT and GANN models is aggregated to yield the final forecast, which encompasses both the linear and non-linear aspects of the time series according to (3). The resulting $\bar{y}_t$ denotes the ultimate forecast value, encapsulating the interplay of linearity and non-linearity in the data. Additionally, the non-linear forecast of the time series using GANN, denoted as $\bar{N}_t$, is isolated for further analysis.

$$\bar{y}_t = \bar{L}_t + \bar{N}_t \tag{3}$$

To assess the forecast performance of the models, a range of accuracy metrics have been selected. These include the MAE, MAPE, and RMSE, which are defined by the formulas based on the (4)-(6). In this context, N_{es} represents the total number of observations designated for forecasting, while σ_i signifies the relative percentage error associated with each prediction.

$$MAE = \frac{1}{N_{es}}\sum_{i=1}^{N_{es}}|\bar{y}_t - y_t| \tag{4}$$

$$MAPE\% = \frac{1}{N_{es}}\sum_{i=1}^{N_{es}}\sigma_i \tag{5}$$

$$RMSE = \sqrt{\frac{1}{N_{es}}\sum_{i=1}^{N_{es}}\sigma_i^2} \tag{6}$$

4. RESULTS AND DISCUSSION

The proposed hybrid method is applied to the historical monthly vegetable price data between years 2010 and 2021 provided by the FAMA Malaysia. The sample vegetable data chosen in this research are the spinach (*bayam*), round wax gourd (*kundur bulat*), galangal (*lengkuas*), and Holland potatoes (*ubi kentang Holland*). However, only the figures of the *bayam* would be included and explained in detail.

In the beginning, the time series are found to contain the seasonality and annual upward trend after the seasonal decomposition. After the ADF test and the plotting of both the ACF and PACF, the time series are indicated as non-stationary, hence, the data transform needs to be taken to turn the time series into stationary. In this research, the differences are taken for the data stationary. The plot of the stationary time series should be lie between the zero line ideally. The training set from the year 2010 to 2018 which contains 108 observations is used to choose the best SARIMA model using a grid search. The best-fitted SARIMA model with the smallest AIC values is tabulated in Table 1.

For the SARIMA model diagnosis checking, the authors use the Ljung-Box test and the diagnostics plots. The result from the Ljung-Box test indicates that the residuals are white noise, which means the residuals did not include any linear components anymore and that the model retrieved sufficient information from the time series. After the model diagnosis, the best-fitted SARIMA model will be utilized to do linear forecasting for the test set, which is the remaining 36 observations from the overall time series of each selected vegetable. The linear forecasting of the chosen SARIMA model $SARIMA(0, 1, 1)\,(0, 1, 1)_{12}$ for *bayam* is shown in Figure 4.

Table 1. Selected SARIMA models of the sample vegetable time series

Vegetables	Selected SARIMA models
Bayam	SARIMA (0,1,1) $(0,1,1)_{12}$
Kundur bulat	SARIMA (0,1,0) $(2,1,0)_{12}$
Lengkuas	SARIMA (0,1,1) $(2,1,0)_{12}$
Ubi kentang Holland	SARIMA (1,1,0) $(2,1,0)_{12}$

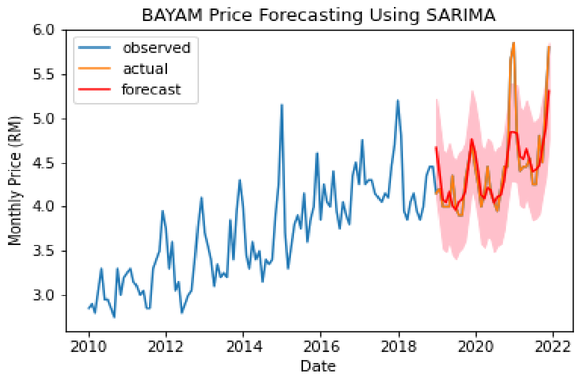


Figure 4. The linear forecast by SARIMA model for *bayam*

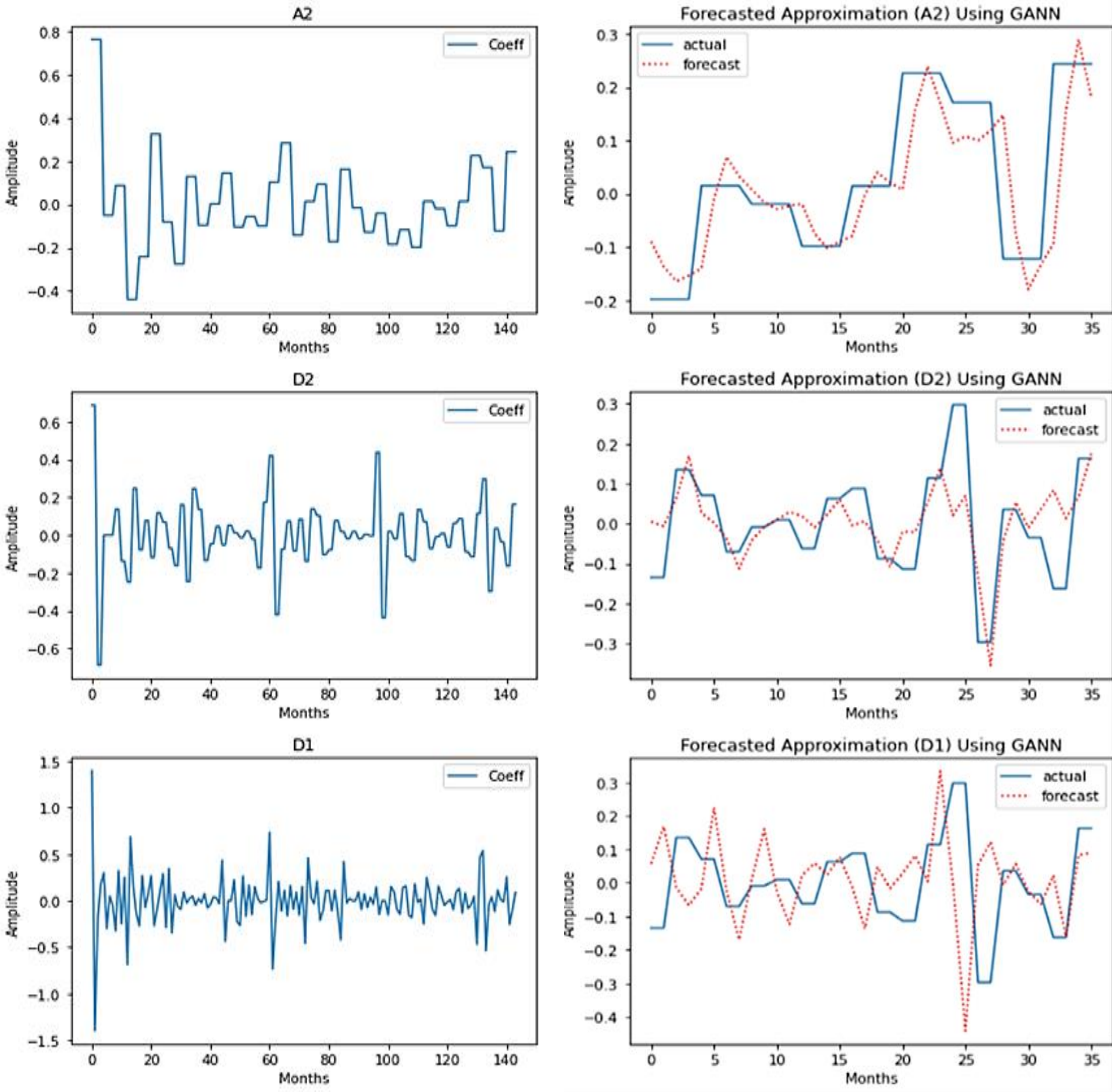
The residuals of the SARIMA model are retrieved and decomposed into the details and approximations using DWT in wavelet decomposition level 2 with the Haar wavelet. Figure 5(a) displays the decomposed coefficients of the *bayam*, including the A_2, D_1, and D_2, while Figure 5(b) illustrates their forecast using the GANN. The forecasted coefficients are subsequently recombined to form the non-linear forecast of GANN with the use of IDWT. The resulting non-linear forecast of the *bayam* time series is shown in Figure 6.

The SARIMA model effectively captures the linear relationship within the time series, while the DWT and GANN models excel at forecasting the non-linear aspects. Performance metrics including MAE, MAPE, and RMSE for both the single SARIMA model and the hybrid SARIMA-DWT-GANN model are detailed in Table 2. A comprehensive comparison of the SARIMA model and the hybrid model for selected vegetables is provided in Figure 7, where Figures 7(a) to (d) focus on *bayam*, *kundur bulat*, *lengkuas*, and *ubi kentang Holland* respectively. The comparison showcases the forecasting accuracy of both models, highlighting the superiority of the hybrid model in capturing the intricate dynamics of these selected vegetables' price trends.

The development of the agricultural E-commerce platform primarily focused on providing a display interface for the forecasted prices of vegetables. However, it is worth noting that this research places an emphasis on enhancing forecasting accuracy, which contributes significantly to the platform's functionality. In order to present the forecasted prices effectively, graphical representations and statistical data are employed, as demonstrated in Figure 8.

Table 2. Accuracy metrics of the SARIMA model and the hybrid model

	SARIMA			SARIMA-GANN		
	MAE	MAPE (%)	RMSE	MAE	MAPE (%)	RMSE
Bayam	0.2003	4.2533	0.2892	0.1471	3.1928	0.2265
Kundur bulat	0.09393	2.6086	0.1097	0.0843	2.3369	0.1067
Lengkuas	0.2219	3.1496	0.2669	0.1368	1.9632	0.175
Ubi kentang Holland	0.05849	1.679	0.0698	0.04605	1.3234	0.05897



(a) (b)

Figure 5. SARIMA model's residual of *bayam* time series, (a) original details and approximations and (b) forecasted details and approximations

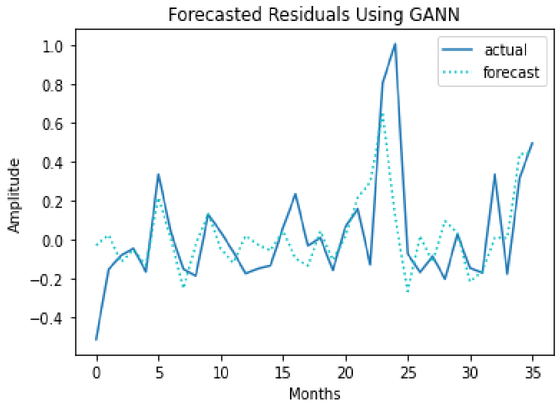


Figure 6. Non-linear forecast of GANN model of *bayam* time series

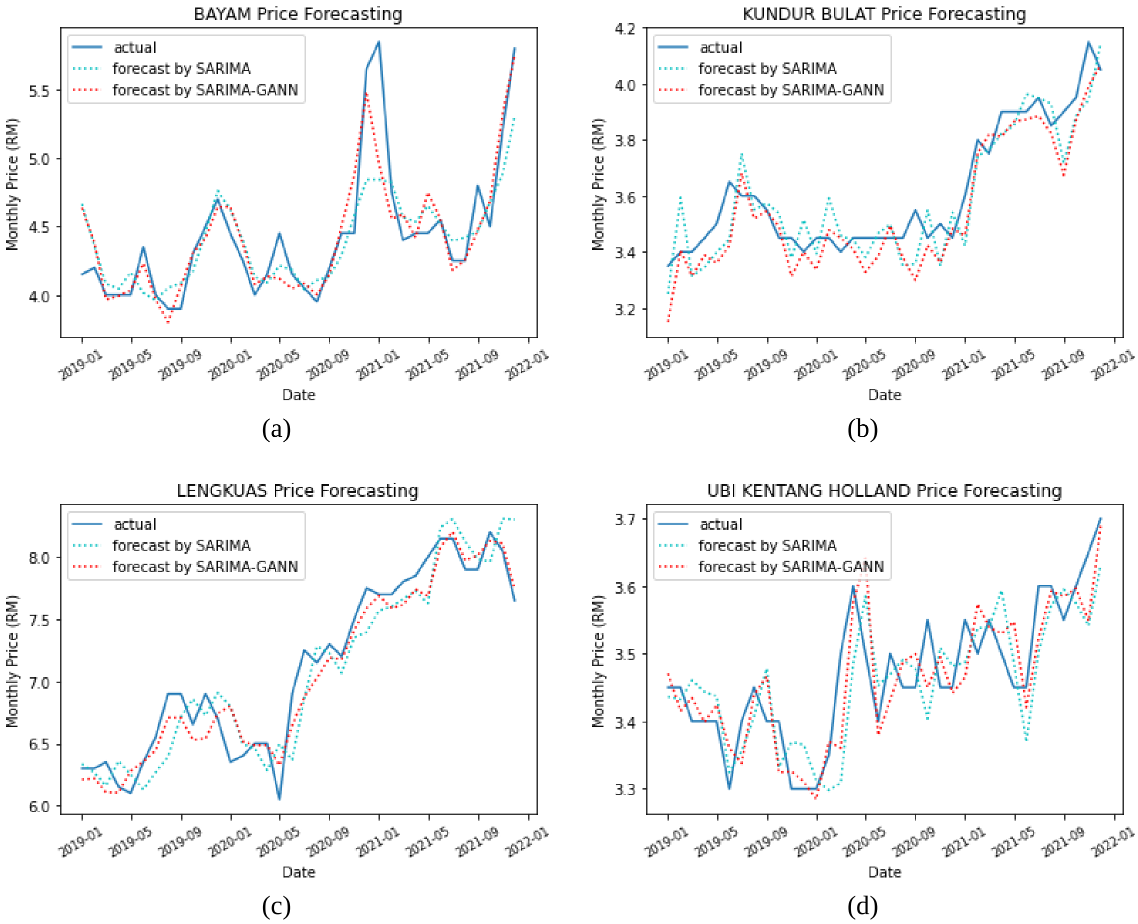


Figure 7. Comparison of the SARIMA and hybrid model of the selected vegetables: (a) *bayam*, (b) *kundur bulat*, (c) *lengkuas*, and (d) *ubi kentang Holland*

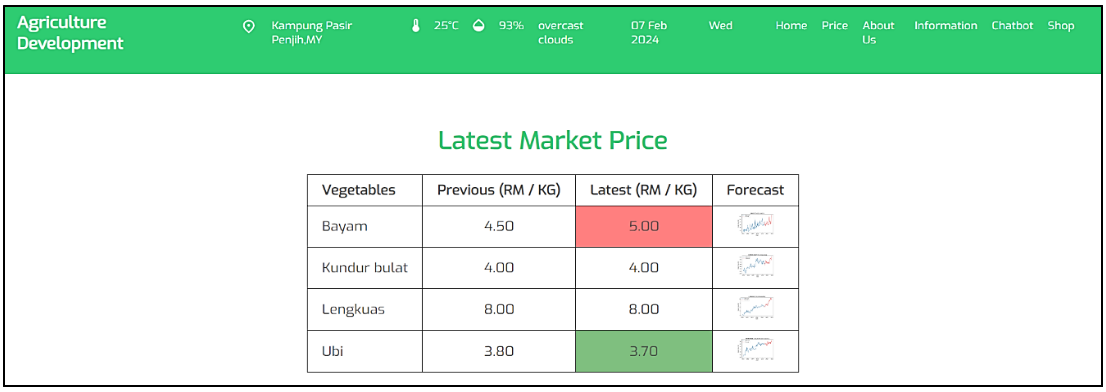


Figure 8. Agricultural E-commerce platform with price forecasting model

5. CONCLUSION

An agricultural E-commerce platform could provide the potential to improve the agriculture sector in Malaysia which would fulfill Malaysia's government policies such as NAP 2.0 and the national E-commerce roadmap. The authors referred the Zhang's hybrid model and developed the SARIMA-DWT-GANN hybrid model in forecasting the vegetable price for *bayam*, *kundur bulat*, *lengkuas*, and *ubi kentang Holland*. The MAPE (%) for the hybrid model for all four chosen vegetables is less than 5%, which means all of them achieved at least 95% forecast accuracy. It even proves that the hybrid model outperformed the use of an individual model. In the future, the authors targeted to utilize the multiplicative hybrid model and compare its performance with the current proposed hybrid method and the other additive hybrid model. Besides, the authors aim to include the other exogenous variables and use multi-step ahead forecasting rather than one-step forecasting to improve the forecast accuracy.

ACKNOWLEDGEMENTS

The authors wish to express their deepest gratitude to the FAMA Malaysia for providing the historical monthly vegetable price data. This research is supported in part by a grant under the Fundamental Research Grant Scheme (FRGS/1/2020/ICT06/UNIMAP/02/1) by the Ministry of Higher Education Malaysia.

REFERENCES

[1] P. J. Brockwell and R. A. Davis, *Introduction to time series and forecasting*. New York: Springer, 2013.
[2] F. Caniato, M. Kalchschmidt, and S. Ronchi, "Integrating quantitative and qualitative forecasting approaches: organizational learning in an action research case," *Journal of the Operational Research Society*, vol. 62, no. 3, pp. 413–424, 2011, doi: 10.1057/jors.2010.142.
[3] R. D. Reid and N. R. Sanders, *Operations management: an integrated approach*, 7th ed. Hoboken, USA: John Wiley & Sons, 2019.
[4] L. Slater *et al.*, "Hybrid forecasting: using statistics and machine learning to integrate predictions from dynamical models," *Hydrology and Earth System Sciences Discussions*, pp. 1–35, 2022, doi: 10.5194/hess-2022-334.
[5] T. Thanavanich, M. Yaibuates, P. Duangtang, and S. Winyangkul, "Hybrid algorithms based on historical accuracy for forecasting particulate matter concentrations," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 4, pp. 1297–1305, 2022, doi: 10.11591/ijai.v11.i4.pp1297-1305.
[6] C. K. Chin, D. A. B. A. Mat, and A. Y. Saleh, "Hybrid of convolutional neural network algorithm and autoregressive integrated moving average model for skin cancer classification among Malaysian," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 3, pp. 707–716, 2021, doi: 10.11591/ijai.v10.i3.pp707-716.
[7] A. D. Rozhan, Z. M. Hafizuddin, and A. H. Noorlidawati, "Marketing channels for fresh vegetables and fruits in Malaysia," *Acta Horticulturae*, vol. 1012, no. 1012, pp. 1371–1377, Nov. 2013, doi: 10.17660/ActaHortic.2013.1012.185.
[8] I. F. M. Yunus, K. Supian, A. Ahmad, S. H. S. Abdullah, and N. Sabli, "Consumers' perception of agricultural market products in Selangor: a preliminary study," *Proceedings of the International Conference on Sustainable Practices, Development and Urbanisation (IConsPADU 2021), 16 November 2021, Universiti Selangor (UNISEL), Malaysia*, vol. 3, pp. 141–154, 2022, doi: 10.15405/epms.2022.10.14.
[9] A. Z. Abideen, V. P. K. Sundram, and S. Sorooshian, "Scope for sustainable development of small holder farmers in the palm oil supply chain - a systematic literature review and thematic scientific mapping," *Logistics*, vol. 7, no. 1, pp. 1-24, 2023, doi: 10.3390/logistics7010006.
[10] "Raising income of farmers tops FAMA chairman's agenda," *Sun Media Corporation*, 2019. [Online]. Available: https://thesun.my/business/raising-income-of-farmers-tops-fama-chairman-s-agenda-FD984520. Access date: 5 July 2023.
[11] N. Singh and S. R. Mohanty, "A review of price forecasting problem and techniques in deregulated electricity markets," *Journal of Power and Energy Engineering*, vol. 03, no. 09, pp. 1–19, 2015, doi: 10.4236/jpee.2015.39001.
[12] ATKearney, "National ecommerce strategic roadmap overview," *Ministry of International Trade and Industry*, pp. 1-26, 2018. [Online]. Available: https://www.miti.gov.my/miti/resources/Gallery_Walk.pdf. Access date: 5 July 2023.
[13] A. A. Reddy, "Price forecasting of tomatoes," *International Journal of Vegetable Science*, vol. 25, no. 2, pp. 176–184, 2019, doi: 10.1080/19315260.2018.1495674.
[14] R. M. Mutwiri, "Forecasting of tomatoes wholesale prices of Nairobi in Kenya: time series analysis using SARIMA model," *International Journal of Statistical Distributions and Applications*, vol. 5, no. 3, pp. 46–53, 2019, doi: 10.11648/j.ijsd.20190503.11.
[15] V. M. Vibas and A. R. Raqueño, "A mathematical model for estimating retail price movements of basic fruit and vegetable commodities using time series analysis," *International Journal of Advance Study and Research Work*, vol. 2, no. 7, pp. 1-5, 2019, doi: 10.5281/zenodo.3333529.
[16] R. M. Y. L. Rathnayake, A. M. Razmy, and M. C. Alibuhtto, "A comparison of price fluctuations for the green chilli and tomato in Dambulla dedicated," in *5th Annual Science Research Sessions*, pp. 144–156, 2019.
[17] K. M. Sabu and T. K. M. Kumar, "Predictive analytics in agriculture: forecasting prices of arecanuts in Kerala," *Procedia Computer Science*, vol. 171, pp. 699–708, 2020, doi: 10.1016/j.procs.2020.04.076.
[18] R. K. Paul *et al.*, "Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India," *PLOS ONE*, vol. 17, no. 7, Jul. 2022, doi: 10.1371/journal.pone.0270553.
[19] S. K. Purohit, S. Panigrahi, P. K. Sethy, and S. K. Behera, "Time series forecasting of price of agricultural products using hybrid methods," *Applied Artificial Intelligence*, vol. 35, no. 15, pp. 1388–1406, 2021, doi: 10.1080/08839514.2021.1981659.
[20] Sanjeev and N. Bhardwaj, "Time series prediction using hybrid ARIMA-ANN models for sugarcane," *International Journal of Plant & Soil Science*, vol. 34, no. 23, pp. 772–782, Oct. 2022, doi: 10.9734/ijpss/2022/v34i232488.
[21] Y. N. Chi and O. Chi, "Time series forecasting global price of bananas using hybrid ARIMA-NARNN model," *Data Science in Finance and Economics*, vol. 2, no. 3, pp. 254–274, 2022, doi: 10.3934/dsfe.2022013.
[22] M. Subhasree and C. A. Priya, "Forecasting vegetable price using time series data," *International Journal of Advanced Research*, vol. 3, pp. 535–541, 2016.
[23] M. Inthachot, V. Boonjing, and S. Intakosum, "Artificial neural network and genetic algorithm hybrid intelligence for predicting Thai stock price index trend," *Computational Intelligence and Neuroscience*, vol. 2016, pp. 1–8, 2016, doi: 10.1155/2016/3045254.
[24] K. Drachal and M. Pawłowski, "A review of the applications of genetic algorithms to forecasting prices of commodities," *Economies*, vol. 9, no. 1, pp. 1-22, 2021, doi: 10.3390/economies9010006.
[25] C. Luo, Q. Wei, L. Zhou, J. Zhang, and S. Sun, "Prediction of vegetable price based on neural network and genetic algorithm," in *Computer and Computing Technologies in Agriculture IV*, Berlin, Heidelberg: Springer, 2011, pp. 672–681, doi: 10.1007/978-3-642-18354-6_79.

BIOGRAPHIES OF AUTHORS

Kar Yan Choong is currently pursuing the Master of Science (M.Sc.) in Computer Engineering in Faculty of Electronic Engineering and Technology, Universiti Malaysia Perlis, Malaysia. She received her Bachelor of Engineering (B.Eng.) in Computer Network Engineering from Universiti Malaysia Perlis in 2019. Her research interests include data science, neural network, and network security. She has been working as a research assistant at Universiti Malaysia Perlis since 2020. She can be contacted at email: karyan@studentmail.unimap.edu.my.

Suhizaz Sudin has an experience of more than 20 years, in the field of computing, networking and system development. He had been working in a computing environment in various industries locally and internationally before joining UniMAP in 2004. He was also appointed as a member of the National Cyber Security Task Force by the Malaysia Science Academy in 2016. Currently, he is the Lead researcher of the Computing Group in the Sports Engineering Research Centre in UniMAP. The area that he is working on at the moment is IoT, big data, and system development in sports. Despite his passion for sports technology, he has a few works in the field of information systems, artificial intelligence, system development techniques & methodologies, and computer security. He can be contacted at email: suhizaz@unimap.edu.my.

Rafikha Aliana A. Raof obtained her first degree in Electronics Engineering from Multimedia University (MMU), Cyberjaya, Malaysia, with honors, majoring in Computer Engineering and graduating in 2002. She received the Masters of Science (M.Sc.) in Intelligent Knowledge-Based System (IKBS) in 2003 from Universiti Utara Malaysia (UUM), Sintok, Malaysia. Her Ph.D. was obtained in 2014 from Universiti Malaysia Perlis (UniMAP), Perlis, with research in microbiological image diagnosis and artificial intelligence. She is currently working as an associate professor at Faculty of Electronic Engineering & Technology (FKTEN), Universiti Malaysia Perlis. Her research interests include medical diagnosis system, intelligent system, fuzzy logic, and embedded system. She can be contacted at email: rafikha@unimap.edu.my.

Rhui Jaan Ong is currently pursuing the Master of Science (M.Sc.) in Computer Engineering in Faculty of Electronic Engineering and Technology, Universiti Malaysia Perlis, Malaysia. He received his Bachelor of Engineering (B.Eng.) in Computer Network Engineering from Universiti Malaysia Perlis in 2019. His research interests include internet of things (IoT), web development, and chatbot development. He has been working as a research assistant at Universiti Malaysia Perlis since 2020. He can be contacted at email: rhuijaan@studentmail.unimap.edu.my.