

Time Series and Machine Learning Hybrid Models for Food Condiment Demand Forecasting: A Case Study in Thailand

Naragain Phumchusri* and Woratanon Sirimak

Department of Industrial Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand
Email: naragain.p@chula.ac.th (N.P.); 6270244721@student.chula.ac.th (W.S.)

*Corresponding author

Manuscript received June 2, 2023; revised August 12, 2023; accepted September 28, 2023; published January 16, 2024

Abstract—Food industry is one of the most important industries in Thailand. The case-study company is a condiment manufacturer that needs to efficiently manage and plan for their business. One of the most important issues is demand forecasting. The company should precisely forecast their product demands, which will be used for operation planning. This study proposes forecasting models for both short-term and long-term planning for the company's main condiment products. The proposed models are time series, machine learning and hybrid forecasting models which will be compared with pure time series and machine learning methods. Unlike previous work, this study proposes an innovative hybrid model, i.e., Holt-Winters exponential smoothing and Seasonal Autoregressive Integrated Moving average hybrid with Artificial neural network, which has never been considered previously. The accuracy is measured by mean absolute percentage error (MAPE) where the results are also compared to the method currently used in the company. The results show hybrid forecasting model provides the lowest overall error for both short-term and long-term forecast. The most accurate model from this paper can provide MAPE of 2.07% from short-term forecast and MAPE of 2.20% for long-term forecast (6 months in advance). When comparing with the company's existing MAPE of 20.05%, the proposed model can increase forecast accuracy effectively.

Keywords—hybrid forecasting models, artificial neural network, food demand forecasting, machine learning

I. INTRODUCTION

The food industry uses processed agricultural products from both plants and animals as the primary raw materials. The production technologies get the products that are convenient to consume or use in the next step. In addition, it can help extend the shelf life of those products by processing them into semi-finished or finished products. In the global food industry, Thailand is promoted as 'Thai Kitchen to the World Kitchen'. It is known as the kitchen of the world due to abundant agricultural resources. There are many agricultural products to be used as raw materials and many talented people working in Thailand's food industry. The value of the industry comes from domestic consumption and exports. Thailand is one of the world's leading food exporters and one of the Association of Southeast Asian Nations (ASEAN)'s top ranks, with a food trade balance of \$21.1 billion in 2021 [1]. The ASEAN food industry is a high-growth potential sector resulting from the expansion of the new generation causing the increased purchasing of various raw materials from other countries, and ASEAN's food cultures that affect the consumption trend. The average growth rate of ASEAN's food market is 28.4% [1].

The Case study company is a leading manufacturer of food

condiment products in Thailand. Its main target is to produce and supply the products to both the domestic and abroad consumer groups. The company plans production using customer demand. Business plans are mostly aimed to reduce costs and maximize profit. Hence, accurate predicting of customer demand is a key for those objectives. Currently, the company uses the 'Sale force composite forecasting' method to predict their product demand. The current mean absolute percentage error (MAPE) values of the 2018 and 2019 sample products are 14.00% and 21.68%, respectively. In 2020, there was a Covid-19 impact to sales, thus the data after 2020 is not appropriate for comparing the forecasting method. The high forecasting error problem affects the company in many aspects: production planning, production scheduling, purchasing raw materials, inventory costs, and other costs. If the product demand forecast provides more accurate results, it will be beneficial to the management team for further production, marketing, and sales strategies planning. The product selection for this study is based on the company's product sales. The three products having major impact on the company's sales are product A, B and C with the sales percentage of 22.55, 20.93, and 7.87, respectively. The sum of these three are more than half of the company's total sales.

There are only a few studies on food industry forecasting since the data is confidential. The disclosure of this confidential information for the research purpose may benefit the competing companies. However, forecasting is an essential factor of the food business, which helps reduce costs and organizational management [2]. In this industry, food can be divided into several categories depending on many factors. One of them is the shelf-life of the long shelf-life food and short shelf-life [3]. Aburto and Weber [4] developed a forecasting model to predict products in the supermarket based on daily data. Single models used in the forecasting are Seasonal Auto-Regressive Integrated Moving Average (SARIMA), Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX), A multilayer perceptron (MLP), and the hybrid model integrating the nonlinear (MLP) and linear model (SARIMA). The results demonstrated that the hybrid model provided the lowest Mean Absolute Percentage Error (MAPE) value compared to other models. As performance measures, metric analysis of the Mean Absolute Percentage Error (MAPE) is used. The HW and Decomposition models was found to obtain better results regarding the performance metrics.

For short shelf-life food, [5] aimed to use forecasting models to predict dairy products: Holt-Winters Exponential Method, Auto Regressive Moving Average (ARMA), and neural network. The neural network method was found to

predict the demand with the lowest error percentage. This neural network model was performed by integrating a genetic algorithm technique and a Radial Basis Function (RBF). Co and Boosarawongse [6] investigated Auto-Regressive Integrated Moving Average (ARIMA), Holt-Winters, and Artificial Neural Network (ANN) model to forecast monthly rice export in Thailand. The study revealed that the ANN model is the most accurate method, compared to the ARIMA and Holt-Winters models. After the hybrid model was developed, forecasting could provide more accurate results. Huang [7] used machine learning algorithm to forecast the food demand of healthy food company in the market. The research used backpropagation neural network developed from combining particle swarm optimization algorithm (PSOBPN). The result came out that PSOBPN gave lower MAPE compared to the method that healthy food company used.

Hybrid forecasting is different from combined forecasting, which is a method that uses a combination of forecast values from a single model. However, hybrid forecasting is implementation techniques or methods combined with forecasting models such as linear and nonlinear combination forecasting models. Data transformation complied with forecasting model and heuristic optimization using improvement forecasting model's parameter. The hybrid method was first proposed by [8], who described a hybridization of ARIMA and MLP. The hybridization is obtained by adding the prediction of the linear model (ARIMA) and nonlinear model (MLP). In supply chain management, demand forecasting is an important factor of this system. Aburto and Weber [4] presented a new intelligent method for the forecasting models, which was the hybridization between SARIMA and neural network. This model showed more accuracy than a single model.

In 2017, [9] improved the accuracy of the hybrid forecasting model which consists of a linear model (SARIMA) and a nonlinear model (ANN). The combination of forecasting models was selected from 13 SARIMA models and 14 ANN models. The machine learning method had recently adopted the XGboost, which integrated with a traditional forecasting model, such as ARIMA. According to the experiment, it was found that the proposed model has the best accuracy [10]. In the food industry, demand for raw materials was a primary factor affecting the production sector in the manufacturer. Zhu et al. [11] proposed the forecasting models: traditional model, machine learning model, and hybrid forecasting model. The result of accuracy showed that the hybrid (HW and SVM) model outperformed the original ones.

From the literature, ARIMA and SARIMA are the most common models for time series forecasting. For Machine learning, ANN has shown to outperform others. The hybrid forecasting model which is a combination of a linear model and a nonlinear model have been considered for the past decade. Unlike previous work, this study also proposes an innovative hybrid model (Holt-Winters exponential smoothing and Seasonal Autoregressive Integrated Moving average, hybrid with Artificial Neural Network), which has never been considered previously. Those models are explored in this paper with a goal to find the most suitable method for the case-study company's products in food industry.

II. METHOD

The proposed models are time series, machine learning and hybrid forecasting models. Details of each method are presented in this section.

A. Holt-Winters Exponential Smoothing Method

Holt-Winters exponential smoothing method is a method of constructing a forecasting equation for time series with trend and influence of the seasonal developed from Holt original method. The Holt-Winters exponential smoothing method performs a prediction equation and three smoothing equations with smoothing parameters α , β , and γ . The seasonal pattern can be divided into two categories: The additive seasonal variation, the variance constant with the time change, and the multiplicative seasonal variation, which fluctuates as the time changes [12].

B. Seasonal Autoregressive Integrated Moving Average (SARIMA)

The Auto-regressive Integrated Moving Average (ARIMA) model is a statistical technique and reasonable model suitable for time series data. The ARIMA model uses the autocorrelation to present the time series data. ARIMA model contains three main parameters (p, d, q): p is the number of autoregressive terms, d is the number of differencing to stationary, and q is the number of moving averages [13]. Seasonality parameters can be applied to the model in order to use it with seasonal data. Therefore, the ARIMA model is adapted to the Seasonal Autoregressive Integrated Moving Average (SARIMA), which added parameters (P, D, Q, s): P is the number autoregressive of seasonal terms, D is the number differencing to stationary of seasonal terms, Q is the number of moving averages of seasonal terms, and s is the seasonal length.

C. Seasonal Autoregressive Integrated Moving Average with Exogenous (SARIMAX)

Forecasting values can normally be explained by the relationship of explanatory variables in a causal method such as linear regression. At the same time, the SARIMA model uses autocorrelations to present the time series data. SARIMAX model combines the SARIMA model and a causal method that added explanatory variables into the model [14]. The parameters are (p, d, q) x (P, D, Q) s X: X is an explanatory variable.

D. Artificial Neural Network (ANN)

The artificial neural network (ANN) is artificial intelligence. It is also considered a mathematical model that attempts to simulate the structure and function of biological neural networks. The structural and functional patterns are similar to the process of living brains, which is to modify themselves to respond the input according to learning principles. The model has three basic principles: multiplication, summation, and activation. Firstly, the neural network model is feed by weighting each input data. The weighted input data is then put together with bias through the transfer function and comes out as the output [15].

E. Hybrid Forecasting Models

The concept of a hybrid forecasting model is a combination of a linear model and a nonlinear model, which can capture

both prediction values from two type models. In the hybrid model, linear is a traditional forecasting model. The residuals are calculated from the linear model, which put into a nonlinear model (a machine learning model). The formula of the hybrid forecasting model is:

$$\hat{Y}_t = L_t + N_t \quad (3)$$

where \hat{Y}_t is hybrid forecast value at period t, L_t forecast value from linear model at period t, and N_t is forecast value from non-linear model at period t [8]. This research focuses on two types of hybrid forecasting models as follows.

1) SARIMA hybrid with ANN model

This hybrid model defines SARIMA as a linear model and ANN as a non-linear model with the following steps.

- 1) Analyzing time series data with the SARIMA model to forecast the data in the linear function (L_t)
- 2) Calculating the residual values of the SARIMA model, where this data will be used for the non-linear function
- 3) In order to forecast the non-linear form in the ANN forecasting process, it is required two methods:
 - a) Using residual values from the SARIMA model as input data for the ANN forecasting (N_t).
 - b) Using residual values from the SARIMA model and adding external variables as input data for the ANN forecasting (N_t).
- 4) Combining the SARIMA forecast values with the ANN forecast values by following this equation:

$$\hat{Y}_t = L_t + N_t$$

- 5) The prediction value of the hybrid model will be taken to evaluate with the test set using MAPE.

2) The proposed model (HW + SARIMA hybrid with ANN model)

This hybrid model defines averaging of Holt-Winters and SARIMA as linear model and ANN as a non-linear model with the following steps.

- 1) Analyze time series data with Holt-Winter model to forecast data in a part of a linear function
- 2) Analyze time series data with SARIMA model to forecast data in a part of a linear function
- 3) Combine forecast values from Holt-Winters and SARIMA using a simple averaging technique, called HW+SARIMA, which is linear function (L_t)
- 4) Calculate the residual values of the HW+SARIMA model, where this data will be used to forecast the non-linear function
- 5) In the ANN forecasting process, in order to forecast the non-linear form, it is divided into two methods:
 - a) Use the residual values from the HW+SARIMA model as input data forecasting us ANN (N_t)
 - b) Use the residual values from the HW+SARIMA model and add external variables as input data forecasting us ANN (N_t)
- 6) Combine the HW+SARIMA forecast values with the ANN forecast values following this equation: $Y_t = L_t + N_t$
- 7) The prediction value's hybrid model will be taken to evaluate with a test set using MAPE. HW+SARIMA hybrid ANN workflow is shown in Fig. 1.

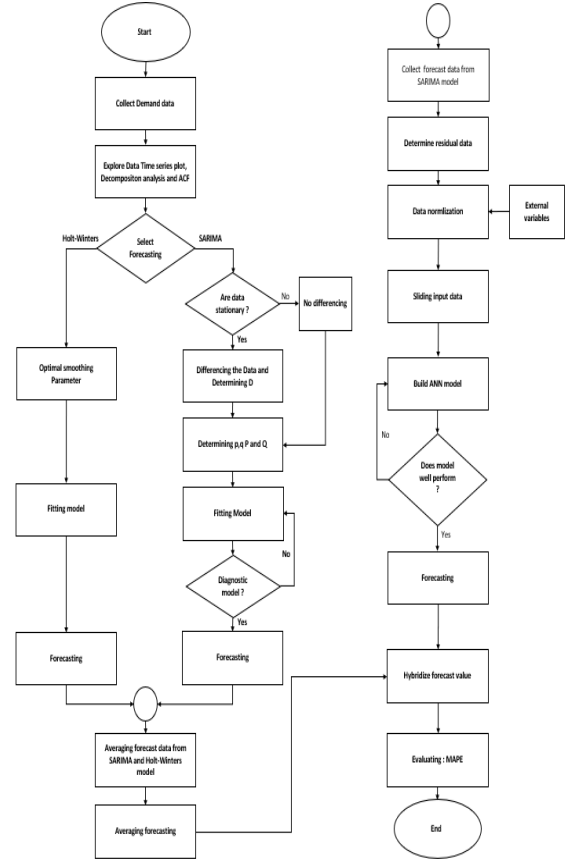


Fig. 1. Summary of the proposed HW+ SARIMA Hybrid with ANN model methodology.

This proposed model is different compared to those in the literature since they considered SARIMA-ANN model [8].

III. DATA AND PREPARATION

The product demand obtained from the case study company is monthly data from January 2013 to December 2019, presented in Fig. 2. As mentioned in Section 1, there had been Covid-19 impact to sales since 2020. So, the data after 2020 is not appropriate for comparing the forecasting method. The data use for forecasting analysis is from product A, B, and C, which are divided into three parts. First, training sets are the demand data from January 2013 to December 2018 (60 months). Second, validation sets are the demand data from January 2019 to June 2019 (6 months). The last one is test sets, which are from the demand data from July 2019 to December 2019.

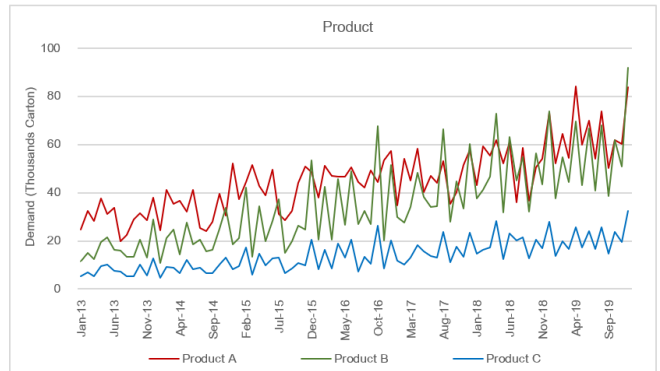


Fig. 2. Time series sales data of each product from January 2013 to December 2019.

A. Selection Categorical Explanatory Variables

Categorical explanatory variables focus on the study of the demand data of the product monthly variables. Categorical explanatory variables are coded with variables or binary numbers. This type of variable is known as dummy variables. One of the seasonal dummy variables is set as number one to serve as equation reference or baseline to prevent redundancy. This selection use product demand data to create descriptive stat using boxplot analysis. Monthly variables are selected from a minimum demand of each product's twelve-month data

B. Stepwise Regression

Data from the explanatory variables are used on the selection process with stepwise regression, a combination of the forward selection and the backward elimination [16]. First, input the dependent variable that is most correlated with the independent variable using forward selection, examine the P-value of the variable influence and eliminate it with the backward elimination [17]. The next step is to enter the dependent variable, which is secondarily correlated with the independent variable using the forward selection, examine the P-value of the parent variable influence, and eliminate it with the backward elimination. That repeatedly until the end of the process is when all remaining explanatory variable influences in the equations are statistically significant [18]. The Minitab is used in the research for stepwise regression. Table 1 demonstrates the explanatory variables chosen by stepwise regression performed with hypothesis testing at alpha 0.05. The explanatory variables select in product A, B, and C are dummy monthly variables and population variables.

Monthly demand data has a high variance. This method selects a month that contains high or low demand that is important and explanatory for information. In product A, the selected months are February, March, April, June, July, October, and December. In product B, the months selected are April, September, and December. Moreover, the selected months are April, June, October, and December in product C. In respect of the population variables, it is the number of populations in Thailand. The positive relationship is explained as the population growth increases that influence increasing food consumption, which product A, B, and C are part of the cooking process. Therefore, the product demands increase accordingly. This procedure can choose necessary and suitable explanatory variables for each demand to perform with a forecasting model.

Table 1. Explanatory Variables Chosen for Each Product

Explanatory	Product A	Product B	Product C
Monthly dummy variables	February, March, April, June, July, September, December	April, October, December	April, June, October, December
Economics indicator	Not chosen	Not chosen	Not chosen
Population	Chosen	Chosen	Chosen
Setindex	Not chosen	Not chosen	Not chosen
R2	79.85%	64.74%	65.09%
R2 (adj)	77.29%	62.64%	62.45%

IV. RESULTS AND DISCUSSION

This section presents the forecasting results using time series models, machine learning models, and hybrid

forecasting model. Evaluation of these prediction models is based on the evaluation of mean absolute percentage error (MAPE) since this is also the main measurement used by the case-study company who provided us demand data. The validation sets of January 2019 to June 2019 (6 months) and the testing set of July 2019 to December 2019 (6 months) present by the MAPE evaluation are used for model comparison.

A. Sale Force Composite Forecasting (Existing Model Used by the Case Study Company)

Sale force composite forecasting is a method that requires forecasts from each department. For example, the sales leaders of each sector will estimate the sales volume. All data are then put together to become the company's total sales forecast. It is a technique used in the case study company at present. The accuracy of forecasting evaluates using the validation set. The evaluated result of the current method shows the MAPE of product A, B, and C, which are 17.76%, 20.48%, and 21.92%, respectively.

B. Holt-Winters Exponential Smoothing Method

Holt-Winters exponential smoothing method can predict data values that include trend and seasonal patterns. The key factor of this method is parameters optimized by R programming. Optimal smoothing parameters (alpha, beta, and gamma) minimize forecast errors. The accuracy of forecasting evaluates using the validation set, which indicates each product using MAPE. The evaluated result of the Holt-Winters exponential smoothing method shows the MAPE of product A, B, and C, which are 10.60%, 6.95%, and 8.71%, respectively.

C. SARIMA and SARIMAX

This section shows the result of the predict value using SARIMA and SARIMAX methods. SARIMA is a time series method that is described using only historical data in an autocorrelation term, while SARIMAX uses both autocorrelation and explanatory variables. The evaluated result of SARIMA shows the MAPE of product A, B, and C, which are 7.70%, 6.51%, and 4.66%, respectively. The evaluated result of SARIMAX shows the MAPE of product A, B, and C, which are 8.49%, 6.08%, and 6.49%, respectively. The results indicated that SARIMAX shows the lowest MAPE on the product B which is 6.08%. On the other hand, the SARIMA shows the lowest MAPE values, which are 7.70% and 4.66% on product A and C, respectively. The explanatory variables in SARIMAX may not properly fit the demand data of product A and C leading to a higher error percentage compared to SARIMA.

D. Artificial Neural Network

In this study, ANN forecasting is explored in different conditions. First, there are two types of forecasting scope: one-step ahead and multi-step ahead, which are short-term and long-term forecast. Also, hyperparameters are varied consisting of the number of hidden layers, number of hidden units, batch size, epochs, and activation functions. All of these conditions are applied with the python programming to minimize predictive errors

The evaluated result of the one-step ahead artificial neural network shows the lowest MAPE of product A, B and C are

5.09% 3.56% and 5.67%, respectively. The multi-step ahead artificial neural network shows the lowest MAPE of product A, B and C at 3.98, 1.48% and 1.19%, respectively.

E. Hybrid FORECASTING model

1) SARIMA-ANN

SARIMA is a linear model and ANN is a non-linear model in this hybridization. Forecasting product demanding using multi-step ahead SARIMA-ANN indicates low accuracy and a slight fluctuation since this technique is relatively less stable when using it for a wide period of time. In addition, the forecast data does not have a clearly defined pattern. The forecasting in the ANN part uses residual, which is in the random pattern, as the input. It is different from one-step ahead forecasting, which gradually shifts the input values one-step at a time. Even if the input data is a residual value, it demonstrates good predictable results

2) Proposed model (HW + SARIMA hybrid with ANN model)

The proposed model is developed from a conventional hybrid forecasting model. An additional part of the developed hybrid forecasting model is the simple averaging of the time series model. In this study, Holt-Winters is selected to combine with SARIMA. The pre-hybrid value of each product is acquired before the process. It is classified as a part of linear model evaluated. When the result of the combination is completed, it determined the residual obtained from the pre-forecasting values, which are the input data in the ANN model. The same process is applied to all three products.

The evaluated result of one-step ahead proposed model in original input shows the MAPE of product A, B, and C, which are 4.33%, 3.03%, and 0.80%, respectively. The original form with all external variable input shows the MAPE of product A, B, and C, which are 5.77%, 3.85%, and 4.83%, respectively. The results of multi-step ahead proposed model using only original input shows the MAPE of product A, B, and C, which is 6.15%, 3.48%, and 3.30%, respectively. The original form with all external variable input shows the MAPE of product A, B, and C, which are 3.92%, 6.39%, and 5.48%, respectively.

3) Forecasting model selection

The accuracy results of validation set from all models are compared in this using MAPE as shown in Table 2 and 3

Table 2. Measurement error of forecasting models for each product

Model	MAPE
Product A	
<i>Company</i>	
Current method	17.76%
<i>Time series model</i>	
SARIMA	7.70%
<i>Machine learning model</i>	
ANN (OG-RRR-OS)	5.09%
ANN (OG-RSR-MS)	3.98%
<i>Hybrid forecasting model</i>	
HW-SARIMA-ANN (OG-RRR-OS)	4.33%
HW-SARIMA-ANN (OGEX-RR-MS)	3.92%
Product B	

Model	MAPE
Product A	
<i>Company</i>	
Current method	20.48%
<i>Time series model</i>	
SARIMAX	6.08%
<i>Machine learning model</i>	
ANN (OGEX-RSR-OS)	1.52%
ANN (OGSW-SSR-MS)	1.42%
<i>Hybrid forecasting model</i>	
SARIMA-ANN (OGEX-RR-OS)	2.27%
HW-SARIMA-ANN (OG-SRR-MS)	3.48%
Product C	
<i>Company</i>	
Current method	21.92%
<i>Time series model</i>	
SARIMA	4.66%
<i>Machine learning model</i>	
ANN (OGEX-RRR-OS)	1.45%
ANN (OGSW-RSR-MS)	1.19%
<i>Hybrid forecasting model</i>	
SARIMA-ANN (OG-RR-OS)	0.36%
SARIMA-ANN (OG-SSR-MS)	1.94%

Table 3. The data testing result from the best model for each product

Model	MAPE	
	Validate set	Test set
Product A		
Current method	17.76%	10.14%
HW-SARIMA-ANN (OG-RR-OS)	4.33%	5.15%
HW-SARIMA-ANN (OGEX-RR-MS)	3.92%	5.90%
Product B		
Current method	20.48%	23.01%
ANN (OGEX-RSR-OS)	1.52%	15.39%
ANN (OGSW-SSR-MS)	1.48%	11.34%
Product C		
Current method	21.92%	20.48%
SARIMA-ANN (OG-RR-OS)	0.36%	14.07%
ANN (OGSW-RSR-MS)	1.19%	16.05%

Note: OS = one-step, MS = multi-step, OG = original data without external variables, OGSW = original data with external variables selected from Stepwise method and OGEX = original data with all external variables.

V. CONCLUSIONS

This research proposed forecasting models to forecast product demand for a case-study company in food industry. The models were divided into three groups: time series model, machine learning model, and hybrid forecasting model. The results initially showed that the multi-step forecasting technique had higher accuracy for product A, B and C. The results of all experimental conditions studied for ANN, which were single forecasting and hybrid forecasting model, showed that the most accurate model for short-term planning

for product A was the proposed model (HW + SARIMA Hybrid with ANN, original data) at 4.33% MAPE.

For short-term forecast, ANN (original data with external variables selected by stepwise method) is most accurate for product B at 1.52% MAPE, The SARIMA-ANN hybrid model (original data) is best for product C at 0.36% MAPE. For long-term forecast (6 months in advance), the most accurate model for product A was the proposed model (HW + SARIMA Hybrid with ANN, original data with all external variables) at 3.92% MAPE. For product B and C, the most accurate model was ANN (original data with external variables selected by stepwise method) at MAPE of 1.48% and 1.19%, respectively. The results showed the proposed hybrid forecasting model provides the lowest overall error for both short-term and long-term forecast. The most accurate model from this paper can provide MAPE of 2.07% from short-term forecast and MAPE of 2.20% for long-term forecast (6 months in advance). When comparing with the company's existing MAPE of 20.05%, the proposed model can effectively increase forecast accuracy.

There are many possibilities for future research extension, additional explanatory variables can be considered to test if they can improve model accuracy. Other hybrid techniques are also interesting for future work. For example, creating pre-hybrid forecast values by another method before building the hybrid forecasting model may be possible for accuracy improvement.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Naragain Phumchusri conducted the research and wrote the paper; Woratanon Sirimak performed computational experiment and wrote the paper; all authors had approved the final version.

REFERENCES

- [1] Food Intelligence Center Thailand. (December 2021). [Online]. Available: http://fic.nfi.or.th/foodindustry_situation_detail.php?smid=2450
- [2] Thailand Board of Investment. (December 2021). [Online]. Available: <https://www.statista.com/statistics/1023244/thailand-gdp-from-accommodation-and-food-services/>

- [3] R. Fildes, S. Ma, and S. Kolassa, "Retail forecasting: Research and practice," *International Journal of Forecasting*, vol. 38, pp. 1283–1318, 2022.
- [4] L. Aburto and R. Weber, "Improved supply chain management based on hybrid demand forecasts," *Applied Soft Computing.*, vol.7, no.1, pp. 136–44, 2007.
- [5] P. Doganis, A. Alexandridis, P. Patrinos, and H. Sarimveis, "Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing," *Journal of Food Engineering*, vol. 75, no. 2, pp. 196–204, 2006.
- [6] H. C. Co and R. Boosarawongse, "Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks," *Computers & Industrial Engineering*, vol. 53, no. 4, pp. 610–627, 2007.
- [7] H.C. Huang, "Construction of a health food demand prediction model using a back propagation neural network," *Advance Journal of Food Science and Technology*, vol. 5, no. 7, pp. 896–899, 2013.
- [8] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [9] H. Moeeni and H. Bonakdari, "Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model," *Stochastic Environmental Research and Risk Assessment*, vol. 31, no. 8, pp. 1997–2010, October 2017.
- [10] P. Li, and J.S. Zhang, "A new hybrid method for China's energy supply security forecasting based on ARIMA and XGBoost," *Energies*, vol. 11, no. 7, p. 1687, 2018.
- [11] Y. Zhu, Y. Zhao, J. Zhang, N. Geng, and D. Huang, "Spring onion seed demand forecasting using a hybrid Holt-Winters and support vector machine model," *PLoS one*, vol. 14, no. 7, e0219889, July 2019.
- [12] R.J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, OTexts, 2018.
- [13] J. Fattah, L. Ezzine, Z. Aman, H. Moussami, and A. Lachhab, "Forecasting of demand using ARIMA model," *International Journal of Engineering Business Management*, vol. 10, 1847979018808673, 2018.
- [14] S. I. Vagropoulos, G. I. Chouliaras, E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting," in *Proc. the IEEE Energy Conference (ENERGYCON)*, IEEE, Leuven, 2016.
- [15] K. Suzuki editor. *Artificial Neural Networks: Methodological Advances and Biomedical Applications. BoD-Books on Demand*, April 2011.
- [16] G. Smith, "Step away from stepwise," *Journal of Big Data*, vol. 5, no. 1, pp. 1–2, December 2018.
- [17] I.M. Ghani, and S. Ahmad, "Stepwise multiple regression method to forecast fish landing," *Procedia-Social and Behavioral Sciences*, vol. 8, pp. 549–554, January 2010.
- [18] M. Fritz, and P.D. Berger, *Improving the User Experience Through Practical Data Analytics: Gain Meaningful Insight and Increase Your Bottom Line*, Morgan Kaufmann, 2015.

Copyright © 2024 by the authors. 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 ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).