Demand forecasting for a food condiment manufacturer in Thailand

Mr. Woratanon Sirimak

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Industrial Engineering
Department of Industrial Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2020

การพยากรณ์ความต้องการสินค้าสำหรับโรงงานเครื่องปรุงรสอาหารแห่งหนึ่งในประเทศไทย

นายวรฐนนท์ ศิริมาก

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมอุตสาหการ ภาควิชาวิศวกรรมอุตสาหการ
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2563

Thesis Title	Demand forecasting for a food condiment manufacturer in Thailand
By	Mr. Woratanon Sirimak
Field of Study	Industrial Engineering
Thesis Advisor	Associate Professor NARAGAIN PHUMCHUSRI, Ph.D.

Accepted by the FACULTY OF ENGINEERING, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

.... Dean of the FACULTY OF ENGINEERING

(Professor SUPOT TEACHAVORASINSKUN, D.Eng.)

THESIS COMMITTEE

.... Chairman

(Associate Professor Daricha Sutivong, Ph.D.)

.... Thesis Advisor

(Associate Professor NARAGAIN PHUMCHUSRI, Ph.D.)

.... Examiner

(NANTACHAI KANTANANTHA, Ph.D.)

.... External Examiner

(Assistant Professor Siravit Swangnop, Ph.D.)

วรฐนนท์ ศิริมาก : การพยากรณ์ความต้องการสินค้าสำหรับโรงงานเครื่องปรุงรสอาหารแห่งหนึ่งในประเทศไทย. (Demand forecasting for a food condiment manufacturer in Thailand) อ.ที่ปรึกษาหลัก : รศ. ดร.นระเกณฑ์ พุ่มชูศรี

อุตสาหกรรมอาหารเป็นหนึ่งในอุตสาหกรรมที่มีความสำคัญของประเทศไทย บริษัทกรณีศึกษาเป็นผู้ผลิตเครื่องปรุงรสอาหารที่มีความจำเป็นต้องจัดการและวางแผนธุรกิจอย่างมีประสิทธิภาพ ประเด็นที่สำคัญประการหนึ่งคือการพยากรณ์อุปสงค์ บริษัทควรที่จะสามารถพยากรณ์อุปสงค์ของผลิตภัณฑ์อย่างแม่นยำเพื่อนำไปใช้ในการวางแผนการดำเนินงานต่างๆ การศึกษานี้เสนอแบบจำลองการพยากรณ์สำหรับพยากรณ์ระยะสั้นและระยะยาวของผลิตภัณฑ์เครื่องปรุงรสอาหาร 3 ผลิตภัณฑ์หลัก โดยข้อมูลที่นำมาศึกษาคืออุปสงค์รายเดือนของผลิตภัณฑ์เครื่องปรุงรสตั้งแต่เดือนมกราคม 2556 ถึง ธันวาคม 2563 แบบจำลองการพยากรณ์ที่นำเสนอได้แก่ วิธีอนุกรมเวลา วิธีการเรียนรู้ของเครื่อง วิธีพยากรณ์ร่วม และ วิธีพยากรณ์แบบผสม ผลความแม่นยำของแบบจำลองการพยากรณ์ที่นำเสนอได้ถูกเปรียบเทียบกันและนำไปเปรียบเทียบกับวิธีการปัจจุบันของบริษัทกรณีศึกษาด้วยการวัดผลค่าเฉลี่ยของร้อยละความผิดพลาดสัมบูรณ์ (MAPE) จากผลการวิจัยพบว่าแบบจำลองโครงข่ายประสาทเทียม (ANN) เป็นแบบจำลองที่ให้ค่าคลาดเคลื่อนโดยรวมต่ำที่สุดสำหรับการพยากรณ์ทั้งระยะสั้นและระยะยาว โดยให้ค่าเฉลี่ย MAPE เท่ากับ 4.44 สำหรับการพยากรณ์ระยะสั้นของผลิตภัณฑ์ และ ค่าเฉลี่ย MAPE เท่ากับ 4.64 สำหรับการพยากรณ์ระยะยาวล่วงหน้า 12 เดือน เมื่อเทียบกับค่าเฉลี่ย MAPE ปัจจุบันของบริษัทกรณีศึกษาอยู่ที่ 18.97 ชี้ให้เห็นว่าแบบจำลองจากงานวิจัยนี้สามารถเพิ่มความแม่นยำได้อย่างมีประสิทธิผล

สาขาวิชา	วิศวกรรมอุตสาหการ	ลายมือชื่อนิสิต ....
ปีการศึกษา	2563	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

6270244721 : MAJOR INDUSTRIAL ENGINEERING
KEYWORD: Forecasting, Machine learning, Hybrid forecasting model, Artificial neural network, Food demand forecasting

Woratanon Sirimak : Demand forecasting for a food condiment manufacturer in Thailand. Advisor: Assoc. Prof. NARAGAIN PHUMCHUSRI, Ph.D.

The food industry is one of the most important industries in Thailand. The case study company is a condiment manufacturer that needs to manage and plan for their business efficiently. 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 three main condiment products. The data used are monthly condiment product demands from January 2013 to December 2020. The forecasting methods explored in the study are time series, machine learning, combined forecasting, and hybrid forecasting models. The accuracy of those models are measured by mean absolute percentage error (MAPE) where the results are also compared to the method currently used in the case-study company. The results show that artificial neural network (ANN) model provides the lowest overall MAPE for both short-term and long-term forecast. ANN model's MAPE from short-term forecast is 4.44, while MAPE from long-term forecast (12 months in advance) is 4.64. When comparing with the company's existing MAPE of 18.97, the proposed model can increase forecast accuracy effectively.

Field of Study: Industrial Engineering		Student's Signature ....
Academic Year: 2020		Advisor's Signature ....

ACKNOWLEDGEMENTS

This thesis was successfully completed with generousness and assistance from Associate Professor Dr. Naragain Phumchusri, my thesis advisor, who gave me the opportunity to be her advisee and provided knowledge, advice, as well as guidance for this thesis. The solution suggests from her immense knowledge and experience has encouraged me to complete this thesis. I would like to express my sincere gratitude to my advisor.

I would like to express my deepest appreciation to my committee Associate Professor Dr. Daricha Sutivong, Dr. Nantachai Kantanantha and Assistant Professor Dr. Siravit Swangnop, who gave advice on the problems in order that the thesis was improved to a more accurate process.

I appreciate the assistance and cooperation of the managing director of the case study company, who provided the data and information that was very necessary for this thesis.

Finally, I would like to express my special thanks to my parents for their encouragement and support for me to study for a master's degree and thank you to all of my friends for their support throughout the duration of this thesis study.

Woratanon Sirimak

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

Chapter 1 Introduction

1.1 Thailand's food industry overview

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 ASEAN's top ranks, with a food trade balance of $ 16.7 billion in 2016 (Thailand Board of Investment, 2019).

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 current average growth rate of ASEAN's food market is 28.4%.

Thailand's food industry had a GDP value of 922,835 million baht, or 5.5% of Thailand's GDP. In 2019, 20.6% of the industry GDP was from this section. There are 128,137 domestic food businesses. A million employees were calculated as 19.7% of industrial employment. Export contraction occurred in the food industry in the previous year due to the economic slowdown of trading partners, which declined Thai food export demand. The overvaluation of the Thai baht was also the cause of the higher exported food price than the competitive countries. According to Thai food exporters in the global market, the most exported products were tuna and cassava, followed by rice, chicken, and sugar, while the condiment product was ranked sixth. Moreover, the percentage of exporters demonstrates that Thailand was ranked eleventh with 2.51% in the global market. The largest share of the worldwide market was China, with 15.0%. Figure 1 showed the ratio of food exporters in the global market (Food Interligence Center Thailand, 2019), (GSB Research, 2019).

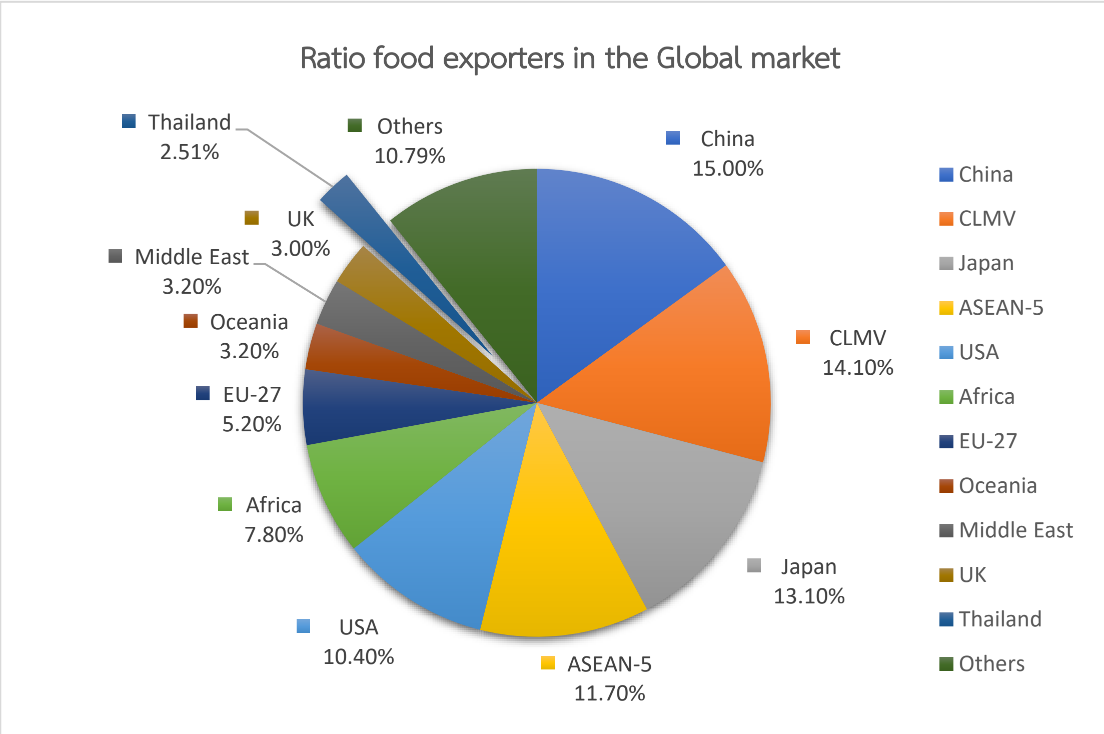

Figure 1: Proportion of exporters in the global market

In 2020, the decline of Thailand's GDP in Q1 and Q2 caused by the crisis of Covid-19 affected the manufacturing sector's activities. The food industry, on the other hand, still was growing when entering Q3. Production was improved due to the cancellation of the lockdown in Thailand and abroad. As a result, most economic activities returned to normal encouraging an increase in domestic consumption. However, the hotel, restaurant, and catering businesses, which were the main distributions, were not recovered that also result in a shrink of the food production sector, especially meat, seafood, and condiments.

Despite the effects of economic problems and crises, the essential role of food manufacturing is producing food to satisfy consumer demands. One of the important parts of the food is flavor enhancer by food condiment products. This secret of deliciousness causes the food condiment business to increase by leaps and bounds every year (Figure 2).

The current value of instant food was increased by 5%. Its sales volume was also increased by 2% in 2020 reaching a record of 48 billion baht and 442,000 tons. The average unit price of condiments rose by 2%. One with the strongest growth was sauces. During the forecast period, the condiments present value was expected to be 6% (4% increased for the fixed value of 2020), and the 2025 sales were expected to reach 60.8 billion baht (Euromonitor International, 2020).

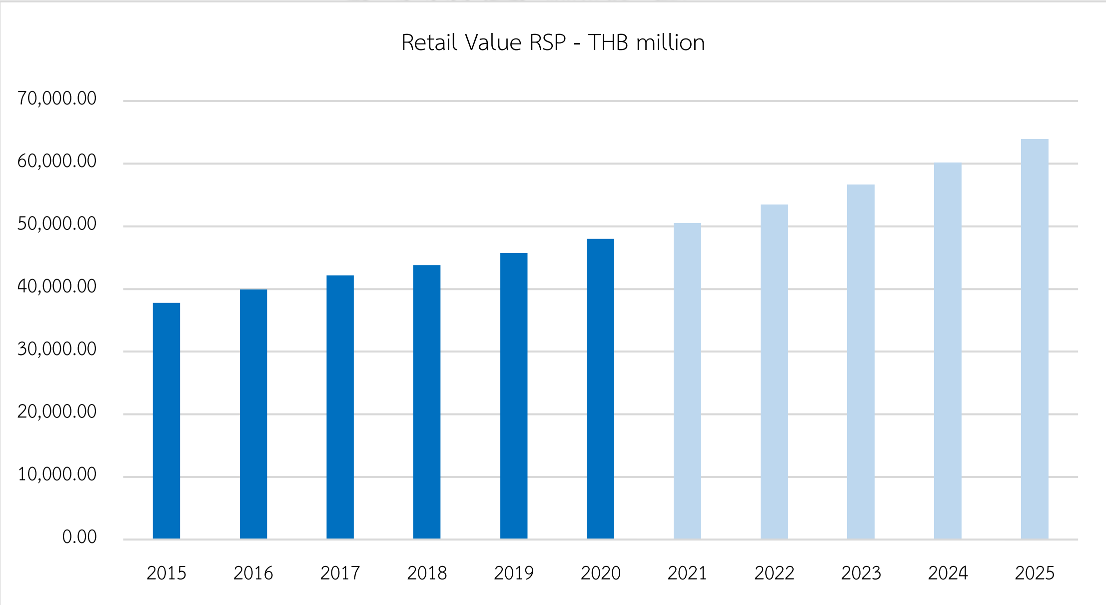

Figure 2: Condiment retail values in Thailand from 2015 to 2025

Source : (Euromonitor International, 2020)

1.2 Company overview

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 case study company focuses on research and product development. As a result, many modern technologies are employed. The quality control system is applied during every step of the production, ingredient selection, packaging, and final production. Every product produced by the company has been certified by the GMP system (Good Manufacturing Practice). The National Food Institute developed this system. It is also guaranteed by HACCP (Hazard Analysis Critical Control Point), which is accepted worldwide.

1.2.1 Sale segments

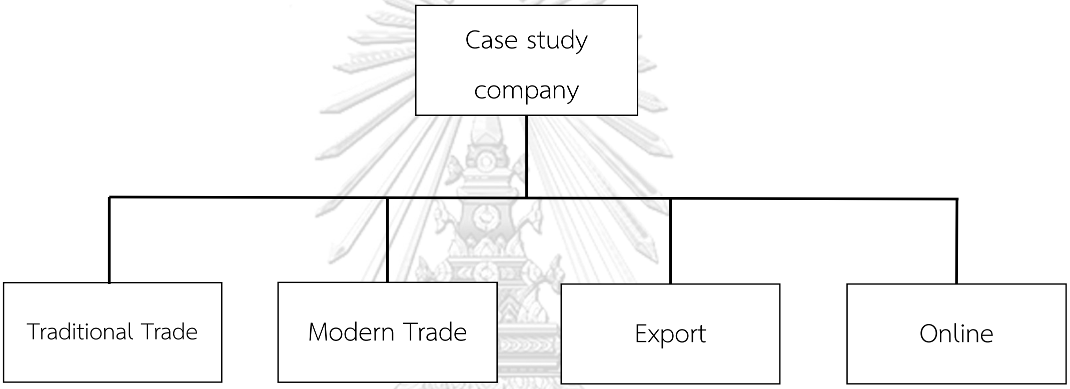

Figure 3: Sale segment in the case study company

The case study company has divided the segment of product sales into four parts (Figure 3).

- Traditional trade: retail shops, fresh markets, and wholesale shops
- Modern trade: convenience store, supermarket, and hypermarket
- Export: Europe, Asia, North America, and Africa
- Online: Lazada and Shopee application

1.2.2 Sale region

The case study company produces products and distributes to all over the country. There are five sales regions (Figure 4): Bangkok, Central-East, North, North-East, and South. The Distribution Center (DC) of the case study company is located in every region to make it easier to distribute and supply the products to each region as follows:

- Manufacturing Plant and Central Distribution Center
- North-East Distribution Center
- South Distribution Center
- North Distribution Center
- Central-East Distribution Center

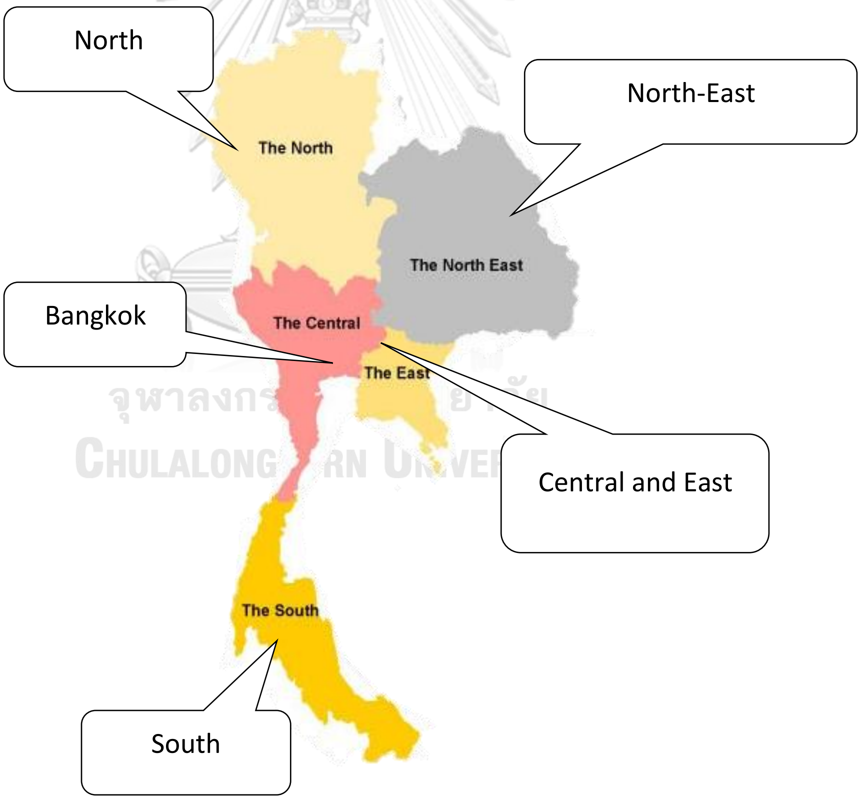

Figure 4: Sale region of the case study company

1.2.3 Demand data flow chart

The flow chart of product demand information starts from customers forwarded to sales and marketing departments to collect and keep it as product demand data. The demand data will be sent to two departments: the planning department and the sales and marketing department.

The planning department will analyze the data and deal with the purchase of raw materials used in production, employee's working schedules, and production planning. The sales and marketing department will analyze the data for further promotion and marketing plans. The product demand flow chart is demonstrated in figure 5.

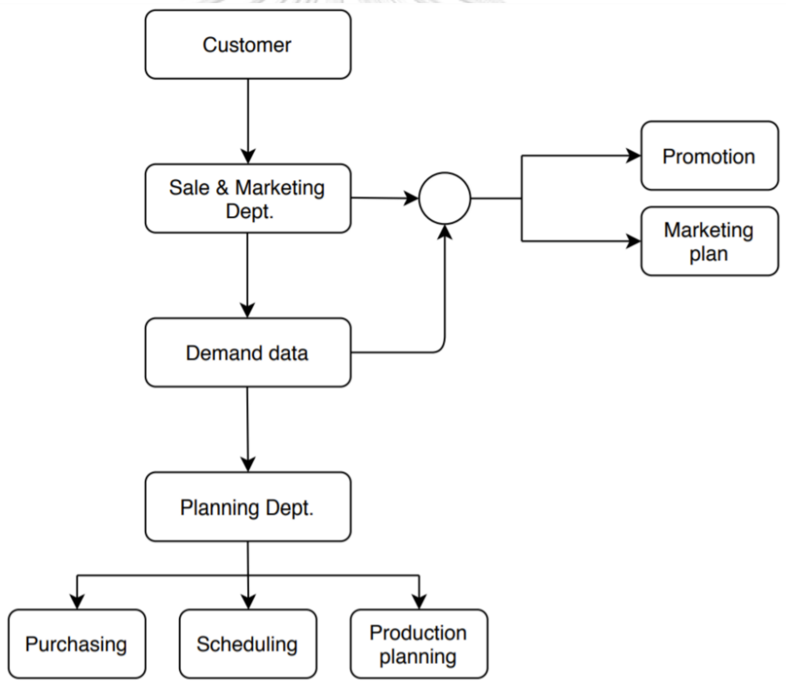

Figure 5: Demand data flow chart of the case study company

1.3 Problem statement

The case study company plans production using customer demand. Business plans are mostly aimed to reduce costs and maximize profit. Hence, accurately predicting customer demand is a key for those objectives.

Currently, the company uses the 'Sale force composite forecasting' method to predict the product demand. This method uses the sales demand, which the sales manager or sales representative staff of each department estimates and combines all the data into a company's total demand forecast. The MAPE values of the 2018 and 2019 sample products are 14.00% and 21.68%, respectively (Table 1). It can be inferred that this method has high errors in forecasting the demand values and may not be suitable enough.

This 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 accurate results, it will be beneficial to the management team to use it for further production, marketing, and sales strategies planning.

Table 1: Measurement error of the case study company's forecasting method in 2018 and 2019

Month	2018			2019		
	Actual	Forecast	Error	Actual	Forecast	Error
January	52,377	61,982	18.34%	37,789	54,300	43.69%
February	64,594	74,047	14.63%	54,653	54,707	0.10%
March	54,361	81,952	50.76%	44,300	56,762	28.13%
…	…	…	…	…	…	…
October	62,177	56,989	8.34%	62054	60371	2.71%
November	60,395	58,269	3.52%	50891	60650	19.18%
December	83,999	65,350	22.20%	91838	74061	19.36%
Evaluation	MAPE		14.00%	MAPE		21.68%

Problem statements indicate the error of the product forecasting. Since the case study company has 63 SKUs, the ABC analysis is applied to analyze the product that will most impact the company. The product selection is based on a high rate of product sales. The results provide in Figure 6 shows product rankings using ABC analysis. The Y-axis represents the proportions of product sales, and the X-axis represents the names of the product.

The three products having major impact on the company's sales, which is more than 50% of the total sales, are product A, B, and C with the sales percentage of 22.55, 20.93, and 7.87, respectively.

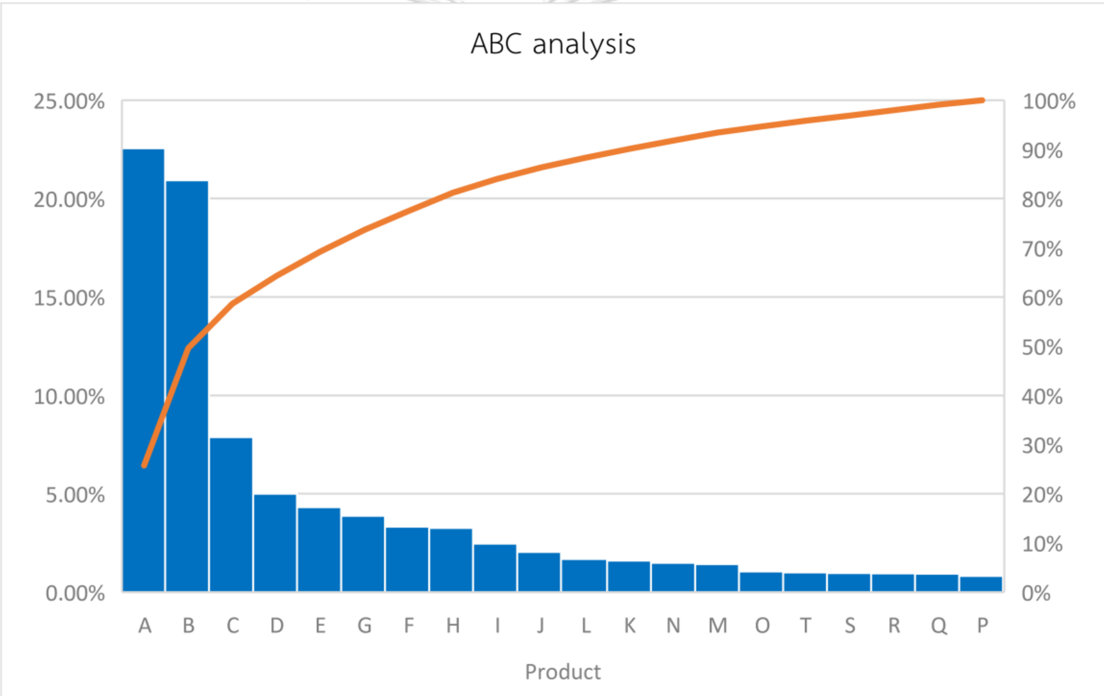

Figure 6: ABC analysis of the case study company

Figure 7 shows the time-series plot of monthly demands of product A, B, and C from January 2013 to December 2019. The result shows that the demand for each product increases every year. Product A has the highest sales percentage with product demand up to 80,000 cartons. Product B also has a high growth rate and high product demand at the same level as product A. However, product C has the lowest demand with 20,000 cartons as the maximum value.

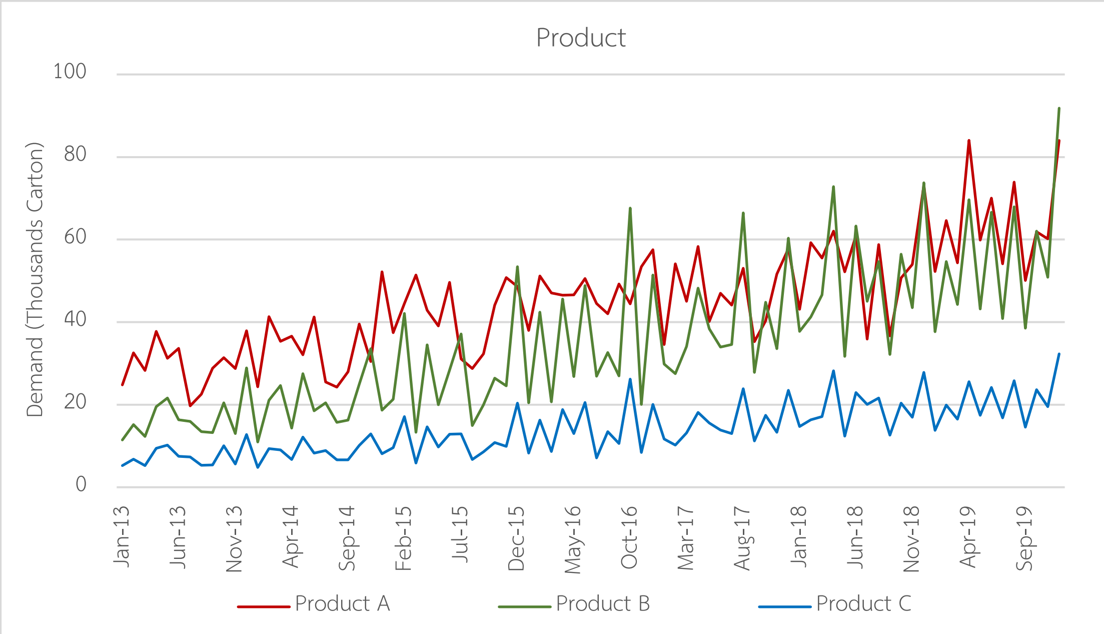

Figure 7: Time-series plot of monthly demand for products from January 2013 to December 2019

The analysis determines the accuracy of the company's forecasting techniques applied to all three products evaluate by MAD and MAPE. The results of the evaluation are presented in table 2.

Table 2: Measurement errors of product A, B, and C

Company method	Product A		Product B		Product C	
	2018	2019	2018	2019	2018	2019
MAD	8,539	8,776	10,122	10,125	4,236	3,695
MAPE	18.28%	14.00%	23.21%	21.69%	22.14%	21.31%

1.4 Research objective

This research aims to improve the forecasting accuracy of food condiment demand for the case study company using time series models, machine learning model, combining forecasting model, and hybrid forecasting model.

1.5 Research scope

- The major products are chosen for this study according to the last-year product sales. The considered products are accounted for 50% of all sales. The product selection criteria are the top three products with a high percentage of total product sales using the ABC analysis.
- The demand data of this thesis are the monthly case study product demands from January 2013 to December 2020, which are divided into three parts:
    - Training datasets are used to perform forecasting models for each product demand from January 2013 to December 2018 (84 months).
    - Validating datasets are used to evaluate forecasting models for each product demand from January 2019 to December 2019 (12 months).
    - Testing datasets are used to evaluate the final forecasting model for each product demand from January 2020 to December 2020 (12 months).
- The considered explanatory variables are Economic Indicator, e.g., Gross domestic product (GDP), Unemployment rate, Consumer Price Index, Exchange rate, Set Index, and the dummy variables for the month.
- This research considers time series models, which are the Holt-Winters method, SARIMA, and SARIMAX, machine learning model which is artificial neural network, combined forecasting model, and hybrid forecasting model.
- The accuracy of the forecasting models is evaluated by measuring the mean absolute percentage Error (MAPE).

1.6 Expected results

- External factors that can be used to describe food product demands.
- Suggested accurate forecasting model that suits each product demand of the case study company.

1.7 Expected benefits

- The data analysis will provide insight information that the company can use to develop their products and for business strategy.
- Accurate forecasting helps the case study company in purchasing, worker's scheduling, and production plans.
- Framework developed in this research can be extended to other products in the future.

Chapter 2 Literature Reviews

2.1 Forecasting review

Forecasting is a statistical method used in the business field as background information for organizational management, such as production schedule, transportation, and long-term strategic planning. However, business forecasting usually gives poor results and easily confuses the business planning and goals (Hyndman & Athanasopoulos, 2018). There are still several problems happening in industrial forecasting. Predictions that are only based on external factors such as advertising campaigns, holidays, weather, and other events may be insufficient. In addition, this data is usually stored separately or may not be completely collected. Current methods tend to give inaccurate results for both short and long-term predictions. Many forecasting techniques require deep knowledge for each specific field. There are still problems such as incorrect time-series data, messy stored data, and methods used for each specific data type in the old forecasting model (Fildes et al., 2019).

2.2 Food Industry forecasting

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 rival companies. However, forecasting is an essential factor of the food business, which helps reduce costs and organizational management (Tsoumakas, 2019).

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 (Kilcast & Subramaniam, 2011). In short shelf-life food, Doganis et al. (2006) was expected to use forecasting models to predict dairy products: Holt-Winters Exponential Method, ARMA, and neural network. The neural network method was approved 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). Furthermore, Da Veiga et al. (2014) also studied the short shelf-life food using Holt-Winters and ARIMA method to estimate monthly data of the dairy product. They found that the Holt-Winters method gave better results compared to the ARIMA. Aburto and Weber (2007) developed a forecasting model to predict products in the supermarket based on daily data. Single models used in the forecasting are SARIMA, SARIMAX, MLP, and the hybrid model, integrating the nonlinear (MLP) and linear model (SARIMA). The results demonstrated that the hybrid model indicates the lowest Mean Absolute Percentage Error (MAPE) value compared to other models. In the same year, Co and Boosarawongse (2007) investigated ARIMA, Holt-Winters, and ANN model to forecast monthly rice export in Thailand. They calculated prediction errors with various measurements. The study revealed that the ANN model worked effectively compared to the ARIMA and Holt-Winters models. After the hybrid model was developed, forecasting could provide more accurate results. Arunraj and Ahrens (2015) studied daily product forecasting by the ANN, SARIMA, and hybrid models. A single model was combined with statistical methods in this study. The statistical method used in the study was Quartile (QR). The combination of the SARIMA model and the QR model was used to compare with other single models. It was found that the SARIMA-QR gave more accurate predictions when measured with MAPE (Table 3).

2.2.1 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 (Hyndman & Athanasopoulos, 2018).

Additive seasonal variation

$$L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1}) \tag{1}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{2}$$

$$S_t = \gamma(Y_t - L_t) + (1-\gamma)S_{t-s} \tag{3}$$

$$\widehat{Y}_{t+p} = L_t + pT_t + S_{t-s+p} \tag{4}$$

Multiplicative seasonal variation

$$L_t = \alpha\left(\frac{Y_t}{S_{t-s}}\right) + (1-\alpha)(L_{t-1} + T_{t-1}) \tag{5}$$

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{6}$$

$$S_t = \gamma\left(\frac{Y_t}{L_t}\right) + (1-\gamma)S_{t-s} \tag{7}$$

$$\widehat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p} \tag{8}$$

$\widehat{Y}_{t+p}$ = forecast value at time t + p

L_t = level smoothing value at t

T_t = trend smoothing value at t

S_t = seasonal smoothing value at t

α = level smoothing constant; $0 \leq \alpha \leq 1$

β = trend smoothing constant; $0 \leq \beta \leq 1$

γ = seasonal smoothing constant; $0 \leq \gamma \leq 1$

p = the period to be forecast into future

s = seasonal period

t = time

2.2.2 Seasonal Autoregressive Integrated Moving average (SARIMA)

The Autoregressive Integrated Moving Average model is a statistical technique and reasonable model suitable for time series data (Box et al., 2011). It is widely used as much as the Holt-Winters model. The ARIMA model uses the autocorrelation to present the time series data (Hyndman & Athanasopoulos, 2018). ARIMA model is a parameter model (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 (Fattah et al., 2018). Seasonality parameters can be applied to the model in order to use it with seasonal data (Arunraj & Ahrens, 2015). Therefore, the ARIMA model is adapted to the Seasonal Autoregressive Integrated Moving Average (SARIMA), which added a parameter. (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. The formula of SARIMA is (Cools et al., 2009) :

$$\phi_p(B)\Phi_p\left(B^S\right)(1\text{-}B)^d(1\text{-}B^s)^D Y_t = c + \theta_q(B)\Theta_Q(B^s)\epsilon_t \tag{9}$$

Y_t = forecast value at time t

$\phi_p(B)$ = regular autoregressive factor

$\theta_q(B)$ = moving average factor

$(1\text{-}B)^d$ = remove non-seasonal differencing

$(1\text{-}B^s)^D$ = seasonal differencing

$\Phi_p\left(B^S\right)$ = the seasonal autoregressive factor

$\Theta_Q(B^s)$ = the seasonal moving average factor

ϵ_t = residuals

c = constant

2.2.3 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 (Vagropoulos et al., 2016). The parameters are (p, d, q) x (P, D, Q) s X: X is an explanatory variable. The formula of SARIMAX is (Cools et al., 2009) :

$$Y_t=\beta_0+\beta_1 X_{1,t}+\ldots+\beta_k X_{k,t}+\frac{\theta_q(B)\Theta_Q(B^s)\epsilon_t}{\phi_p(B)\Phi_p\left(B^S\right)(1-B)^d(1-B^s)^D} \tag{10}$$

$X_{i,\,t}$ = observations of explanatory variables i at time t

β_k = regression coefficients of explanatory variable k

Table 3: Summary of forecasting in food industry

Author	Type of data	Method
Doganis et al. (2006)	Demand of short shelf-life food products in food manufacturer	Holt-Winters ARMA ANN
Aburto and Weber (2007)	Demand of oil in supermarket	SARIMAX ANN Hybrid SARIMAX-ANN
Co and Boosarawongse (2007)	Thailand's rice export	ARIMA Holt-Winters ANN
Da Veiga et al. (2014)	Demand of dairy products in food retail	ARIMA Holt-Winters
Arunraj and Ahrens (2015)	Demand of banana in food retail	ANN SARIMA SARIMA-MLR / QR model
Güzin Tirkeş (2017)	Demand of sherbet and jam in food company	Trend analysis Decomposition model Holt-Winters

2.3 Machine learning forecasting models

Machine learning is a computer application that can make the systems learn automatically, which is to convert input data into algorithms. Machine learning is similar to human learning. For example, when humam receive data, they can learn to identify and analyze that information. Machine learning works the same way: by feeding data sets and instructions into the computer. It is used for information classification or analysis training. Learning algorithms learn from data and training. The algorithms are developed to achieve high efficiency and can be interpreted with accuracy machine learning including supervised, unsupervised, and reinforced learning techniques (Figure 8) (Shalev-Shwartz & Ben-David, 2014).

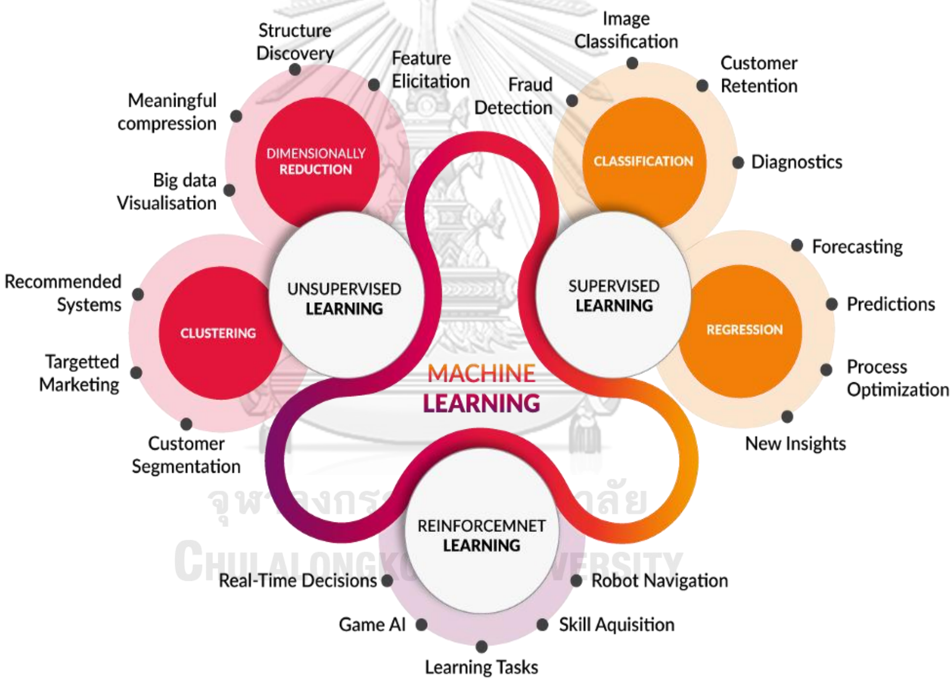

Figure 8: Type of machine learning

Source : (https://towardsdatascience.com)

Supervised learning is one of machine learning systems. It allows computers to find and solve problems by themselves after learning from an input dataset. When the data is put into the model, it will adjust weights through a learning process for efficiency to ensure that the model is suitable. The most famous model for supervised learning is artificial neural network algorithm (Mitra et al., 2016).

Huang (2013) 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. The management of the port system affected the port city. Demir and Akkaş (2018) studied on forecasting models and compared the performance of several forecasting models, including time series forecasting model and machine learning models, such as support vector regression (SVR) and artificial neural network, in order to forecast the demand feed of the case-study company. The comparison of forecasting model performance indicated that support vector regression gave the lowest MAPE. In 2019, Chan et al. (2019) demonstrated a comparison between statistical time series models and machine learning forecasting models i.e., support vector regression and artificial neural network. The accuracy result showed that the machine learning models were more efficient than the time-series models. The prediction of container throughput was important in the port management system. In food industry, the product demand is fluctuated, especially seasoning products. This study focused on the method of sales forecasting which compared between traditional model and machine learning model i.e., artificial neural network. The comparison of accuracy result demonstrated that the neural network is the best forecasting model in the study (Elinta, 2019). Fruits and vegetables are perishable food which is a huge concern for food supply chain. Forecasting model is an important part of product inventory management in retail stores. The study selected time series and machine learning models to predict demanding products. The accuracy result showed that LSTM and SVR were more efficient than the others (Priyadarshi et al., 2019) (Table 4).

2.3.1 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 (Figure 9).

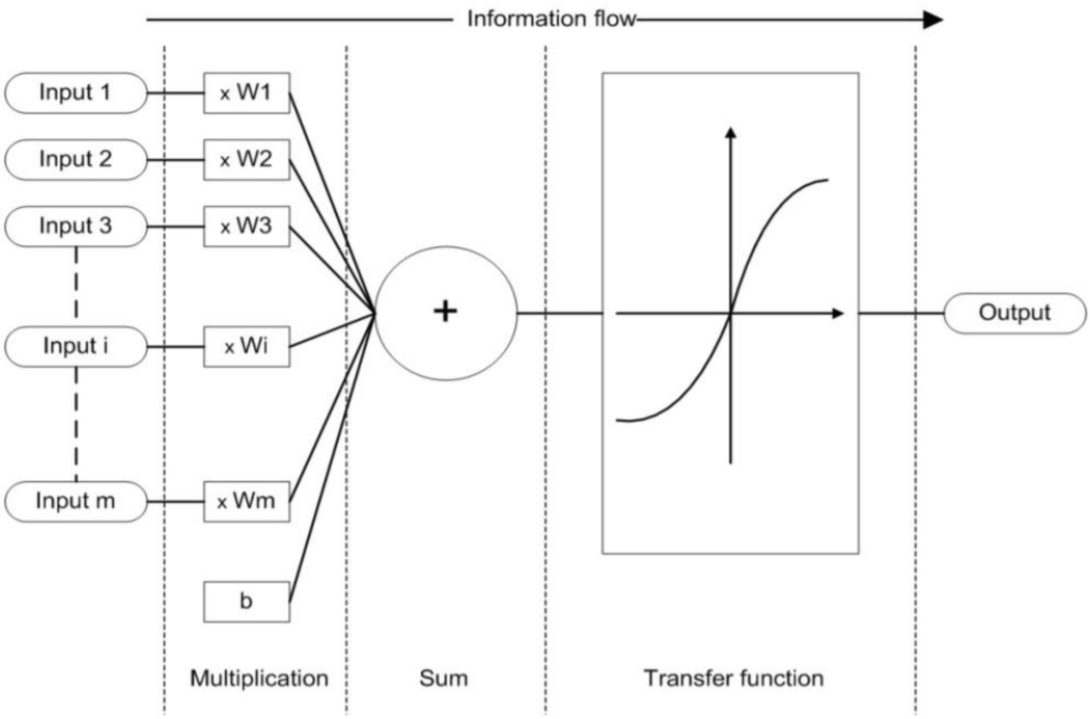

Figure 9: Framework of artificial neural network

Source : (Suzuki, 2011)

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 (Suzuki, 2011). The structure of an artificial neural network processes the data in a sub-processor called a node. A node imitates behaviors from connected node signaling. The function that sets the output signal is called the activation function, which is a process in the neural network. The neural network consists of the following components:

- Number of input nodes

The input number of nodes is equal to that of variables. The input variables will be used to predict forecast values in the future. Input numbers are generally easy to understand and tend to relate to each other.

- Number of hidden layers and hidden nodes

Hidden layers are an important parameter in neural network algorithm. This layer is a middle layer that can affect the model's learning efficiency, which can have many hidden layers. Each layer can have any number of hidden nodes. Additional layers and nodes can affect the learning pattern of the model in the hidden layer.

- Number of output nodes

Output nodes take out the sum of the hidden node values, which is a determiner of the overall values. The number of nodes depends on the output used in the research. There are two types of output forecasts: one-step ahead value and multi-step ahead value. Lijuan and Guohua (2016) studied tourism industry forecast. The forecasting models studied were machine learning models. They predicted the tourist numbers by using these two types of forecasting models. The result showed that the one-step ahead model gives lower MAPE compared to the multi-step. In 2018, the researcher studied electric power forecasting using horizontal time comparison in hybrid forecasting models. The 3-step forecasting gives the highest MAPE values.

- Activation function

The activation function controls the neuron's output in the next range, which calculates the total input sum in one neuron and determines whether it will be forwarded as an output. The hidden layer uses the function to determine the output range that will be interpreted as the answer. The type of activation function using in this study is: (Sharma & Sharma, 2017)

ReLU

The Rectified Linear Activation Function or ReLU, a modified linear function, is the most widely used activation function. This function makes training faster than the Sigmoid. Moreover, it can solve the Vanishing Gradient Problem since the function slope is constant at 1 (Xu et al., 2015). It can be defined in figure 10.

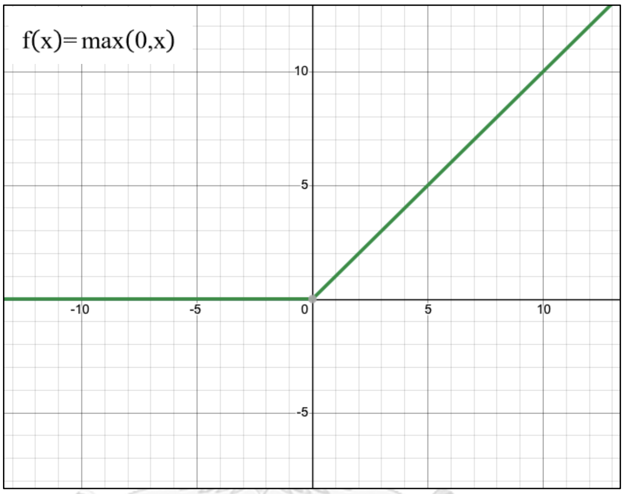

Figure 10: Graph of ReLU function

Source : (https://deeplearninguniversity.com)

Swish

Swish is a function developed by Google researchers in 2017 (Sharma & Sharma, 2017). It is resulted from taking the Sigmoid equation into the calculation and indication, a process similar to ReLU. The team of researchers experimented with images dataset and found that the performance of this function is 0.9% more accurate than the ReLU (Ramachandran et al., 2017). Furthermore, Forex Market Forecasting studied the advantages of several functions in 2019. The results of using Swish in the experiment showed less predictive error than the ReLU (Munkhdalai et al., 2019). It can be defined in figure 11.

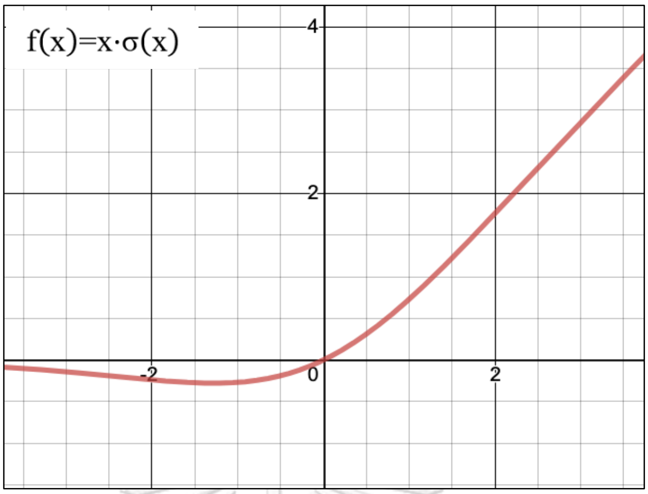

Figure 11: Graph of Swish function

Source : (https://deeplearninguniversity.com)

The mathematical formular of artificial neural network is (Moreno, 2011):

$$b_{pj} = f_L\left(\theta_j + \sum_{i=1}^{N} w_{ij} \cdot x_{pi}\right) \tag{11}$$

b_{pj} = Output value

f_L = activation function of layer L

w_{ij} = weight of neuron i to Hidden neuron j

x_{pi} = Input value

θ_j = bias

$$\hat{Y}_{pk} = f_M\left(\theta_k + \sum_{J=1}^{L} v_{jk} \cdot b_{pj}\right) \tag{12}$$

$\hat{Y}_{pk}$ = Output value by neuron k

f_M = activation function of layer M

v_{jk} = weight of neuron j to Hidden neuron k

θ_k = bias

Table 4: Summary of machine learning forecasting models

Author	Type of data	Method
Huang (2013)	Demand of health food product in food company	Company method Artificial neural network
Demir and Akkaş (2018)	Demand of feed product in feed company	Moving Average Holt-Winters ANN SVR
Chan et al. (2019)	Demand of container throughput in port	Moving Average ARIMA SVR ANN
Elinta (2019)	Demand of seasoning product in food ingredients manufacturer	Trend regression Holt-Winter's ANN
Priyadarshi et al. (2019)	Demand of vegetable product in retail stage	ARIMA RFR XGboost LSTM SVR

2.4 Combined forecasting models

Combined forecast (two or more methods) is an alternative method to replace an individual method. Maaß et al. (2014) suggested that combined forecasting is a widely accepted and practiced method that provides more accurate forecasted values than a single forecasting model. The most important thing is each single method of combined forecasting must have high accuracy. This combination method is claimed to reduce errors caused by biases in the data.

Zou et al. (2007) proposed ARIMA and artificial neural network, which was used in the food industry to predict wheat price. The forecasting model results were combined by the equal weight method. The evaluated accuracy of all models indicated that the combined forecasting model showed the lowest error compared to all measurement methods. Riansut (2016) studied combined forecasting in rubberwood and furniture export values in Thailand. This model used the least square method combined with the ARIMA and Holt-Winters models. The combined forecasting models gave lower MAPE compared with the single forecasting model. In 2017, a new technique was proposed in the electricity demand forecast. The new forecast method used a simple weight average combined with the single models that provided the best accuracy (Laouafi et al., 2017). In the food industry, product demand forecasting was developed by using a combined method. The new method used the company's averaging predictive models: Exponential smoothing and ARIMA (Silva et al., 2019) (Table 5).

It is combining predicted values from each method. All forecast values are combined using weights combination: $w_{1,}$ $w_{2,}$ ..., w_i to build the combined forecast as follow:

$$\hat{Y}_c = w_1\hat{Y}_1 + w_2\hat{Y}_2 + \ldots + w_i\hat{Y}_i \tag{13}$$

$\hat{Y}_c$ = Combined forecast value

w_i = Weight for model i

$\hat{Y}_i$ = Forecast value in model i

Table 5: Summary of combined forecasting models

Author	Type of data	Method
Zou et al. (2007)	Chinese food grain price in market	Simple averaging method
Maaβ et al. (2014)	Demand of short-lifecycle consumer products	Simple averaging method
Riansut (2016)	Export values of rubber wood and furniture in Thailand	Least square method
Laouafi et al. (2017)	Online electricity demand	Simple averaging method and Inverse error method
Silva et al. (2019)	Demand of meat product in the food Industry	Simple averaging method

2.5 Hybrid forecasting models

Hybrid forecasting is different from combined forecasting, which is a method that uses a combination of forecast values from a single model. The concept of hybrid and combined forecasting is shown in figure 12. 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.

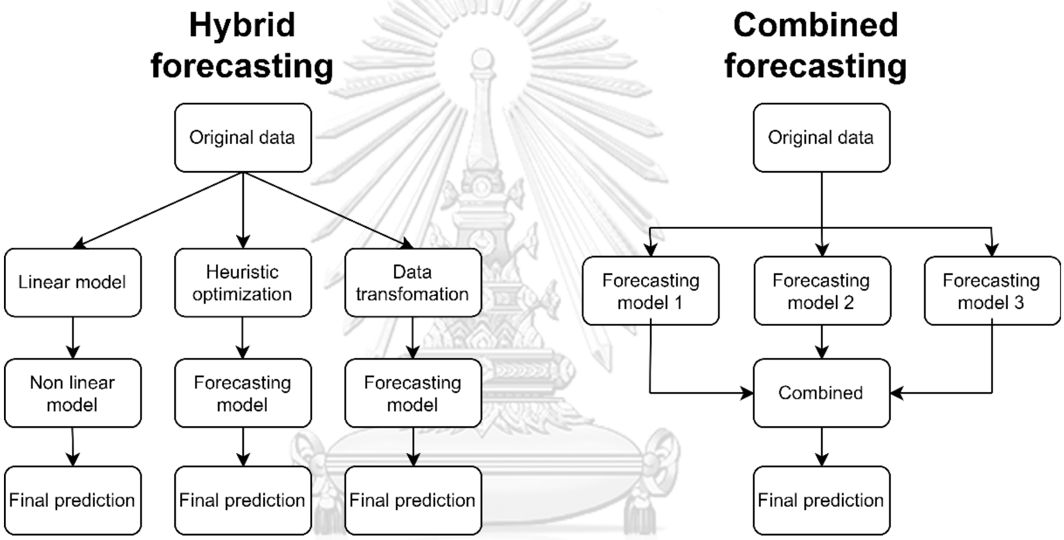

Figure 12: Concept of hybrid and combined forecasting

The hybrid method was first proposed by (Zhang, 2003), 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 (2007) 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, Moeeni and Bonakdari (2017) 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. These combinations were compared with Zhang's model. The result showed that the proposed hybrid forecasting model was more accurate than Zhang's model. The study of hybrid forecasting models was suggested to merge different methods or models. The ML 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 (Li & Zhang, 2018). In the food industry, demand for raw materials was a primary factor affecting the production sector in the manufacturer. Zhu et al. (2019) 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 other models (Table 6).

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:

$$\hat{Y}_t = L_t + N_t \tag{14}$$

$\hat{Y}_t$ = Hybrid forecast value at period t

L_t = Forecast value from linear model at period t

N_t = Forecast value from non-linear model at period t

Table 6: Summary of hybrid forecasting models

Author	Type of data	Method
Zhang (2003)	Sunspot data Exchange rate data Canadian lynx data	Hybrid ARIMA-ANN
Aburto and Weber (2007)	Demand product in supermarket	Hybrid SARIMAX-ANN
Moeeni and Bonakdari (2017)	Monthly inflow to the dam	Hybrid SARIMA-ANN
Li and Zhang (2018)	Demand energy in China	Hybrid ARIMA-XGboost
Zhu et al. (2019)	Demand spring onion seed in seed company	Hybrid Holt-Winters-SVM

2.6 Explanatory variable in demand forecasting

Demand is the ability to purchase a product or service in the quantity that consumers need at that time. Furthermore, demand data is the data that fluctuates with some factors.

As a result, the demand data of the past, present, and future can be found by explanatory variables that represent the movement of the data. In the tourism industry, Athanasopoulos et al. (2011) used explanatory variables to predict the tourism numbers using the income level of origin, which represents the tourist's relative cost of living in the destination country, relative CPI, and exchange rates. Besides, incidents such as terrorism, epidemic, or country-related events, including seasonal events are considered as explanatory variables or are also known as dummy variables. The explanatory variables can improve the industry. For example, textile industries are required to reach the demand that is a variation of trend and seasonality to forecast sales data affected by external factors such as product features, strategies, sales areas, and sales promotions (Thomassey et al., 2002). Domestic air travel industries consider air transport demand as an important factor. Sivrikaya and Tunç (2013) studied factors that affect air travel passengers. Geo-Economics factors, prices, travel time, dummy variables, and other explanatory variables are used in the semi-logarithmic regression model to predict passenger's numbers.

Product demand forecasting or customer service is a key element that impacts the business. Kandananond (2011) studied electricity consumption in Thailand using explanatory variables in the multiple regression model. Population number, stock exchange index, GDP, and export volume are factors used to predict electricity consumption. The study found that those factors can predict electricity consumption, but the method has no significant differences compared to others. Food product demand forecasting is driven by the customers, sales promotions, weather, product characteristics, prices, and holidays. Arunraj and Ahrens (2015) found that explanatory variables can be divided into several types. Two of those are internal factors such as prices and product characteristics, which are controllable and external factors such as events, holidays, and weather conditions, which are uncontrollable also seasonality, which is a day of the week, a month of the year, and a quarter of the year.

Wittayaorn-puripunpinyoo et al. (2017) emphasized the importance of factors affecting ready-to-drink milk consumption in Thailand. Retail milk price, income per capita, population number, and advertising expenditure are explanatory variables of the milk demand prediction study using a multiple regression model. Demand forecasting is a key element in the supply chain that improves performance. Feizabadi (2020) studied forecasting methods to predict the steel demand. ARIMAX and ANN are the demand forecasting methods whose model added economic indicators as explanatory variables. Table 7 summarizes explanatory variables studied for demand forecasting

2.6.1 Economic Indicator

Economic indicators are economic data commonly considered as macroeconomic indexes. It helps indicate the overall state of the present and future economic trends. It is used as an economic overshadow alarm to guide both government and private sector's planning and economic policy. They are used in the analysis and interpretation of investment possibilities (The Economist, 2007).

Moreover, they are used as a tool for fundamental analysis to assess the country's economic condition. There are many different types of economic indicators, including each country's specific one.

- Gross domestic product

The Gross domestic product is the market value of the final products and services from domestic and abroad in a country at a specific time. The GDP consists of products and services manufactured to supply into markets and some Non-Value-Added Activities, such as education. The Gross Domestic Product is used as an indicator of the population's standard of living in that country (Callen, 2020).

- Consumer price index

The Consumer price index (CPI) is a price index used to measure the change in retail prices of consumer goods and services such as shipping, food, and healthcare and calculate each item's price changes that may happen in the future.

The CPI is used to estimate the price changes associated with the cost of living. Moreover, it is one of the most frequently used statistics for determining periods of inflation or deflation. It is also used to measure the real income level of different groups of people. It can measure the changes in product prices sold in each category. Finally, it is used to indicate the country's economic states (Yamarone, 2017).

- Exchange rate

Exchange rates have both positive and negative impacts on the Thai economy. The imbalance of the exchange rate can lead to the country's inability to compete on the price. Furthermore, it can reduce the amount of export income converted into baht and affect worker wages in exporting businesses and the purchasing power causing a slowdown in the economy.

However, this is only one aspect that the exchange rate affects the Thai economy since the healthy exchange rate can reduce the cost of imported raw materials and slow the increase in the living cost, especially during the time when energy prices in the world market rose rapidly (Bank of Thailand, 2018).

- Unemployment rate

The ratio of the unemployment to the total labor force multiplied by 100 indicates the labor market and household income, which is an important variable linked to economic conditions. In addition, this information indicates people's living conditions since their income is from wages and salaries. More purchasing consumption means an increase in the demand (The Economist, 2007).

- Terms of trade

Terms of trade is the export price compared to the price of each country's import. Higher trade means more benefits from international trade. Moreover, if the export rises, prices are likely to stimulate gross demand, which can lead to an increase in investments thanks to profitability increases. Overall consumption is most likely to increase if some of the high domestic demand comes from high domestic products (Gruss & Kebhaj, 2019).

- International reserve

International reserves are liquid assets and regulated by a central bank to settle the debt. It is an international payment to control the balance of payments. Payments are made indirectly through foreign exchange intervention for the country's stability when facing unexpected changes. It is also a tool of economic growth in terms of investment reserves, which can provide credibility to the country. Some export-promoting countries may use reserves as a part of industry policy for the benefit of international trade. By keeping the currency lower than the exported product's base, the products can be more competitive on the price (Vimolchalao, 2003).

2.6.2 Stock market

A stock index is a statistical calculation to track changes in measured things or as a tool to indicate different situations. The stock index reflects all securities price's movement by calculating from the listed shared (The Stock Exchange of Thailand, 2015). A study of electricity use in Thailand (Kandananond, 2011) uses the SET index and other values as the variables to predict the demand. Error shown while testing the model was 0.996%

2.6.3 Population

A population is a group of the same living things, who live together in the same territory at a particular period. Each group's interactions and shared activities create each specific characteristic such as population growth, which affects many aspects, e.g., economic growth. The increase in employment results from an increase in the population (The Economist, 2007). According to the Challenge of Feeding the World, the growth in the population also impacts production changes, especially in the increasing food demand which the production is increased to meet the rising demand (Fróna et al., 2019).

2.6.4 Dummy variable

The dummy variable is characterized as a qualitative variable or categorical explanatory variables. This type of information such as gender, month, year, interested sample, an incident is typically not mathematically meaningful. However, variables can be assigned in terms of numbers or codes. For example, one represents female while zero represents male. This variable is called ‘dummy variable’, which is applied to the linear regression equation for relativity or forecasting (Mirasgedis et al., 2006). Time units, days, and months share explanatory variables in predicting electricity demand. The results show that the predictive value is high when a monthly seasonal dummy variable is added.

Table 7: Summary of explanatory variables studied in demand forecasting

Author	Type of data	Explanatory variables
Kandananond (2011)	Electricity demand in Thailand	GDP Export value Set index Population
Arunraj and Ahrens (2015)	Demand of banana in food retail	Weather Holiday Dummy variable (month of year)
Wittayaorn-puripunpinyoo et al. (2017)	Drink milk consumption of Thailand	Price GDPH Population Advertising expenditure
Feizabadi (2020)	Demand of steel in manufacturing company	Macro-economic indicator

2.7 Measurement error

The evaluated forecasting model is an essential part of forecasting because its error indicates the model's performance. An evaluation of accuracy represents the quality or suitability of the model (Hyndman & Athanasopoulos, 2018). Forecasting in an experiment divides the dataset into two parts: a training set and a test set. A training set is applied to a fitting model to predict the value, which will later be evaluated by a test set (Tsoumakas, 2019), (Klimberg et al., 2010) (Table 8).

2.7.1 Mean absolute deviation: MAD

Mean absolute deviation is a precision measurement technique that measures the accuracy from the forecast and the actual values. The calculation uses the absolute value. The MAD is defined as:

$$\mathrm{MAD} = \frac{1}{n}\sum_{t=1}^{n}\left|Y_t - \hat{Y}_t\right| \tag{15}$$

$\hat{Y}_t$ = Forecast value at the period t

Y_t = Actual data value at period t

n = Number of data

2.7.2 Mean squared error: MSE

Mean squared error (MSE) measures the error from the forecast and the actual value. If MSE value is high, it can be concluded that the data of the model is abnormal. The unit of error value will be different from the forecast value. Root mean squared error (RMSE) is the error rate by the square root of MSE. Therefore, RMSE is similar to MSE, but the unit RMSE is not shown in a square root form to make it easier to analyze as the RMSE unit is the same unit as the forecasting model.

The MSE is defined as:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2 \tag{16}$$

The RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2} \tag{17}$$

2.7.3 Mean absolute percentage error: MAPE

Mean absolute percentage error (MAPE) is a value used to measure the accuracy of the forecasting model based on the forecasted error compared to the actual value. The MAPE unit is in the percentage form. Therefore, it can be used to compare the accuracy of multiple forecasts with different data units. The MAPE is defined as:

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \tag{18}$$

Table 8: Summary of forecasting accuracy measurement in food industry

Author	Topic	Measurement
Doganis et al. (2006)	Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing	MAPE
Co and Boosarawongse (2007)	Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks	MAE MAPE MSE RMSE
Da Veiga et al. (2014)	Demand forecasting in food retail: a comparison between the Holt-Winters and ARIMA models	MAPE
Arunraj and Ahrens (2015)	A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting	MAPE RMSE
Güzin Tirkeş (2017)	Demand forecasting: a comparison between the Holt-Winters, trend analysis, and decomposition models	MAPE

The main contributions of this thesis are as follows

1) Unlike previous work, this study proposes an innovative hybrid model (HW-SARIMA-ANN), which has never been considered before.
2) Most literatures focused on either short-term or long-term forecast at a time, while this thesis explores both aspects together.
3) Activation functions for ANN considered in this thesis are both Swish and ReLU. To the best of our knowledge, Swish activation function had never been studied in food demand forecasting before.

Chapter 3 Methodology

The methodology used in this study is based on the objectives of the thesis. The first is to study the time series model that is suitable for the case study company and could reduce errors. The second is to improve the prediction accuracy of data using forecasting model. The third is to examine and compare the main performance of the forecasting model.

The case study data is prepared for the analysis in this chapter, which is used to predict and study the explanatory variables related to the food product demands that are thoroughly analyzed. Furthermore, this research studies input formation, hyperparameter, and forecasting horizons in machine learning model. The process of forecasting model performance comes out, and results are interpreted.

3.1 Data preparation

The product demand of the case study company is monthly data from January 2013 to December 2020 (Figure 13-15). 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 December 2019 (12 months). The last one is test sets, which are from the demand data from January 2020 to December 2020 (12 months).

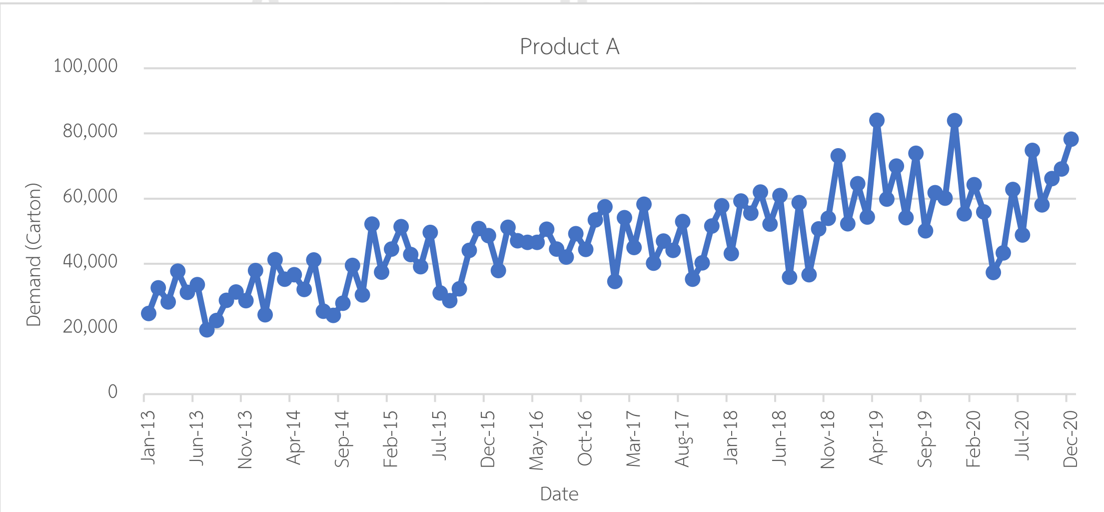

Figure 13: Time series data of product A

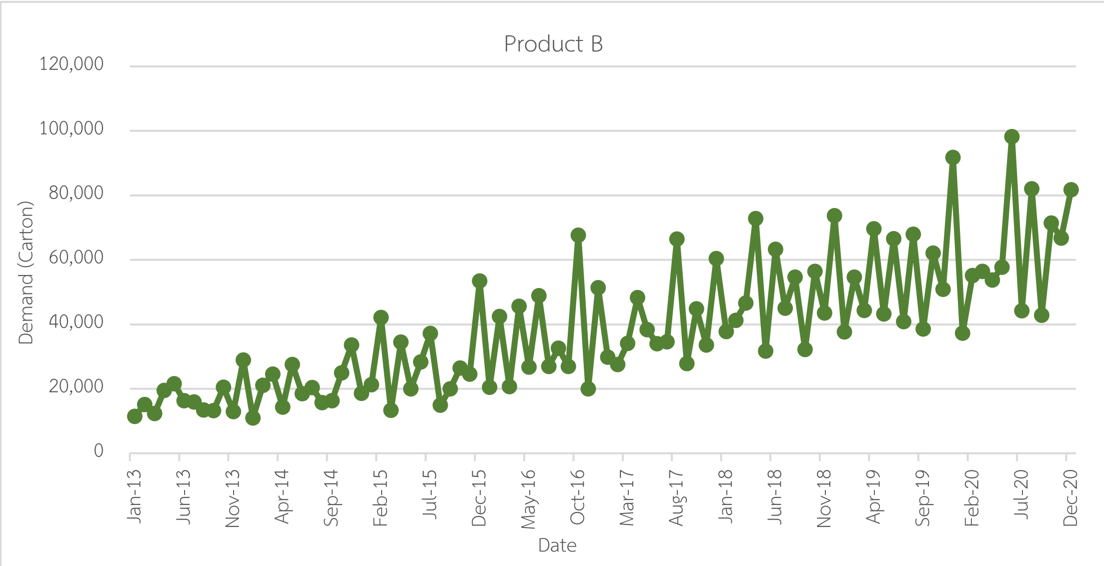

Figure 14: Time series data of product B

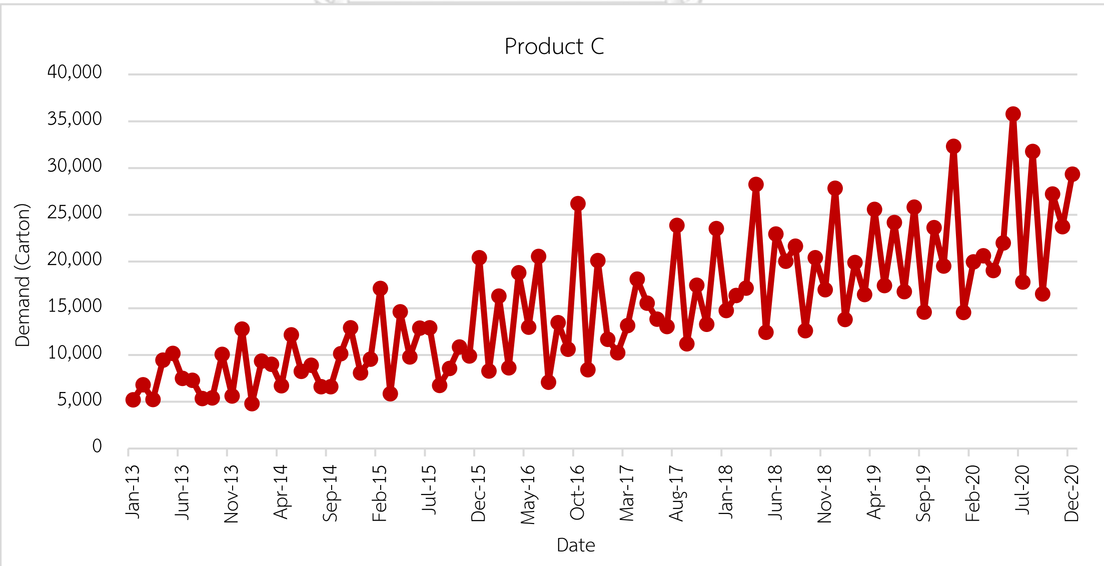

Figure 15: Time series data of product C

3.2 Data analysis

There are three sets of data. The training set is analyzed by the decomposition analysis and separated into three components of the time series data: trend, seasonal, and irregular to analyze the data patterns. Furthermore, it is analyzed by a correlation of the data and time. This correlation is displayed in a graphical tool called correlogram, which is used to provide autocorrelation values with different time periods. Time is demonstrated in the horizontal axis, while the correlation coefficient is in the vertical axis. The difference in data patterns can be indicated by these figures present below (Hanke & Wichern, 2009). Once the patterns are identified, the data are applied to appropriate technique performance. Figure 16, 17, and 18 are data patterns of all three products using decomposition analysis.

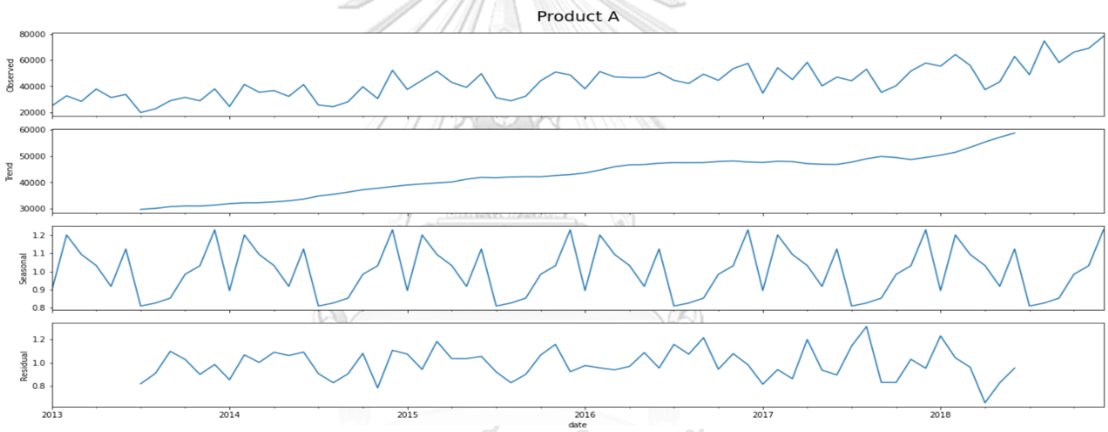

Figure 16: Time series decomposition analysis of product A

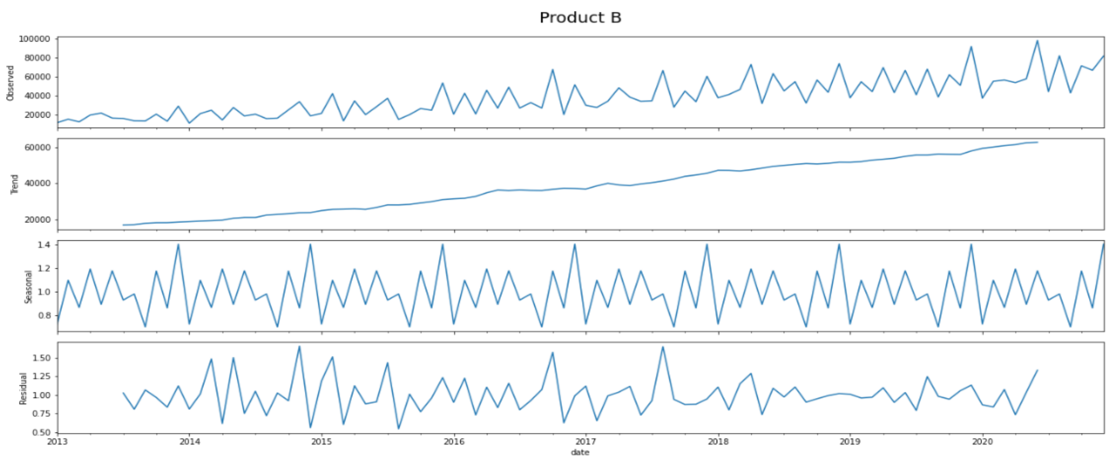

Figure 17: Time series decomposition analysis of product B

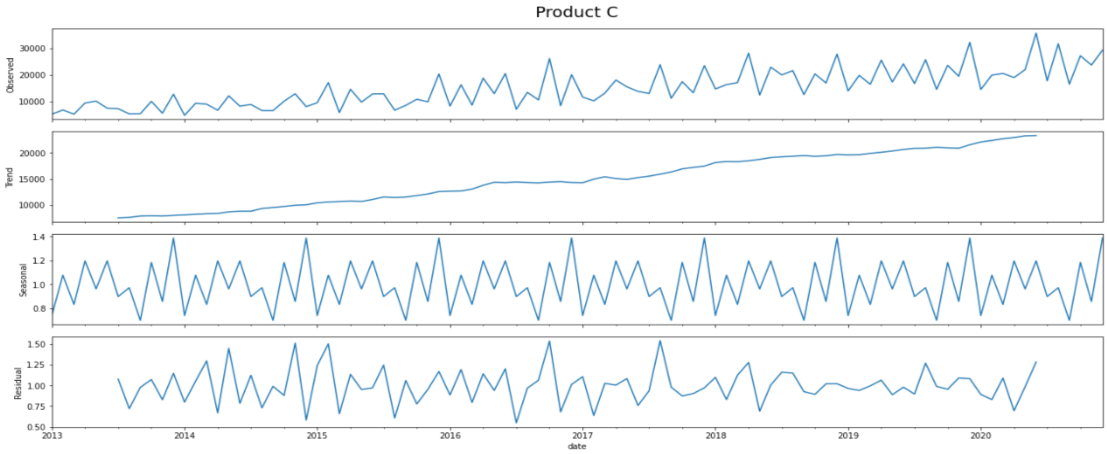

Figure 18: Time series decomposition analysis of product C

Those figures show the increasing data patterns from January 2013 to December 2020, which can be inferred that all products are trend-pattern data. Furthermore, the 12-month time of all three products shows seasonal patterns. The seasonal pattern of product A shows that the product demand is increased at the beginning of the year and is decreased to the lowest level in June before increasing to the maximum level in December. Product B and C share similar seasonal patterns in which the demand of each month alternately increases and decreases. In September, the product demands are at the lowest level. On the other hand, the highest demand levels are in December. Figure 19, 20, and 21 are graphical data using autocorrelation analysis for all three products.

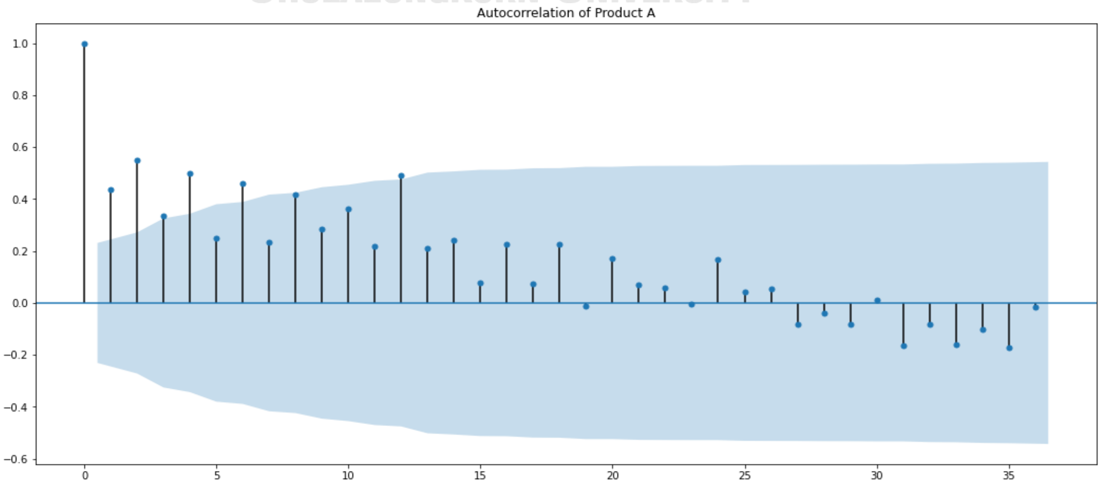

Figure 19: Autocorrelation plot of product A

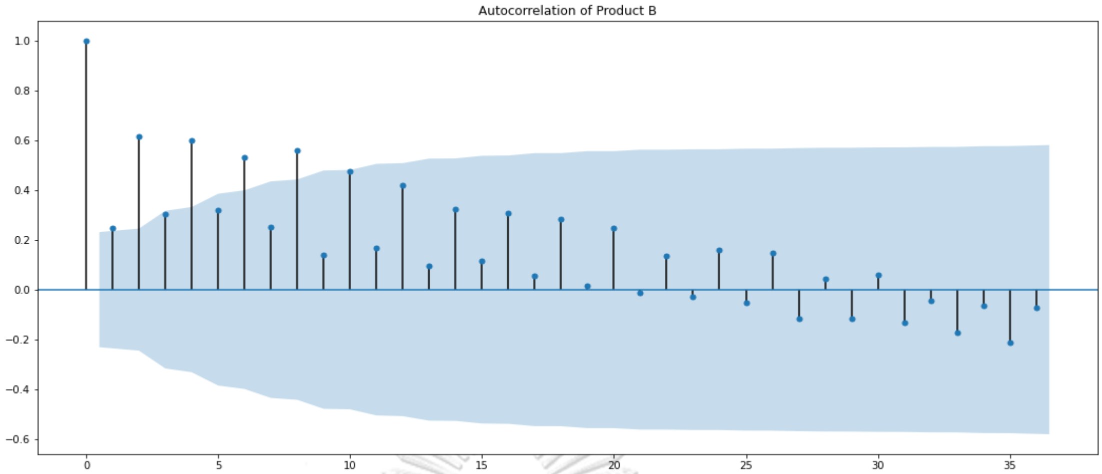

Figure 20: Autocorrelation plot of product B

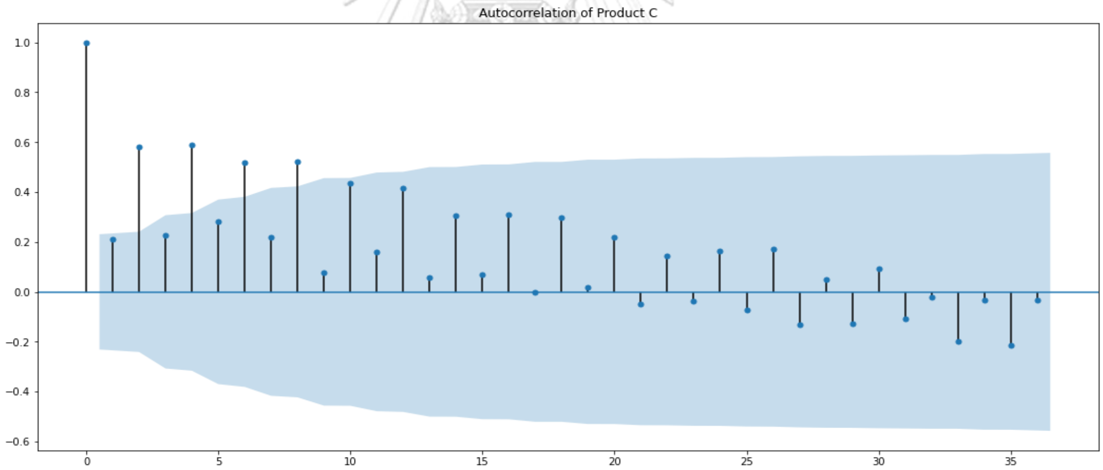

Figure 21: Autocorrelation plot of product C

The figures show autocorrelation values compare to the time lag in the correlogram, which present all product trend data. According to the autocorrelation coefficients, the values are significantly different for the first several time lags and then decline as the number of time lags increases. These findings can be concluded that data is non-stationary. Thus, parameters should be set differencing.

3.3 Explanatory variables

This study uses explanatory variables mentioned in the literature review on demand forecasting. The details of each variable are represented as Gross Domestic Product (GDP), Consumer Price Index (CPI), Unemployment Rate (UR), Exchange rate (EX), Terms of Trade (TT), International reserve (IR), Population (POP), Set index (SET), and Dummy variables.

The study of explanatory variables is based on the relationship of variables and product demand data. They are selected from the literature review that is similar to this research. These explanatory variables categorize the data into two groups: continuous explanatory variables use to store the same data as the training set in January 2013 to December 2018 and categorical explanatory variables, which is the monthly data.

3.3.1 Selection of continuous explanatory variables

This research focuses on the continuous explanatory variables related to product demand using an analysis of each factor’s correlation coefficient with the training set. Figure 22 shows the correlation matrix of all variables. Table 9 shows the correlation coefficient of each variable from the demand of product A, B, and C.

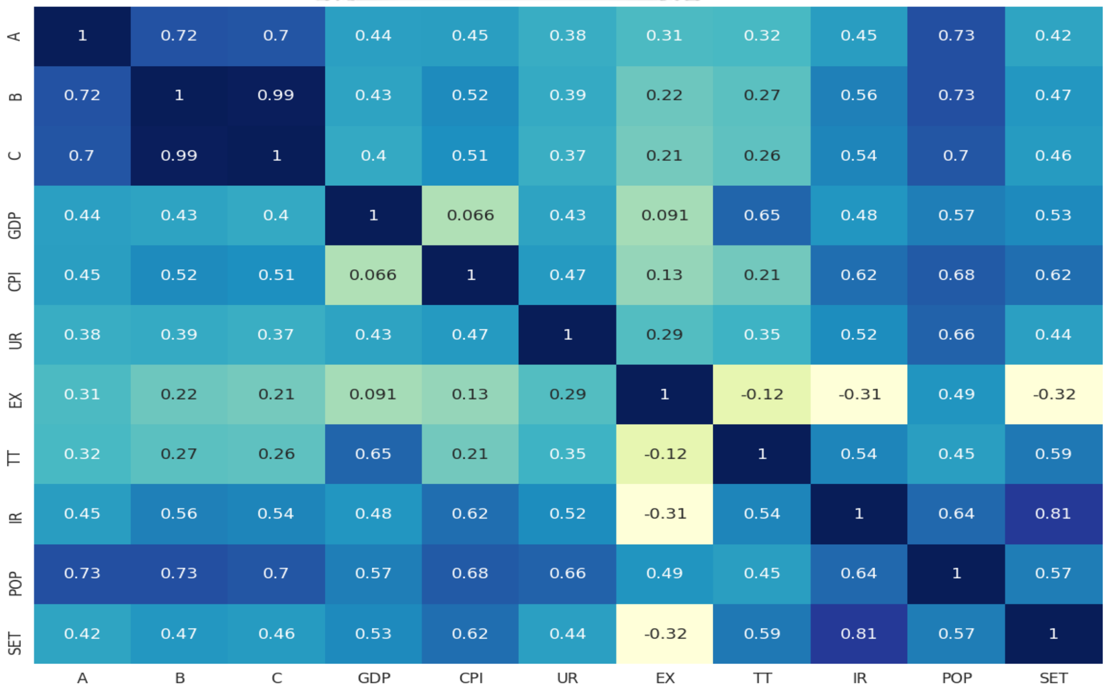

Figure 22: Correlation coefficient matrix of all variables

Table 9: Correlation coefficient of each product

No.	Variables	Product A	Product B	Product C
1	GDP	0.44	0.43	0.40
2	CPI	0.45	0.52	0.51
3	UR	0.38	0.39	0.37
4	EX	0.31	0.22	0.21
5	TT	0.32	0.27	0.26
6	IR	0.45	0.56	0.54
7	POP	0.73	0.73	0.70
8	SET	0.42	0.47	0.46

3.3.2 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. Figure 23, 24, and 25 show a summary of product demand using boxplot analysis.

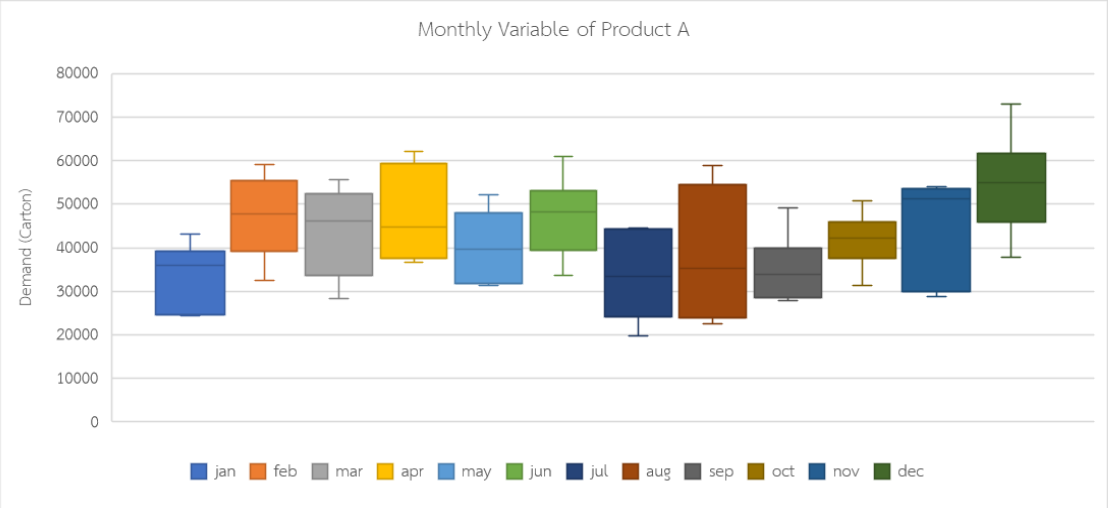

Figure 23: Product A's monthly demand boxplot

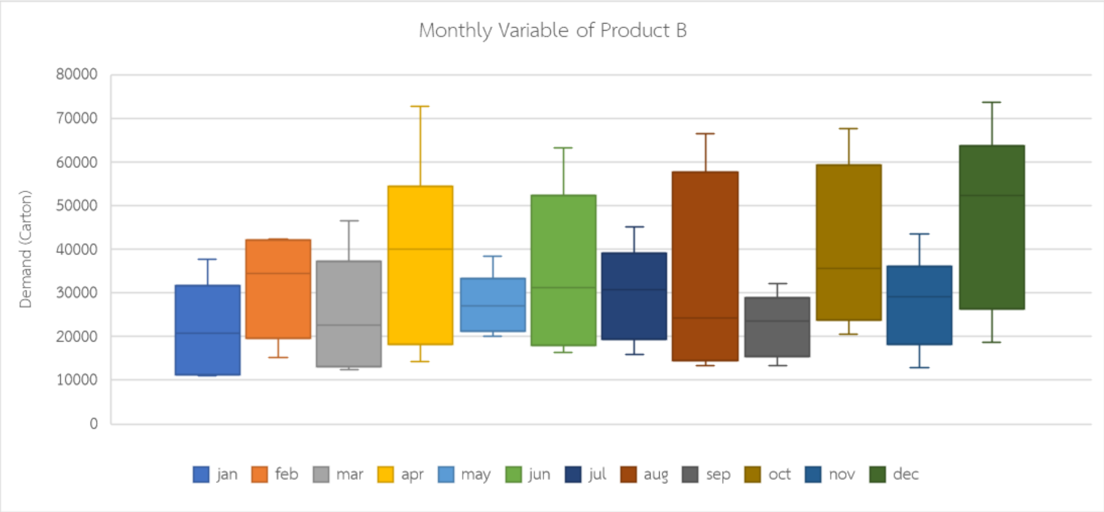

Figure 24: Product B's monthly demand boxplot

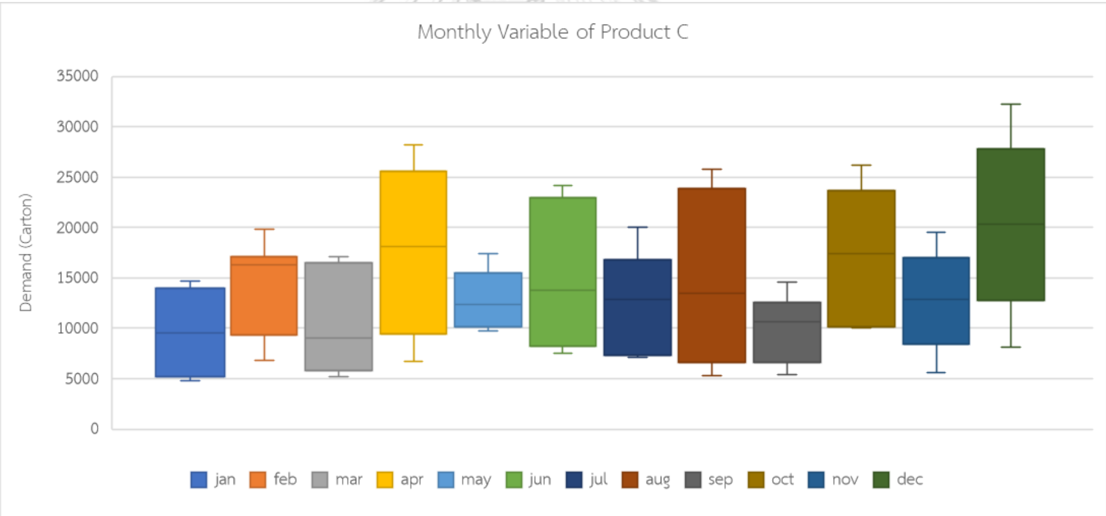

Figure 25: Product C's monthly demand boxplot

Monthly variables of product A are February, March, April, May, June, July, August, September, October, November, and December. The monthly variables of product B and C are January, February, March, April, May, June, July, August, October, November, and December.

3.3.3 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 (Smith, 2018). 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 (Ghani & Ahmad, 2010). 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 (Fritz & Berger, 2015). The Minitab is used in the research for stepwise regression.

The table 10 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 10: Explanatory variables chosen for each product

Explanatory variables	Product A	Product B	Product C
Month dummy variables	February March April June July September December	April October December	April June October December
Economic indicator	Not chosen	Not chosen	Not chosen
Population	Chosen	Chosen	Chosen
Set index	Not chosen	Not chosen	Not chosen
R^2	79.85%	64.74%	65.09%
R^2(adj.)	77.29%	62.64%	62.45%

3.4 Data preprocessing

Data preprocessing is an important step after data cleaning, which is preparing the dataset, such as selection feature or transformation data in order to be ready for using in the model. Demand product data need a preprocess by the processing training set, validate set, and test set before putting them into the machine learning forecasting model. The most used main preprocessing is the normalization and sliding window method.

3.4.1 Normalization

First, all of the data should be transformed into normalized data since the performance in the training model, and forecasting accuracy will be improved (Alameer et al., 2020). In this study, the range of normalization values is 0 – 1, which makes the dataset homogenous (Lijuan & Guohua, 2016) :

The normalization formular is defined as:

$$x_{i(norm)} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{19}$$

$x_{i(norm)}$ = normalized data

x_{max} = maximum data

x_{min} = minimum data

3.4.2 Sliding window method

Sliding window method is a technique of data preprocessing using the actual value of the time series data. The history data that can predict the next step is called the sliding window method (Hota et al., 2017). It is called a lag method in statistics. The dataset in this study is processed using two types of forecasting methods: one-step ahead and multi-step ahead forecasting methods.

- One-step ahead forecasting

In time series data, one-step ahead is a technique used to preprocess the machine learning forecasting model. This technique makes forecasts relatively stable and indicates less error. The one-step ahead forecasting is one of the slide window methods. It is also known for using previous data to predict the future value on one time step, where the input is shifted one by one to continuously forecast the next step as shown in the table 11.

Table 11: Sliding window for one-step ahead forecasting (original data only)

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
Jan 1	Feb 1	Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2
Feb 1	Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2	Feb 2
Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2	Feb 2	Mar 2
…												
Nov 6	Dec 6	Jan 7	Feb 7	Mar 7	Apr 7	May 7	Jun 7	Jul 7	Aug 7	Sep 7	Oct 7	Nov 7
Dec 6	Jan 7	Feb 7	Mar 7	Apr 7	May 7	Jun 7	Jul 7	Aug 7	Sep 7	Oct 7	Nov 7	Dec 7

- One-step ahead forecasting with external variables

In addition to forecasting, time series data is used one-step ahead. It is also needed to add related values, such as the external variables used in the data preprocessing, in the same format as the one-step ahead. These variables will be integrated. The external values must be the same time period as the output, which is shown in the table 12.

Table 12: Sliding window for one-step ahead forecasting (original data with external variables)

X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	EX1	EX2	EX3	Y
Jan 1	Feb 1	Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	EX1 (Jan)	EX2 (Jan)	EX3 (Jan)	Jan 2
Feb 1	Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2	EX1 (Feb)	EX2 (Feb)	EX3 (Feb)	Feb 2
Mar 1	Apr 1	May 1	Jun 1	Jul 1	Aug 1	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2	Feb 2	EX1 (Mar)	EX2 (Mar)	EX3 (Mar)	Mar 2
...															
Nov 6	Dec 6	Jan 7	Feb 7	Mar 7	Apr 7	May 7	Jun 7	Jul 7	Aug 7	Sep 7	Oct 7	EX 1(Nov)	EX2 (Nov)	EX3 (Nov)	Nov 7
Dec 6	Jan 7	Feb 7	Mar 7	Apr 7	May 7	Jun 7	Jul 7	Aug 7	Sep 7	Oct 7	Nov 7	EX1 (Dec)	EX2 (Dec)	EX3 (Dec)	Dec 7

- Multi-step ahead forecasting

Using the multi-step technique is similar to the one-step method, but the difference is that it uses the arrangement of past data to forecast more than one value in the future, such as 3, 6, 12 steps. If external variables are used in the forecasting model, the same process technique uses in one-step forecasting will be applied and integrates with the previous time series call Multivariate multi-step ahead forecasting. This method is quite complicated and tends to have high errors. On the contrary, the advantage of this technique is that it can predict value, which has more than one horizon period time, as shown in table 13.

Table 13: Sliding window for multi-step ahead forecasting

X1	X2	X3	X4		X9	X10	X11	X12	Y1	Y2	Y3	Y4		Y9	Y10	Y11	Y12
Jan 1	Feb 1	Mar 1	Apr 1	...	Sep 1	Oct 1	Nov 1	Dec 1	Jan 2	Feb 2	Mar 2	Apr 2	...	Sep 2	Oct 2	Nov 2	Dec 2
Feb 1	Mar 1	Apr 1	May 1		Oct 1	Nov 1	Dec 1	Jan 2	Feb 2	Mar 2	Apr 2	May 2		Oct 2	Nov 2	Dec 2	Jan 3
⋮																	
Jan 6	Feb 6	Mar 6	Apr 6	...	Sep 6	Oct 6	Nov 6	Dec 6	Jan 7	Feb 7	Mar 7	Apr 7	...	Sep 7	Oct 7	Nov 7	Dec 7

3.5 Forecasting

3.5.1 Holt-Winters exponential smoothing method

Holt-Winters exponential smoothing method is developed from the Holt exponential smoothing method. It is used with trend data. The developed part can be used to forecast seasonal patterns. There are three smoothing parameters in this study. The study uses R programming to optimize the values to create the model.

The process of the Holt-Winters method starts with data analysis by the time series plot and decomposition analysis to determine whether it is a trend or seasonal data. The next step is to optimize the smoothing parameter, forecast the value, and take it to evaluate the model with a validate set using MAPE. Holt-Winters method workflow is shown in figure 26:

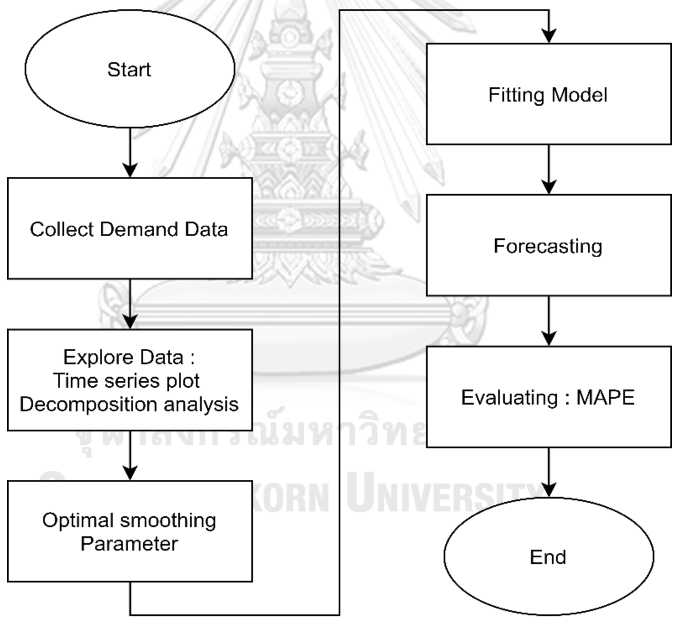

Figure 26: Holt–Winters exponential smoothing method flow chart

3.5.2 SARIMA

SARIMA is a technique used in time series forecasting. The technique is based on ARIMA, in which seasonal variables are added. Initially, the ARIMA model had only p, q, and d parameters. Its consideration methods are similar to the ARIMA model. The only difference is seasonal length parameter integration, which is set as 12 in this study.

The forecasting process using SARIMA starts with the data analysis using a time series plot, decomposition analysis determines a trend or seasonal pattern, and autocorrelation to identify whether the data is stationary or non-stationary, which will be used to determine d and D. The next step is to optimize parameter p, q, P, and Q for the SARIMA model performance. Once the process is done, the model will get a residual check with a diagnostic. The forecasted value will be taken to evaluate with a validate set using MAPE. The study uses python programming to optimize the parameters to create the SARIMA model. SARIMA workflow is shown in figure 27:

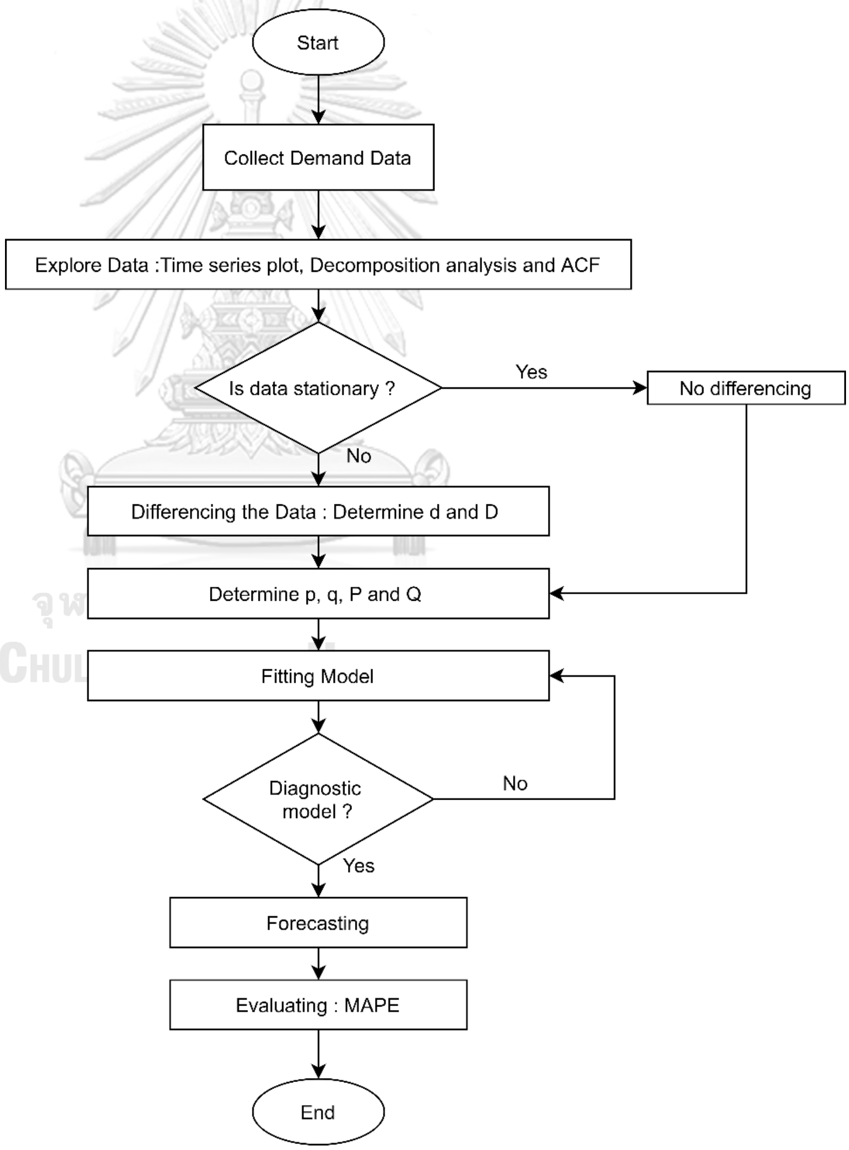

Figure 27: SARIMA flow chart

3.5.3 SARIMAX

SARIMAX is a forecasting technique developed by adding explanatory variables to the SARIMA model. As a consequence, forecasting data can be explained by the causal method. These added explanatory variables result in the addition of an exogenous variable defined as a parameter X.

The SARIMAX forecasting process is similar to the SARIMA, in which the additional explanatory variables are added to the fitting model process. The SARIMAX model uses python programming to optimize the parameter. SARIMAX workflow is shown in figure 28.

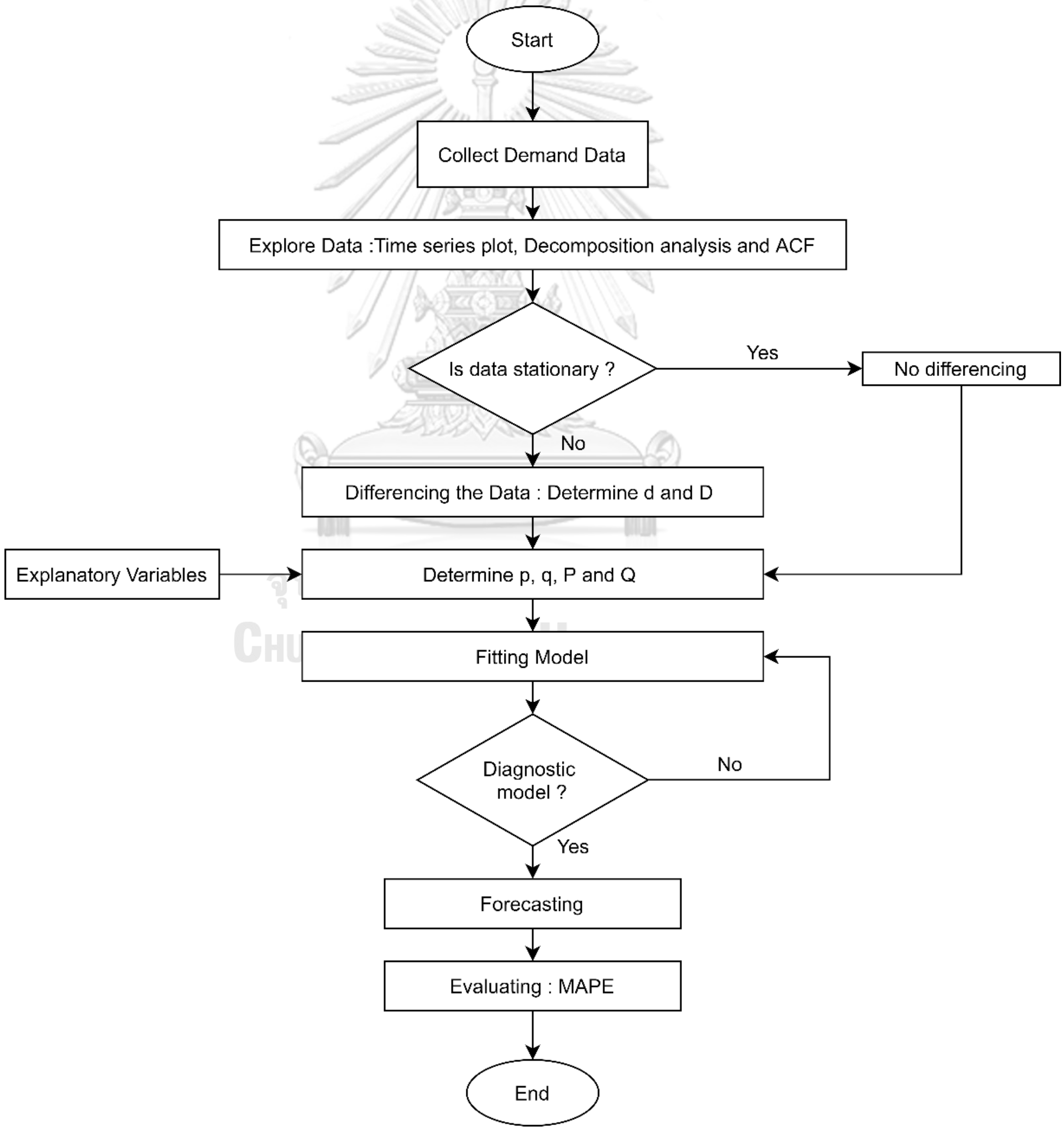

Figure 28: SARIMAX flow chart

3.5.4 Artificial neural network (ANN)

Artificial neural network is a machine learning model, which is adapted in forecasting. This model works similarly to the human nervous system. In this study, the forecasting model will be constructed in a supervised learning form. To creating a model, the data is required for the algorithm to learn the pattern so that it can be used to predict future values. The python programming is used to optimize the hyperparameter and create ANN model. The variables that will affect the model are as follows (Table 14):

- Number of input units

The number of input units is the starting point for model construction. The number of input units depends on the input data used to create the model and the procedures in the data processing.

- Number of hidden layers

The number of hidden layers can contain multiple layers. Each layer has processing units called Neuron. This study explores to create two types of forecasting models: one hidden layer and two hidden layers. The data used in the analysis is not very much. So, this study chooses to use no more than two hidden layers. With more than two hidden layers, the model will turn into deep learning, which will lead to longer computational time.

- Number of hidden units

The number of neurons or units in the hidden layer will process the input into the output of each layer. In this study, the hidden units are varied between 50-150 units

- Number of epochs

The number of epochs is the number of training in one dataset. The determination of the epoch affects the training of the model, where an excessive number of epochs can lead to overfitting. In the experiment, the number of epochs are varied at 25, 50, 100, 150, and 200.

- Number of batch sizes

The batch size is the number of datasets use to update the model's learning. The optimal number of batch sizes should not exceed 32 (Masters & Luschi, 2018). The batch size numbers in this experiment are varied at 2, 4, 8, 16, and 32.

- Activation functions

The activation function is a function that takes sum of all operations from every input and determines the output of each class. In this research, ReLU and Swish functions will be used in training models. For structure of ANN, the calculation of the Swish activation function can result in negative values. To prevent this problem, the output layer provided was ReLU, which will only compute positive values. These will be created into six types: Relu-Relu, Swish-Relu, Swish-Relu-Relu, Swish-Swish Relu, Relu-Swish-Relu, and Relu-Relu- Relu.

Table 14: The hyperparameter for artificial neural network in all products

Name	Hyperparameters	Levels
Number of Hidden layers	1-2	2
Number of Hidden units	50-150	101
Number of epochs	25, 50, 100, 150, 200	5
Number of Batch sizes	2, 4, 8, 16, 32	5
Activation functions	SR, RR, SRR, SSR, RSR, RRR	6
Number of possible combinations		30,300

The forecasting process using ANN starts with the basic data analysis, which is the same as the time series processing. The next step is data preprocessing, such as data normalization and sliding data. In the data normalization process, the value of input data is in the range of 0-1. The input data is taken into the machine learning model. The format of the input data is the product historical demand data. There are historical data, which is combined with step wised external variables, and historical data, which is combined with all external variables. The sliding window is used as a technique to handle forecast boundaries, which are divided into two types: short-term (one-step ahead) and long-term (multi-step ahead). Building artificial neural network models will choose parameters in the following table 14. The ANN algorithm is trained and validated depending on the respective parameter ranges. The best parameters are selected, which gives the lowest MAPE, will be set with the data validation set. The prediction value will be taken to evaluate with the test set using MAPE. ANN workflow is shown in figure 29:

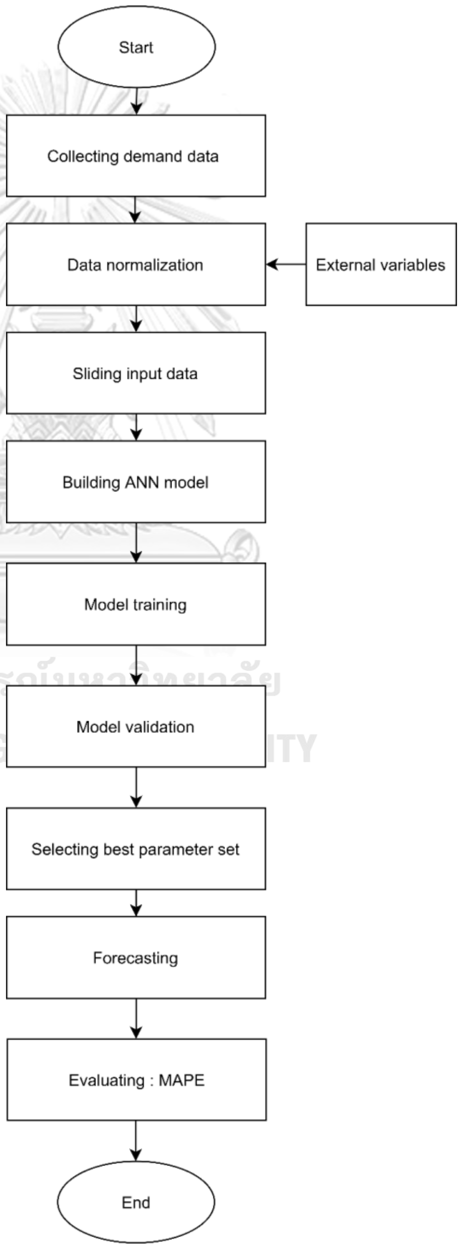

Figure 29: Artificial neural network flow chart

3.5.5 Combined forecasting model

The combined forecasting model is to calculate the forecast value from each method to find the combined value. This model will adjust the weight of the method, which is calculated together.

- Simple Weight Average

A simple weight average is a calculation of the prediction value from the forecasting method that will be combined into the form of averaging forecast data. The simple weight average combined with the forecasting process starts with single forecasting, such as Holt-Winters, Box-Jenkins, and artificial neural network. The next step is bringing the prediction values of each model to calculate using the simple weigh average. Combined forecasting workflow is shown in figure 30.

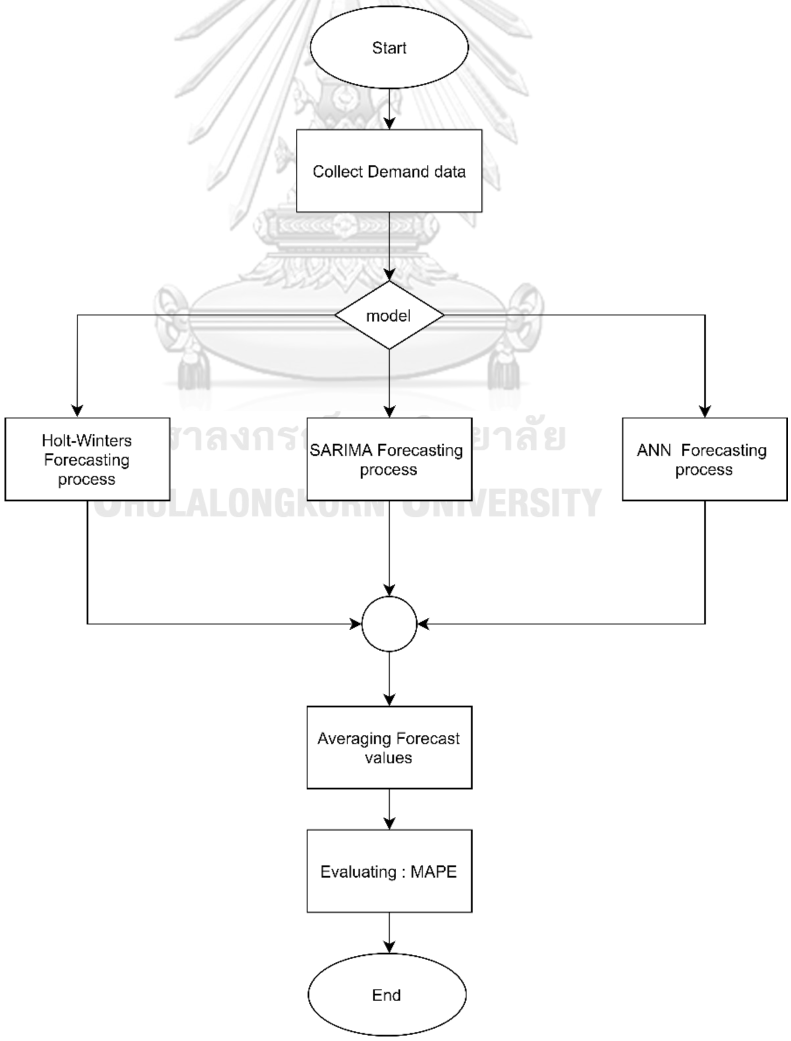

Figure 30: Combined forecasting flow chart

3.5.6 Hybrid forecasting model

The hybrid forecasting model is a technique improved to increase forecasting accuracy by developing from traditional models that perform hybridized linear and non-linear models. The traditional forecasting model is classified as a linear model, while the machine learning model is classified as a non-linear model. The study uses python programming to optimize the hyperparameters and create the hybrid forecasting model. This research studies two types of hybrid forecasting models as follows.

- SARIMA-ANN

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: $Y_t = L_t + N_t$
5. The prediction value of the hybrid model will be taken to evaluate with the test set using MAPE. SARIMA-ANN workflow is shown in figure 31:

- Proposed model (HW combine SARIMA hybrid ANN)

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 call 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-ANN workflow is shown in figure 32:

To the best of our knowledge, this study is the first to propose this model since the previous studies considered SARIMA-ANN model other methods which is different from this model.

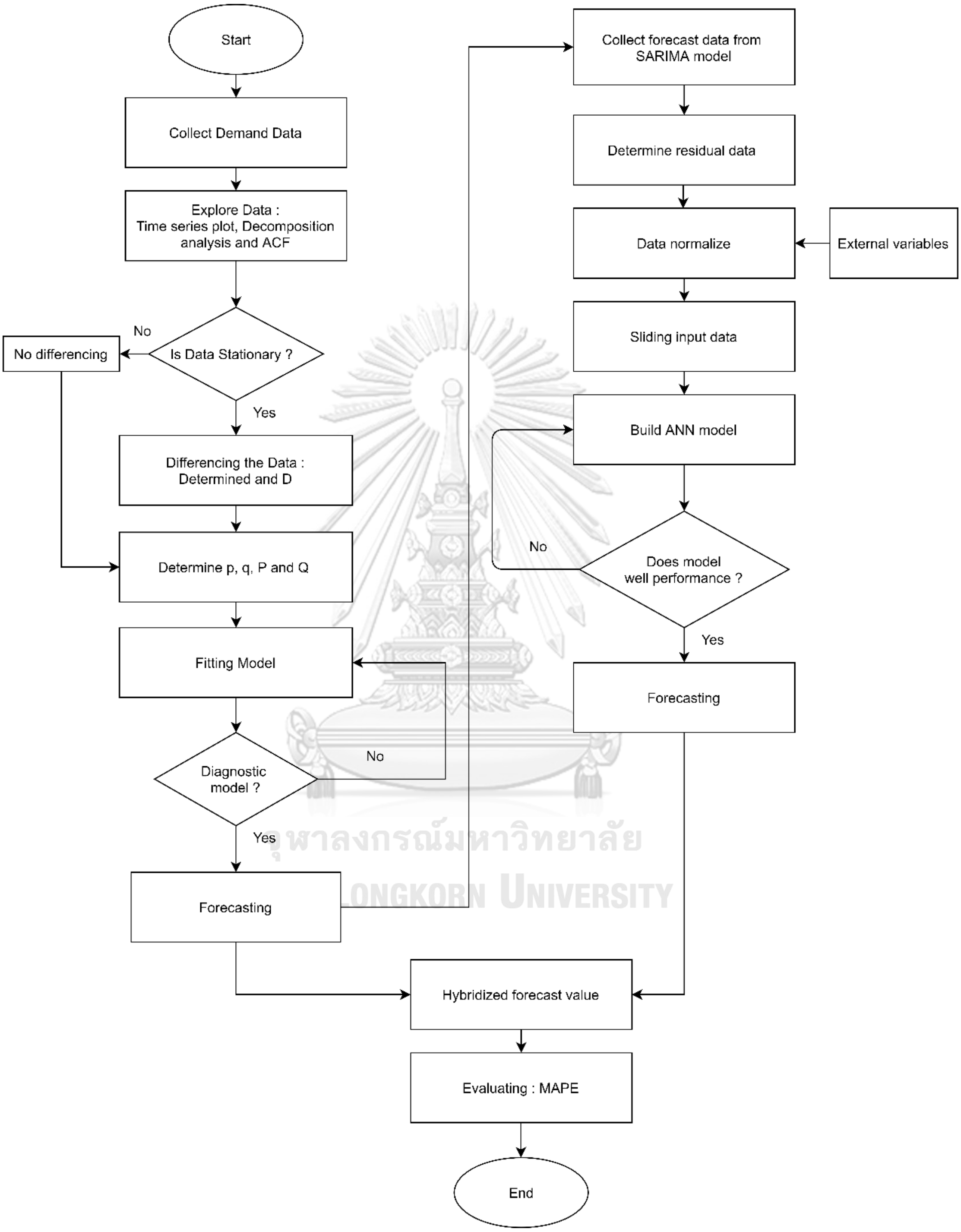

Figure 31: SARIMA-ANN flow chart

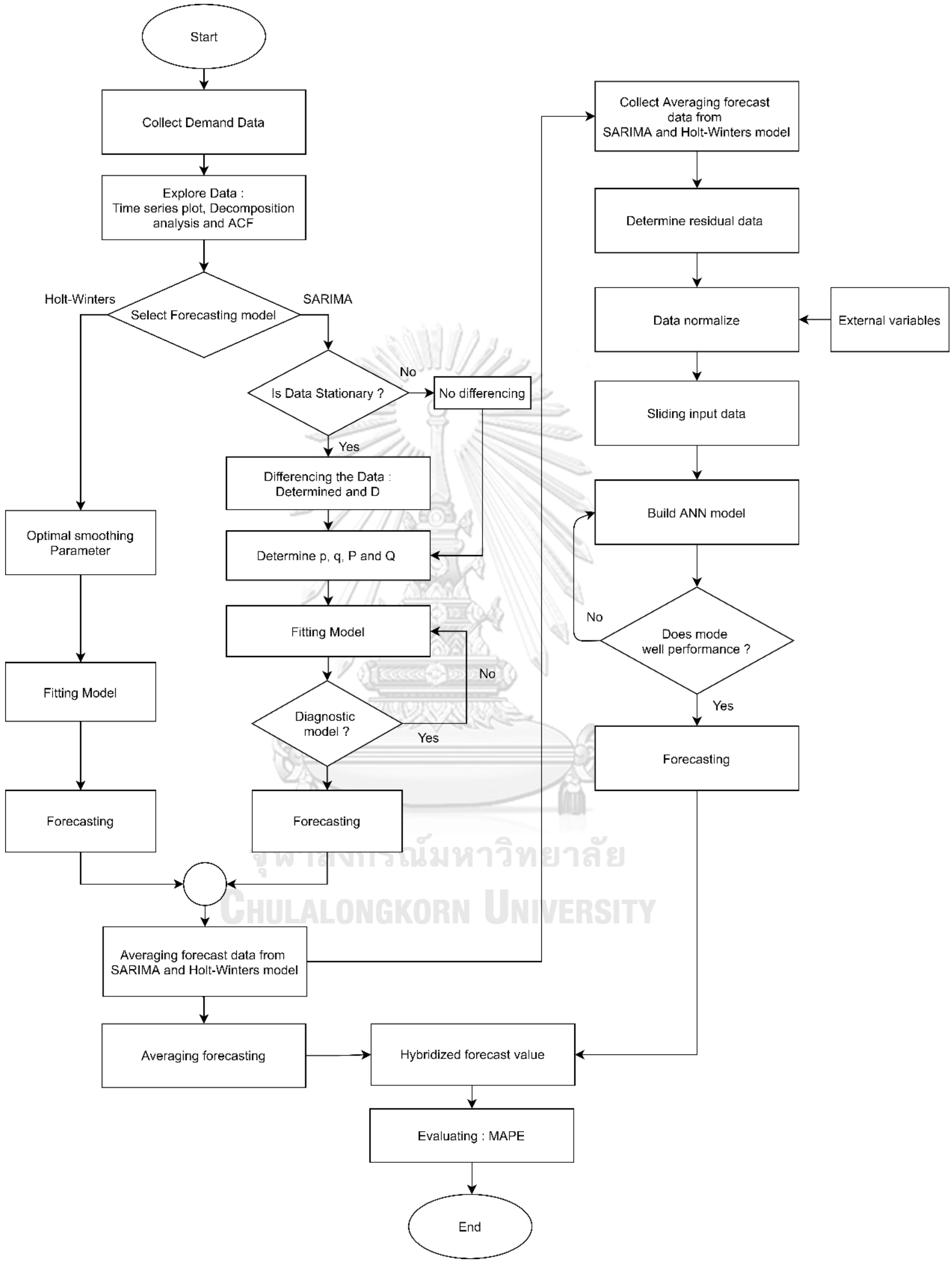

Figure 32: Proposed model flow chart

Chapter 4 Results and Discussion

4.1 Results

This chapter presents the forecasting results using the time series models, machine learning models, combined forecasting models and hybrid forecasting model. Evaluation of these prediction models is based on the evaluation of the error and actual demand data. The validation sets of January 2019 to December 2019 (12 months) and the testing set of January 2020 to December 2020 (12 months) present by the MAPE evaluation are used in showing results.

4.1.1 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. Figure 33, 34, and 35 demonstrate prediction values from the current forecasting method of product A, B, and C compared with the actual demand data of each product.

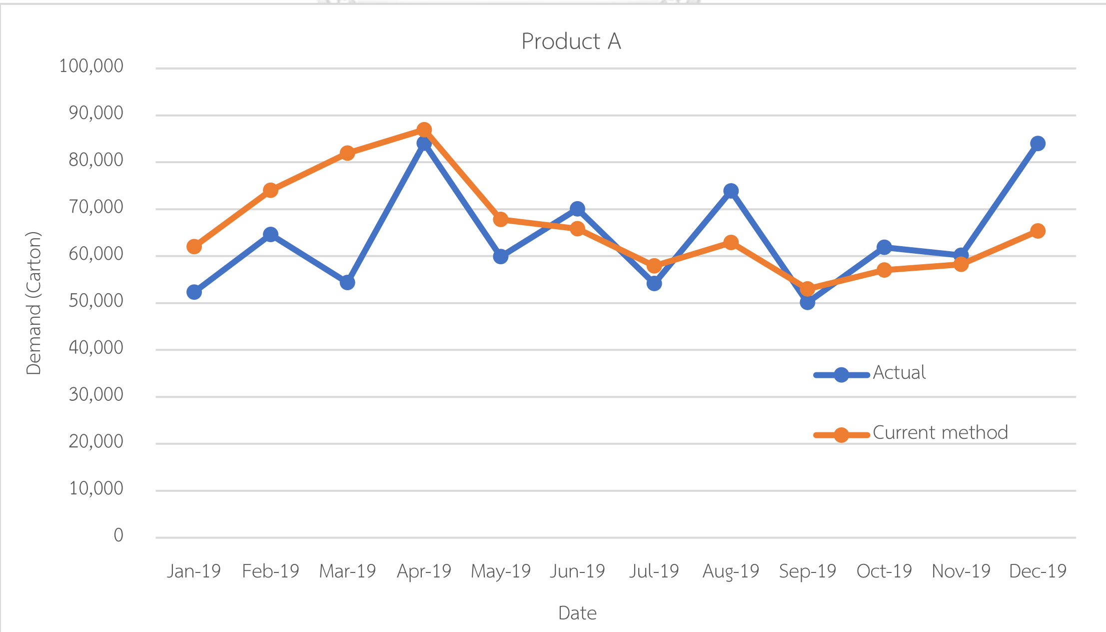

Figure 33: Sale force composite forecasting compared with actual values of product A

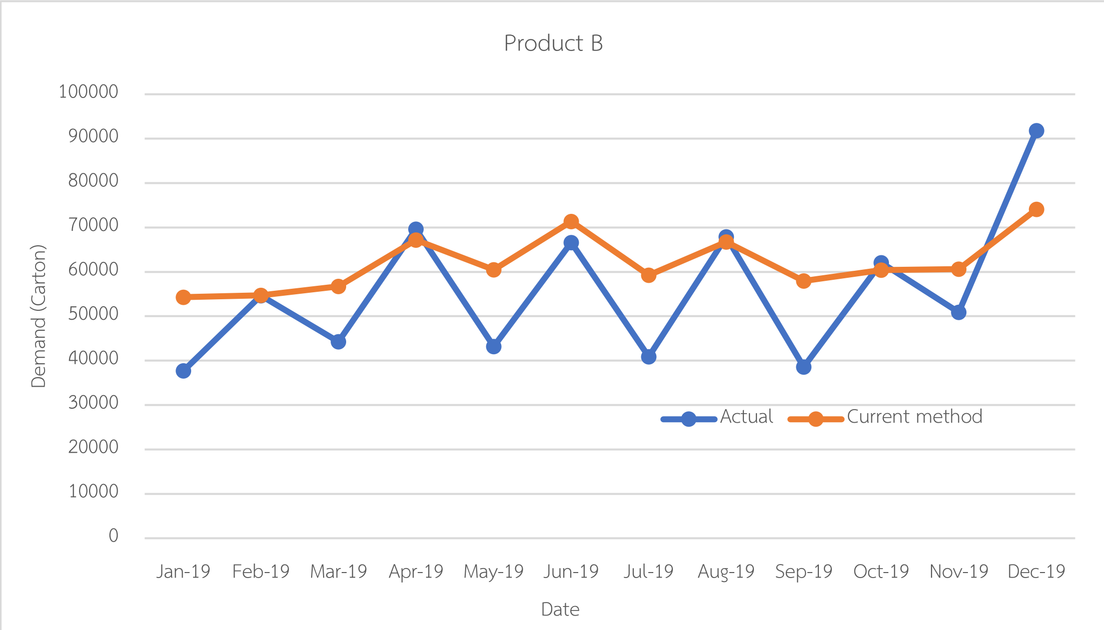

Figure 34: Sale force composite forecasting compared with actual values of product B

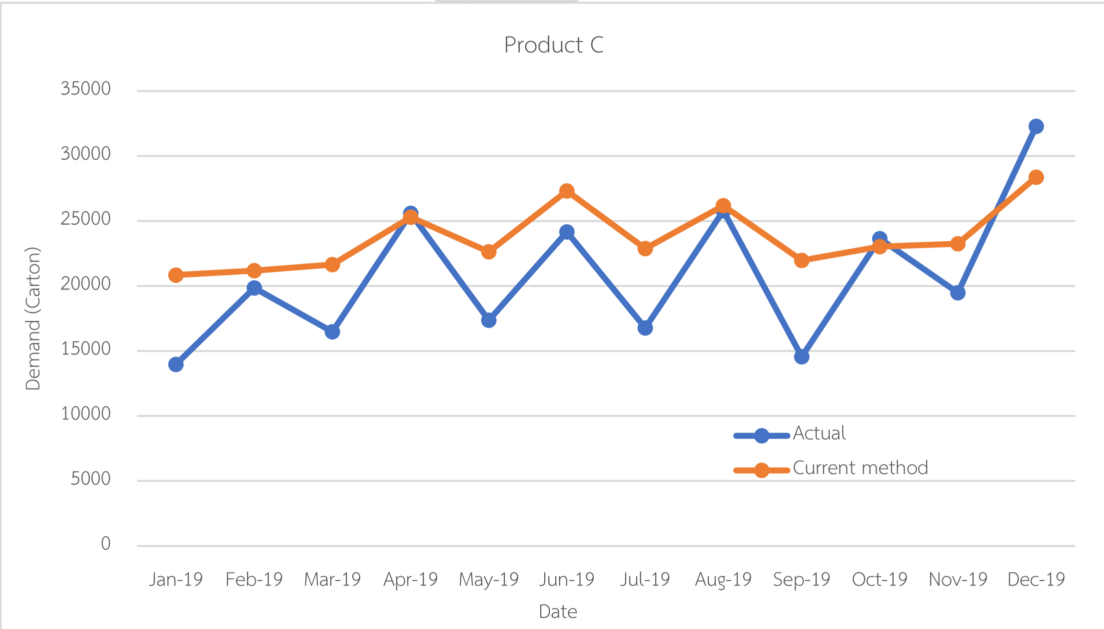

Figure 35: Sale force composite forecasting compared with actual values of product C

The accuracy of forecasting evaluates using the validation set, which indicates each product using MAPE in table 15. The evaluated result of the current method shows the MAPE of product A, B, and C, which are 13.95%, 21.75%, and 21.20%, respectively.

Table 15: Measurement error of sale force composite forecasting

Measurement	Product A	Product B	Product C
MAPE	13.95%	21.75%	21.20%

4.1.2 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 shown in the table 16. Figure 36, 37, and 38 demonstrate predicted values using the Holt-Winters exponential smoothing method for product A, B, and C compared with the actual demand data of each product.

Table 16: Smoothing parameters of each product

Product	Alpha	Beta	Gamma
A	0.090	0.0061	0.32
B	0.0043	0.45	0.37
C	0.0058	0.38	0.38

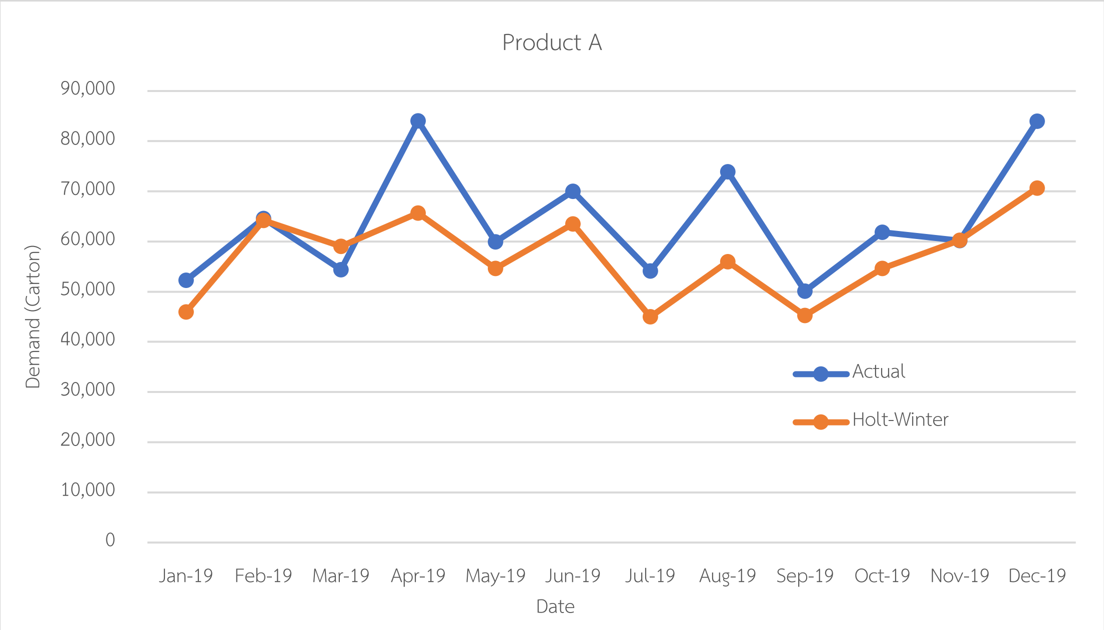

Figure 36: Holt-Winters forecasts compared with actual values of product A

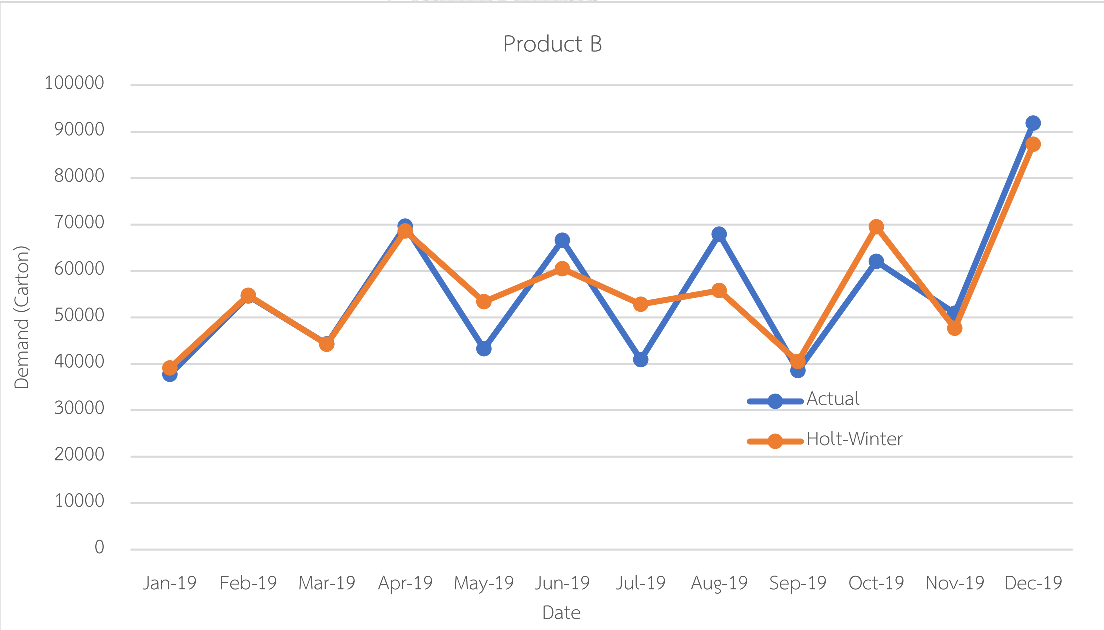

Figure 37: Holt-Winters forecasts compared with actual values of product B

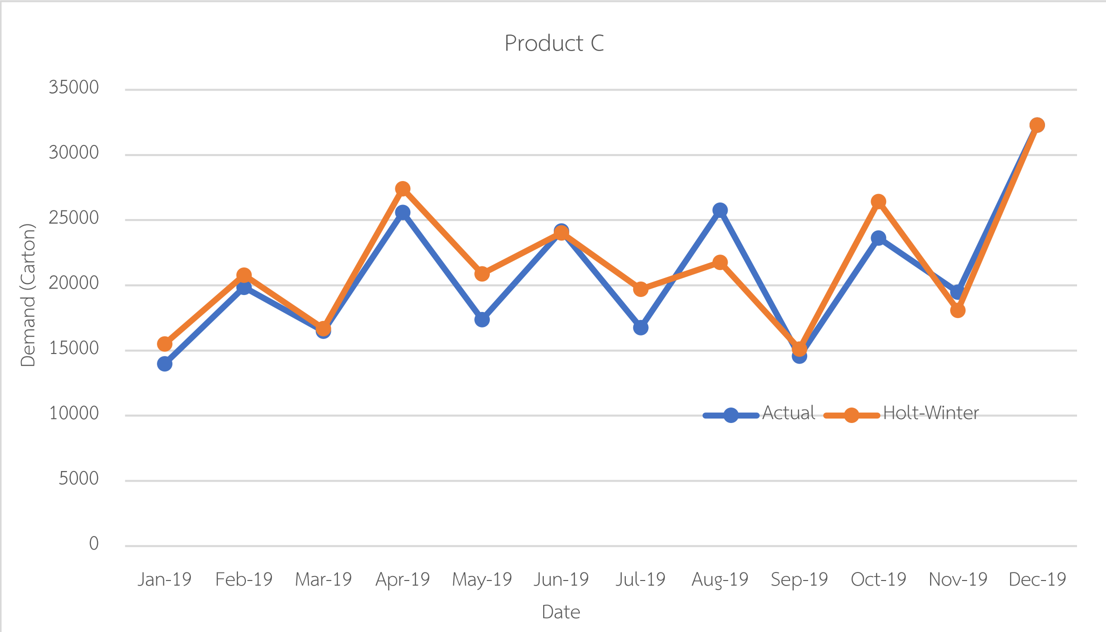

Figure 38: Holt-Winters forecasts compared with actual values of product C

The accuracy of forecasting evaluates using the validation set, which indicates each product using MAPE in table 17. The evaluated result of the Holt-Winters exponential smoothing method shows the MAPE of product A, B, and C, which are 10.75%, 9.03%, and 8.52%, respectively.

Table 17: Measurement errors of Holt-Winters Method

Measurement	Product A	Product B	Product C
MAPE	10.75%	9.03%	8.52%

4.1.3 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 results of estimating ARIMA order are set according to the minimum AIC. The parameters of SARIMA and SARIMAX are shown in table 18. Figure 39, 40, and 41 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product using SARIMA and SARIMAX methods.

Table 18: SARIMA and SARIMAX parameters of each product

Product	SARIMA	SARIMAX
A	(3,2,2) x (1,0,0,12)	(2,2,1) x (1,1,1,12)
B	(2,2,3) x (0,1,1,12)	(0,0,1) x (0,1,1,12)
C	(3,1,0) x (2,1,0,12)	(0,0,1) x (0,1,1,12)

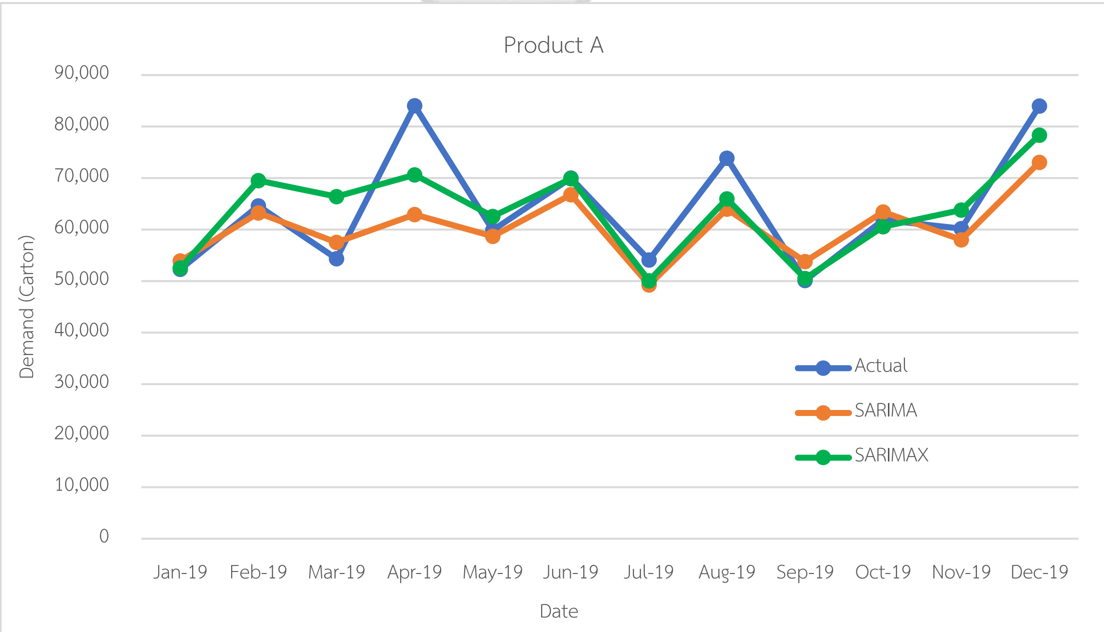

Figure 39: SARIMA and SARIMAX forecasts compared with actual values of product A

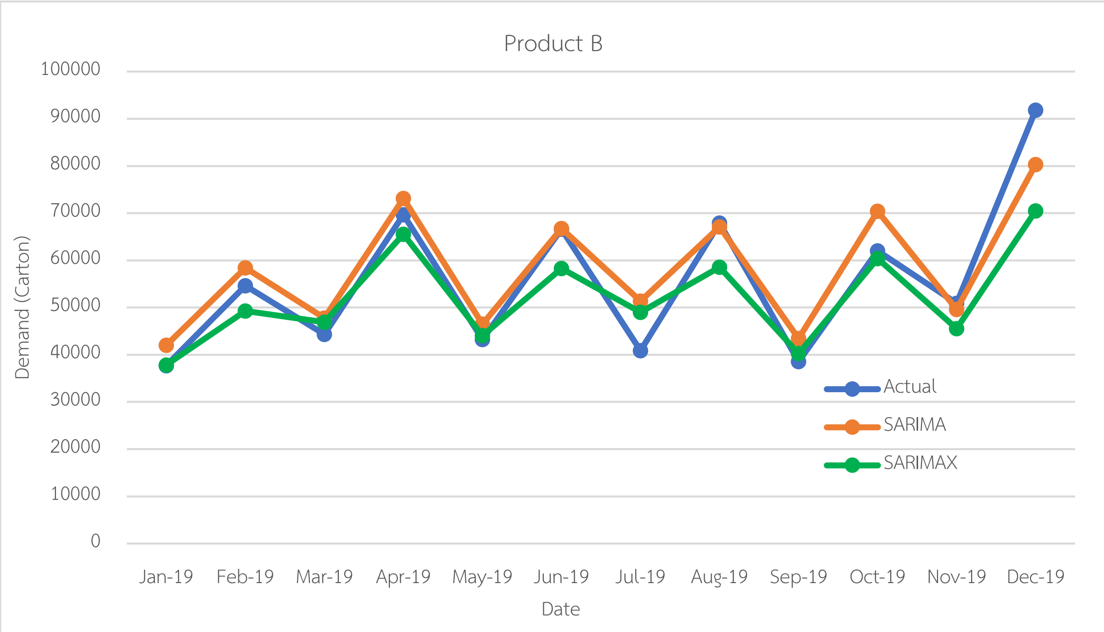

Figure 40: SARIMA and SARIMAX forecasts compared with actual values of product B

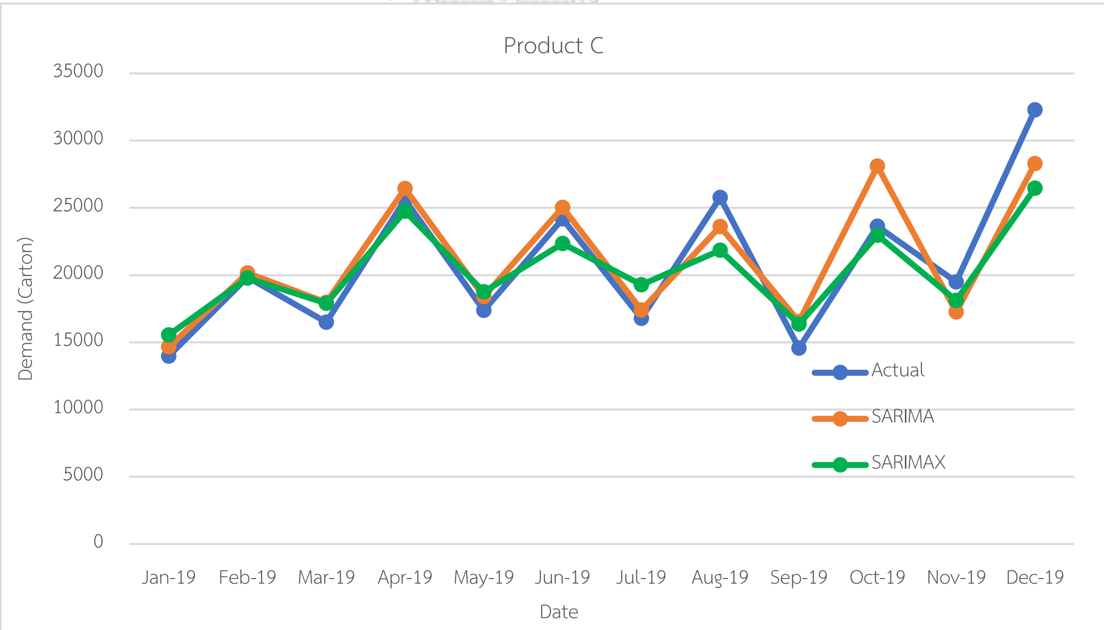

Figure 41: SARIMA and SARIMAX forecasts compared with actual values of product C

The accuracy of forecasting evaluates using validation set, which indicates each product using MAPE in table 19.

Table 19: Measurement error of SARIMA and SARIMAX

Measurement	Method	Product A	Product B	Product C
MAPE	SARIMA	7.64%	8.94%	8.05%
	SARIMAX	7.06%	9.25%	9.12%

The evaluated result of SARIMA shows the MAPE of product A, B, and C, which are 7.64%, 8.94%, and 8.05%, respectively. The evaluated result of SARIMAX shows the MAPE of product A, B, and C, which are 7.06%, 9.25%, and 9.12%, respectively.

The results indicated that SARIMAX shows the lowest MAPE on the product A which is 7.06%. On the other hand, the SARIMA shows the lowest MAPE values, which are 8.94% and 8.05% on product B and C, respectively. The explanatory variables in SARIMAX may not properly fit the demand data of product B and C leading to a higher error percentage compared to SARIMA.

4.1.4 Artificial neural network

In this study, ANN forecasting is explored in three conditions. First, there are three forms of input data: original data, original data with stepwise external variables, and original data with all external variables. Second, there are two types of forecasting scope: one-step ahead and multi-step ahead, which are short-term and long-term forecast. Finally, 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, which will be displayed in table 20 and 21 as follows:

Table 20: ANN's one-step hyperparameters and MAPE of each product (Short-term forecasting)

Parameter	Product A	Product B	Product C
Input data format	Original	OG+all external	OG+all external
Model	Model 6	Model 16	Model 14
Hidden layers	2	2	1
Hidden Units	135	55	113
Epochs	25	100	100
Batch size	2	16	4
Activation function	RRR	SSR	SR
MAPE	4.37%	5.43%	4.11%

Table 21: ANN's multi-step hyperparameters and MAPE of each product (12 months in advance forecasting)

Parameter	Product A	Product B	Product C
Input data format	Original	OG+all external	OG+stepwise external
Model	Model 5	Model 16	Model 7
Hidden layers	2	2	1
Hidden Units	142	52	130
Epochs	150	50	100
Batch size	4	8	16
Activation function	RSR	SSR	RR
MAPE	5.65%	4.27%	3.40%

The evaluated result of the one-step ahead artificial neural network shows the lowest MAPE of product A which is 4.37% from the ANN original data model 1. The ANN original data with all external variables shows the MAPE of product B from model 16 as 5.43% and product C from model 14 as 4.11%.

The multi-step ahead artificial neural network shows the lowest MAPE of product A from ANN original data model 5 as 5.65%. ANN original data with all external variables shows the MAPE of product B from model 16 as 4.27% and the MAPE of product C from ANN original data with stepwise external variables from model 7 as 3.40%. Figure 42, 43, and 44 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product for artificial neural network (one-step and multi-step forecasting).

The forecasting results using artificial neural network from all conditions are quite variant. Hyperparameters such as ReLU and Swish activation functions, which help implement accurate models in all products, are studied. The input data of product A shows that using only the demand data can predict more accurate results compare to when external variables are added. On the contrary, the result of product B and C provide more accurate results with external variables. This situation may be resulted from the selection process. External variables selected may not correlate with the demand data and may not be suitable for the input data in the ANN. Thus, the result indicates more forecasting error when a lot of input data is added. It is arguable that one-step ahead is more effective than multi-step because it produces less predictive error values in product A. However, in product B and C, the multi-step forecasting model shows the higher performance of forecast results, which may be due to the pattern of the data used. The data is mostly from the repeated pattern used every year. This led to the outperformance of multi-step sliding window in both products.

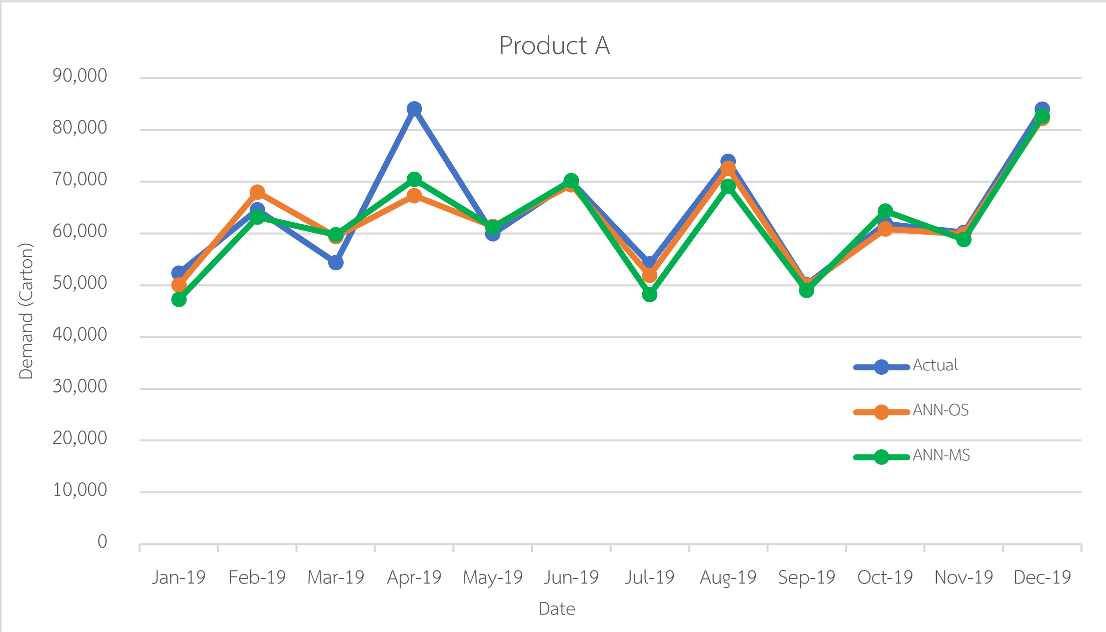

Figure 42: ANN-OS and ANN-MS forecasts compared with actual values of product A

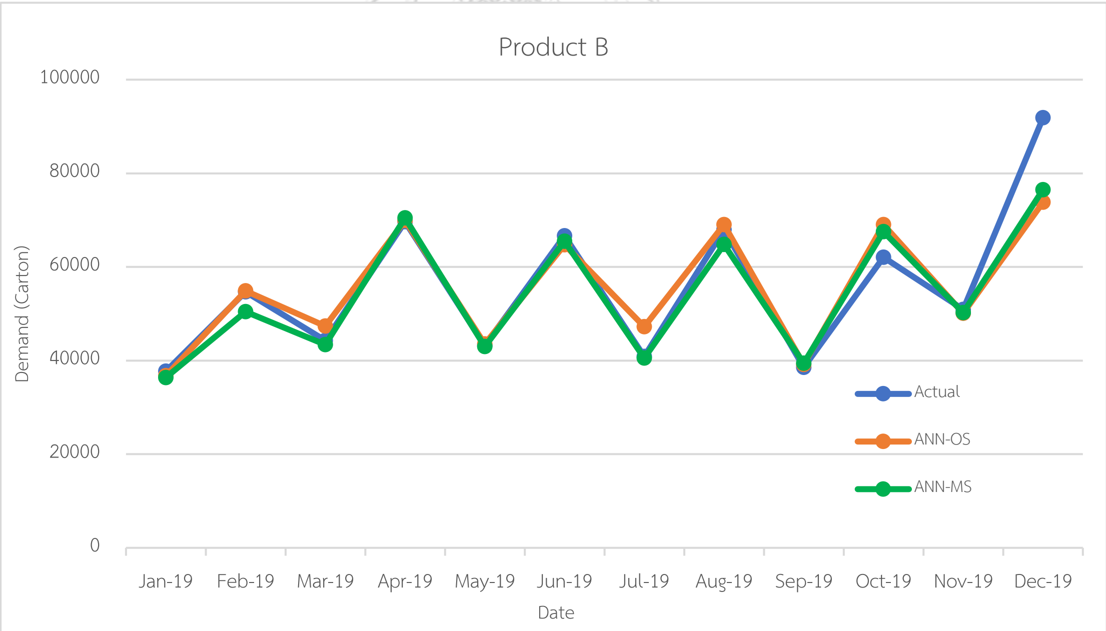

Figure 43: ANN-OS and ANN-MS forecasts compared with actual values of product B

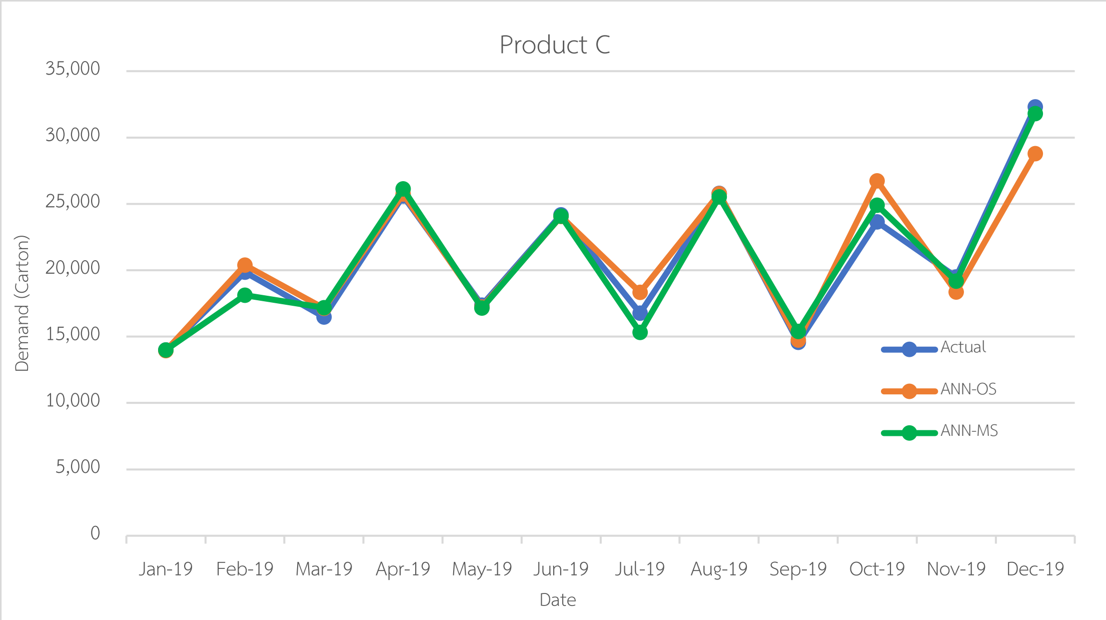

Figure 44: ANN-OS and ANN-MS forecasts compared with actual values of product C

4.1.5 Combined forecasting model

The combined forecasting method is a technique to increase efficiency in forecasting, which uses a combination of a single forecasting model to be calculated and the weight of each method in the experiment (Winkler & Makridakis, 1983). In this study, the simple average method is used. The calculations are also separated according to the scope of forecast: one-step and multi-step. According to the combination techniques, the model selection is made by choosing the best forecasting model from each group. They can be classified as follows: Exponential smoothing method, Box-Jenkins method, and Machine learning model. The MAPE of forecasting result from validation set, from each product are shown in table 22. Figure 45, 46, and 47 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product using combined forecasting model for one-step and multi-step forecasting.

Table 22: Measurement error of combined forecasting models

Measurement	Forecast period	Product A	Product B	Product C
MAPE	Short-term	6.80 %	7.55 %	6.80 %
	Long-term	6.77 %	6.59 %	5.69 %

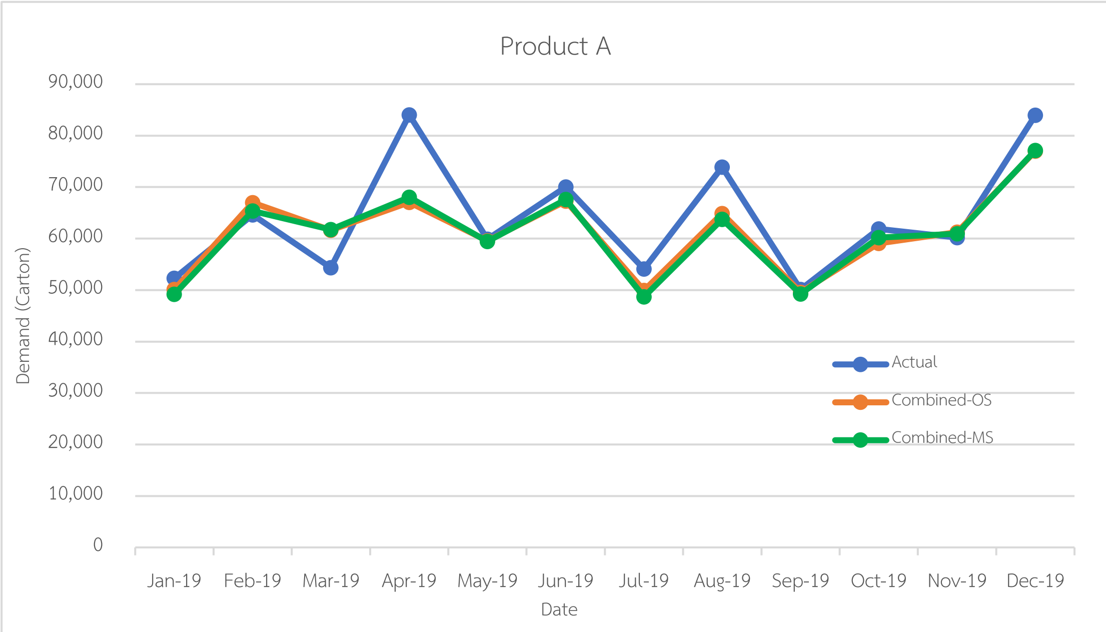

Figure 45: Combined-OS and Combined-MS forecasts compared with actual values of product A

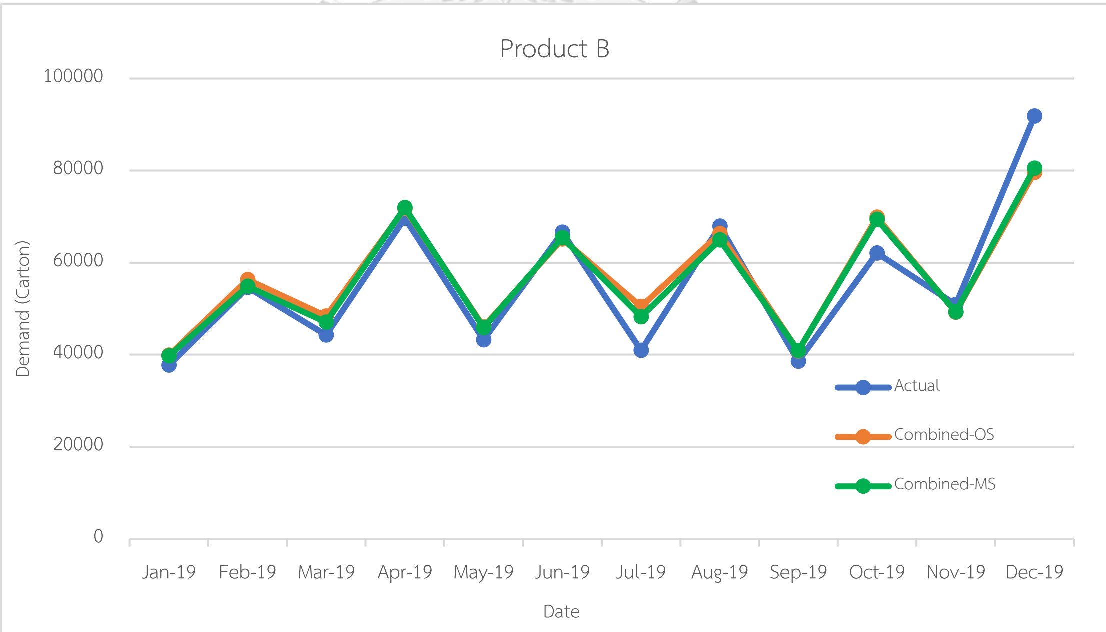

Figure 46: Combined-OS and Combined-MS forecasts compared with actual values of product B

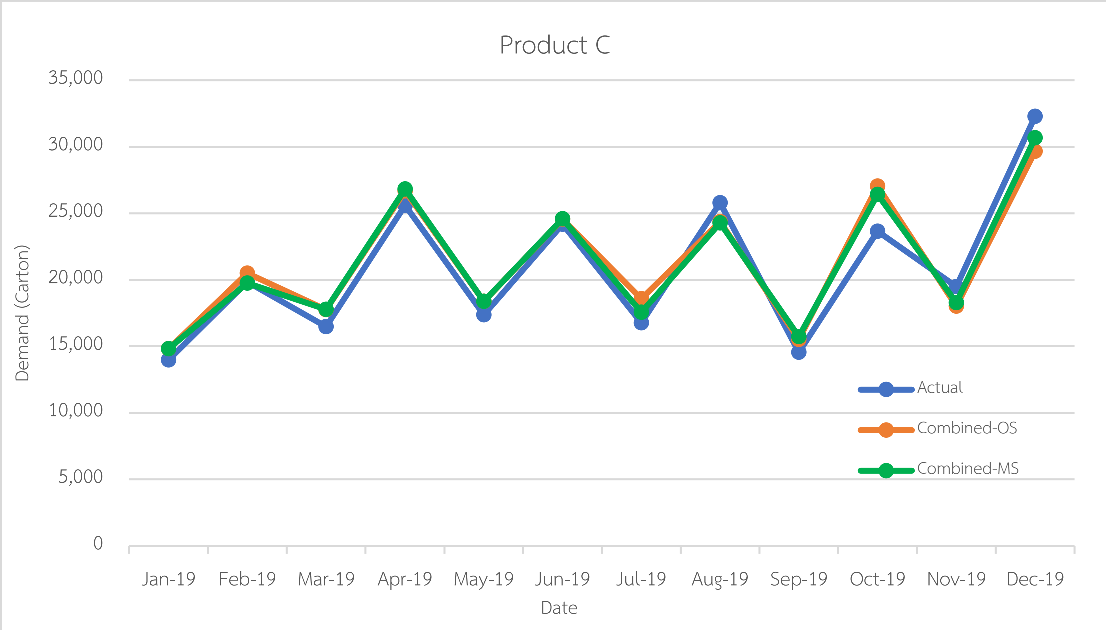

Figure 47: Combined-OS and Combined-MS forecasts compared with actual values of product C

The evaluated result of one-step ahead combined forecasting shows that the MAPE of product A, B, and C are 6.80%, 7.55%, and 6.80%, respectively. The multi-step ahead combined forecasting shows the MAPE of product A, B, and C at 6.80%, 6.59%, and 5.69%, respectively.

The results of combined forecasting indicate that the multi-step ahead forecast is more accurate since it shows lower MAPE. When it is combined with the forecast values, it can reduce error.

4.1.5 Hybrid forecasting model

Hybrid forecasting is another technique developed to increase forecasting efficiency using hybridization of a linear model and a non-linear model. This research studies two hybrid forecasting models, which are SARIMA-ANN and the proposed model (HW+SW-ANN). It is also added three experimental conditions used in the non-linear forecasting section: input data, hyperparameter, and scope of forecasting.

- SARIMA-ANN

SARIMA is a linear model and ANN is a non-linear model in this hybridization. Experimental results from variant conditions are shown in table 23 and 24.

Table 23: SARIMA-ANN's one-step hyperparameters and MAPE of each product (Short-term forecasting)

Parameter	Product A		Product B		Product C	
Input data	OG	OGEX	OG	OGEX	OG	OGEX
Model	1	7	1	12	1	7
Hidden layers	1	1	1	2	2	2
Hidden Units	137	112	85	148	129	54
Epochs	25	50	100	25	25	25
Batch size	4	32	8	32	32	32
Activation function	RR	RR	RR	RRR	RRR	RSR
MAPE	4.51%	5.58%	8.09%	6.85%	5.66%	6.02%

Table 24: SARIMA-ANN's multi-step hyperparameters and MAPE of each product (12 months in advance forecasting)

Parameter	Product A		Product B		Product C	
Input data	OG	OGEX	OG	OGEX	OG	OGEX
Model	5	10	1	10	3	11
Hidden layers	2	2	2	2	2	2
Hidden Units	112	107	132	149	84	82
Epochs	150	25	150	100	200	25
Batch size	32	2	32	2	4	16
Activation function	RSR	SSR	RRR	SSR	SRR	RSR
MAPE	7.07%	8.67%	9.02%	13.87%	4.17%	7.35%

The result of one-step SARIMA-ANN hybrid model using only original input indicates the MAPE of product A, B, and C at 4.51%, 8.09% and 5.66%, respectively. When this model is added with all external variables, the MAPE of product A, B, and C are at 5.58%, 6.85%, and 6.68%, respectively.

The results of the multi-step SARIMA-ANN using only original input show the MAPE of product A, B, and C at 7.07%, 8.60%, and 4.17%, respectively. Furthermore, the results of the original model with all external variable show the MAPE of product A, B, and C at 8.67%, 13.87% and 7.35%, respectively. Figure 48, 49, and 50 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product using SARIMA-ANN (one-step and multi-step forecasting).

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.

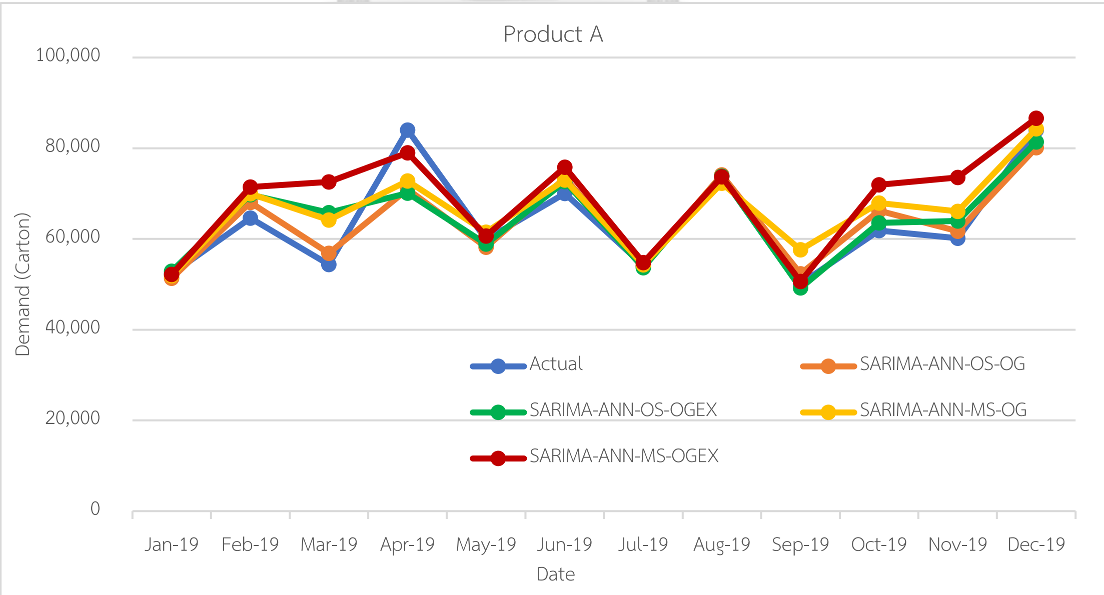

Figure 48: SARIMA-ANN-OS and SARIMA-ANN-MS forecasts compared with actual values of product A

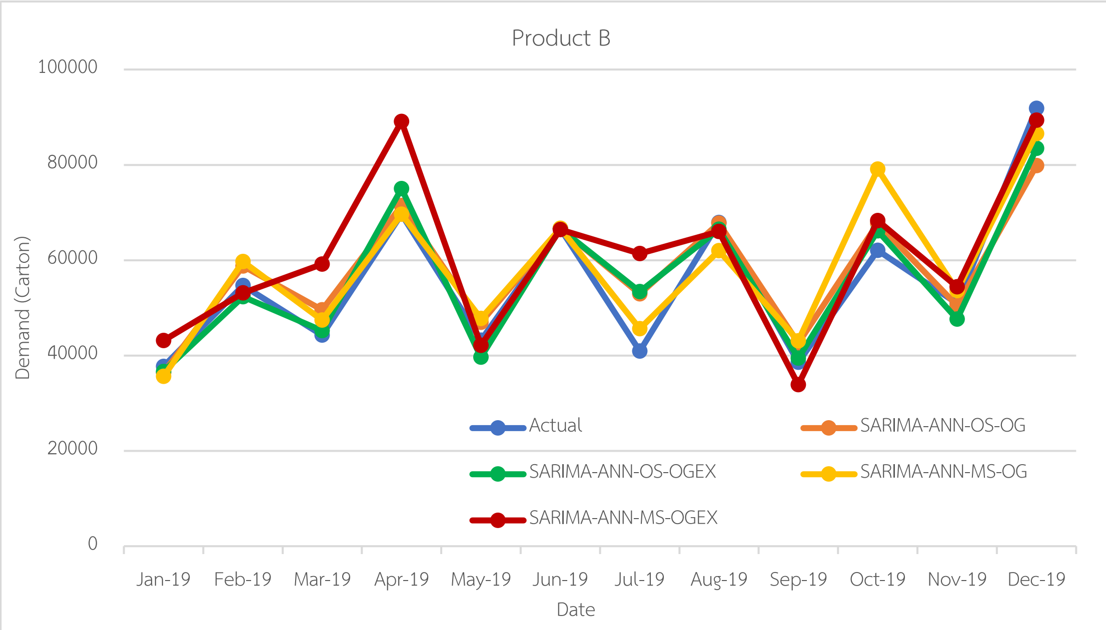

Figure 49: SARIMA-ANN-OS and SARIMA-ANN-MS forecasts compared with actual values of product B

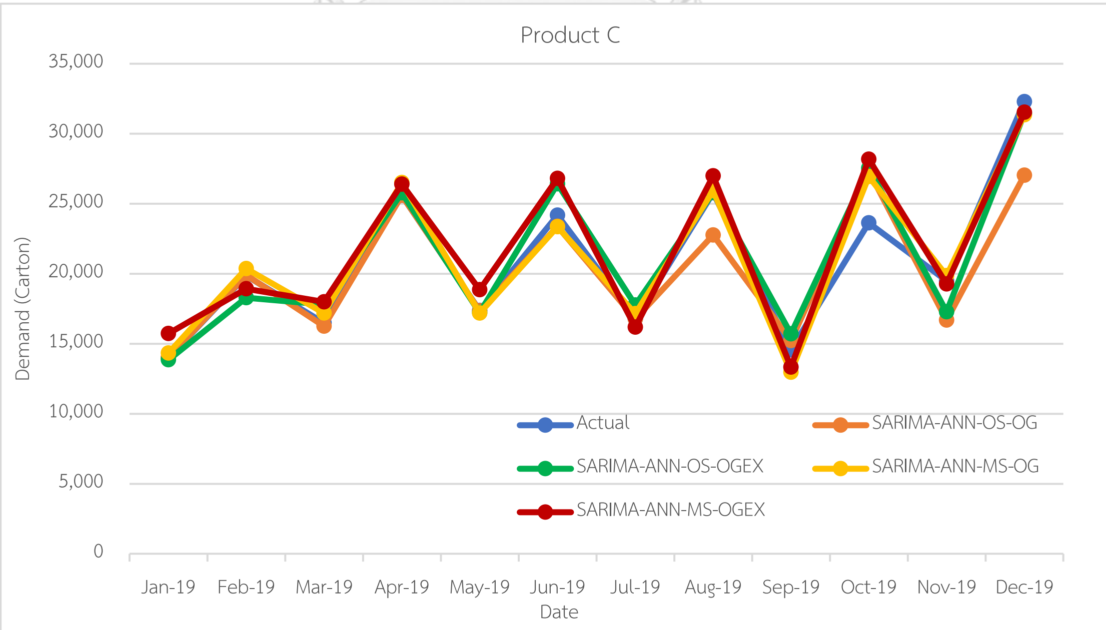

Figure 50: SARIMA-ANN-OS and SARIMA-ANN-MS forecasts compared with actual values of product C

- Proposed model (HW+SARIMA – ANN)

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 products. The forecasting results are shown in table 25 and 26.

Table 25: Proposed model's one-step hyperparameters and MAPE of each product (Short-term forecasting)

Parameter	Product A		Product B		Product C	
Input data	OG	OGEX	OG	OGEX	OG	OGEX
Model	1	7	3	12	2	10
Hidden layers	1	1	2	2	1	2
Hidden Units	66	79	147	69	52	55
Epochs	50	50	25	25	50	25
Batch size	16	32	4	32	32	32
Activation function	RR	SR	SRR	RRR	SR	SSR
MAPE	4.52%	4.96%	6.84%	5.96%	4.46%	5.67%

Table 26: Proposed model's multi-step hyperparameters and MAPE of each product (12 months in advance forecasting)

Parameter	Product A		Product B		Product C	
Input data format	OG	OGEX	OG	OGEX	OG	OGEX
Model	5	7	1	7	4	15
Hidden layers	2	1	1	1	2	2
Hidden Units	112	77	113	50	143	77
Epochs	100	50	100	25	25	100
Batch size	32	16	2	32	16	32
Activation function	RSR	SR	RR	SR	RSR	SRR
MAPE	5.70%	5.28%	7.91%	11.80%	6.58%	9.62%

The evaluated result of one-step ahead proposed model in original input shows the MAPE of product A, B, and C, which are 4.52%, 6.84%, and 4.46%, respectively. The original form with all external variable input shows the MAPE of product A, B, and C, which are 4.96%, 5.96%, and 5.67%, respectively.

The results of multi-step ahead proposed model using only original input shows the MAPE of product A, B, and C, which is 5.70%, 7.91%, and 6.58%, respectively. The original form with all external variable input shows the MAPE of product A, B, and C, which are 5.28%, 11.88%, and 9.62%, respectively.

Figure 51, 52, and 53 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product using the proposed model (one-step and multi-step forecasting).

The result of the proposed model indicates that there is a slight limitation in the forecasting effect since the use of hybrid techniques only enhances the accuracy of the linear model, i.e. time series model. It can be seen in the experiment results that form pre-hybrid values of product A, B, and C from the combination, may be factor affecting to this forecasting method. Therefore, the forecast value of the one-step proposed method can be low in MAPE when applying with the Hybrid. On the other hand, the multi-step proposed method does not bring about high-precision results. It is the long-term result of the forecasting when the input data is random pattern residuals, which are slightly difficult to achieve accurate results.

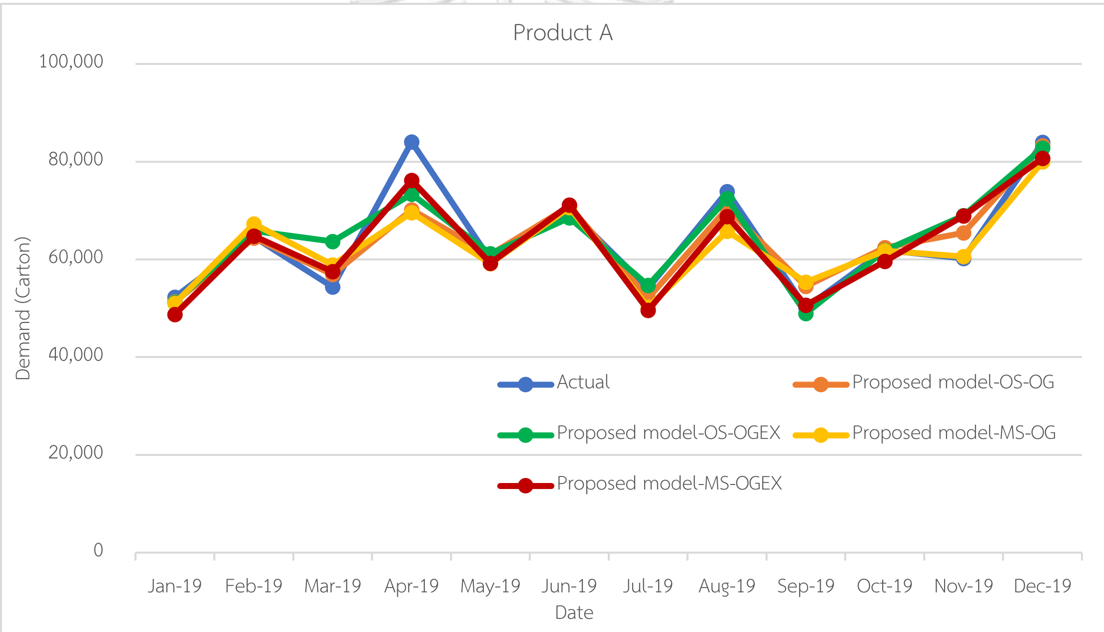

Figure 51: Proposed model-OS and Proposed model-MS forecasts compared with actual values of product A

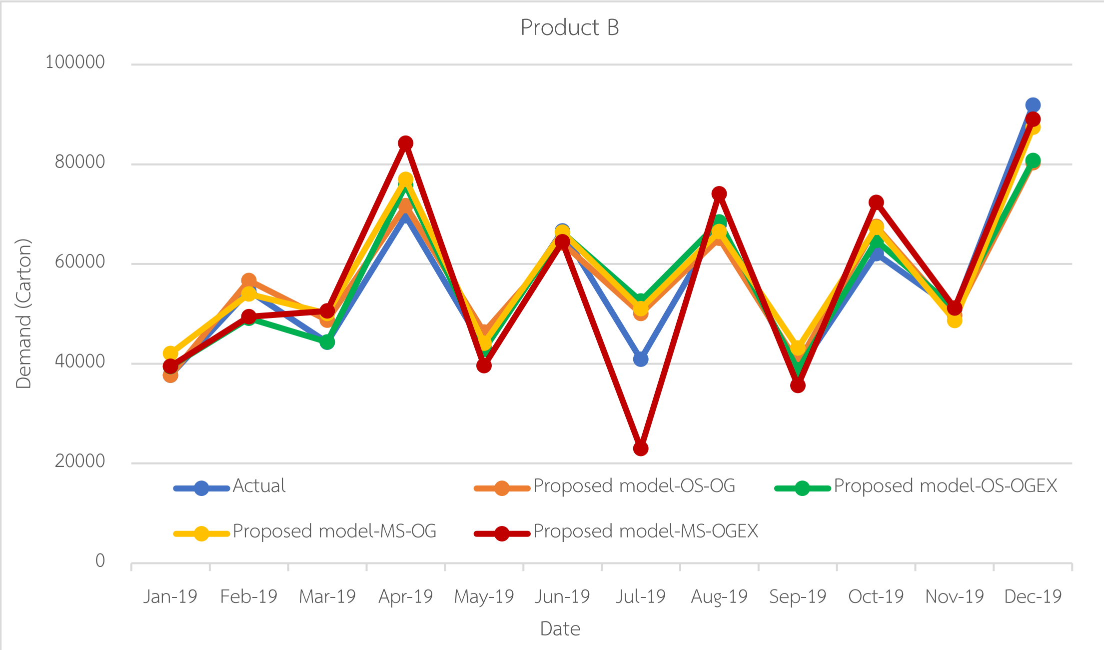

Figure 52: Proposed model-OS and Proposed model-MS forecasts compared with actual values of product B

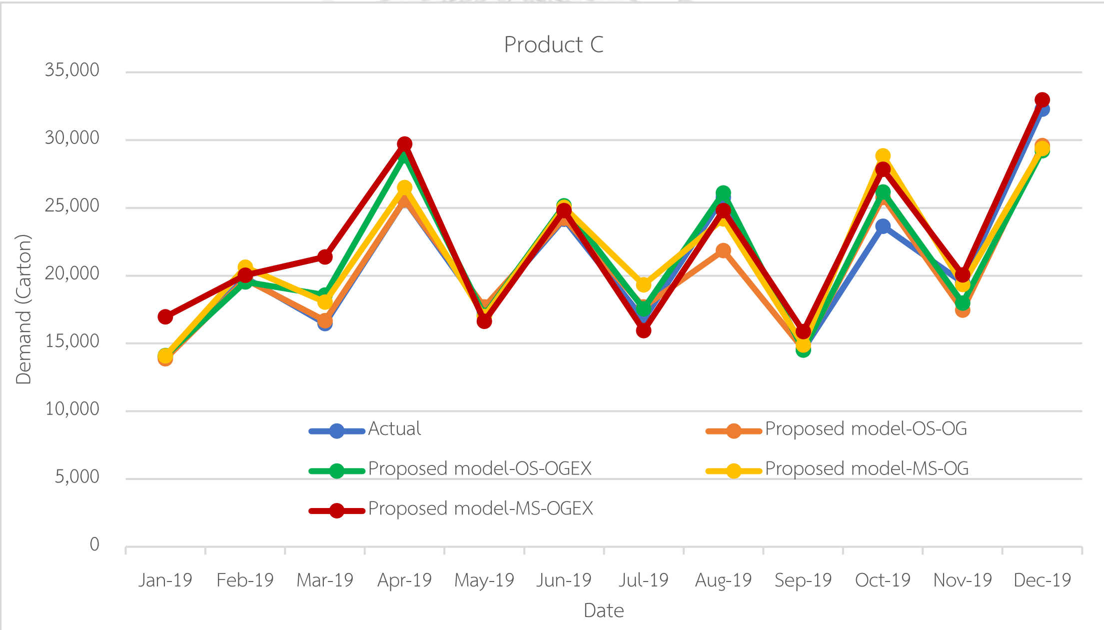

Figure 53: Proposed model-OS and Proposed model-MS forecasts compared with actual values of product C

4.2 Best Forecasting model selection

All accuracy result from all models are collected in this section including measuring the data validate using mean absolute percentage error as shown in table 27, 28, and 29.

Table 27: Measurement error of best forecasting model for product A

Group of forecasting model	Model	Result
Company	Current method	13.95%
Time series model	SARIMAX	7.06%
Machine learning model	ANN (OG-RRR-OS)	4.37%
	ANN (OG-RSR-MS)	5.65%
Combined model	Simple average (OS)	6.80%
	Simple average (MS)	6.77%
Hybrid model	SARIMA-ANN (OG-RR-OS)	4.51%
	HW+SARIMA-ANN (OGEX-SR-MS)	5.28%

Table 28: Measurement error of best forecasting model for product B

Group of forecasting model	Model	Result
Company	Current method	21.75%
Time series model	SARIMA	8.94%
Machine learning model	ANN (OGEX-SSR-OS)	5.43%
	ANN (OGEX-SSR-MS)	4.27%
Combined model	Simple average (OS)	7.55%
	Simple average (MS)	6.59%
Hybrid model	HW+SARIMA-ANN (OG-RSR-OS)	5.96%
	HW+SARIMA-ANN (OG-RR-MS)	7.91%

Table 29: Measurement error of best forecasting model for product C

Group of forecasting model	Model	Result
Company	Current method	21.20%
Time series model	SARIMA	8.94%
Machine learning model	ANN (OGEX-SR-OS)	4.11%
	ANN (OGEX-RR-MS)	3.40%
Combined model	Simple average (OS)	6.80%
	Simple average (MS)	5.69%
Hybrid model	HW+SARIMA-ANN (OG-SR-OS)	4.46%
	SARIMA-ANN (OG-SRR-MS)	4.17%

Note: OS = one-step, MS= multi-step, OG = original data and OGEX=original data with all external variables.

Table 27, 28, and 29 demonstrate all mean absolute percentage error of the studied forecasting models. The model accuracy comparison of all products shows that the artificial neural network using one-step technique in the original format give the most accurate result in product A, at 4.37% MAPE. For product B and C, the artificial neural network using one-step technique with all external variables shows MAPE at 5.43% and 4.11%, respectively. For multi-step, the proposed model is the most accurate model for product A at MAPE of 5.28%. The artificial neural network using multi-step technique in the original format with external variables also shows the most accurate results for product B and C at 4.27% and 3.40% MAPE, respectively.

Most results from the analysis using machine learning techniques are related to data such as data pattern and data relation. The results show that machine learning model is the most suitable forecasting model for the three products in general. The data pattern of all products has both trend and seasonality. When considering the demand of each year, it is found that the characteristics of data B and C were similar since these two products are sold together according to the company's marketing plan. These two data also have similar behaviors every year.

When ANN is applied to both products. it is capable of learning through the previous data. The result shows better accuracy of product B and C compared to product A, which was different pattern from the others.

It is noticeable, that the hybrid forecasting model only provides the best results for product A, which is arguably resulted from the residual data forecasting. This data pattern was random. The data with less randomity or fluctuation results in satisfying results for hybrid model. However, the residual data in product B and C was extremely random leading to inaccurate results of the ANN part in the hybrid forecasting model.

It can be seen that ANN forecasts are most accurate for overall products since this forecasting method directly utilizes the actual demand data. The inputs are added into the algorithm to make calculations based on actual demand, while the outputs are based on the algorithm. It is different from other models that require data to be calculated in other parts, which occurs in some cases as a limitation leading to less accurate results. However, the accuracy of each model is not significantly different.

4.3 Testing result from test data set

The most promising forecasting models are performed to predict product demands in 2020. Data testing uses test set, which is extracted from the data preparation. The result of evaluated accuracy is measured by the mean absolute percentage error. The results are compared with the company method. Both data validation and data testing results are shown in the table 30. Figure 54, 55 and 56 demonstrate predicted values of product A, B, and C compared with the actual demand data of each product with current method, short term and long term forecasting model in 2020.

Table 30: The data testing result in all products

Product	Forecasting Model	validate set	test set
A	Current method	13.95%	48.92%
	ANN (OG-RRR-OS)	4.37%	22.55%
	HW-SARIMA-ANN (OGEX-SR-MS)	5.28%	23.53%
B	Current method	21.75%	21.45%
	ANN (OGEX-SSR-OS)	5.43%	16.00%
	ANN (OGEX-SSR-MS)	4.27%	15.43%
C	Current method	21.20%	20.18%
	ANN (OGEX-SR-OS)	4.11%	19.21%
	ANN (OGEX-RR-MS)	3.40%	14.07%

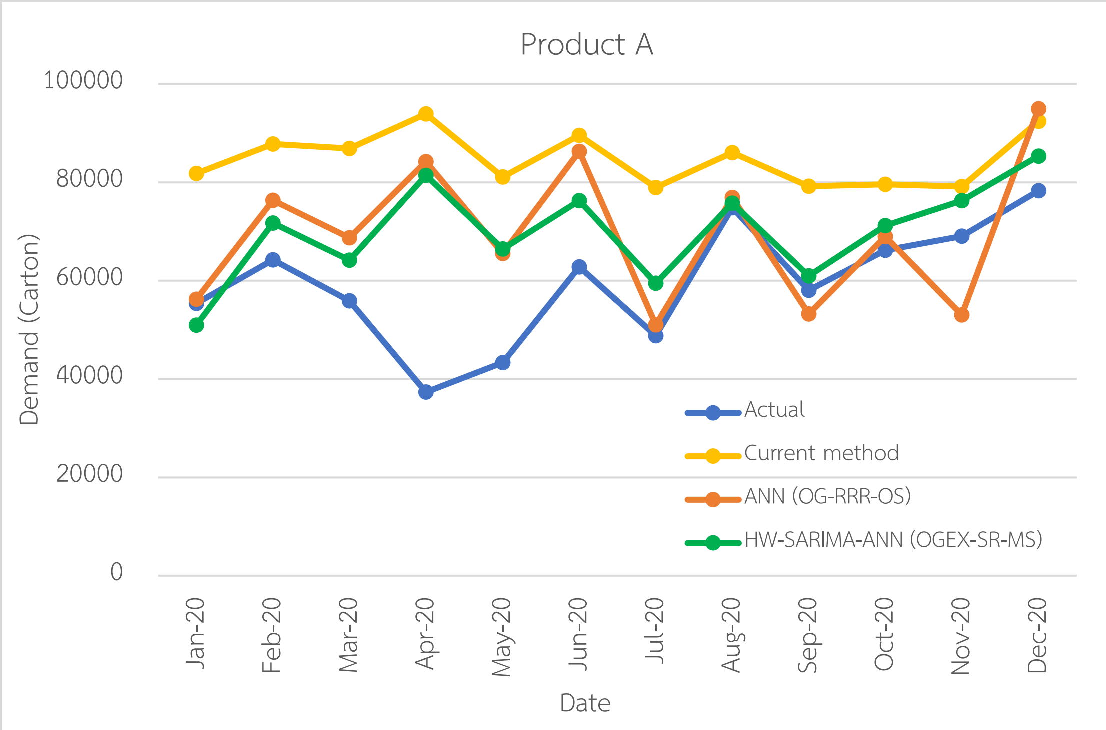

Figure 54: Comparison actual demand of product A with forecasting model in 2020

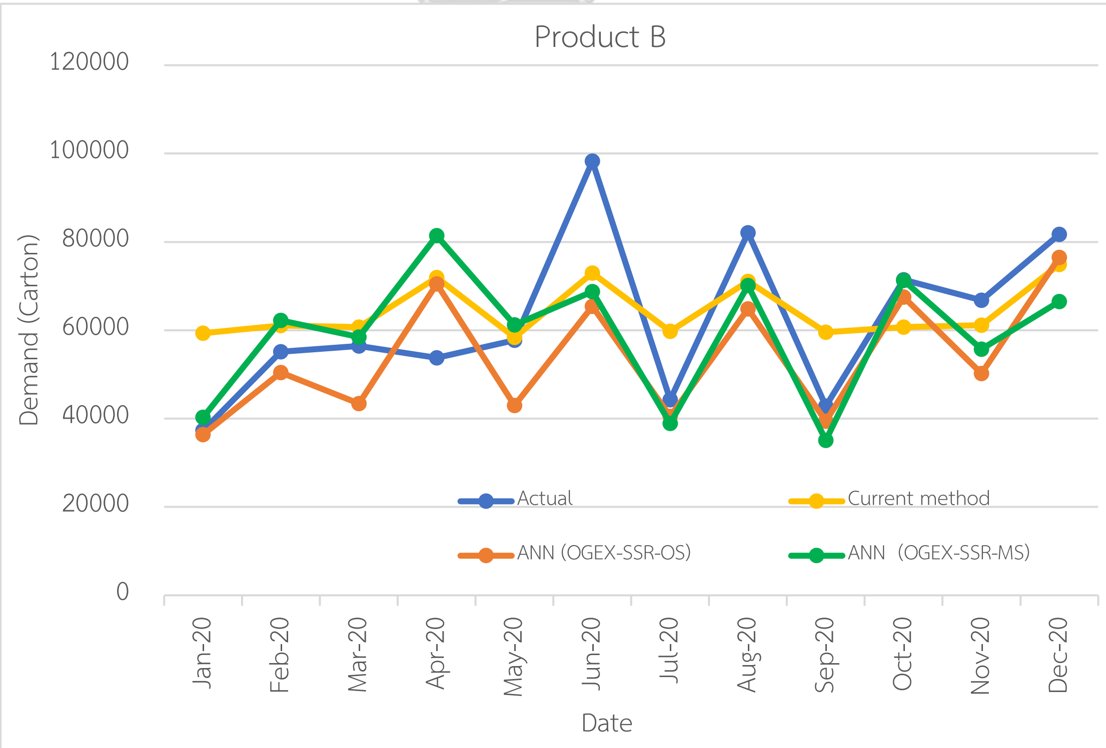

Figure 55: Comparison actual demand of product B with forecasting model in 2020

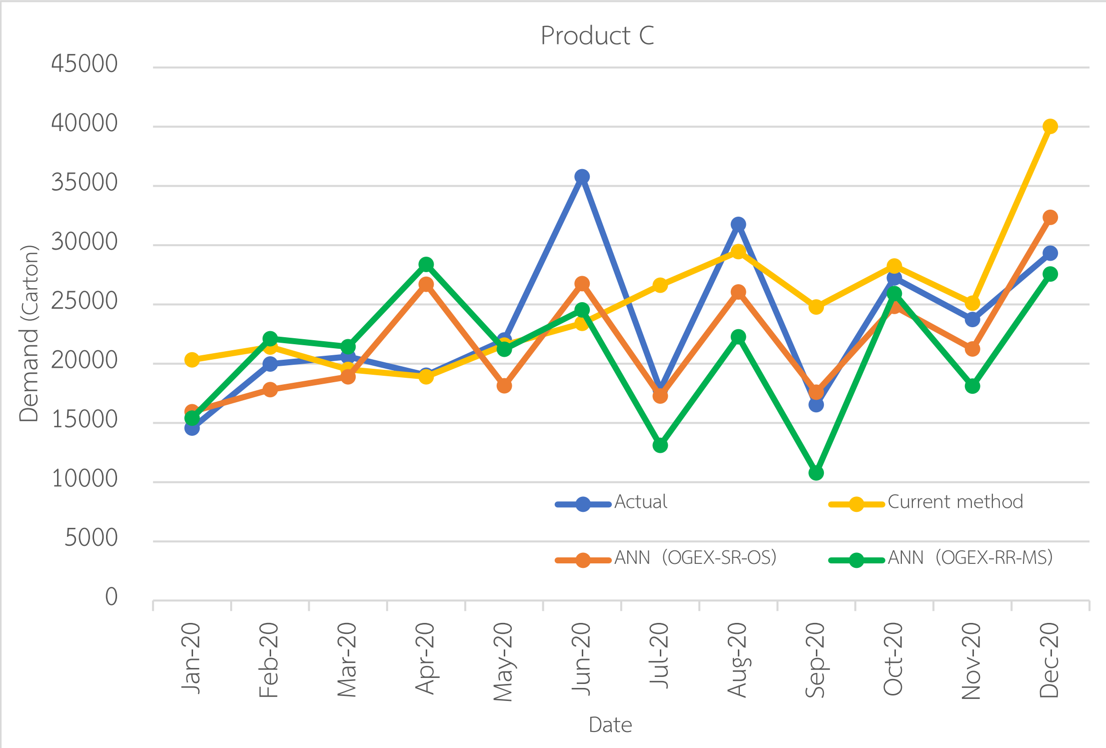

Figure 56: Comparison actual demand of product C with forecasting model in 2020

One-step ahead forecasting by the artificial neural network shows the most accurate result for product A with the MAPE at 22.55%. For product B and C, multi-step artificial neural network indicates the MAPE at 15.44% and 14.07%, respectively. When compared with the method currently used by the company, the development of forecasting models in the research yields better accuracy in all products. When compared the results of data testing to the data validation, the analysis shows that the MAPE is higher for every product due to the 2020 Covid-19 crisis occurring all over the world, including Thailand. This crisis is responsible for the change in behavior or data pattern. It is recommend that when there are other factors that only happen in the testsing data set. Addtitional adjustment from uses are required since those factors were never been seen in the training and validating data. However, the studied models can still provide much more accurate results compared to the current method of the case studied company.

Chapter 5 Conclusion and Recommendation

This research proposed forecasting models to predict the product demand for the case study company. The models were divided into four model groups: time series model, machine learning model, combined forecasting model, and hybrid forecasting model. The result from the models were compared among each other and also with the current company's method.

5.1 Conclusion

The food industry system is divided into several parts while all parts are nearly connected as called a supply chain. Most task planning in the factory is driven by the product demand, which depends on customer requirements. If the demand can be responded effectively, it will lead to a smooth operation in the next step. Response to demands on time requires experience and understanding of data. Another choice is to use forecasting principles to predict demands, which will happen in the future. If the forecast is accurate, it will affect the entire supply chain system to work more efficiently.

The forecastings of the three products: A, B, and C, were studied and compared. The data from 2013 to 2020 were used in this study. The first group of forecasting techniques was the time series model, which indicated that the SARIMAX model with external variables provided the most accurate forecast for product A. External variables were analyzed, screened, and selected using a stepwise technique. The selected variables were the month of year and population since they could be used with all products in the SARIMA. On the other hand, the time series model that gave the most accurate forecast for product B and C was SARIMA.

For machine learning model, this research selected an artificial neural network, which the three conditions were put into the model. Three types of input data were original, original with stepwise external variables, and original with all external variables. Original in this context means its own previous demand in the past. The second condition was hyperparameter. The last one was the scope of forecast, which was subdivided into one-step and multi-step. The study results initially showed that the one-step forecasting technique had higher accuracy for product A compared to multi-step. Conversely, for product B and C, multi-step forecasting provided better results. Furthermore, the hyperparameter, which was the main factor of the ANN was the activation function. Using ReLU and Swish, they provided low errors forecasting. The addition of Swish activation can increase accuracy in certain product or forecasting models. The results of all experimental conditions studied for ANN, which were single forecasting, combined forecasting, and hybrid forecasting model, showed that the most accurate model for short-term planning for product A was ANN (original data) at 4.37% MAPE. For short-term forecast, ANN (original data with all external variables) is most accurate for product B at 5.43% MAPE, ANN (original data with all external variables) is best for product C at 4.11% MAPE. For long-term forecast (12 months in advance), the most accurate model for product A is HW-SARIMA-ANN (original data with all external variables) at 5.28% MAPE. For product B and C are ANN (original data with all external variables) at MAPE of 4.27% and 3.40%, respectively.When comparing the data to the company method, it could be concluded that the artificial neural network model was an overall most accurate forecasting model. Even though the proposed hybrid model (HW-SARIMA-ANN) is most accurate for multi-step forecast for product A, it is not very different compared to ANN model. Unlike previous studies, this thesis considered both short-term and long-term forecasting for the same data setting.

5.2 Recommendations for future work

There were knowledge gap to develop in order to improve forecasting models. There were many research trying to study for deeper knowledge in this field. This research could be useful to those who are interested to learn more about this field of study. Possible extension are as follows.

1. Adding experiments on the hidden units by changing each the difference of the layers.
2. Adding other activation functions in the machine learning model.
3. Adding a combined model format using other techniques.
4. Trasformating input data to different values other than 0-1.
5. Finding both internal and external variables that may related to the data.
6. Creating pre-hybrid forecast values by another method before building the hybrid forecasting model.
7. Studying the combined forecasting using optimized weight techniques.

5.3 Recommendations for case study

Due to the COVID-19 crisis, A reason is that the error significantly increased in the data testing part. the changes in people's behavior and government policies amending the Covid-19 crisis also affected the data pattern making it different from what it could have been.

Owing to the large variations in data testing results, it was quite difficult to draw a conclusion. Therefore, it was suggested to select forecasting model from the results evaluated from the data validation set. The result of all forecasts in this study indicated in the same conclusion for all product demand data that the machine learning model demonstrated the most accurate result. It was mentioned that the results of this study would be applied to different departments of the companies, whether marketing planning or production planning, while these operations often require forecasting values in the time scope, which is farther than one-step or one month ahead can provide. Thus, this study suggested that the multi-step ahead artificial neural network accurate forecasting. Although it showed more error compared to one-step ahead forecasting, the advantage of this method was that it could forecast values for several months in advance. The artificial neural network was the most accurate forecasting model among others.

REFERENCES

Aburto, L., & Weber, R. (2007). Improved supply chain management based on hybrid demand forecasts. *Applied Soft Computing,* 7(1), 136-144. https://doi.org/10.1016/j.asoc.2005.06.001

Alameer, Z., Fathalla, A., Li, K., Ye, H., & Jianhua, Z. (2020). Multistep-ahead forecasting of coal prices using a hybrid deep learning model. *Resources Policy,* 65, 101588.

Arunraj, N. S., & Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting. *International Journal of Production Economics,* 170, 321-335. https://doi.org/10.1016/j.ijpe.2015.09.039

Athanasopoulos, G., Hyndman, R. J., Song, H., & Wu, D. C. (2011). The tourism forecasting competition. *International Journal of Forecasting,* 27(3), 822-844.

Bank of Thailand. (2018). *The Impact of exchange rates on the Thai Economy*. Monetary policy report. https://www.bot.or.th/Thai/MonetaryPolicy/MonetPolicyComittee/MPR/BOX_MPR/BOX_MPR_22561_2.pdf

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control* (Vol. 734). John Wiley & Sons.

Callen, T. (2020). *Gross Domestic Product: An Economy's All*. https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm

Chan, H. K., Xu, S., & Qi, X. (2019). A comparison of time series methods for forecasting container throughput. *International Journal of Logistics Research and Applications,* 22(3), 294-303.

Co, H. C., & Boosarawongse, R. (2007). Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks. *Computers & Industrial Engineering,* 53(4), 610-627. https://doi.org/10.1016/j.cie.2007.06.005

Cools, M., Moons, E., & Wets, G. (2009). Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: assessing the effect of holidays on two site locations. *Transportation research record,* 2136(1), 57-66.

Da Veiga, C. P., Da Veiga, C. R. P., Catapan, A., Tortato, U., & Da Silva, W. V. (2014). Demand forecasting in food retail: A comparison between the Holt-Winters and ARIMA models. *WSEAS transactions on business and economics,* 11(1), 608-614.

Demir, L., & Akkaş, S. (2018). A comparison of sales forecasting methods for a feed company: A case study. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi,* 24(4), 705-712.

Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. *Journal of Food Engineering,* 75(2), 196-204. https://doi.org/10.1016/j.jfoodeng.2005.03.056

Elinta, S. (2019). *FORECAST SEASONING PRODUCT AT PT. SA USING TIME SERIES METHOD AND NEURAL NETWORK* President University].

Euromonitor International. (2020). *SAUCES, DRESSINGS AND CONDIMENTSIN THAILAND -*

ANALYSIS. https://www.portal.euromonitor.com/portal/analysis/related

Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management,* 10, 1847979018808673.

Feizabadi, J. (2020). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, 1-24. https://doi.org/10.1080/13675567.2020.1803246

Fildes, R., Ma, S., & Kolassa, S. (2019). Retail forecasting: Research and practice. *International Journal of Forecasting*.

Food Interligence Center Thailand. (2019). *Thailand Food Industry Profile* 2019. http://fic.nfi.or.th/info_graphic_detail.php?id=28

Fritz, M., & Berger, P. D. (2015). Can you relate in multiple ways? Multiple linear regression and stepwise regression. In *Improving the User Experience Through Practical Data Analytics* (pp. 239-269). https://doi.org/10.1016/b978-0-12-800635-1.00010-0

Fróna, D., Szenderák, J., & Harangi-Rákos, M. (2019). The Challenge of Feeding the World. *Sustainability,* 11(20). https://doi.org/10.3390/su11205816

Ghani, I. M. M., & Ahmad, S. (2010). Stepwise Multiple Regression Method to Forecast Fish Landing. *Procedia - Social and Behavioral Sciences,* 8, 549-554. https://doi.org/10.1016/j.sbspro.2010.12.076

Gruss, B., & Kebhaj, S. (2019). Commodity Terms of Trade: A New Database.

GSB Research. (2019). *Food and beverage industry*. https://www.gsbresearch.or.th/gsb/economics/6520/

Güzin Tirkeş, C. G., Neş'e Çelebi. (2017). Demand forecasting: a comparison between the Holt-Winters, trend analysis and decomposition models. *Tehnicki vjesnik - Technical Gazette,* 24(Supplement 2). https://doi.org/10.17559/tv-20160615204011

Hanke, J. E., & Wichern, D. W. (2009). Business forecasting 9th ed. *New Jersey*.

Hota, H., Handa, R., & Shrivas, A. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research,* 13(5), 1145-1156.

Huang, H.-C. (2013). Construction of a Health Food Demand Prediction Model Using a Back Propagation Neural Network. *Advance Journal of Food Science and Technology,* 5(7), 896-899.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Kandananond, K. (2011). Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. *Energies,* 4(8), 1246-1257. https://doi.org/10.3390/en4081246

Kilcast, D., & Subramaniam, P. (2011). *Food and beverage stability and shelf life*. Elsevier.

Klimberg, R. K., Sillup, G. P., Boyle, K. J., & Tavva, V. (2010). Forecasting performance measures—What are their practical meaning? *Advances in business and management forecasting,* 7, 137-147.

Laouafi, A., Mordjaoui, M., Haddad, S., Boukelia, T. E., & Ganouche, A. (2017). Online electricity demand forecasting based on an effective forecast combination methodology. *Electric Power Systems Research,* 148, 35-47.

Li, P., & Zhang, J.-S. (2018). A new hybrid method for China's energy supply security forecasting based on arima and xgboost. *Energies,* 11(7), 1687.

Lijuan, W., & Guohua, C. (2016). Seasonal SVR with FOA algorithm for single-step and multi-step ahead forecasting in monthly inbound tourist flow. *Knowledge-Based Systems,* 110, 157-166.

Maaß, D., Spruit, M., & de Waal, P. (2014). Improving short-term demand forecasting for short-lifecycle consumer products with data mining techniques. *Decision analytics,* 1(1), 1-17.

Masters, D., & Luschi, C. (2018). Revisiting small batch training for deep neural networks. *arXiv preprint arXiv:*1804.07612.

Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Lalas, D., Moschovits, M., Karagiannis, F., & Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy,* 31(2-3), 208-227. https://doi.org/10.1016/j.energy.2005.02.016

Mitra, P., Ray, R., Chatterjee, R., Basu, R., Saha, P., Raha, S., Barman, R., Patra, S., Biswas,

S. S., & Saha, S. (2016). Flood forecasting using Internet of things and artificial neural networks. 2016 IEEE 7th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),

Moeeni, H., & Bonakdari, H. (2017). Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model. *Stochastic environmental research and risk assessment,* 31(8), 1997-2010.

Moreno, J. J. M. (2011). Artificial neural networks applied to forecasting time series. *Psicothema,* 23(2), 322-329.

Munkhdalai, L., Munkhdalai, T., Park, K. H., Lee, H. G., Li, M., & Ryu, K. H. (2019). Mixture of activation functions with extended min-max normalization for forex market prediction. *IEEE Access,* 7, 183680-183691.

Priyadarshi, R., Panigrahi, A., Routroy, S., & Garg, G. K. (2019). Demand forecasting at retail stage for selected vegetables: a performance analysis. *Journal of Modelling in Management*.

Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. *arXiv preprint arXiv:*1710.05941.

Riansut, W. (2016). Forecasting Model for the Export Values of Rubber Wood and Furniture of Thailand/ตัว แบบ พยากรณ์ มูลค่า การ ส่ง ออก ไม้ ยางพารา แปรรูป และ เฟอร์นิเจอร์ ของ ประเทศไทย. *Naresuan University Journal: Science and Technology (NUJST),* 24(3), 108-122.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory* *to algorithms*. Cambridge university press.

Sharma, S., & Sharma, S. (2017). Activation functions in neural networks. *Towards Data Science,* 6(12), 310-316.

Silva, J. C., Figueiredo, M. C., & Braga, A. C. (2019). Demand forecasting: A case study in the food industry. International Conference on Computational Science and Its Applications,

Sivrikaya, O., & Tunç, E. (2013). Demand forecasting for domestic air transportation in Turkey. *The Open Transportation Journal,* 7(1).

Smith, G. (2018). Step away from stepwise. *Journal of Big Data,* 5(1), 32.

Suzuki, K. (2011). *Artificial neural networks: methodological advances and biomedical applications*. BoD–Books on Demand.

Thailand Board of Investment. (2019). *Thailand's Food Industry*. https://www.boi.go.th/upload/content/Food%20industry_5aa7b40bd758b.pdf

The Economist. (2007). *The Economist Guide to Economic Indicators: Making Sense of Economics* (6 ed.). Profile Books.

The Stock Exchange of Thailand. (2015). *Stock index*. https://www.set.or.th/education/th/begin/stock_content07.pdf

Thomassey, S., Happiette, M., & Castelain, J. M. (2002). An automatic textile sales forecast using fuzzy treatment of explanatory variables. *Journal of Textile and*

Apparel, Technology and Management, 2(4), 1-12.

Tsoumakas, G. (2019). A survey of machine learning techniques for food sales prediction. *Artificial Intelligence Review,* 52(1), 441-447.

Vagropoulos, S. I., Chouliaras, G., Kardakos, E. G., Simoglou, C. K., & Bakirtzis, A. G. (2016). Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting. 2016 IEEE International Energy Conference (ENERGYCON),

Vimolchalao, S. (2003). Optimal Level of International Reserves: The Case of Thailand, Presentation at the Bank of Thailand Monthly Workshop.

Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society: Series A (General),* 146(2), 150-157.

Wittayaorn-puripunpinyoo, A., CHANDASIRI, V., & KONGKAEW, S. (2017). Ready to Drink Milk as a Nutritional Tool for Human Capital Development of Thailand. *International Journal of Environmental and Rural Development,* 8(2), 117-121.

Xu, B., Wang, N., Chen, T., & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:*1505.00853.

Yamarone, R. (2017). *The economic indicator handbook: How to evaluate economic trends to maximize profits and minimize losses*. John Wiley & Sons.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing,* 50, 159-175.

Zhu, Y., Zhao, Y., Zhang, J., Geng, N., & Huang, D. (2019). Spring onion seed demand forecasting using a hybrid Holt-Winters and support vector machine model. *PloS one,* 14(7), e0219889.

Zou, H., Xia, G., Yang, F., & Wang, H. (2007). An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. *Neurocomputing,* 70(16-18), 2913-2923.

APPENDIX A

List of abbreviation

Abbreviations	Terms
ANN	Artificial neural network
CPI	Consumer index price
DC	Distribution center
EX	Exchange rate
GDP	Gross domestic product
HW	Holt-Winters exponential smoothing method
IR	International reserve
MAD	Mean absolute deviation
MS	Multi-step
MSE	Mean squared error
MAPE	Mean absolute percentage error
OG	Original data
OGEX	Original data with all external variables
OS	One-step
POP	Population
RMSE	Root mean squared error
RR	Activation function ReLU-ReLU
RRR	Activation function ReLU-ReLU-ReLU
RSR	Activation function ReLU-Swish-ReLU
SARIMA	Seasonal Autoregressive Integrated Moving average
SARIMAX	Seasonal Autoregressive Integrated Moving average with exogenous
SET	The stock exchange of Thailand
SR	Activation function Swish-ReLU

Abbreviations	Terms
SRR	Activation function Swish-ReLU-ReLU
SSR	Activation function Swish-Swish-ReLU
TT	Terms of trade
UR	Unemployment rate

APPENDIX B

Artificial neural network – one-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	80	100	2	RR	5.19%
Original	Model 2	1	91	25	32	SR	7.07%
Original	Model 3	2	126	25	2	SRR	5.23%
Original	Model 4	2	135	25	32	SSR	5.90%
Original	Model 5	2	135	25	2	RSR	4.42%
Original	Model 6	2	135	25	2	RRR	4.37%
OG+stepwise external	Model 7	1	131	25	2	RR	6.40%
OG+stepwise external	Model 8	1	116	25	2	SR	6.82%
OG+stepwise external	Model 9	2	127	25	2	SRR	5.98%
OG+stepwise external	Model 10	2	73	25	16	SSR	6.65%
OG+stepwise external	Model 11	2	149	25	4	RSR	5.73%
OG+stepwise external	Model 12	2	124	25	4	RRR	6.10%
OG+all external	Model 13	1	73	100	2	RR	6.10%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+all external	Model 14	1	85	25	2	SR	6.99%
OG+all external	Model 15	2	141	150	32	SRR	6.38%
OG+all external	Model 16	2	61	25	2	SSR	6.76%
OG+all external	Model 17	2	73	25	2	RSR	6.01%
OG+all external	Model 18	2	93	100	2	RRR	6.65%

Artificial neural network – multi-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	71	200	2	RR	6.74%
Original	Model 2	1	107	50	32	SR	9.61%
Original	Model 3	2	112	150	2	SRR	6.12%
Original	Model 4	2	112	25	16	SSR	8.19%
Original	Model 5	2	142	150	4	RSR	5.65%
Original	Model 6	2	101	100	2	RRR	7.07%
OG+stepwise external	Model 7	1	65	25	16	RR	8.36%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+stepwise external	Model 8	1	101	25	16	SR	7.75%
OG+stepwise external	Model 9	2	143	200	4	SRR	9.11%
OG+stepwise external	Model 10	2	58	25	32	SSR	9.56%
OG+stepwise external	Model 11	2	134	200	4	RSR	8.00%
OG+stepwise external	Model 12	2	74	200	2	RRR	8.31%
OG+all external	Model 13	1	105	100	4	RR	8.05%
OG+all external	Model 14	1	93	100	4	SR	7.49%
OG+all external	Model 15	2	73	200	16	SRR	8.80%
OG+all external	Model 16	2	112	200	16	SSR	7.88%
OG+all external	Model 17	2	91	200	4	RSR	8.09%
OG+all external	Model 18	2	96	50	4	RRR	7.88%

Artificial neural network – one-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	68	50	2	RR	7.76%
Original	Model 2	1	50	25	8	SR	8.34%
Original	Model 3	2	94	100	32	SRR	8.25%
Original	Model 4	2	50	25	16	SSR	8.55%
Original	Model 5	2	93	50	4	RSR	7.80%
Original	Model 6	2	59	25	4	RRR	8.43%
OG+stepwise external	Model 7	1	53	100	32	RR	6.99%
OG+stepwise external	Model 8	1	64	50	16	SR	7.57%
OG+stepwise external	Model 9	2	110	25	32	SRR	7.79%
OG+stepwise external	Model 10	2	67	25	16	SSR	7.87%
OG+stepwise external	Model 11	2	75	25	16	RSR	7.38%
OG+stepwise external	Model 12	2	69	50	4	RRR	7.57%
OG+all external	Model 13	1	111	50	32	RR	5.78%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+all external	Model 14	1	139	100	4	SR	5.99%
OG+all external	Model 15	2	105	25	16	SRR	5.47%
OG+all external	Model 16	2	55	100	16	SSR	5.43%
OG+all external	Model 17	2	128	25	32	RSR	5.60%
OG+all external	Model 18	2	119	25	16	RRR	5.77%
Original	Model 1	1	68	50	2	RR	7.76%

Artificial neural network –multi-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	113	25	2	RR	7.88%
Original	Model 2	1	110	25	2	SR	8.25%
Original	Model 3	2	112	25	32	SRR	7.25%
Original	Model 4	2	134	200	8	SSR	7.80%
Original	Model 5	2	149	50	8	RSR	8.40%
Original	Model 6	2	147	25	32	RRR	7.37%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+stepwise external	Model 7	1	146	100	8	RR	5.42%
OG+stepwise external	Model 8	1	137	25	4	SR	5.82%
OG+stepwise external	Model 9	2	85	25	16	SRR	5.19%
OG+stepwise external	Model 10	2	84	25	4	SSR	5.74%
OG+stepwise external	Model 11	2	53	50	4	RSR	5.41%
OG+stepwise external	Model 12	2	147	25	8	RRR	5.86%
OG+all external	Model 13	1	150	100	2	RR	5.42%
OG+all external	Model 14	1	95	25	2	SR	5.94%
OG+all external	Model 15	2	98	25	4	SRR	6.94%
OG+all external	Model 16	2	52	50	8	SSR	4.27%
OG+all external	Model 17	2	52	25	2	RSR	6.91%
OG+all external	Model 18	2	105	200	2	RRR	6.10%

Artificial neural network – one-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	61	150	4	RR	6.55%
Original	Model 2	1	70	100	2	SR	8.23%
Original	Model 3	2	97	150	32	SRR	7.01%
Original	Model 4	2	94	100	4	SSR	8.28%
Original	Model 5	2	70	100	4	RSR	6.61%
Original	Model 6	2	76	50	16	RRR	7.24%
OG+stepwise external	Model 7	1	92	25	32	RR	6.31%
OG+stepwise external	Model 8	1	73	25	8	SR	7.16%
OG+stepwise external	Model 9	2	84	25	16	SRR	7.23%
OG+stepwise external	Model 10	2	140	150	2	SSR	7.13%
OG+stepwise external	Model 11	2	97	25	32	RSR	6.44%
OG+stepwise external	Model 12	2	78	25	16	RRR	7.04%
OG+all external	Model 13	1	106	25	8	RR	4.91%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+all external	Model 14	1	113	100	4	SR	4.11%
OG+all external	Model 15	2	105	25	16	SRR	4.82%
OG+all external	Model 16	2	51	100	16	SSR	4.74%
OG+all external	Model 17	2	93	25	16	RSR	4.58%
OG+all external	Model 18	2	119	25	16	RRR	4.74%

Artificial neural network – multi-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	132	50	32	RR	7.76%
Original	Model 2	1	114	50	8	SR	8.84%
Original	Model 3	2	132	50	16	SRR	8.82%
Original	Model 4	2	83	100	32	SSR	9.94%
Original	Model 5	2	139	50	32	RSR	8.46%
Original	Model 6	2	73	50	32	RRR	8.62%
OG+stepwise external	Model 7	1	130	100	16	RR	3.40%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
OG+stepwise external	Model 8	1	91	25	8	SR	4.70%
OG+stepwise external	Model 9	2	65	25	4	SRR	4.64%
OG+stepwise external	Model 10	2	126	25	8	SSR	4.42%
OG+stepwise external	Model 11	2	97	50	32	RSR	3.86%
OG+stepwise external	Model 12	2	95	25	8	RRR	4.55%
OG+all external	Model 13	1	123	200	32	RR	5.91%
OG+all external	Model 14	1	95	25	2	SR	5.39%
OG+all external	Model 15	2	56	25	4	SRR	5.51%
OG+all external	Model 16	2	52	50	8	SSR	4.58%
OG+all external	Model 17	2	55	50	8	RSR	6.90%
OG+all external	Model 18	2	103	200	8	RRR	6.98%

Hybrid SARIMA - ANN – one-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	137	25	4	RR	4.51%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	95	25	4	SR	4.96%
Original	Model 3	2	71	25	4	SRR	4.92%
Original	Model 4	2	141	25	32	SSR	5.04%
Original	Model 5	2	61	150	16	RSR	5.50%
Original	Model 6	2	149	25	8	RRR	5.38%
OG+all external	Model 7	1	112	50	32	RR	5.58%
OG+all external	Model 8	1	65	200	32	SR	7.94%
OG+all external	Model 9	2	66	25	32	SRR	7.50%
OG+all external	Model 10	2	66	25	32	SSR	7.06%
OG+all external	Model 11	2	104	25	32	RSR	5.62%
OG+all external	Model 12	2	128	50	32	RRR	7.62%

Hybrid SARIMA - ANN –multi-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	108	150	4	RR	8.13%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	71	150	2	SR	7.41%
Original	Model 3	2	107	150	16	SRR	8.88%
Original	Model 4	2	112	25	4	SSR	7.57%
Original	Model 5	2	112	150	32	RSR	7.07%
Original	Model 6	2	112	25	4	RRR	10.89%
OG+all external	Model 7	1	77	200	32	RR	10.53%
OG+all external	Model 8	1	107	25	4	SR	9.15%
OG+all external	Model 9	2	111	100	32	SRR	12.18%
OG+all external	Model 10	2	107	25	2	SSR	8.67%
OG+all external	Model 11	2	92	50	8	RSR	10.12%
OG+all external	Model 12	2	64	100	32	RRR	10.72%

Hybrid SARIMA - ANN – one-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	85	100	8	RR	8.09%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	148	200	2	SR	8.80%
Original	Model 3	2	121	25	2	SRR	8.16%
Original	Model 4	2	123	100	8	SSR	8.98%
Original	Model 5	2	72	25	8	RSR	9.12%
Original	Model 6	2	64	25	4	RRR	8.19%
OG+all external	Model 7	1	141	50	32	RR	9.02%
OG+all external	Model 8	1	125	50	32	SR	7.30%
OG+all external	Model 9	2	90	25	32	SRR	7.71%
OG+all external	Model 10	2	117	25	32	SSR	7.05%
OG+all external	Model 11	2	80	25	32	RSR	7.27%
OG+all external	Model 12	2	148	25	32	RRR	6.85%

Hybrid SARIMA - ANN – multi-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	108	200	8	RR	9.02%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	116	200	32	SR	11.15%
Original	Model 3	2	137	100	2	SRR	9.19%
Original	Model 4	2	114	50	16	SSR	11.86%
Original	Model 5	2	123	25	16	RSR	10.15%
Original	Model 6	2	132	150	32	RRR	8.60%
OG+all external	Model 7	1	68	150	32	RR	20.88%
OG+all external	Model 8	1	102	100	4	SR	20.18%
OG+all external	Model 9	2	106	150	2	SRR	14.47%
OG+all external	Model 10	2	149	100	2	SSR	13.87%
OG+all external	Model 11	2	107	200	8	RSR	16.44%
OG+all external	Model 12	2	122	200	8	RRR	19.58%

Hybrid SARIMA - ANN – one-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	145	50	32	RR	5.66%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	117	25	16	SR	5.78%
Original	Model 3	2	63	50	32	SRR	5.81%
Original	Model 4	2	64	50	32	SSR	5.82%
Original	Model 5	2	54	100	32	RSR	5.73%
Original	Model 6	2	129	25	32	RRR	5.66%
OG+all external	Model 7	1	100	25	32	RR	6.68%
OG+all external	Model 8	1	64	25	16	SR	7.48%
OG+all external	Model 9	2	85	150	2	SRR	8.07%
OG+all external	Model 10	2	72	25	32	SSR	7.20%
OG+all external	Model 11	2	54	25	32	RSR	6.02%
OG+all external	Model 12	2	102	25	32	RRR	7.86%

Hybrid SARIMA - ANN – multi-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	132	50	32	RR	6.64%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	108	50	32	SR	8.44%
Original	Model 3	2	84	200	4	SRR	4.17%
Original	Model 4	2	124	200	2	SSR	7.73%
Original	Model 5	2	73	150	2	RSR	7.28%
Original	Model 6	2	115	150	4	RRR	10.07%
OG+all external	Model 7	1	50	50	16	RR	7.91%
OG+all external	Model 8	1	129	50	16	SR	10.54%
OG+all external	Model 9	2	51	200	2	SRR	8.91%
OG+all external	Model 10	2	64	25	8	SSR	7.89%
OG+all external	Model 11	2	82	25	16	RSR	7.35%
OG+all external	Model 12	2	85	25	8	RRR	9.51%

Hybrid proposed model – one-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	66	50	16	RR	4.52%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	82	25	16	SR	5.24%
Original	Model 3	2	66	25	16	SRR	6.34%
Original	Model 4	2	140	25	16	SSR	6.09%
Original	Model 5	2	140	25	32	RSR	6.15%
Original	Model 6	2	75	25	16	RRR	6.07%
OG+all external	Model 7	1	82	100	32	RR	5.10%
OG+all external	Model 8	1	79	50	32	SR	4.96%
OG+all external	Model 9	2	90	25	8	SRR	5.36%
OG+all external	Model 10	2	94	25	32	SSR	5.47%
OG+all external	Model 11	2	84	50	16	RSR	5.45%
OG+all external	Model 12	2	76	25	16	RRR	5.56%

Hybrid proposed model – multi-step ahead forecasting (Product A)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	138	100	32	RR	13.02%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	71	25	2	SR	6.74%
Original	Model 3	2	112	50	32	SRR	9.63%
Original	Model 4	2	112	50	32	SSR	5.83%
Original	Model 5	2	112	100	32	RSR	5.70%
Original	Model 6	2	67	50	32	RRR	8.59%
OG+all external	Model 7	1	107	25	8	RR	6.13%
OG+all external	Model 8	1	77	50	16	SR	5.28%
OG+all external	Model 9	2	65	50	8	SRR	5.99%
OG+all external	Model 10	2	85	50	32	SSR	5.39%
OG+all external	Model 11	2	111	50	32	RSR	5.37%
OG+all external	Model 12	2	111	50	32	RRR	6.32%

Hybrid proposed model – one-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	113	100	2	RR	8.16%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	150	100	2	SR	7.85%
Original	Model 3	2	147	25	4	SRR	6.84%
Original	Model 4	2	97	100	2	SSR	8.12%
Original	Model 5	2	77	100	2	RSR	8.62%
Original	Model 6	2	80	100	2	RRR	8.64%
OG+all external	Model 7	1	125	25	32	RR	8.48%
OG+all external	Model 8	1	131	100	8	SR	7.10%
OG+all external	Model 9	2	99	25	32	SRR	7.08%
OG+all external	Model 10	2	58	25	32	SSR	6.71%
OG+all external	Model 11	2	94	25	16	RSR	6.66%
OG+all external	Model 12	2	69	25	32	RRR	5.96%

Hybrid proposed model – multi-step ahead forecasting (Product B)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	113	100	2	RR	7.91%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	116	25	32	SR	8.14%
Original	Model 3	2	128	25	32	SRR	10.58%
Original	Model 4	2	86	25	32	SSR	8.42%
Original	Model 5	2	86	25	8	RSR	9.95%
Original	Model 6	2	101	150	8	RRR	8.12%
OG+all external	Model 7	1	102	25	32	RR	12.47%
OG+all external	Model 8	1	50	25	32	SR	11.80%
OG+all external	Model 9	2	90	25	32	SRR	14.08%
OG+all external	Model 10	2	60	25	16	SSR	12.46%
OG+all external	Model 11	2	50	25	32	RSR	12.44%
OG+all external	Model 12	2	114	50	32	RRR	13.26%

Hybrid proposed model – one-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	55	25	4	RR	4.84%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	52	50	32	SR	4.46%
Original	Model 3	2	83	25	2	SRR	6.20%
Original	Model 4	2	132	150	16	SSR	6.11%
Original	Model 5	2	67	25	8	RSR	5.75%
Original	Model 6	2	85	100	32	RRR	6.31%
OG+all external	Model 7	1	51	25	16	RR	6.99%
OG+all external	Model 8	1	133	25	32	SR	6.64%
OG+all external	Model 9	2	99	25	16	SRR	7.20%
OG+all external	Model 10	2	55	25	32	SSR	5.67%
OG+all external	Model 11	2	60	25	32	RSR	6.76%
OG+all external	Model 12	2	70	25	32	RRR	6.32%

Hybrid proposed model – multi-step ahead forecasting (Product C)

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 1	1	81	25	16	RR	7.01%

Input data format	Model	Hidden layers	Hidden Units	Epochs	Batch size	Activation function	MAPE
Original	Model 2	1	116	25	32	SR	8.34%
Original	Model 3	2	102	100	16	SRR	9.99%
Original	Model 4	2	114	25	32	SSR	8.45%
Original	Model 5	2	143	25	16	RSR	6.58%
Original	Model 6	2	123	100	32	RRR	9.21%
OG+all external	Model 7	1	58	25	8	RR	10.57%
OG+all external	Model 8	1	102	25	32	SR	11.62%
OG+all external	Model 9	2	77	100	32	SRR	9.62%
OG+all external	Model 10	2	55	25	16	SSR	10.05%
OG+all external	Model 11	2	140	50	32	RSR	11.51%
OG+all external	Model 12	2	114	25	32	RRR	11.04%

VITA

NAME	Woratanon Sirimak
DATE OF BIRTH	21 March 1994
PLACE OF BIRTH	Chon buri, Thailand
INSTITUTIONS ATTENDED	Bachelor of Science from the Faculty of Agro-Industry, Kasetsart University Master of Engineering from the Faculty of Engineering Chulalongkorn University
HOME ADDRESS	11/4 Huay Yai Banglamung Chon buri 20150
AWARD RECEIVED	First class honor in Bachelor Degree