DEMAND FORECASTING FOR EXPORT AND IMPORT WAREHOUSE BY USING ARTIFICIAL NEURAL NETWORK: A CASE STUDY OF PHARMACEUTICAL WAREHOUSE

WANLOP FUANGFOO

**A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF ENGINEERING
(INDUSTRIAL ENGINEERING)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2014**

Thesis
entitled

DEMAND FORECASTING FOR EXPORT AND IMPORT WAREHOUSE BY USING ARTIFICIAL NEURAL NETWORK: A CASE STUDY OF PHAMARCEUTICAL WAREHOUSE

…………………………………………
Mr. Wanlop Fuangfoo
Candidate

…………………………………………
Asst. Prof**.** Waressara Weerawat,
Ph.D. (Industrial Engineering)
Major advisor

…………………………………………
Asst. Prof. Thanakorn Naenna,
Ph.D. (Engineering Science)
Co-advisor

…………………………………………
Lect. Jirapan Liangrokapart,
Ph.D. (Industrial Engineering)
Co-advisor

……………………………………
Assoc.Prof. Sombat Thanawan, Ph.D.,
Acting Dean
Faculty of Graduate Studies
Mahidol University

…………………………………………
Asst.Prof. Thanakorn Naenna,
Ph.D. (Engineering Science)
Program Director
Master of Engineering Program
In Industrial Engineering
Faculty of Engineering,
Mahidol University

Thesis
entitled

DEMAND FORECASTING FOR EXPORT AND IMPORT WAREHOUSE BY USING ARTIFICIAL NEURAL NETWORK: A CASE STUDY OF PHAMARCEUTICAL WAREHOUSE

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Engineering (Industrial Engineering)
on
April 09,2014

……………………………………………
Mr. Wanlop Fuangfoo
Candidate

……………………………………………
Lect. Waranyu Wongseree,
Ph.D. (Electrical Engineering)
Chair

……………………………………………
Asst. Prof**.** Waressara Weerawat,
Ph.D. (Industrial Engineering)
Member

……………………………………
Asst.Prof. Thanakorn Naenna,
Ph.D. (Engineering Science)
Member

……………………………………………
Lect. Jirapan Liangrokapart ,
Ph.D. (Industrial Engineering)
Member

……………………………………
Assoc.Prof. Sombat Thanawan, Ph.D.,
Acting Dean
Faculty of Graduate Studies
Mahidol University

……………………………………………
Lect. Worawit Isarangkul,
M.S. (Technical Management)
Dean
Faculty of Engineering
Mahidol University

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and gratefulness to my advisor, Asst. Prof. Waressara Weerawat, for her guidance, supervision, valuable suggestion and encouragement throughout the course of this research. Besides my advisor, it is pleasure to thank to my co-advisors, Lect. Waranyu Wongseree who provided me with many helpful comments on the research as it developed without all of his help this work might not have been completed.

I would like to thank to my friends and staffs at Industrial Engineering department of Mahidol University for their kindness, support and any cooperation during this study.

Finally, I would like to deeply express my most gratefulness to my family for their encouragement and moral support that have inspired me to achieve this research.

Wanlop Fuangfoo

DEMAND FORECASTING FOR EXPORT AND IMPORT WAREHOUSE BY USING ARTIFICIAL NEURAL NETWORK: A CASE STUDY OF PHARMACEUTICAL WAREHOUSE

WANLOP FUANGFOO 5236761 EGIE/M

M.Eng. (INDUSTRIAL ENGINEERING)

THESIS ADVISORY COMMITTEE: WARESSARA WEERAWAT, Ph.D., THANAKORN NAENNA, Ph.D., JIRAPAN LIANGROKAPART, Ph.D.

ABSTRACT

Import or Export Warehouses are comparative distribution centers of international trade. In order to be managed, these warehouses face plenty of challenges and complications, above typical domestic ones. In the receiving and put away process, it takes a long time of approximately 70% of the total operation time. This is due to the fact that a lack of information about the actual amount of incoming products being received. Hence, the main purpose of this research study is to find an appropriate forecasting model in order to estimate the actual incoming products received, that is suitable for an import and export warehouse. This study highlights only medical devices or pharmaceutical products as the representative case samplings because in terms of market sectors, medical or pharmaceuticals contribute 1 out of 6 sections of the total national consumer goods. This research used 3 methodologies comparatively; Time Delay Neural Network (TDNN), Box-Jenkins Model (ARIMA), and Hybrid Model. The study defined that the TDNN model provides the best accuracy in forecasting, indicated by the least deviation. The model provides the most accurate results against the other models. Therefore, with efficiency forecasting model development, the processing time in a warehouse is reduced by 22%, and the cost of the operation of the activities is reduced by 20%.

KEY WORDS: TIME DELAY NEURAL NETWORK/ HYBRID MODEL/FORECASTING/ IMPORT WAREHOUSE

90 pages

การพยากรณ์ความต้องการสินค้าสำหรับคลังสินค้านำเข้า และส่งออก โดยใช้โครงข่ายใยประสาทเทียม กรณีศึกษาคลังยาและเวชภัณฑ์

DEMAND FORECASTING FOR EXPORT AND IMPORT WAREHOUSE BY USING ARTIFICIAL NEURAL NETWORK: A CASE STUDY OF PHARMACEUTICAL WAREHOUSE

วัลลภ เฟื่องฟู 5236761 EGIE/M

วศ.ม. (วิศวกรรมอุตสาหการ)

คณะกรรมการที่ปรึกษาวิทยานิพนธ์: วเรศรา วีระวัฒน์, Ph.D., ธนกรณ์ แน่นหนา, Ph.D., จีรพรรณ เลี่ยงโรคาพาธ, Ph.D.

บทคัดย่อ

คลังสินค้าเพื่อการนำเข้าและส่งออก เปรียบเสมือนได้กับเป็นหัวใจหลักในการค้าขายระหว่างประเทศ การบริหารจัดการคลังสินค้ามีความท้าทายและยุ่งยากกว่าคลังสินค้าภายในประเทศทั่วไปโดยเฉพาะอย่างยิ่งการรับสินค้าเข้าจัดเก็บในคลังสินค้าซึ่งเกิดปัญหาในการดำเนินงานเนื่องจากการขาดข้อมูลปริมาณสินค้าที่แท้จริงล่วงหน้า ดังนั้นในงานวิจัยชิ้นนี้มีวัตถุประสงค์ที่จะศึกษาหาตัวแบบพยากรณ์ที่เหมาะสม และแม่นยำ มาพยากรณ์ปริมาณสินค้านำเข้าและส่งออกที่ถูกส่งมาเก็บในคลังสินค้า การศึกษาครั้งนี้ได้เลือกใช้กลุ่มข้อมูลจากสินค้ากลุ่มยาและเวชภัณฑ์ซึ่งมีสัดส่วน 1 ใน 6 ของปริมาณสินค้าอุปโภคบริโภคทั้งประเทศมาเป็นข้อมูลในการศึกษา โดยเลือกใช้วิธีการทางโครงข่ายใยประสาทแบบ Time delay (Time delay neural network :TDNN) เปรียบเทียบกับตัวแบบบ็อกเจนกินส์ (Box-Jenkins) ซึ่งเป็นวิธีการทางอนุกรมเวลา และตัวแบบผสม (Hybrid) ซึ่งเป็นวิธีการที่ผสมวิธีการทั้งสองเข้าไว้ด้วยกัน จากผลการศึกษาพบว่าตัวแบบพยากรณ์ TDNN สามารถพยากรณ์ให้ค่าพยากรณ์ที่มีความแม่นยำมากที่สุดโดยวัดจากค่าความคลาดเคลื่อนที่น้อยที่สุดทั้งในช่วงการพยากรณ์ระยะสั้น 1 เดือน และระยะยาว 3 เดือน จากการพยากรณ์ที่แม่นยำขึ้นส่งผลให้เกิดการลดเวลาในกระบวนการทำงานโดยเวลาได้ 22% ของเวลาการดำเนินงานการรับสินค้าเข้าจัดเก็บภายในคลังสินค้าทั้งสิ้น หรือ สามารถลดค่าใช้จ่ายในการดำเนินงานได้เฉลี่ย 20%

90 หน้า

CONTENTS

CONTENTS (cont.)

CONTENTS (cont.)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (cont.)

LIST OF FIGURES (cont.)

CHAPTER I
INTRODUCTION

1.1 Background and Problem statement

Thailand as the medical hub of Asia since 2004, the consumption rate of pharmaceutical products in Thailand has been increasing continuously due to the increasing of population and the increment of chronic disease impact to human health. Especially the imported products for example pharmaceutical of the antibiotics, cardiovascular system, and digestive system are highly needs for patient demand (http://www.hsri.or.th/news/804). According to the statistic recorded from Food and Drug Administration, Ministry of Public Health, regarding to pharmaceutical consumption and medical device usage rate shows that the comparison of domestic manufactured compared to imported medicine from the past 20 years, the imported pharmaceutical results in higher consumption rate rather than domestic manufacturing product. The graph in figure 1.1 shows that the consumption rate in Thailand in year 1990, the percentage amount of imported medicine is represented only 20% from the total market compared to year 2011 the amount of imported medicine is increased to 63% from the total market.

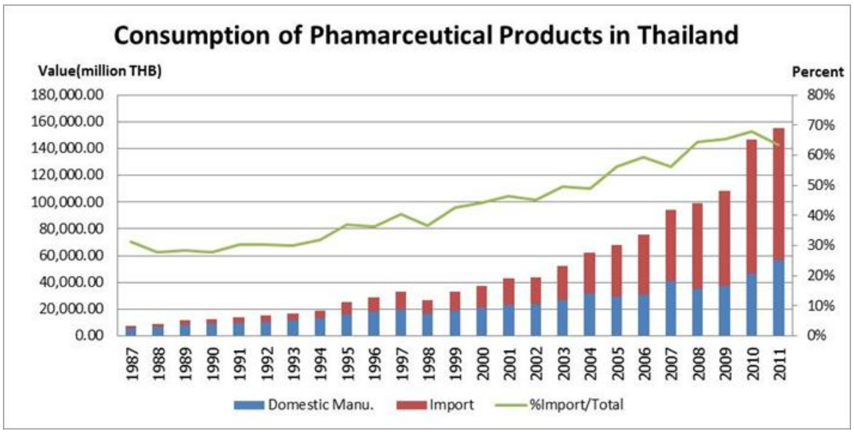
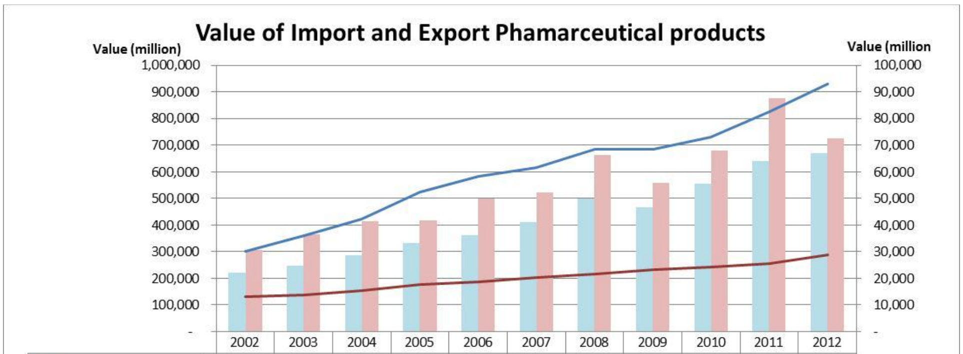


Figure 1.1 Consumption of Pharmaceutical products in Thailand

Source: Food and Drug Administration, Ministry of Public Health

However, in figure 1.2, the graph shows the value of import and export pharmaceutical products is increasing every year in different ratios approximately 10% each year for import product, whereas exporting product increased 8% each year.



	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Import_ConsumerGoods	222,14	247,58	287,62	331,27	362,24	409,02	498,29	464,73	555,85	640,68	670,56
Export_ConsumerGoods	305,418	365,038	414,523	418,070	499,675	522,532	662,229	559,459	679,719	875,661	724,266
Import_Pharma Products	29,992	36,022	42,173	52,238	58,417	61,384	68,374	68,303	72,991	82,591	92,849
Export_Pharma Products	13,021	13,619	15,209	17,505	18,752	20,226	21,439	23,346	24,314	25,544	28,632

Figure 1.2 Consumption of Pharmaceutical products in Thailand

Source: Department of Foreign Trade, Ministry of Commerce

In healthcare supply chain perspective, warehouse or distribution center is the most important member responsible for importing or exporting product. Warehouse is the connection between customers which consists of retail pharmacies, hospitals and clinics throughout the country link with pharmacies' manufacturer. The process starts with a customer order to purchase drug through sale representatives. Sale representative will create a sale order ordering their drug through the warehouse that has several channels for ordering such as sending an ordering document, phone ordering, or send an order online via website. Then warehouse is delivering to customers upon their ordered.

Meanwhile, pharmaceutical companies are collect historical sales data and forecasting sale in the future according to lead time for order processing lead time and delivery lead time. Most of forecasting data is forecast data within 3-5 months and inform the order in advance to the manufacturing department in order to planning the production order and ship to Thailand through the international shipping by air or ocean as shown in Figure 1.3.

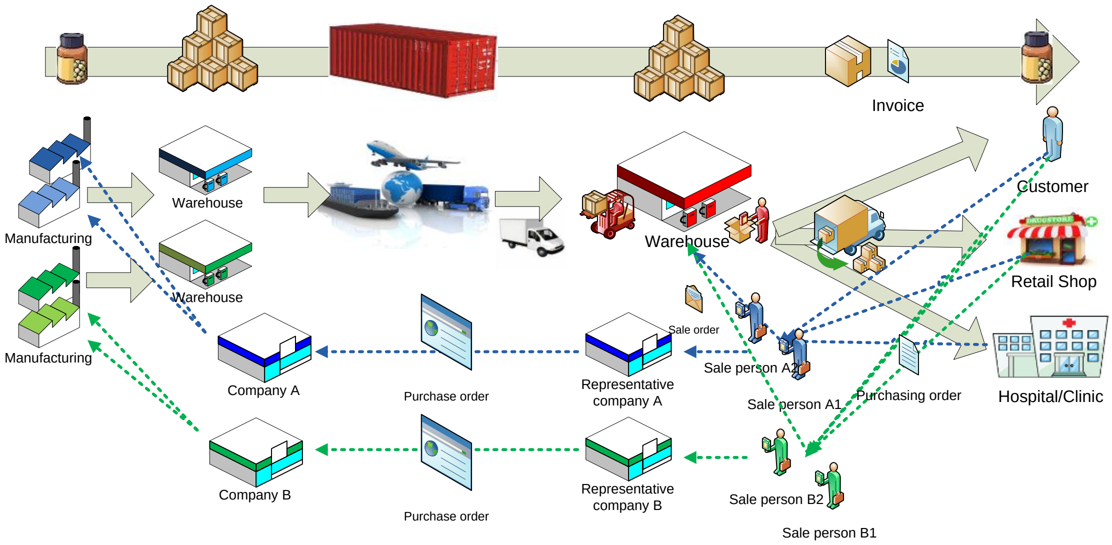


Figure 1.3 Supply chain of pharmaceutical warehouse

The challenge of warehouse management for importing and exporting phamarcuetical products is much higher than a typical warehouse management for consumer goods, e.g.

- Fragility from a container of the products, normally contained in vial or glass, therefore it is possible to break
- Examination before acceptance, checking whether the products are damaged or contaminated by germs or fungi in every cartons.
- Strict temperature control or even more challenging with various conditions of acceptable ranges, for instance; 2-8, 20-25, or below -2 degree celcius
- GMP and company standard in storaging products are categorized by serial numbers and separated by categories, for example, Cytotoxic or Narcotics must be handled with special care

Warehouse new incoming shipment process is to accommodate the large volume of incoming shipments from multiple lots of manufacturing in each company and consolidate to delivery at one time to warehouse because of a minimal shipping charge. Therefore, it has many products together in a single pallet or container. When goods are delivered to warehouses, warehouse had to unload the shipment and make repallet from shipment pallet to warehouse pallet. Sometime new shipment come only 2-3 pallets, warehouse may have to spread on to new pallets up to 20 pallets before receiving process and update in system and put away goods to storage location.

In this put-away process, the inventory will put away product to designated areas. If space is not sufficient, a management team dissolved pallet space by other products or goods to other areas to make space for the new product. Or sometimes replenish team have to pick and replenish goods to picking area to make the available space as details in work flow diagrams in figure 1.4.

Therefore, the preparation for the delivery of new shipment for receiving process is a guideline to prevent delays in the put away process. But for the supply chain perspective warehouse and manufactures company are lack of IT system linking the information together. They usually do not know the shipment product in advance in each time. Or sometimes they only inform just the quantity of a particular product or number of pallets used to transport. The unknown of the SKU number and batch number of each product that necessary to use to calculate the amount of pallets required and space for storage reservation.

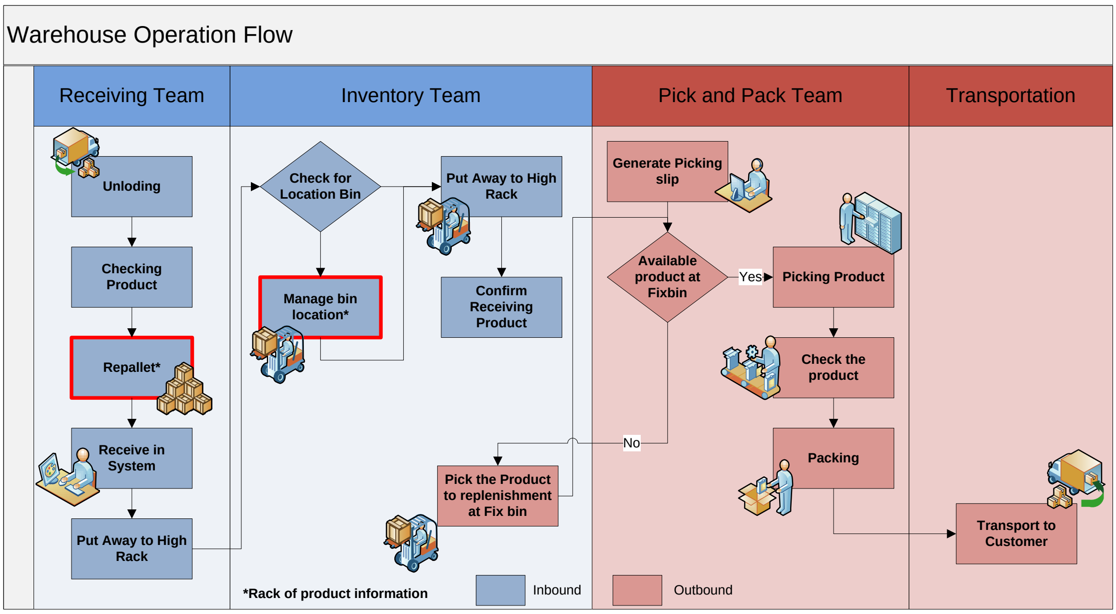


Figure 1.4 Operation flow in pharmaceutical warehouse

From the problem, the biggest pharmaceutical warehouse's company in Thailand is selected as a case study in this research. This warehouse is capable to receive 11,500 products per month or around 25,000 pallets per month from 250 pharmaceutical companies from both Thailand and worldwide shipment. The most important problem occur in this receiving process in the warehouse is the bottleneck problem at receiving area which takes long processing lead time. The reason is that the lack of incoming shipment information for example the real quantity of the incoming shipment, so that warehouse cannot be able to calculate the actual pallet needs in the warehouse. Unfortunately, the proportion of put way product to storage location time used 72% from total operation time in warehouse compared to the whole process time used.The average put away time per pallet is approximately 3.5 minute ,therefore warehouse need to operate for total 26.2 hour per shipment which is very time consumed as shown in figure 1.5.

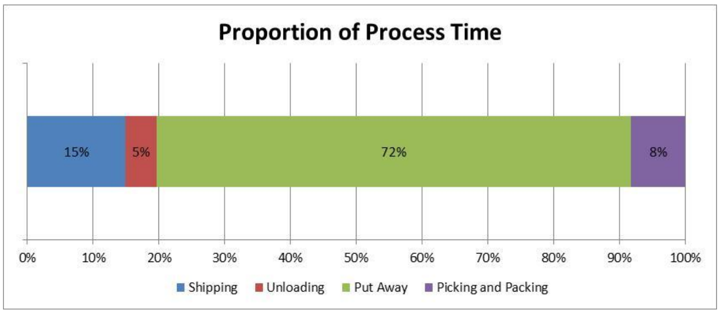


Figure 1.5 Proportion of process time

The precisely forecasting should be applied and implemented in warehouse operation to improve the efficiency of warehouse management. In the past, forecast was used for monitor the trend of sales value or monitors the growing of market which used for making the decision for product ordering. Warehouse did not use the forecasting technique to forecast the quantities of product for preparing. Moreover, warehouses hardly know the actual amount of products especially the detail of product batch because many international pharmaceutical companies usually produced in many countries and will be submitted to some country before forward shipment to Thailand. This issue causes the problems in management of warehouse area. If warehouse know the number of separate products at high precision and accuracy in advance, warehouse should be able to prepare the resources in advance and improve the warehouse operation process more efficiency.

Therefore, the propose of this research study is to create a proper forecasting model for the actual incoming products in the case study warehouse in order to increase the efficiency of process time and space management.

1.2 Research Objective

The objective of this research is to create the appropriate forecasting model for warehouse.

1.3 Scope of Work

In this study, the import warehouse for pharmaceutical and medical devices product is selected to be a case study. The total number of pallet monthly which is put away into storage location is collected monthly since July 2008 until October 2013 is used to investigate the forecasting model.

From total pharmaceutical company, the top three ranging for the volume of pallet is used to be the example data for modeling in this study which is approximately 15% from total amount of monthly pallet.

1.4. Expected Result

The most appropriate forecasting model used to forecasting incoming product in warehouse.

CHAPTER II
BACKGROUND THEORY AND LITERATURE REVIEW

This chapter will show summary discussion about background theory and literature review for this research

2.1 Time series forecasting techniques

Time series is sequence of data which is collect by the period of time such as daily, monthly, quarterly or annually. Example of time series are the daily oil price, the quarterly sale values of company, or the monthly number of patient of Thailand etc. Time series are very frequently plotted via line charts. Time series are used in statistics, signal processing or pattern recognition in many field of study such as, econometrics, mathematical finance, weather forecasting, earthquake prediction, control engineering, astronomy, and industrial etc.

Time series has 4 main components which are trend, seasonal, cyclical and irregular. Trend refers to the movement of long term time series. This could be a upward or downward. Seasonal component is the movement of time series that repeat occur in the particular time for example the number of patient always rise in June to September every year because of rainy season. Cyclical is similar to the seasonal component but this component will happen in a long period of many years. Lastly, irregular is unusual nature of the data which cannot be known in advance, such as natural disasters, disease outbreaks, war and so on. Of the above mentioned factors, the trend and seasonal factors tend to be used in determining a forecasting model which is appropriate to the data (Taesombut.S,1996).

Time series forecasting is techniques to predict future values based on previously observed values. Forecasting method was invented and developed by many techniques. Each method has difference difficulty, accuracy or suitability data. The basic strategy for finding appropriate model is to analyze the time series component such as trend and seasonal. From the literature review in field of time series forecasting, there are many researcher study in time series forecasting as table 2.1.

Samer et al. (2001) investigated different univariate modeling in forecasting electric power demand in Lebanon which was increasing. Three univariate models , the autoregressive, the autoregressive integrated moving average (ARIMA) and a novel configuration combining an AR(1) with a highpass filter, were used in this work. The forecasting performance of each model is assessed using different measures such as SSE, MSE, and SAE. The AR(1) with highpass filter model yielded the best forecast for this unusual set of electrical energy data.

Volkan et al. (2006) forecasted the domestic fossil fuel production of Turkey because of rapidly growing on imported fossil fuels. The researcher developed a decision support system for forecasting fossil fuel production by applying a regression, ARIMA and SARIMA method to the historical data. In this result , different forecasting models are proposed for different fossil fuel types for Turkey such as the cubic regression model is used for hard coal and lignite, the ARIMA model is used for asphalted and natural gas, and the SARIMA model is used for oil.

Jan and Rob (2006) review the past 25 years of research into time series forecasting. Autoregressive integrated moving average (ARIMA) models is the most popular of their extensions in many areas of science. Often their studies were of an empirical nature, using one or more benchmark methods as a comparison. The almost result show that ARIMA is more successful than others.

Leo et al (2009) researched on customer demand and stock control which are non-stationary by simulation. The ARIMA(0,1,1) was use to generate the demand . Simple exponential smoothing and simple moving average were use to estimate the fill rate inventory. The result shown the appropriate parameter value of each forecast method for the generated data . In conclusion ,the correct method and forecast parameter values is important in stock control performance.

Table 2.1 Application using time series forecasting techniques

No.	First author	Year	Objective	Methodology	Result and Conclusion
1	Reyman	1995	To find the accuracy model for pharmaceutical forecast	Simple moving average, Exponential moving average, Holt-winter exponential smoothing and Box-Jenkins	Box-Jenkins provide the highest accuracy output which can reduce inventory cost 650,000 US dollars in average.
2	Samer et al.	2001	To investigated different univariate modeling in forecasting electric power demand	Three univariate models , the autoregressive, the autoregressive integrated moving average (ARIMA)	The AR(1) with highpass filter model yielded the best forecast for this unusual set of electrical energy data.
3	Volkan et al.	2006	To forecasted the domestic fossil fuel production of Turkey	By applying a regression, ARIMA and SARIMA method to the historical data	Cubic regression model is used for hard coal and lignite, the ARIMA model is used for asphaltite and natural gas, and the SARIMA model is used for oil.
4	Jan and Rob	2006	Review forecasting research between 1986-2005	All time series technique	ARIMA was widely used in latest year because the high accuracy and used in many field of business.
5	Leo et al.	2009	To extend the research to non-stationary demands, by simulation	ARIMA(0,1,1) was used to generate the demand . Simple exponential smoothing and simple moving average were used to estimate the fill rate inventory	The correct method and forecast parameter values is important in stock control performance.

Reymann (1995) have studied the forecasting model to forecast demand for drugs in the medical center because medical center has very high inventory cost. By selected statistical methods such as simple moving average, exponential smoothing, holt-winter exponential smoothing and Box-Jenkins methodology , the result shown that Box-Jenkins method provides the highest accuracy forecasting because Box-Jenkins can use to forecast in data which has trend and seasonal component . Finally , the high accuracy forecasting technique can reduce the cost of inventory average 750,000 US dollars per year.

From various research mentioned above, we can conclude that forecasting methods is appropriate for different pattern of data. Moving average fit to the data with the trend or simple linear but it does not fit for the seasonal data. Holt-Winter or Box-Jenkins (ARIMA) can be precisely forecasting in case of data contained both trend and seasonality. (Ching-Wu and Guoqiang,2003) If considered suitable for adoption in industry can be seen that the method of Box-Jenkins would be suitable for use as possible because the method is highly accurate forecasts. Theory of the Box-Jenkins will be described in the next section.

2.2 Box-Jenkins Method

Box-Jenkins methodology is a technically sophisticated way of forecasting a variable by looking only at the past pattern of time series data. The model of Box-Jenkins method is ARIMA which stands for autoregressive method integrated moving average. The first step in developing a Box–Jenkins model is to determine if the time series is stationary ,constant in mean and variance, or non-stationary. (Henry and Rujirek, 2007 ; Samer et al, 2001)

2.2.1 Time series components

2.2.1.1 Trend

Trend refers to the movement of long term time series. This could be a upward or downward. Fabio and Gianna (2005) suggest that Run test is the easy non-parametric test and also can use to check trend component of time series. Run test has step of testing as following.

1) Define null and alternative hypothesis

   H_0: Time series has no trend component.

   H_a: Time series has trend component.

2) Calculates the Run of time series (R): First ,we need to define positive (+) to the data which is more than median of all data , and negative (-) to the data which is less than median . After that, Run is a sequence of consecutive of sign.

3) Statistic test

$$z = \frac{R - E(R)}{S(R)} \quad \text{.........................(2.1)}$$

   Which $E(R) = \frac{n+2}{2}$ and $S(R) = \sqrt{\frac{n(n-2)}{4(n-1)}}$

   While R is the number of run , n is the total number of data.

4) Decision : The run test will reject null hypothesis if $|z| > z_{1-\frac{\alpha}{2}}$ at the significant level of α

2.2.1.2 Seasonal component:

Seasonal is the movement of time series that repeat occur in the particular time of year for example by month or by quarter. To check time series has seasonal component or not, we can use Kruskal wallist test . Kruskal wallis is non- parametric test which is can detect that the each group of population have the same or different mean. This principle can be used to detect seasonal components. The test has step of testing as following.

1) Define null and alternative hypothesis

   H_0 : The mean of k groups of data are not significant different or data has no seasonal component

   H_a : The mean of k groups of data are significant different. or data has seasonal component

2) Statistic test

$H = \left[\frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}\right] - 3(N+1)$ ……….(2.2)

where

k is the number of season.

R_i is the summation of rank in season i

n_i is the number of data in season i

N is the total number of time series

3) Decision : The test will reject null hypothesis if $H > \chi^2{}_{\alpha,k-1}$ at the significant level of α

2.2.2 Stationary Box-Jenkins models

2.2.2.1 Auto Regressive models (AR)

The model of Auto Regressive order p , AR(p) , is similar to moving average except that the independent variable y depends on its own previous value .The model can be written as equation 2.3.

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \ldots + \emptyset_p Z_{t-p} + e_t \quad \text{…………….(2.3)}$$

or $$\left(1 - \emptyset_1 B - \emptyset_2 B^2 - \ldots - \emptyset_p B^p\right) Z_t = e_t \quad \text{…………….(2.4)}$$

Where $B^n(Z_t) = Z_{t-n}$

Z_t	is the moving average time series.
$\emptyset_1, \emptyset_2, \ldots, \emptyset_p$	are coefficients
$Z_{t-1}, Z_{t-2}, \ldots, Z_{t-p}$	is lagged values of the time series.
e_t	is white noise series.

2.2.2.2 Moving Average models (MA)

The moving average model is predicts y as a function of the past errors in predicting. The model of MA order q , MA(q), would take the following form:

$$Z_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} \quad \text{…………….(2.5)}$$

$$Z_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) e_t \quad \text{…………….(2.6)}$$

Where $B^n(e_t) = e_{t-n}$

Z_t is the moving average time series.

$\theta_1, \theta_2, \dots, \theta_p$ are coefficients

$e_{t-1}, e_{t-2}, \dots, e_{t-q}$ are previous value of the white noise.

e_t is the value of the white noise at time t.

2.2.2.3 Auto Regressive Moving Average models (ARMA)

This model could be produced from a white noise series and the previous value . The equation defines a mixed autoregressive moving average of order p and q , ARMA(p,q) , is written as below:

$$\left(1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_p B^p\right) Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t \quad \text{.......(2.7)}$$

2.2.3 Nonstationary Box-Jenkins models

If the series are nonstationary , the data must be modified to be stationary before identify model by differencing. Differencing means to subtracting the previous value from each observation in the data as equation below.

$$Z^*{}_t = Z_t - Z_{t-1} \quad \text{................(2.8)}$$

In some cases the first difference will not remove the trend and it may be necessary to try higher order of differencing as equation below.

$$Z^*{}_t = (1 - B)^d Z_t \quad \text{................(2.9)}$$

2.2.3.1 Autoregressive Integrated Moving Average Models (ARIMA)

The model ARIMA (p,d,q) is selected when differencing is used to make data stationary while the d inside the parentheses referred to the degree of differencing. The model of ARIMA(p,d,q) is written as follow:

$$\emptyset_p(B)(1 - B)^d Z_t = \theta_0 + \theta_q(B) e_t \quad \text{................(2.10)}$$

2.2.3.2 Seasonality Autoregressive Integrated Moving Average Models (SARIMA)

When the time series to be forecast have seasonal factor, some pattern of series that recurs with the same period, the model ARIMA is not suitable for these data. The appropriate model is SARIMA (p,d,q)(P,D,Q)$_S$ which the second set of P,Q,D value represents the order of seasonal AR , seasonal MA , and the degree of seasonal differencing.

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(b^s)e_t \quad .....(2.11)$$

Where $\quad b^s(Z_t) = Z_{t-s}$

2.2.4 Model identification

Before modeling , the index of model should be determined. ARIMA has 3 indexes which are p,d, and q for non-seasonal model , but has 6 indexes which are p,d,q,P,D, and Q for seasonal model. However, all of the index can combine into 3 groups. First , d and D is the rank or number of differencing to make stationary data for trend and seasonal consequently. Second, p and P is the rank of auto regressive process (AR) . Lastly ,q and Q is a rank of moving average process (MA). Rank of AR and MA can be considered by plot the Autocorrelation (ρ_k) and the partial correlation coefficient (Φ_{kk}) graph or correlogram.

2.2.4.1 Autocorrelation function (ρ_k) or ACF is correlation of a time series with its own past and future values or refers to the correlation between members of a series of numbers arranged in time. The ACF represent the covariance and correlation between y_t and y_{t+k} from the same process,separated only by k time lags.

$$\rho_k = \frac{\sum_{t=1}^{n-k} (Y_t-\bar{Y})(Y_{t+k}-\bar{Y})}{\sum_{t=1}^{n} (Y_t-\bar{Y})^2} \quad ....(2.12)$$

2.2.4.2 Partial correlation coefficient (Φ_{kk}) or PACF refers to the autocorrelation of time series between y_t and y_{t+k} after their mutual linear dependence on intervening variable y_{t+1} and y_{t+k-1} has been removed.

$$\Phi_{kk} = \begin{cases} \rho_1 & ;\ k = 1 \\ \dfrac{\rho_k - \sum_{j=1}^{k-1} \Phi_{k-1,j}\rho_{k-j}}{1-\sum_{j=1}^{k-1} \Phi_{k-1,j}\rho_j} & ;\ k = 2,3,4,\ldots \end{cases} \quad ....(2.13)$$

where $\quad \Phi_{kj} = \Phi_{k-1,j} - \Phi_{kk}\Phi_{k-1,k-j} \quad$ for $k = 1,2,3,\ldots$

and $j = 1,2,3,\ldots,k-1$

2.2.4.3 Characteristics of ACF and PACF for stationary processes.

The pattern of ACF and PACF correlogram can identify the model which is appropriate to the data. For AR process, ACF tails off as a mixture of exponential decays or damped sine wave, while PACF will vanish after lag p. On the other hand , ACF will cut off after lag q and PACF will decrease like sine wave for MA process. Moreover , the model ARMA combine the property of AR and MA process. For the last model ,the ACF will cut off after lag q-p and PACF will cut off after lag q-p. The order of p,q are usually less than or equal to 3. (William, 1989)

Table 2.2 Characteristics of ACF and PACF for stationary processes. (William, 1989)

Process	ACF	PACF
AR(p)	Tails off as exponential decay or damped sine wave	Cuts off after lag p
MA(q)	Cuts off after lag q	Tails off as exponential decay or damped sine wave
ARMA(p,q)	Tails off after lag q-p	Tails off after lag p-q

2.3 Soft computing to neural network

Since year 2000, soft computing is a new technology that was developed to help human for solving the complex problem in the real world. The research in field of soft computing has increased. Soft computing has been applied to many field of research such as supply chain , manufacturing, order fulfillment, and demand management etc. (Kumer et al., 2004) . Soft computing can use the various uncertain factors for example imprecision, uncertainty as part of the decision making. This is the reason why soft computing can male decision close to the human making. (Dutta,2006)

Zadeh, (1994) suggests that soft computing can be divided in many method such as fuzzy logic, neural networks, and genetic algorithm . Fuzzy logic is mainly concerned with imprecision and approximate reasoning. Fuzzy logic is the way the human brain works and decide. Artificial Neural networks(ANN) are used to simulate the working network of the neurons in the human brain. The concept of the human brain is used to perform computation on computers. Genetic algorithm (GA) uses the concept of Darwin's theory of evolution. Darwin's theory is based on the rule, "survival of the fittest." (Bonissone.,1997 ; Wong et al.,2000)

During the period 1992 -1998, neural network had been applied in many business. The one important result shown that 72.6 % of research conclude that neural networks yield better results than other techniques when compared to them , 22.6 % said that Neural networks yield similar results to other techniques when compared to them , and 4.8 % said that Neural networks yield worse results than other techniques when compared to them. Vellido et al.(1999)

Ko et al, (2010) have presented in his journal that the use of soft computing to deal with the needs of the customers or demand management tend to increase every year. Therefore, the reliable demand forecast can develop a strategy to manage the organization effectively. They found that ANN is the most soft computing tools for demand forecasting around 54% because of ANN advantages as figure 2.1. ANN can learn and adjust its model to fit the data. Compared with other methods, ANN is highly flexible and can apply to all pattern of data.

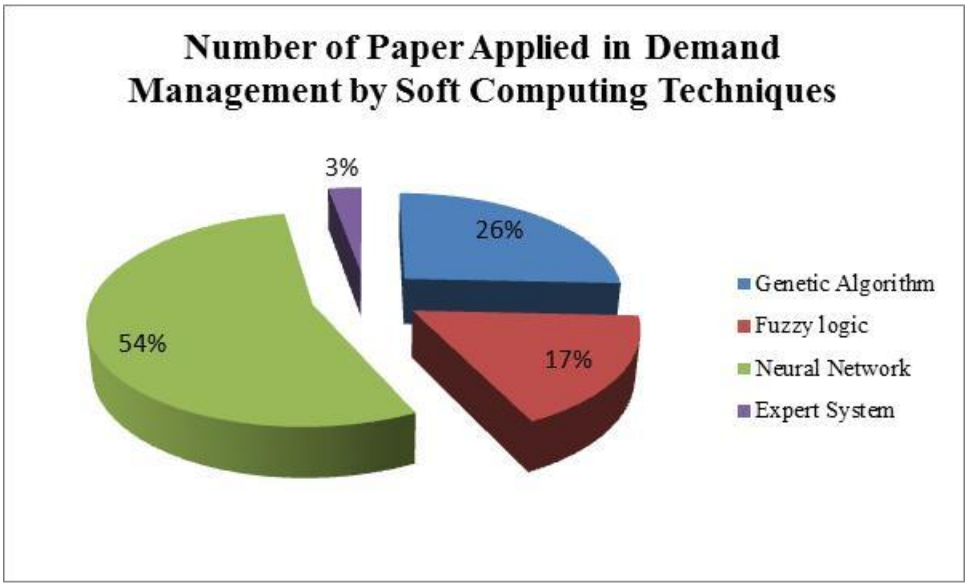


Figure 2.1 Number of paper applied in demand management by soft computing techniques

2.4 Artificial Neural Network

The human brain is composed of millions of cells. Each cell is called a neuron which composed of three parts. First, the tip of the nerve called "Dendrites" , which is the input receptor of cell .Second, the other tail of cell which is used to send nervous impulses to the connected cell called "Axon". The important part of cell is "Nucleus" which controls the cell metabolism. These cells work on electrochemical reactions and transfer the electricity from each cell to others through "Synapse" as figure 2.2. This processes of neurons, humans can distinguish remember or make a decisions for a complex and highly flexible problem . So, Artificial neural network (ANN) has been invented and developed to resolve problems that are complex by mimic the principles of the human brain. (Simon., 1999; Jiawei., 2007)

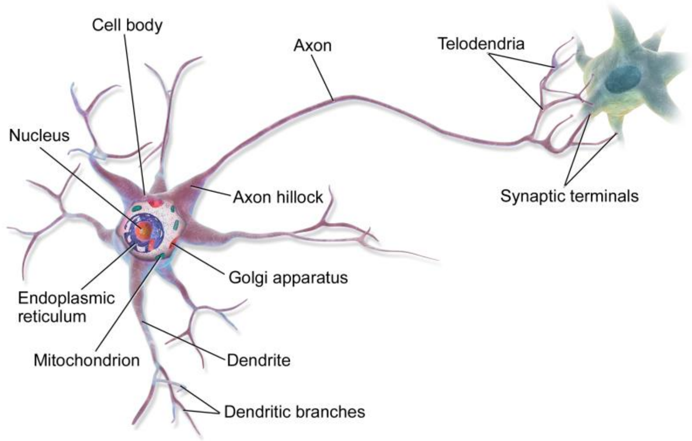


Figure 2.2 Neural cell of human brain.

(Source: http://en.wikipedia.org/wiki/File:Blausen_0657_MultipolarNeuron.png)

2.4.1 Structure of ANN

ANN is a system of non-linear calculations which has ability to learn knowledge from data and adaptability to the variance of data in order to find solutions to the problems. (Flood and Kartam.,1994) ANN is an infrastructure called "Perceptron". Perceptron consists of a 3 layer which are input layer, one or more hidden layer, and an output layer as figure 2.3(a).If many perceptron connected together in many layer ,we call a multi-layer perceptron figure as figure 2.3(b). Each components of perceptron will be describe as follows.

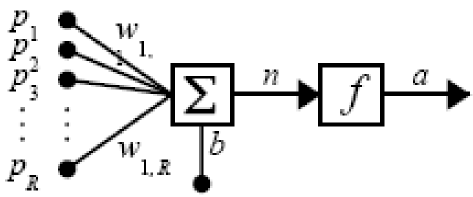


Figure 2.3(a) Single layer perceptron

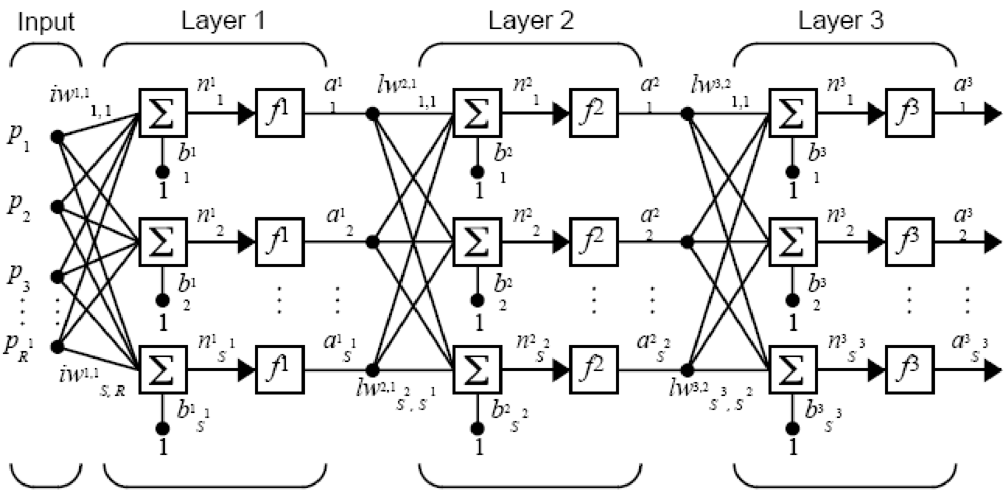


Figure 2.3(b) Multilayer perceptron

2.5.1.1 Input layer

Input layer is the layer that receives input data. Each input corresponds to a single attribute or variable. Wu and Lim (1993) suggest that the number of input for this layer can be choose by trial and error or choose by considering the minimal errors of output. (Anderson et al,1993)

2.5.1.2 Hidden Layer

Hidden Layer is the layer that convert input information into output by 2 steps which are linear combination and non-linear transformation. The number of hidden layer was studied in many research. Roger and Ramarsh (1992) should be equal to the sum of number of input and output node.

1) Linear combination

This step will compute sum of all input by weight for each entering input. Weight express the relative strength of the input data or the many connections that transfer data from layer to layer .The equation of summation function of neuron j for input x_1 to x_n is written as below:

$$summation\ function = x_1 w_{1j} + x_2 w_{2j} + \cdots x_n w_{nj} \quad \ldots\ldots\ldots\ldots\ldots(2.13)$$

2) Non Linear transformation

After linear combination, the summation will be transformed by using function called "Activation function" and release the value to output layer. User can select the function in the form of Linear and Non-Linear function depends on the appropriateness of the information. Sibi et al. (2013) compared the performance of various types of activation function, their study concluded that almost activation function used to the model are non-linear type such Threshold function, piecewise function, and sigmoid function as figure 2.4.

Sigmoid function $\mathrm{Y(t)} = \frac{1}{1+e^{-t}}$ …………….(2.14)

Threshold function $\mathrm{Y(t)} = \begin{cases} o\,; t \geq 0 \\ 1\,; t < 0 \end{cases}$ …………….(2.15)

Piecewise-linear function $\mathrm{Y(t)} = \begin{cases} 1\,; \frac{1}{2} \leq t \\ t\,; -\frac{1}{2} \leq t \leq \frac{1}{2} \\ 0\,; t \leq -\frac{1}{2} \end{cases}$ …………….(2.16)

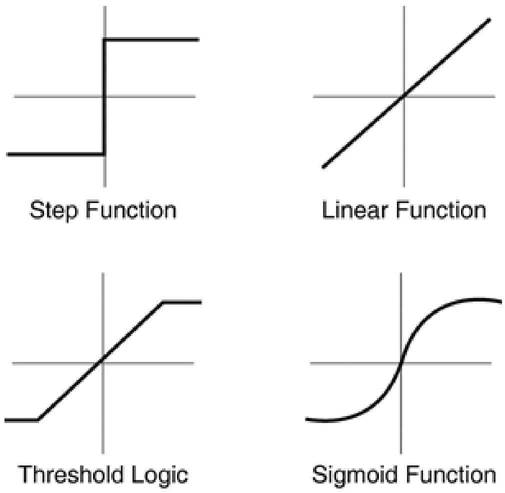


Figure 2.4 Example for activation function

Simon (1999) suggests that sigmoid function is an activation function that is widely used in building ANN because this function has more advantages than other function. Sigmoid function is continuous and return value in range of real number. Moreover ,sigmoid function is easily to differentiate which is use to train the model. Lastly , sigmoid is monotonic increasing

2.5.1.3 Output Layer

Output layer is the layer that relay the final result.

2.4.2 Learning process and parameter estimation

ANN can be used to solve a complex function and highly precision due to the ability to discern patterns and learning the characteristic of data. This capability comes from the learning process . ANN learning process divided into two groups: supervised learning and unsupervised learning. (Lippmann., 1987)

2.5.2.1 Supervised Learning is learning process which consists of input data and target. For ANN, the input feed into the model and calculate for output. The process will compared the output to the actual results or target and sent the error back into the network to adjust the weights for the new iteration. There are many method such as back propagation, hope field network.

2.5.2.2 Unsupervised Learning is learning process without target comparison which was invented by Kohonen (1984), This learning processes are commonly used to classification data for example Kohonen Self-organizing feature maps.

Zhang et al. (1998) has compile the research on forecasting by using ANN. They found that ANN has been applied in various studies , developed and compared to the other model. For the learning algorithm, many papers selected the back propagation for their model. Predictive value will be compared with the target , compute the error, and brings the error to modify the new parameter of model. This method makes the slightest error makes the results very close to reality as figure 2.5.

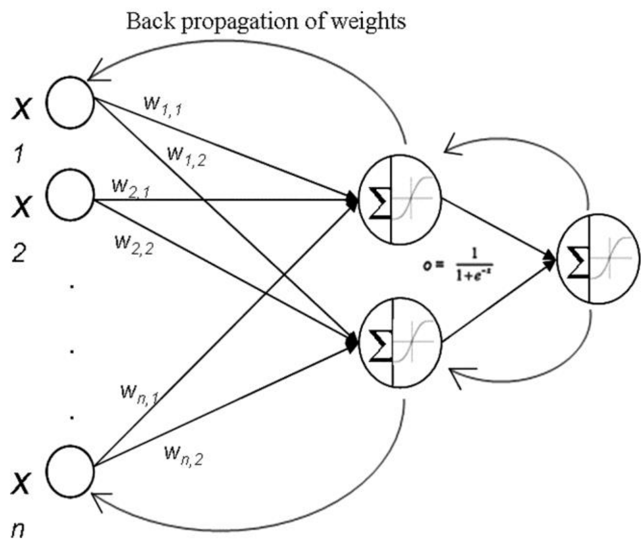


Figure 2.5 Back Propagation process

In the beginning, Algorithm has been built up to adjust the weights which is called Widow-Hoff rule In the year 1986 , Rumelhart, William and Hilton study until they created the Back-propagation which is widely accepted by many researcher. (Lippmann,1987). Back propagation principles can be summarized as follows

1) Forward Computation: $v_j^l(n)$ is the output of node j of layer l while y_i^{l-1} is the input from layer l-1. w_{ji}^l is synaptic weight which is link between node i of layer l-1 and node j of layer l. The forward computation can be written in equation as bellow.

$$v_j^{(l)}(n) = \sum_{i=0}^{m_0} w_{ji}^{(l)}(n)\, y_i^{l-1}(n) \qquad \text{.................(2.17)}$$

Let φ_j^l is an activation function of neural j in layer 1 , the output after transformation will show as below.

$$y_j^{(l)} = \varphi_j^l(v_j^{(l)}(n)) \qquad \text{.................(2.18)}$$

$d_j(n)$ is the target of model and $o_j(n)$ is output of the last layer. The error of model can calculate as the following equation.

$$e_j(n) = d_j(n) - o_j(n) \qquad \text{.................(2.19)}$$

2) Backward computation is the process of weight adjust from error signal. First of all , the local gradients $\delta_j^l(n)$ will define by calculate from these equation.

For node j in layer L: $\delta_j^l(n) = e_j^L(n)\varphi_j'\left(v_j^{(L)}(n)\right)$ ....(2.20)

For node j in layer 1 : $\delta_j^l(n) = \varphi_j'\left(v_j^{(l)}(n)\right)\sum_k \delta_k^{(l+1)}(n) w_{kj}^{l+1}(n)$ ...(2.21)

From the literature above, sigmoid function is the appropriate activation function because of continuity and differentiable. So , the equation show the result of taking sigmoid function in to the model .

$$\varphi_j^l(v_j^{(l)}(n)) = \frac{1}{1+e^{-av_j^l(n)}} \qquad \text{...........(2.22)}$$

$$\varphi_j'\left(v_j(n)\right) = ay_i(n)[1 - y_j(n)] \qquad \text{…………(2.23)}$$

For some problems which is highly complex or large size problems, the model will take a long time for solving the solution or cannot find the solution in finally. Learning rate parameter is the parameter that determines the rate of learning for the model. If learning rate is set in a few values, the model will take a long time running. If we set the value in large, the model will learning fast. Sometime, the model can find the solution but that solution is not optimize or is not the best solution for the problem. In the simplified function of figure 2.6(a) ,the error surface is simple. Any step in a downward direction will take the solution closer to the absolute minimum. For real problems, error surfaces are typically complex as figure 2.6(b). which has a lot of local minimum, the model will trap in one minimum which is not the global. Progress here is only possible by climbing higher before descending to the global minimum. The technique that can help the network out of local minima is the use of a momentum rate. (Khan et al,1993)

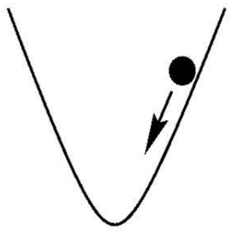


Figure 2.6(a) Problem with absolute minimum

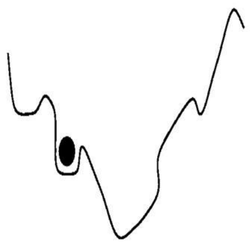


Figure 2.6(b) Problem with many relative minimum

The new adjusted weight of model is calculate from the product of learning rate parameter (η) ,momentum rate(α), local gradient $\delta_j(n)$ และ input of neural j as equation below. For the model which use sigmoid function as activation function will rewrite the adjusted weight equation as equation.

$$\Delta w_{ji}(n) = \eta\delta_j(n)y_i(n) \qquad \text{...........(2.24)}$$

$$w_{ji}^{l}(n+1) = w_{ji}^{l}(n) + \alpha\left[w_{ji}^{l}(n-1)\right] + \eta\delta_j^{l}(n)y_i^{l-1}(n) \qquad \text{...........(2.25)}$$

From the theoretical , there are many researcher who discovered the new high accuracy and high performance method to train and estimate the parameter of model. Aggarwal et al (2005) gather many researches , summary the list of training algorithm as table 2.3 and compare the performance of each learning process. Their study found that Baysian regularization training or Trainbr is the suitable learning algorithm for many problem because of two reason. First, this algorithm can help the model to find the optimal solution fastest because this training combine the back propagation and the Lavenberg-marquardt theory together.(Hirschen and Schafer, 2006) Moreover, the Baysian regularization will prevent the problem of over fitting which is lead to the error when the output are forecasted.

Table 2.3 Different training algorithms. (Aggarwal et al .,2005)

Trn.Fnc	Description
trainb	Trains a network with weight and bias learning rules with batch updates. The weights and biases are updated at the end of an entire pass through the input data.
trainbfg	Updates weight and bias values according to the BFGS quasi-Newton method.
Trainbr	Updates the weight and bias values according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights and then determines the correct combination so as to produce a network that generalizes well. The process is called Bayesian regularization.
Trainc	Trains a network with weight and bias learning rules with incremental updates after each presentation of an input. Inputs are presented in cyclic order.
Train cgb	Updates the weight and bias values according to the conjugate gradient backpropagation with Powell-Beale restarts.

Table 2.3 Different training algorithms. (Aggrawal et al., 2005) (cont.)

Trn.Fnc	**Description**
Train cgf	Updates the weight and bias values according to the conjugate gradient backpropagation with Fletcher-Reeves updates.
Train cgp	Updates the weight and bias values according to the conjugate gradient backpropagation with Polak-Ribiere updates.
Train gd	Updates the weight and bias values according to gradient descent.
Train gda	Updates the weight and bias values according to gradient descent with adaptive learning rate.
Train gdm	Updates the weight and bias values according to gradient descent with momentum.
Train gdx	Updates the weight and bias values according to gradient descent momentum and an adaptive learning rate.
Train lm	Updates the weight and bias values according to Levenberg-Marquardt optimization.
Train oss	Updates the weight and bias values according to the one step secant method.
Train rp	Updates weight and bias values according to the resilient backpropagation algorithm (RPROP).
Train scg	Updates weight and bias values according to the scaled conjugate gradient method.

2.5 Time Delay Artificial Neural Network

Time delay neural network (TDNN) is an architecture type of the feed forward neural network which combines the multi-layer perceptron to the recurrent network. TDNN include the time factor into the model for input layer. The structure of TDNN is not different to simple ANN. The model consists of 3 layers which is input layer, hidden layer and output layer figure 2.6.1. Variable x (t) is the data input into the network at time t and variable d is the time delay, which will determine the amount of historical data to be inserted into the model. For example, if we set d is 3 , the inputs x(t),x(t-1),x(t-2) and x(t-3) will be fed into the network .

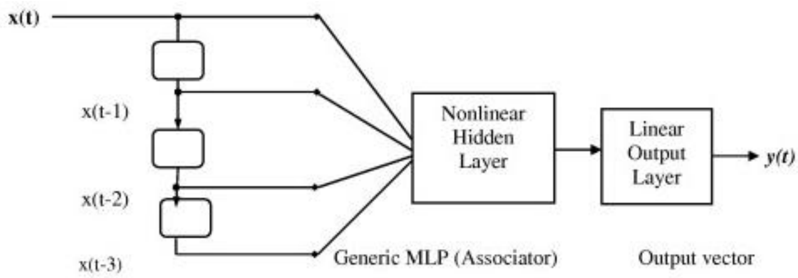


Figure 2.7 Time delay neural network.

TDNN has been used widely in the field of phoneme or speech research. The model was used to distinguish words or sentences of human speech that change over time. TDNN are used for classification of data in many research such as the research of Grayden (1992). He compared the performance between model TDNN and fully connected multilayer perceptron to recognize phoneme. The study founded that TDNN is more effective for classification than the traditional ANN because the input data has correlation in the difference time. Moreover, TDNN is easier than the ANN for modeling.

From the structure of model, TDNN can input the data with time lag. This is the properties which is expected to be use in time series forecasting. (Hamzacebi, 2008). In this research , TDNN will be used to forecasting for time series using the same principle of multi-layer perceptron.

2.6 Hybrid model

Forecasting has been used to solve in many problems in the later year. Many researcher need to find the highest accuracy by comparing many method. The methods which are usually selected are neural network and Box-Jenkins. The example of related research is shows as table 2.4.

Victor et al (2000) forecasted daily maximum ozone concentration because the meteorological variables and photochemical reactions formation were very complex. In this research, a neural network model for forecasting daily maximum ozone levels was developed and compared with two statistical models, regression and Box-Jenkins model (ARIMA). The results shown that the neural network model is superior to the regression and Box-Jenkins models indicated by mean absolute deviations (MAD) , and root mean square (RMS)

Ilan et al. (2001) shown the relative forecasting accuracy of ANNs in comparison to the more traditional time-series methods Winter’s exponential smoothing, ARIMA, and regression models in US aggregate retail sales which had strong trend and seasonal patterns. The result shown that ANN performed the best , followed by ARIMA , Winter’s ,and the poorest is regression because the ANN was able to capture the dynamic nonlinear trend and seasonal patterns more efficiency than others.

Kuo (2001) studied about sales forecasting which was very fluctuate because of the effect of internal and external environments such as promotion method, promotion length, and advertising . Fuzzy neural network with initial weights generated by genetic algorithm (GFNN) was improved and used in this research because EFNN was able to learn the fuzzy IF±THEN rules for promotion provided by marketing experts. The results for a convenience store (CVS) company indicate that the GFNN can perform more accurately than the conventional statistical method such as ARMA and a single ANN.

Chu and Zhang (2003) compared the accuracy of various linear and nonlinear models for forecasting aggregate retail sales because the data of retail sales had a very strong seasonal fluctuations . ARIMA and regression techniques , linear models, were used for forecasting and also neural network was used because neural networks are the most nonlinear models that can represent both nonseasonal and seasonal time series. The result shown that the nonlinear models , neural network, was able to outperform their linear models in forecasting.

Henri and Rujirek (2007) forecasted the demand of four species of rice which were export from Thailand. This research compared the performance of artificial neural networks (ANNs) with exponential smoothing and ARIMA models. The measures of forecast error (MAE, MSE, MAPE, and RMSE) was use to validate process of the models. The results reveal that while the Holt–Winters and the Box–Jenkins models showed satisfactory goodness of fit, the models did not perform as well in predicting unseen data during validation. On the other hand, the ANNs performed relatively well as they were able to track the dynamic non-linear trend and seasonality, and the interactions between them.

Refael et al.(2008) applied neural network (NN) modeling in forecasting lumpy demand of an electronics distributor operating in Monterrey, Mexico. Moreover the researcher compared the performance of NN forecasts to those using three traditional timeseries methods ,single exponential smoothing, Croston’s method, and the Syntetos–Boylan approximation. The result shown that NN models which trained by back-propagation with three hidden layer perform better than the other methods indicated by mean absolute percentage error (MAPE).

Li et al (2008) developed an enhance fuzzy neural network (EFNN) for managing automobile spares part inventory in a central warehouse of China. In this research, the EFNN is used for forecasting the demand for spare parts which very fluctuated by using five factors which are part factor, demand factor, time factor, sale factor and associated factor. The model was adjusted in connection weight activation function. The connection weight was set based on the fuzzy analytic hierarchy process (AHP) method and the activation function was generated according to genetic algorithm. The result show that the EFNN was more accuracy for non linear model than the other forecasting techniques such as exponential smoothing , ARIMA, winter’s exponential smoothing and trend analysis.

From literature review, many time series methodology and neural network were used to forecast the future values for highest accuracy. Although, Box-Jenkins or ARIMA model has the best performance from time series technique , but there are the limitation for some data. William (1998) said that the linear pattern of data is appropriate for ARIMA model. For the real world business, almost of the data is hard to said that which is linear or non-linear. So, ANN's have taken a role in forecasting for these complex data . The advantage of ANN are flexible ,adaptable, and support to the nonlinear input.

Table 2.4 Comparison between times series method and neural network

No.	First author	Year	Objective	Methodology	Result
1	Victor et al	2000	To forecasted daily maximum ozone concentration	neural network compared with two statistical models, regression and Box-Jenkins model (ARIMA)	the neural network model is superior to the regression and Box-Jenkins models
2	Ilan et al.	2001	To forecast US aggregate retail sales	ANNs in comparison to Winter's exponential smoothing, ARIMA, and regression models	ANN was able to capture the dynamic nonlinear trend and seasonal patterns more efficiency than others
3	Kuo	2001	To improve GFNN for sales forecasting compare with simple ANN and ARMA	Fuzzy neural network with initial weights generated by genetic algorithm (GFNN)	GFNN can perform more accurately than the conventional statistical method such as ARMA and a single ANN.
4	Ching and Guoqiang	2003	To compare the accuracy of various linear and nonlinear models for forecasting aggregate retail sales	ARIMA and regression , linear model , were used to forecast and neural network was used in kind of nonlinear model.	Neural network was more efficient with the linear models for seasonal fluctuated data.
5	Henry and Rujirek	2007	To forecasted the demand of four species of rice which were export from Thailand	artificial neural networks (ANNs) , exponential smoothing and ARIMA models.	ANNs performed relatively well as they were able to track the dynamic non-linear trend and seasonality, and the interactions between them.

Table 2.4 Comparison between times series method and neural network (Cont.)

No.	Author	Year	Objective	Methodology	Result
6	Rafael et al	2008	To forecasting lumpy demand of an electronics distributor	applied neural network (NN) modeling and compared the performance , single exponential smoothing,	NN models which trained by back-propagation with three hidden layer perform better than the other methods
7	Li et al	2008	To developed an enhance fuzzy neural network (EFNN) for managing automobile spares part inventory	Fuzzy neural network (EFNN)	The EFNN was more accuracy for non linear model than the other forecasting techniques

Peter (2003) said that the actual data in the business cannot categorize that it is pure-linear or non pure-linear. Hence, the previous model may not produce the best results .He proposed the new model called Hybrid model , which is a combination of ARIMA and ANN. The combining model will increase efficiency to capture pattern of information, and thus increase the performance of the forecasting. From Zhang's assumption, the real world data can be written in term of linear and non linear as equation 2.26.

$$y_t = L_t + S_t \qquad \text{...........(2.26)}$$

Where y_t is the actual data ,

L_t is the linear part ,

S_t is referred to non-linear part of data.

When we create a model by ARIMA, the linear pattern of data will be detected by the model and the output will come from these linear pattern called linear output ($\widehat{L_t}$). Error of model (e_t) calculate from the difference of actual value and forecasting value as equation 2.27. Zhang conclude that these errors are come from the non-linear part of data. If we use some transform function or some method which is appropriate for non-linear such as ANN to forecast the residual, we can forecasting the accuracy non-linear term of future data value as equation 2.28.

$$e_t = y_t - \widehat{L_t} \qquad \text{...........(2.27)}$$

$$\widehat{S_t} = f(e_{t-1}, e_{t-2}, \dots, e_{t-N}) \qquad \text{...........(2.28)}$$

Finally, the linear output and non-linear output will be summarized together to be the real forecasting output as equation 2.29.

$$\widehat{y_t} = \widehat{L_t} + \widehat{S_t} \qquad \text{...........(2.29)}$$

2.7 Forecasting model Performance

In this study, mean absolute deviation (MAD) , mean square error (MSE) , mean percentage error (MPE) and mean absolute percentage error (MAPE) are used to measure the accuracy of the forecasting model. The formulas for calculation of each indicators were described by the following equation.

$$MAD = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n} \qquad \text{...........(2.30)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{n}} \qquad \text{...........(2.31)}$$

$$MPE = \frac{\sum_{i=1}^{n} (\frac{Y_i - \hat{Y}_i}{Y_i} \times 100)}{n} \qquad \text{...........(2.32)}$$

$$MAPE = \frac{\sum_{i=1}^{n} (\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \times 100)}{n} \qquad \text{...........(2.33)}$$

Where Y_i refer to actual value or target

$\hat{Y}_i$ refer to forecasting output

n is the number total input data

2.8 Related Research

Luis and Richard (2007) has studied in retail company which offer many products purchase from a large number of distributor . The retail stores need to competitive their products prices, so they have to decide how much and when to buy.

Solving this problem require an accuracy demand forecast. Researcher choose three method which are Box-Jenkins , ANN and Hybrid model. The result shown that hybrid model perform the best result in demand forecasting because the neural network has learning properties which Box-Jenkins doesn't have as table 2.5. Moreover, the high accuracy demand forecasting can increase the customer satisfaction by reduce shortage and reduce inventory cost.

Table 2.5 Comparison between Box-Jenkins and ANN models

Box-Jenkins	**ANN**
- Linear model	- Non-linear model
- Modeling require the series to be stationary	- Any time series can be analyzed
- Requires interaction with the user	- Requires fewer interaction with the user
- The model provides insight and information through its parameters	- Difficult to interpret the model (black box)
- No over fitting	- Over fitting is possible

Peter and Min (2005) examine the issue of how to use neural networks more effectively in modeling and forecasting a time series with trend and seasonal component. Three time series was used to be input for model which are the data of consumer goods , department store and durable goods. The result show that the neural network can forecast to the deseasonalize and detrended more accuracy than the ordinary data because the nueral network models that ignore these seasonal or trend patterns will results in a high variance thus poor forecasting accuracy. The process of deseasonalizing and detrending removes these large seasonal and trend variations from the raw data, thus helps improving the modeling accuracy.

From background and problem statement which is described in chapter I, this study purpose to find the appropriate forecasting model for product of import and export warehouse. By assumption, the high accuracy forecasting technique will lead to reduce the operation time and cost. From the literature, three method which are Box-Jenkins, Neural Network and Hybrid models are chosen for this research because of high accuracy to forecast the actual amount of product by pallets.

The Box-Jenkins model will use the time series data for input. The data should be in the stationary state which are eliminate trend and seasonal component.

After that the model will be defined by investigate the correlogram of ACF and PACF. For the neural network, TDNN structure will be used in this research with sigmoid function as activation function and Baysian-regularization back propagation as learning algorithm. The historical data with time lag 1,2,3, and 12 will be feed to model as input and the total number of hidden node will be vary from 2-12 to find the lowest error model. For the hybrid model, the step of modelling will be the same as Box-Jenkins and TDNN. The performance of each model will be compared for an accuracy by using three indicator which are RMSE, MAD and MAPE . Then, the appropriate model will be suggest to be the forecasting model for import and export warehouse

CHAPTER III
RESEARCH METHODOLOGY

In this research, the process composes with 5 steps which will be explained in detail.

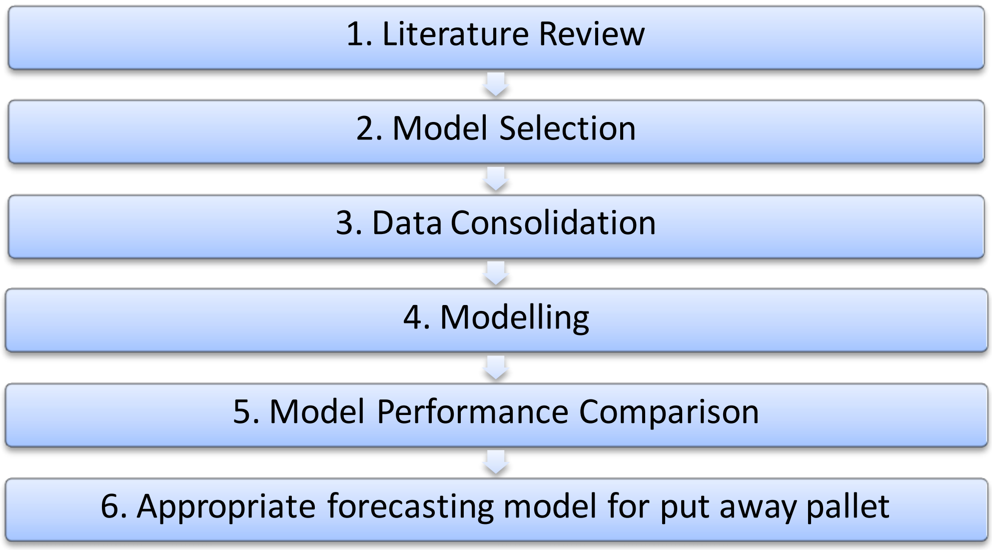


Figure 3.1 Step of research methodology

3.1 Literature review

In order to accomplish a successful research is to investigate the best suitable forecasting model for forecasting real number of pallet which are put away to storage location warehouse. Therefore, the many type of forecasting technique were study according to related research. From the literature review, the forecast techniques can be separated into two groups which are time series technique and artificial neural network. The detail of these methods was described in chapter 2.

3.2 Model Selection

According to the literature review, three methods of forecasting were selected to study in this research. There are Box-Jenkins model (ARIMA), time delay neural network (TDNN), and the hybrid model. The hybrid model is the combination of ARIMA and TDNN. These models can define the data close to the real world situation that are very fluctuated and uncertainty.

3.3 Data Consolidation

The number of actual pallet ready to put away to storage location in warehouse was used for modeling input data. The data was collected since July 2008 until October 2013 by month. In this case study, we use these criteria which are maximum quantity per shipment, frequency of shipment, and pattern of data for screening data. Finally, the top three highest data from different company were selected to this research study.

The data is divided into three groups which are training set, validating set and testing set. The training set is around 80% from total data which is used to train the model and approximate the parameter. The other 20% is used to test the performance of model by comparison the error from forecasting value.

3.4 Modeling Approach

The steps of modeling for each model were described briefly as below.

3.4.1 Box-Jenkins Model

According to the literature review, the processes of modeling for ARIMA can be described into 6 steps as below.

3.4.1.1 Prepare data

Before modeling, the time series should be in stationary state. Stationary data is the data which is stable in mean and variance. In the other word, the data without trend and seasonal factor is in stationary state. To check the trend and seasonal component, we use hypothesis testing which are run test and Kruskal wallist test which is already described in chapter 2. If the data has these components, we need to eliminate them by differencing process as below equation.

Trend Differencing: $Z_t - Z_{t-1}$ or $(1 - \mathrm{B})\mathrm{Z}_\mathrm{t}$

Season Differencing: $Z_t - Z_{t-s}$ or $(1 - \mathrm{B}_\mathrm{s})\mathrm{Z}_\mathrm{t}$

3.4.1.2 Define model structure

ARIMA model has 6 variables that we need to define which are p,d,q,P,D, and Q. The variable d and D is the number of differencing for trend and season consequently. The variable p is the index for autoregressive process and q is the index for moving average process while P and Q are the same in the seasonal time lag. From literature, we can define these variable for model by investigate from pattern of ACF and PACF graph.

3.4.1.3. Parameters Estimation

After we define the 6 variable for model, we need to estimate the parameter of model. In this study, we use Melard's algorithm (Exact likelihood) for estimation by using SPSS program. The estimated parameter need to check for acceptable by using hypothesis testing. If the hypothesis testing shows that the parameter is equal to zero, we need to change the model with other value of variable.

3.4.1.4. Testing Model

When we have the suitable model for the training set of data, we need to test the performance of forecasting model by using the remaining data set.

From all above, we can summary the procedure for Box-Jenkins model by flow chart as figure 3.1 below.

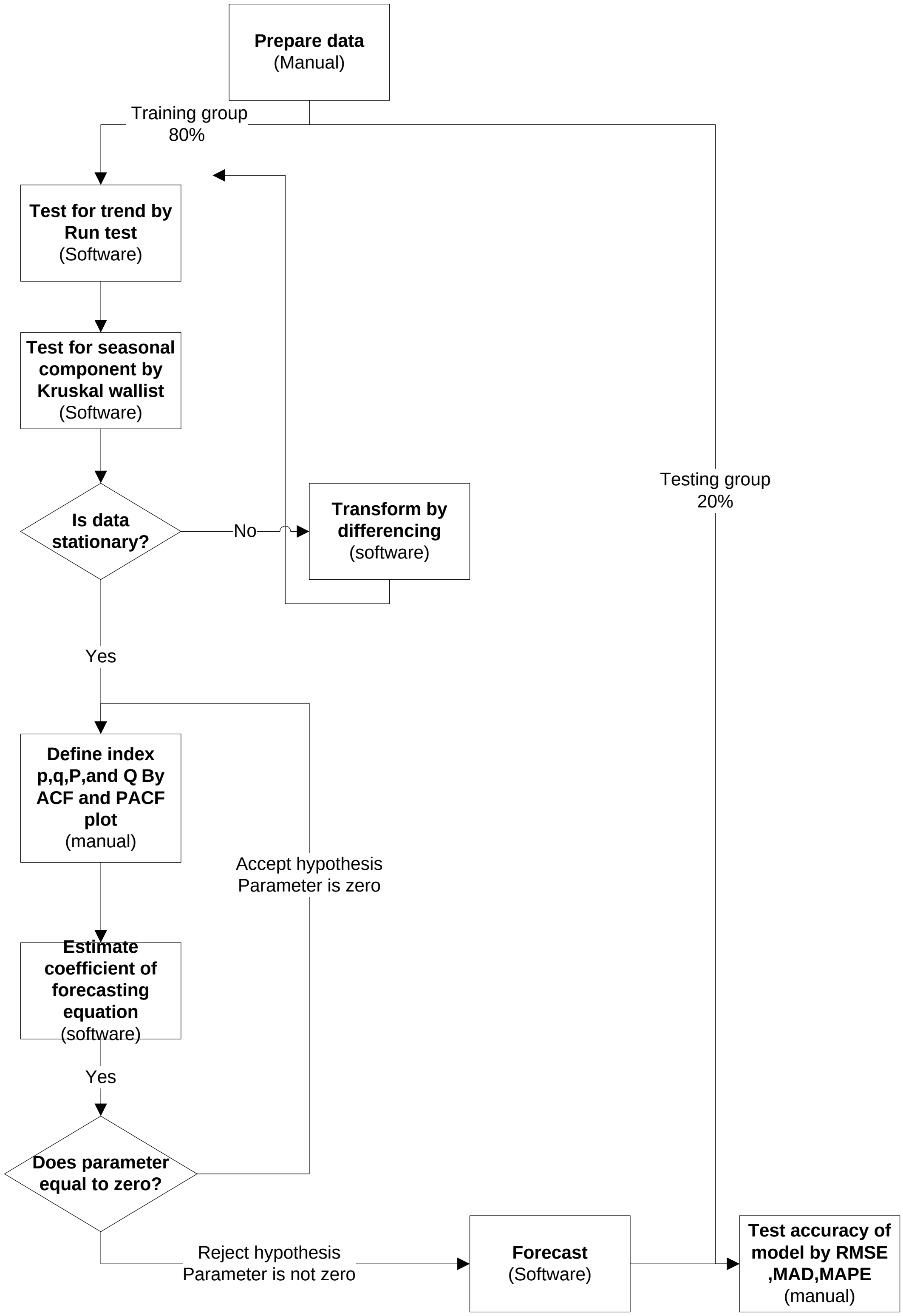


Figure 3.2 ARIMA modeling process

3.4.2 Time delay neural network model (TDNN)

TDNN was adjusted from FFNN by changing input layer. There are five steps for modeling as follow.

3.4.2.1 Data Preparation

Prepared data were separated into two sets which are training set for training and test set for testing the model.

3.4.2.2 Define model structure

TDFN has been use to study in this research which has input layer, hidden layer, and output layer as the figure**.** Input layer was time delay node which has n input nodes, the input were data in 2, 3, and 6 time lags. Hidden layer was varies from 2 to 6 nodes. The sigmoid function was use as activating function. The learning algorithm which was use in this model was lavenberg-marquardt backpropagation. For output layer, there is1 node and use pure-linear function for output transformation.

3.4.2.3 Training model

In process of model training, all of input data was feed to the model. This structure has back propagation as learning algorithm, so the weight of each model was adjusted until the model was optimal. The Gradient descent method was used to find the optimize parameter for the model. The parameter for fitted model will be estimated by Matlab Software.

3.4.2.4 Testing model

When the training process finished and fitted, the data test set was input to evaluate the model. To define the forecasting value, we will run the model 10 time and the output obtained from model used to find the average value for forecasting.

From above, we can summary the procedure for Box-Jenkins model by flow chart as figure 3.2 as below.

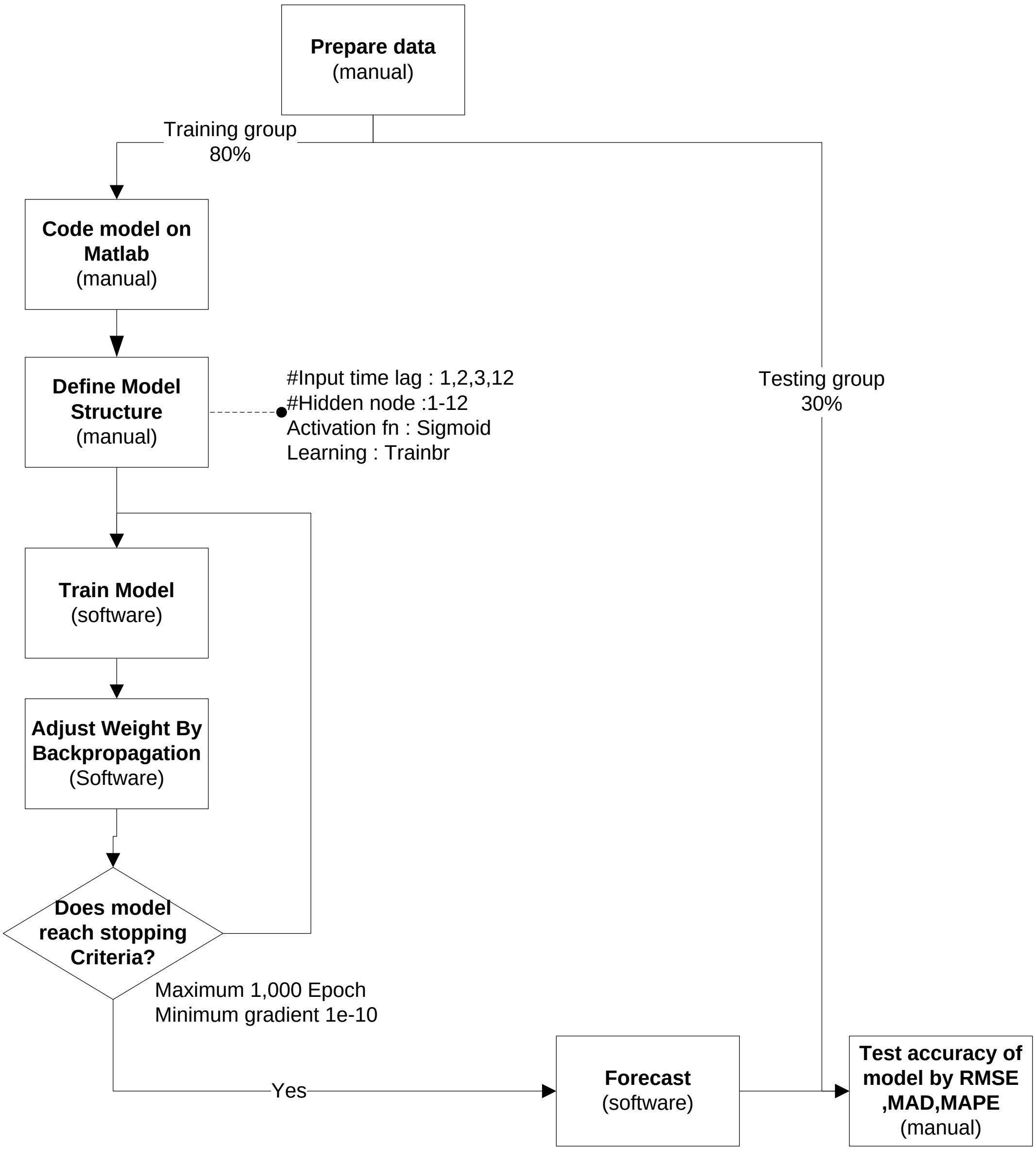


Figure 3.3 TDNN modeling process

3.4.3 Hybrid model

Hybrid model is the combination of ARIMA and TDNN, which has the step of modeling as the follow.

3.4.3.1 Data Preparation

Prepared data were separated into two sets which are training set for training and test set for testing the model as the same as other model.

3.4.3.2 Linear Phase

The process of this phase is similar to 3.4.1.

3.4.3.3 Non-Linear Phase

For this phase, the step of modeling is the same to the 3.4.2 but there are different in input. Residual which is the different of actual and forecasting value are used to be input for this phase.

3.4.3.4 Output calculation

From the literature, the output of forecasting model will calculate by summation the output of linear and non-linear phase together.

3.4.3.5 Testing model

When the training process finished and fitted, the data test set was input to evaluate model. To define the forecasting value, we will run the model 10 time and the output obtained from model used to find the average value for forecasting.

From above, we can summary the procedure for Hybrid model by flow chart as figure below.

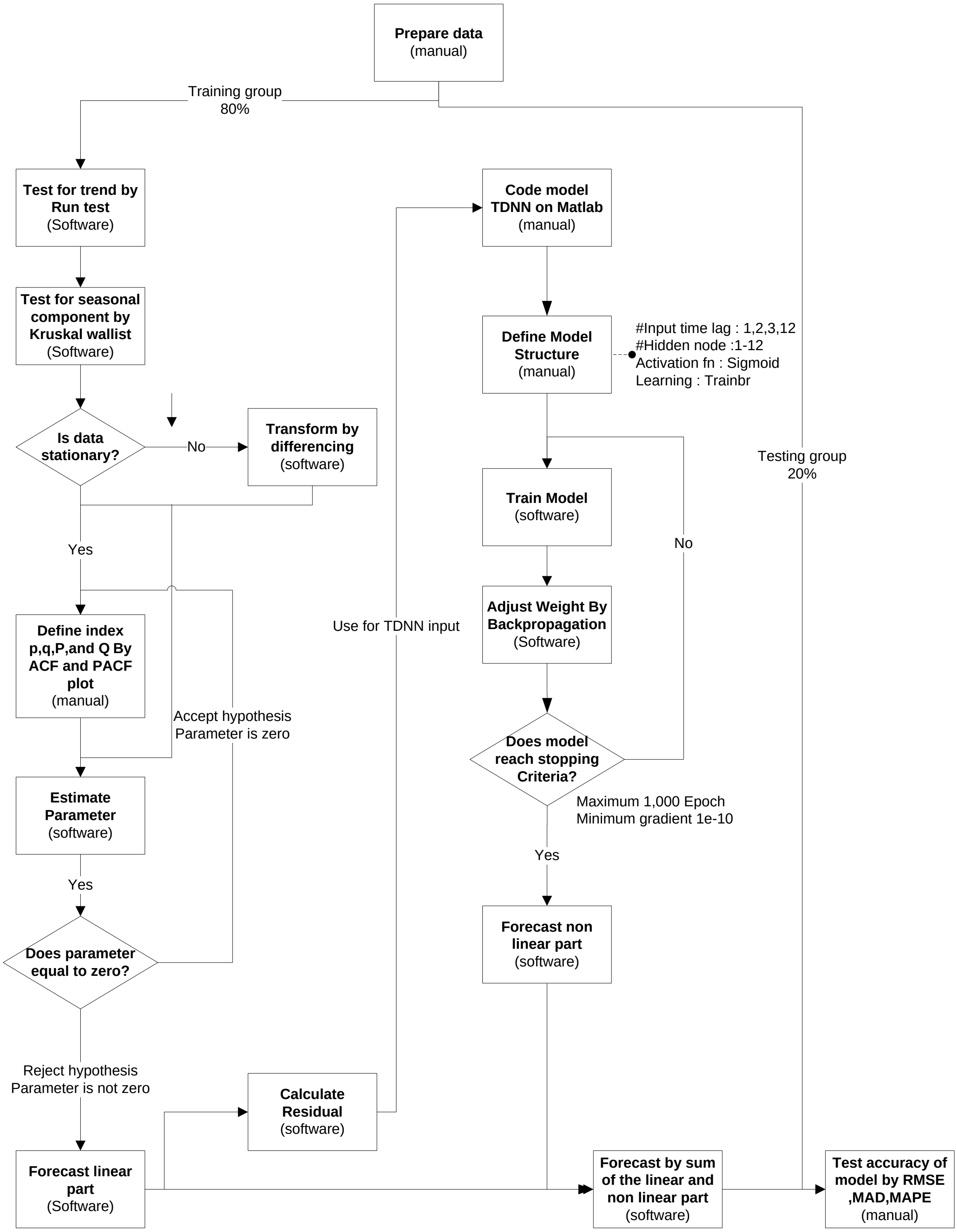


Figure 3.4 Hybrid modeling process

3.5 Model performance Comparison

In this study, we have compare three type of forecasting model which model preform the best in forecasting. The three performance measurements which were selected are, there are mead absolute deviation (MAD), Mean square error (MSE) and Mean absolute percentage error (MAPE). The calculation was shown as below while Y_i is the actual value, $\hat{Y}_i$ is the forecasted value and n is the number of data.

$$MAD = \frac{\sum_{i=1}^{n}|Y_i - \hat{Y}_i|}{n}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}}$$

$$MAPE = \frac{\sum_{i=1}^{n}\left|\frac{Y_i - \hat{Y}}{Y_i} \times 100\right|}{n}$$

3.6 Appropriate forecasting model for put away pallet

In this section will be describe in Chapter IV.

CHAPTER IV
RESULTS AND DISCUSSIONS

This chapter includes four parts according to the research methodology from the previous chapter which are 4.1) Data consolidation 4.2) Modeling 4.3) Model performance comparison 4.4) The contribution of forecast modeling to warehouse management. For Literature review and model selection are described in the chapter 2.

4.1 Data consolidation

This study aims to forecast the number of put away pallets that need to store in import and export warehouse. The time series model and ANN were chosen for forecasting. Therefore, the historical data of put away pallet were collected by month since July 2008 to October 2013. From the scope of this study, three data from pharmaceutical and medical device company were chosen by three criteria. The first criterion is pattern of data which are linear and non-linear characteristic. Second, the product data which deliver consistently on a monthly basis were selected. The last criteria is the top three maximum quantity per shipment. All used data were according to figure 4.1.

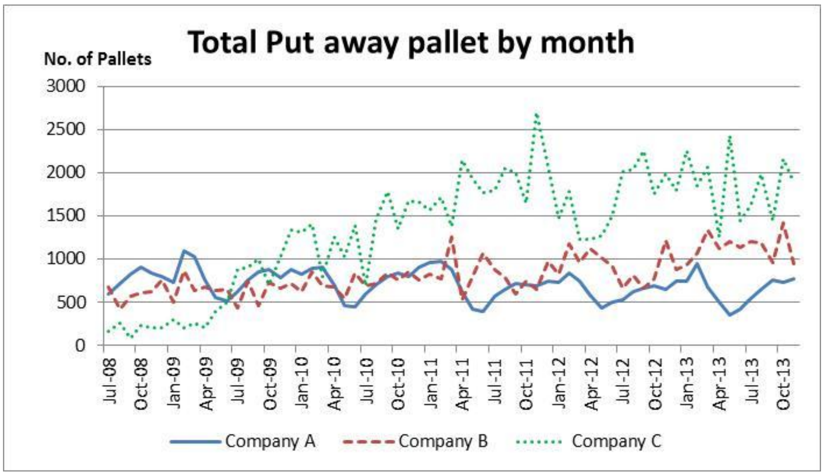


Figure 4.1 Sequence plots of data from case study

For modeling, all of the historical data will be separated into two groups which are training set and testing set. The training set is around 80% of all data and the testing is 20% as figure 4.2. The training set of data will be used for training model and the testing will be used for measure the performance of each model. The step of modeling will be describe in the next session.

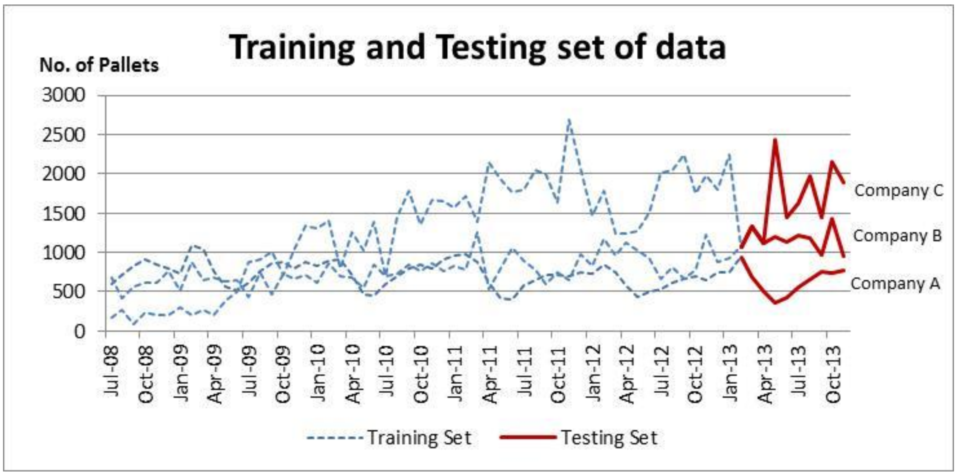


Figure 4.2 Training set and testing set of data

4.2 Modeling

In this section will describe in three types of modeling which are Box-Jenkins , TDNN and Hybrid model.

4.2.1 Box-Jenkins Model

The processes of modeling for Box-Jenkins or ARIMA can be described into 6 steps as below.

4.2.1.1 Data Preparation

All data need to check for trend component and seasonal component first before modeling. If data has these components, the data need to eliminate them by differencing.

1) Trend components

The data was checked the trend by Run test, which has the hypothesis as below, and the values of statistics testing was calculate by the software. The hypothesis testing will accept the null hypothesis when the statistic value doesn't fall in the critical region or the p-value is less than the significant level. In this testing , the significant level was set at 0.05.

H_0 : The data has no trend component.

H_a : The data has trend component.

The testing results from software are shown as figure 4.3. Z-value of historical data from company A,B and C equal to -5.375,-3.374 and -5.871 which fall in the critical area. Moreover, p-value of all companies equal to 0.00 which less than significant level. Thus, the null hypothesis must be rejected for all data , the data was not random. In the other word, the historical data for all company have trend component in their data.

Runs Test

	ProductA
Test Value[a]	729.00
Cases < Test Value	32
Cases >= Test Value	33
Total Cases	65
Number of Runs	12
Z	-5.375
Asymp. Sig. (2-tailed)	.000

a. Median

Figure 4.3(a) Result of Run test from SPSS : Company A

Runs Test

	ProductB
Test Value[a]	775.00
Cases < Test Value	32
Cases >= Test Value	33
Total Cases	65
Number of Runs	20
Z	-3.374
Asymp. Sig. (2-tailed)	.001

a. Median

Figure 4.3(b) Result of Run test from SPSS : Company B

Runs Test

	ProductC
Test Value[a]	1440.00
Cases < Test Value	31
Cases >= Test Value	34
Total Cases	65
Number of Runs	10
Z	-5.871
Asymp. Sig. (2-tailed)	.000

a. Median

Figure 4.3(c) Result of Run test from SPSS : Company C

2) Seasonal Variation components

Kruskal-Wallis test are used to test the season component in this study. The test will investigate that ach month since January until December affect to the pallets data by assumption below.

H_0 : The mean of 12 groups of data are not significant different.

H_a : The mean of 12 groups of data are significant different.

The results from software found that the chi square value (χ^2) for company A equals 48.395, which falls in the critical area. P-value of company A data is 0.000 which is less than significant level. While the chi square value for company B and C equal to 8.383 and 2.671 which don't fall in the critical area and the p-value are 0.679 and 0.994 which more than the significant level as figure 4.4. From the result, the null hypothesis of company A is rejected, while the null hypothesis of company B and C cannot be rejected. In the other word, historical data for company A has seasonal component, while B and C doesn't have this component.

Test Statistics[a,b]

	ProductA
Chi-Square	48.395
df	11
Asymp. Sig.	.000

a. Kruskal Wallis Test

b. Grouping Variable: MONTH, period 12

Figure 4.4 (a) Result of Kruskal-Wallis test from SPSS : Company A

Test Statistics[a,b]

	ProductB
Chi-Square	8.383
df	11
Asymp. Sig.	.679

a. Kruskal Wallis Test

b. Grouping Variable: MONTH, period 12

Figure 4.4 (b) Result of Kruskal-Wallis test from SPSS : Company B

Test Statistics[a,b]

	ProductC
Chi-Square	2.671
df	11
Asymp. Sig.	.994

a. Kruskal Wallis Test

b. Grouping Variable: MONTH, period 12

Figure 4.4 (c) Result of Kruskal-Wallis test from SPSS : Company C

From all hypotheses testing can be concluded that the historical data for all company are not in stationary state which has trend and seasonal component. Data from company A need to be transformed by differencing in both trend and seasonal but data from company B and C need to be transformed by differencing only trend component. After differencing, all historical data will be in a stationary state, which is suitable for Box-Jenkins model.

4.2.1.2 Define model structure

ARIMA model has 6 variables that we need to define which are p,d,q,P,D, and Q. The variable d and D is the number of differencing for trend and season consequently. The variable p is the index for autoregressive process and q is the index for moving average process while P and Q are the same in the seasonal time lag. From literature, we can define these variable for model by investigate from pattern of ACF and PACF graph. The interface of SPSS to create ARIMA model are shown as figure 4.5.

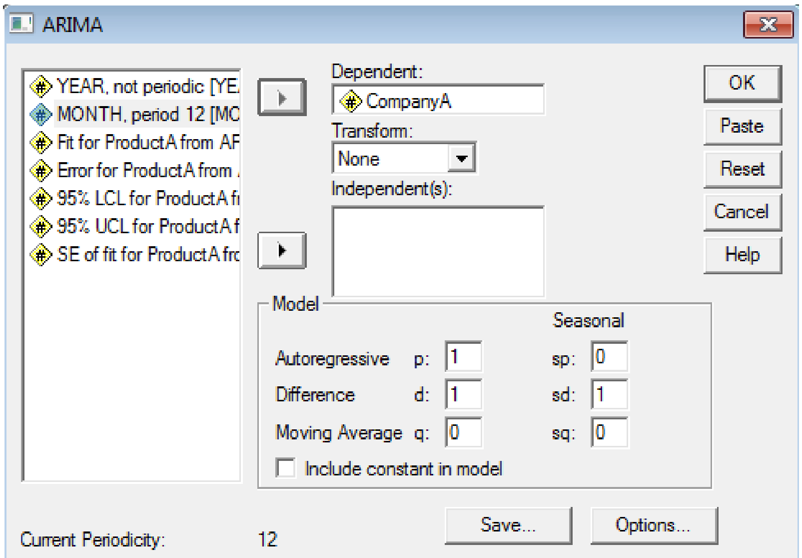


Figure 4.5 SPSS Interface for ARIMA

For company A, the historical data has trend and seasonal component so the d and D index will be assign equal to 1 because of the differencing. The other variables can be identified by using pattern of sample ACF and sample PACF of the properly transformed and differenced series. From the figure 4.6(a) and 4.6(b), the sample ACF tails off as exponential decay and the sample PACF cuts off after lag 1 which is the pattern of autoregressive order 1 , while the ACF and PACF doesn't show the significant value in lag 12 which is the season lag for model. So, the index p can be define as 1 and the others are define as 0. The appropriate model for this time series is ARIMA(1,1,0)(0,1,0)$_{12}$

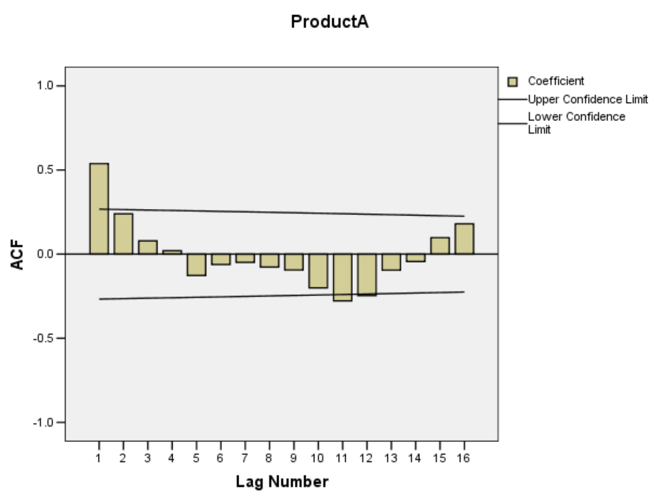


Figure 4.6(a) ACF Plot of data from Company A

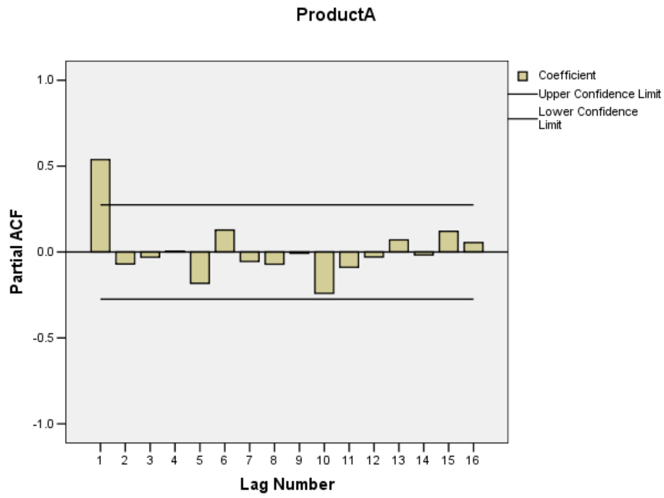


Figure 4.6(b) PACF Plot of data from Company A

For company B, the historical data has only trend elimination, so the d and D can be assign equal to 1 and 0 respectively . The other variables can be identified by using pattern of sample ACF and sample PACF. From the figure 4.7(a) and 4.7(b), both sample PACF has significant value at lag 1 and tails off after since lag 2, the sample ACF tails off as exponential decay and the sample. This pattern can be defined that the index p and q are 0 and 1. Moreover, the ACF has the significant value at lag 12 so the index P and Q will be 0. The appropriate model for this time series is ARIMA(1,1,0)$(0,0,0)_{12}$ or ARIMA(1,1,0)

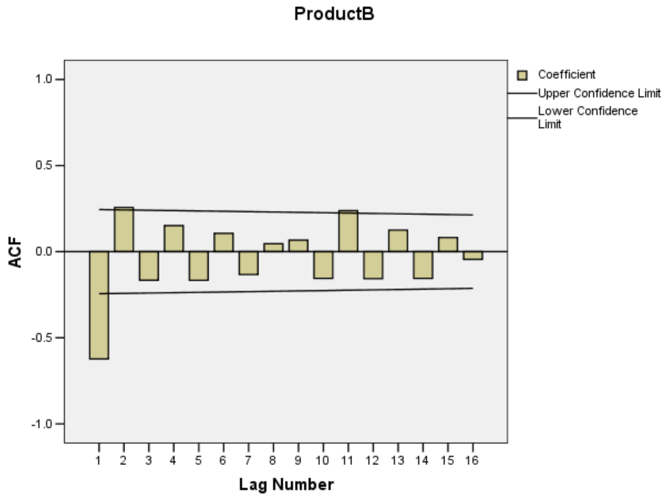


Figure 4.7(a) ACF Plot of data from Company B

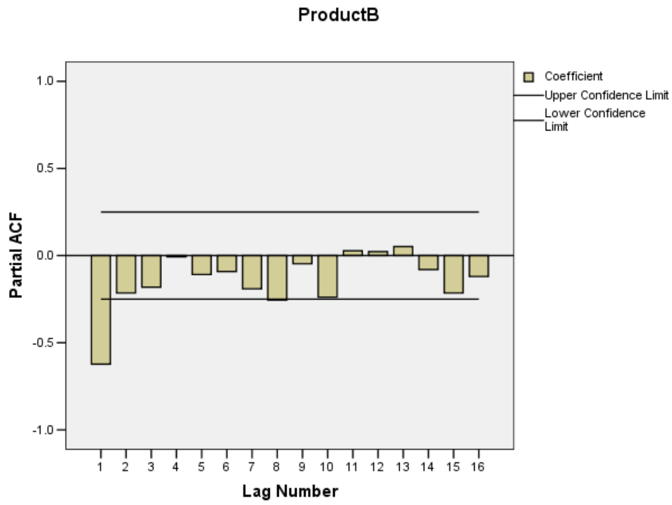


Figure 4.7(b) PACF Plot of data from Company B

For company C, the historical data has only trend elimination, so the d and D can be assign equal to 1 and 0 respectively the same as company B. The other variables can be identified by using pattern of sample ACF and sample PACF of the properly transformed and differenced series. From the figure 4.8(a) and 4.8(b), both sample ACF and PACF has significant value at lag 1 and tails off after since lag 2. This pattern can be defined that the index p and q are 1. The index at seasonal lag has no significant so the index P and Q will be 0. So, the appropriate model for this time series is ARIMA(1,1,0)(0,0,0)$_{12}$

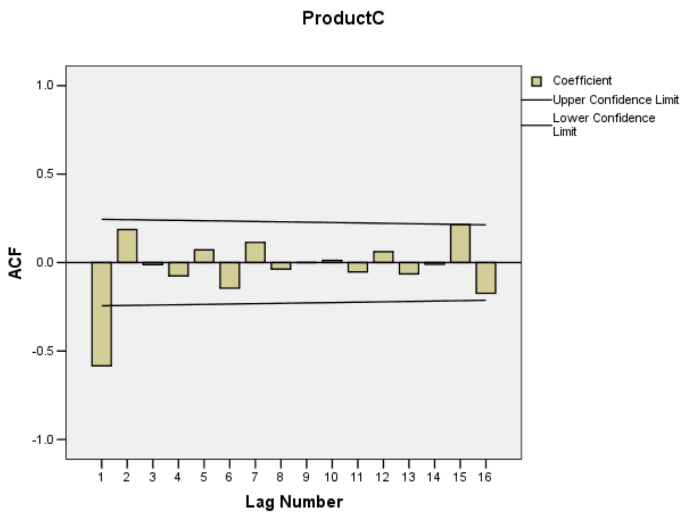


Figure 4.8(a) ACF Plot of data from Company C

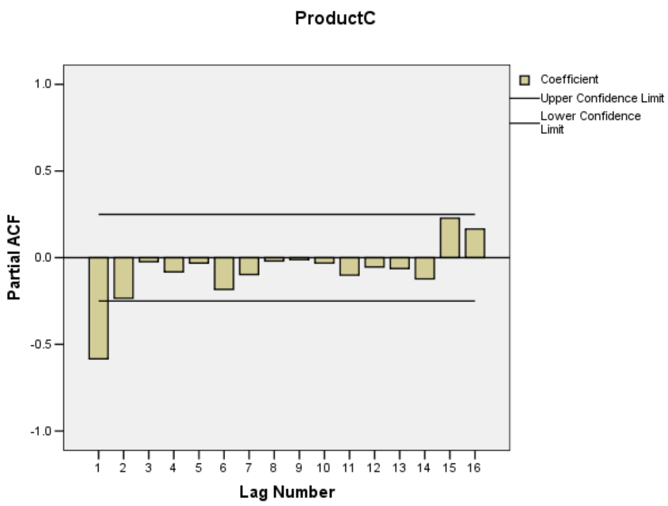


Figure 4.8(b) PACF Plot of data from Company C

4.2.1.3 Parameters Estimation

The appropriate model is ARIMA(1,1,0)(0,1,0)$_{12}$ for company A. The coefficient of autoregressive process will be estimate by software which is 0.635 as figure 4.9(a). The value of this parameter can be use because the p-value from hypothesis testing equal to 0.00 which is less than significant level at 0.05. So, the null hypothesis which is the parameter has value equal to 0 will be rejected. The forecasting equation for company A are shown as below.

Forecasting equation for company A : $0.635(1-B)(1-B_{12})Y_t = e_t$

Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	.635	.118	5.384	.000

Melard's algorithm was used for estimation.

Figure 4.9(a) Parameter estimation of Company A

The appropriate model is ARIMA(1,1,1)(0,0,1)$_{12}$ for company B. The coefficient of autoregressive process will be estimate by software which is 0.635 as figure 4.9(b). The value of this parameter can be used because the p-value from hypothesis testing equal to 0.00 which is less than significant level at 0.05. So, the null hypothesis which is the parameter has value equal to 0 will be rejected. The forecasting equation for company B are shown as below.

Forecasting equation for company B : $-0.643(1-B)Y_t = e_t$

Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	-.643	.101	-6.387	.000

Melard's algorithm was used for estimation.

Figure 4.9(b) Parameter estimation of Company B

The appropriate model is ARIMA(1,1,1)$(0,0,0)_{12}$ for company C. The coefficient of autoregressive process will be estimate by software which is 0.635 as figure 4.9(c). The value of this parameter can be use because the p-value from hypothesis testing equal to 0.00 which is less than significant level at 0.05. So, the null hypothesis which is the parameter has value equal to 0 will be rejected. The forecasting equation for company C are shown as below.

Forecasting equation for company C : $-0.572(1-B)Y_t = e_t$

Parameter Estimates

		Estimates	Std Error	t	Approx Sig
Non-Seasonal Lags	AR1	-.572	.103	-5.567	.000

Melard's algorithm was used for estimation.

Figure 4.9(c) Parameter estimation of Company C

4.2.1.4 Testing Model

After model was trained to be fit for the data, the testing set which was separated in the past will be feed to the model to measure the accuracy of the model. From the result, the forecasting of Company A have the most accurate by the percent as the table 4.1 which has the smallest value of MAD ,MSE and also MAPE. The Company B, and the company C is the second and third respectively. The output from model and the actual data was plot and shown in the graph as figure 4.10.

ARIMA Performance	MAD	RMSE	MAPE
Company A	**65.94**	**74.93**	**11.59%**
Company B	113.15	148.28	12.52%
Company C	379.20	419.74	19.93%

Table 4.1 Performance ARIMA model

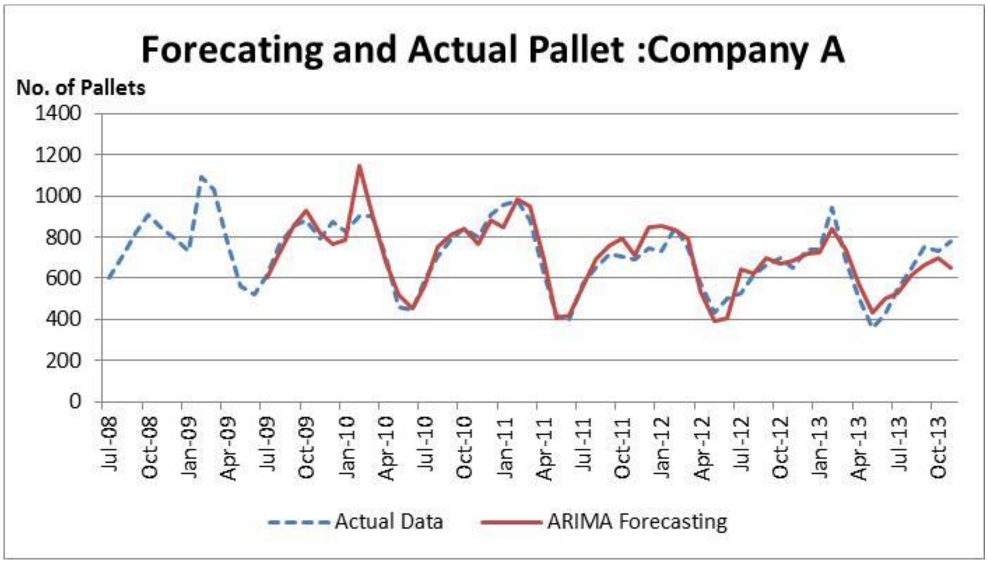


Figure 4.10(a) Actual and Forecasting output of Company A from ARIMA

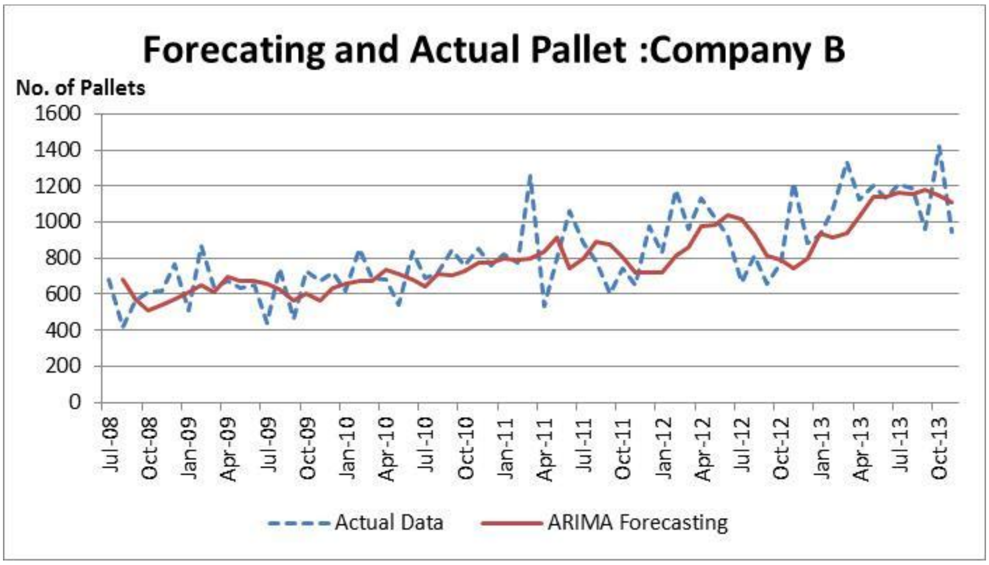


Figure 4.10(b) Actual and Forecasting output of Company B from ARIMA

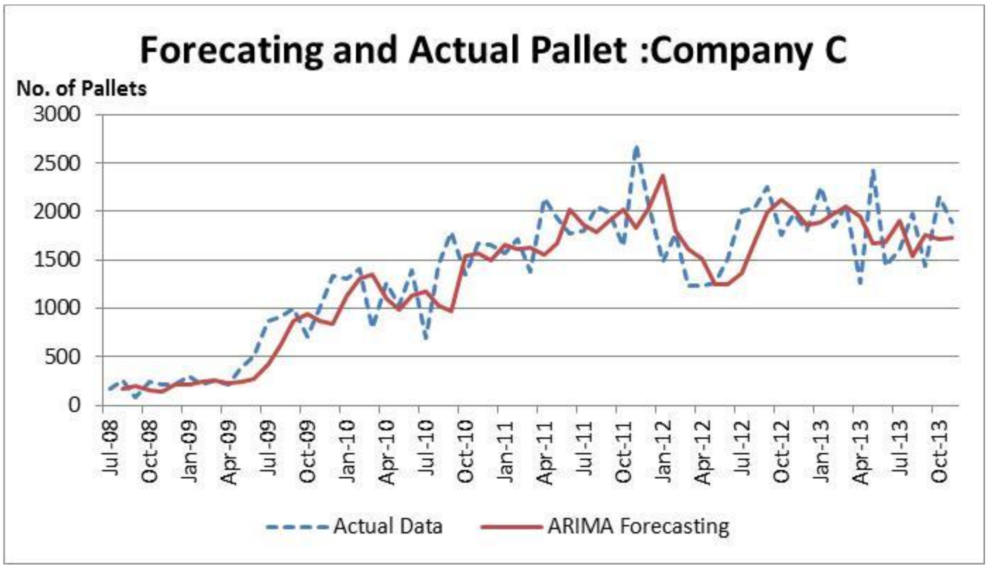


Figure 4.10(c) Actual and Forecasting output of Company C from ARIMA

4.2.2 Time delay neural network (TDNN)

In the process of model building, the model of TDNN was written in the Matlab code, which consists of three parts: Input part, training part and testing part as figure 4.11. The input part has written to upload the input into the model which has two sets of data, training set and testing set. The second part is model training which consists of many important commands. User need to adjust the model by change the parameter in this part such as the number of maximum iteration, the number of input and hidden node , training algorithm etc. The last is the testing and display output. Testing data set will feed into the fitted model and measure the accuracy and finally display the output. For detail and source code of modeling will be describe in the appendix B.

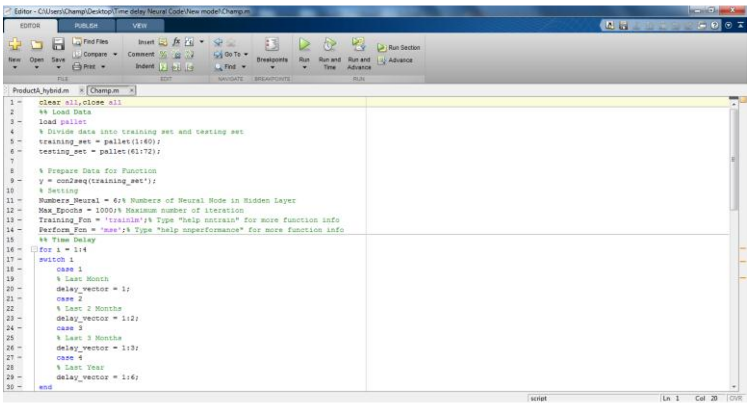


Figure 4.11 Source Code for TDNN model

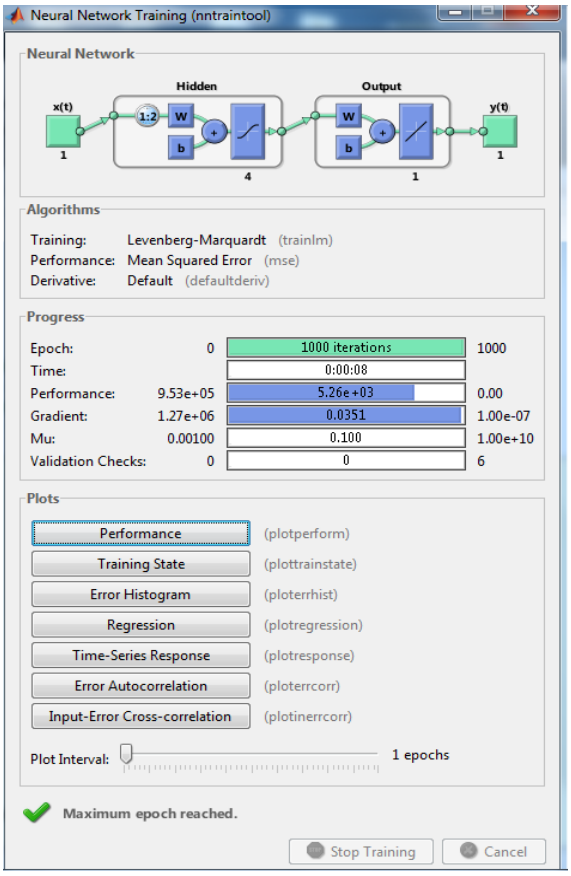


Figure 4.12 The interface of model TDNN

4.2.2.1 Define a network structure and Training model

From literature review, the number of input and hidden node will be selected by trial and error. In this research , the number of input is four which are 1 time lag , 2 time lag ,3 time lag ,12 time lag and 13 time lag. The meaning of 1 time lag is the data from last month will be feed into model ,as same as 2,3,12 and 13 time lag which is the data from last 2 ,3,12 and 13 month will be feed into model. For the hidden node, the number was vary from 1 to 12. The number of input and number of hidden node will be combine and finally, the best performance model with lowest error will be selected. The model was run and simulate for 10 replications and the total result are averaged to the output of model.

From the result of model , the model of company A has the lowest error when the model has input with 13 time lag and the number of hidden node is 12 as table 4.2 (a). For the company B, the selected model is model with 12 time lag input and the number of hidden node is 9 as table 4.2 (b). Finally, the model of company C should have 3 time lag input and hidden node is equal to 8 as table 4.2 (c).

RMSE Company A	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	160.33	209.62	152.82	230.71	171.63	156.42	160.03	201.84	268.53	175.34	250.12
2	180.01	161.49	158.05	201.59	146.22	127.78	137.06	165.62	167.92	166.64	150.34
3	174.70	238.45	138.73	133.15	140.62	154.03	148.00	183.49	184.50	178.59	163.45
12	133.97	115.30	112.75	112.79	131.05	119.64	124.74	127.56	110.48	273.47	299.01
13	133.71	126.73	121.89	145.01	139.33	177.28	150.01	132.23	171.76	124.89	**95.29**

Table 4.2(a) RMSE of Company A vary by time lag and hidden node for TDNN

RMSE Company B	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	215.45	322.97	254.31	230.39	187.33	176.64	209.61	157.31	188.14	212.90	180.60
2	211.76	194.30	270.35	187.16	181.56	180.98	198.17	254.69	215.96	170.14	263.82
3	191.46	176.36	179.22	220.95	169.08	164.87	275.36	221.16	197.31	214.76	179.00
12	153.24	147.52	145.65	162.35	203.63	173.29	204.12	**107.45**	134.50	159.37	268.33
13	215.88	132.89	159.47	196.37	192.24	201.75	272.69	232.41	197.55	281.22	259.89

Table 4.2(b) RMSE of Company B vary by time lag and hidden node for TDNN

RMSE Company C	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	317.89	305.50	312.91	344.42	302.70	310.67	716.55	581.94	383.44	405.41	785.12
2	397.66	484.65	311.76	383.16	303.08	327.39	302.15	459.29	433.78	356.89	447.98
3	414.85	491.10	293.27	331.70	466.65	337.58	**265.17**	501.70	316.71	343.61	314.21
12	311.45	362.00	319.02	424.41	375.33	419.97	323.73	593.67	508.82	515.39	514.76
13	322.63	361.74	290.72	531.35	376.71	536.26	542.62	643.70	551.55	411.24	772.48

Table 4.2(c) RMSE of Company C vary by time lag and hidden node for TDNN

4.2.2.2 Testing model

After testing the model, the result show that TDNN can forecast the data from company C and B in very high accuracy with lowest MAD, RMSE and MAPE while the TDNN was not effective too much. The output from model and the actual data was plot and shown in the graph as table 4.3. The output from model and the actual data was plot and shown in the graph as Figure 4.13.

TDNN Performance	MAD	RMSE	MAPE
Company A	127.21	144.23	25.98%
Company B	87.98	107.11	7.26%
Company C	**77.04**	**97.90**	**6.05%**

Table 4.3 Performance of TDNN model

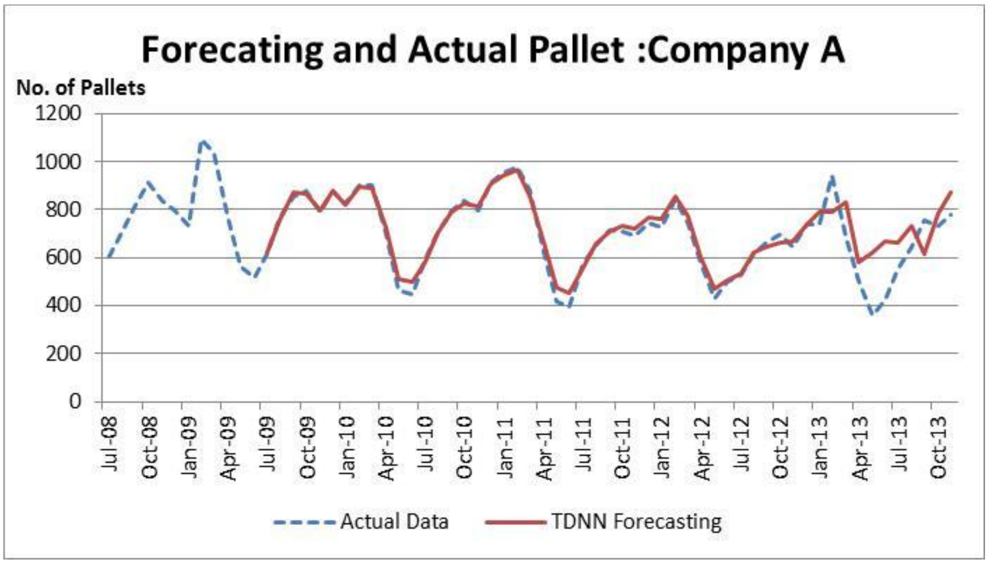


Figure 4.13(a) Actual and Forecasting output of Company A from TDNN

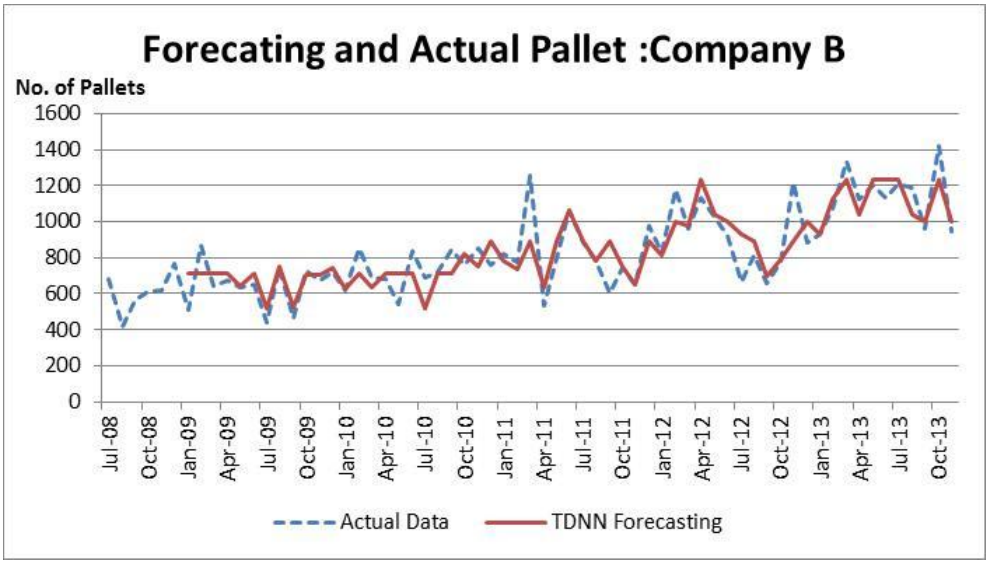


Figure 4.13(b) Actual and Forecasting output of Company B from TDNN

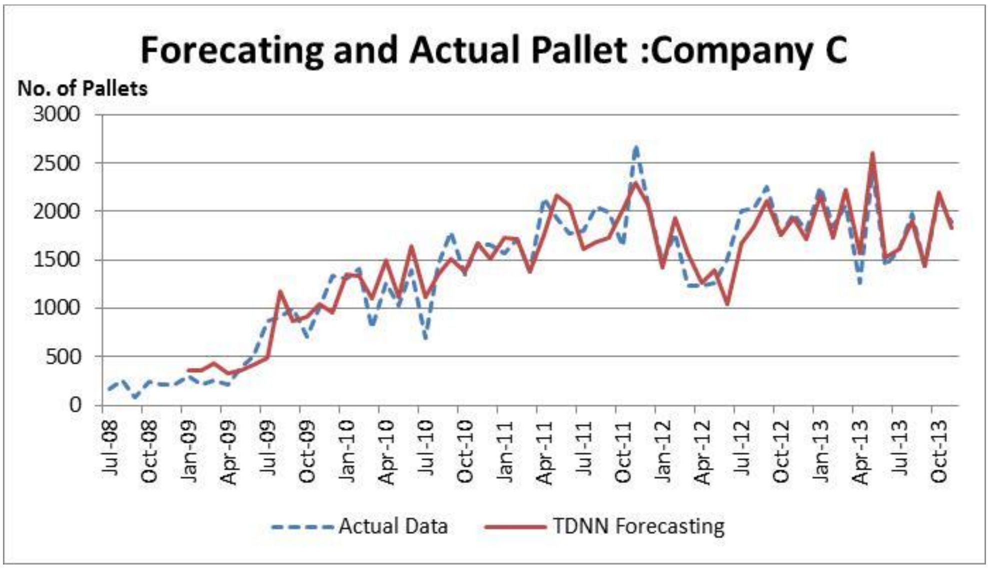


Figure 4.13(c) Actual and Forecasting output of Company C from TDNN

4.2.3 Hybrid model

In the process of model building, the model of Hybrid need two software to support which are SPSS and Matlab. The linear phase will compute in SPSS and the non-linear phase will compute in Matlab. The model consists of three parts: Input part, training part and testing part as figure 4.14. The input part is special than TDNN model because there are 4 inputs. There are training data set, the residual from ARIMA model, the testing set and the result from ARIMA testing set. The training set data was uses to train the model by using the residual from ARIMA be the model target. The ARIMA testing set was use to sum with the output from model and compare with the testing set to measure the performance.

The second part is model training which consists of many important commands. User need to adjust the model by change the parameter in this part such as the number of maximum iteration, the number of input and hidden node , training algorithm etc. The last is the testing and display output. Testing data set will feed into the fitted model and measure the accuracy and finally display the output. The code of model will describe in the appendix B.

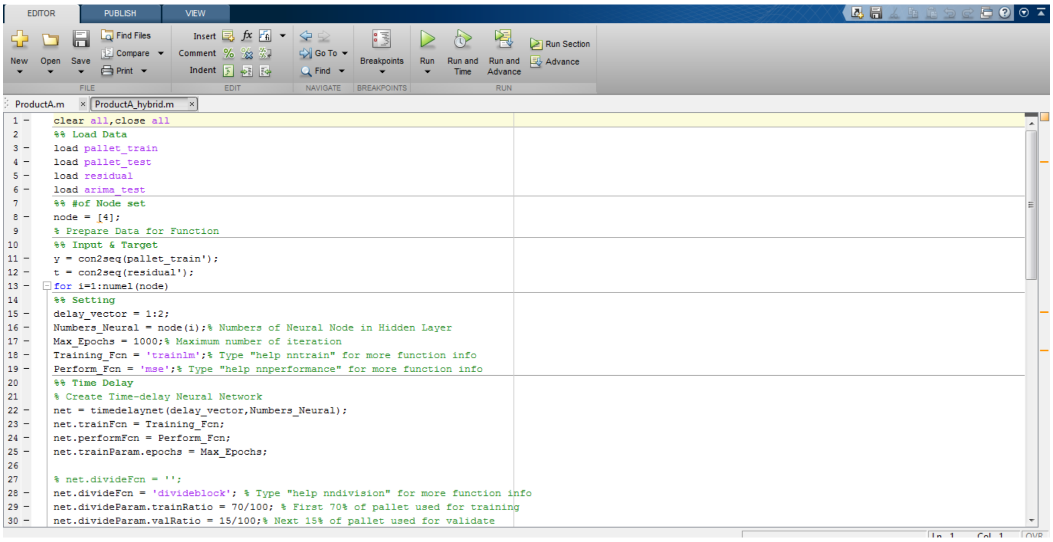


Figure 4.14 Source Code for Hybrid model

4.2.3.1 Define a network structure and Training model

In this research , the number of input is four which are 1 time lag , 2 time lag ,3 time lag and 12 time lag. The meaning of 1 time lag is the data from last month will be feed into model ,as same as 2,3 and 12 time lag which is the data from last 2 ,3 and 12 month will be feed into model. For the hidden node, the number was vary from 1 to 12. The number of input and number of hidden node will be combine and finally, the best performance model with lowest error will be selected. The model was run and simulate for 10 replications and the total result are averaged to the output of model.

From the result of model , the model of company A has the lowest error when the model has input with 12 time lag and the number of hidden node is 5 as table 4.4 (a). For the company B, the selected model is model with 2 time lag input and the number of hidden node is 10 as table 4.4 (b). Finally, the model of company C should have 2 time lag input and the hidden node is equal to 5 as table 4.4 (c).

RMSE Company A	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	238.83	180.74	182.13	170.32	224.08	180.52	256.03	212.92	173.65	241.26	547.07
2	178.57	168.98	180.26	239.25	213.49	171.39	293.39	210.36	200.78	246.02	190.83
3	176.39	176.39	189.36	149.08	369.76	195.97	183.41	237.23	257.65	181.63	323.78
12	173.50	322.14	331.06	76.77	173.51	174.75	572.11	361.48	499.78	155.76	192.73
13	435.49	176.18	186.05	263.40	325.63	179.30	212.02	292.69	327.93	306.02	337.72

Table 4.4(a) RMSE of Company A vary by time lag and hidden node for Hybrid

RMSE Company B	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	168.24	171.21	177.09	186.97	166.84	158.68	168.04	163.46	317.20	151.02	233.34
2	172.73	172.64	161.94	159.73	209.15	165.21	249.11	141.77	**140.44**	145.86	169.72
3	166.58	172.02	190.97	161.13	156.94	168.52	152.66	221.03	169.32	146.47	167.59
12	159.20	272.33	165.71	184.33	254.09	285.02	203.86	308.09	283.54	272.61	180.74
13	206.77	168.39	217.09	229.14	303.75	235.25	318.74	227.36	305.60	217.63	240.73

Table 4.4(b) RMSE of Company B vary by time lag and hidden node for Hybrid

RMSE Company C	Hidden Node										
Time Lag	2	3	4	5	6	7	8	9	10	11	12
1	327.07	337.29	301.75	371.00	325.35	395.87	321.51	310.46	329.66	313.85	290.97
2	362.29	333.04	329.63	**262.08**	342.51	317.08	506.64	322.24	410.73	282.89	311.01
3	354.75	315.81	281.09	383.71	529.18	308.36	610.53	330.07	439.09	365.36	335.51
12	485.21	895.22	682.98	289.58	326.87	433.07	342.85	466.72	489.01	363.45	768.80
13	323.65	269.37	574.69	334.69	552.52	403.13	270.14	423.73	739.25	499.22	342.42

Table 4.4(c) RMSE of Company C vary by time lag and hidden node for Hybrid

4.2.3.2 Testing model

After testing the model, the result show that Hybrid can forecast the data from company C and B in very high accuracy with lowest MAD,RMSE and MAPE while the Hybrid was not effective too much. The output from model and the actual data was plot and shown in the graph as table 4.5. The output from model and the actual data was plot and shown in the graph as Figure 4.15.

Hybrid Performance	MAD	RMSE	MAPE
Company A	**58.13**	**70.85**	**9.21%**
Company B	105.35	135.52	11.10%
Company C	290.78	349.75	14.72%

Table 4.5 Performance of Hybrid model

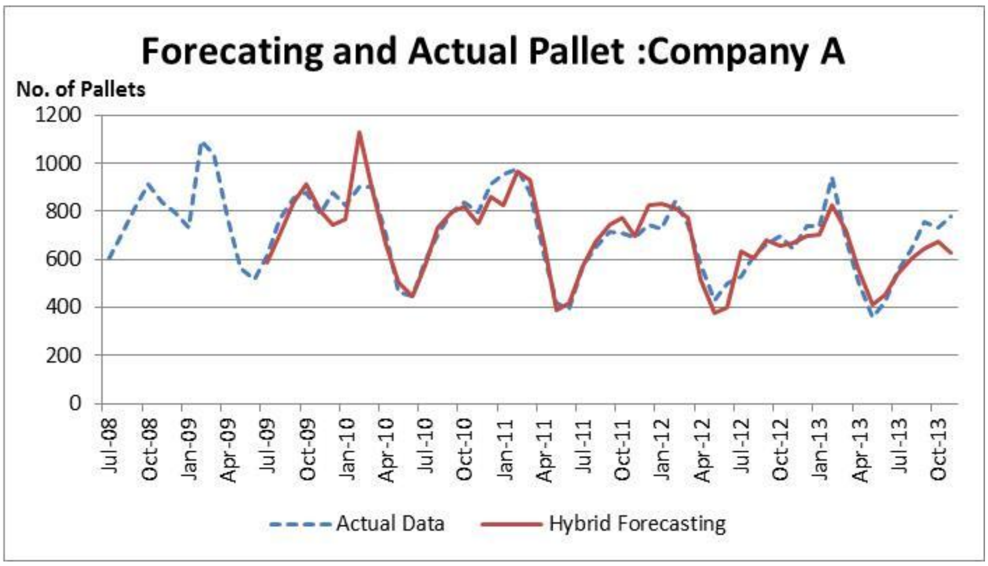


Figure 4.15(a) Actual and Forecasting output of Company A from Hybrid

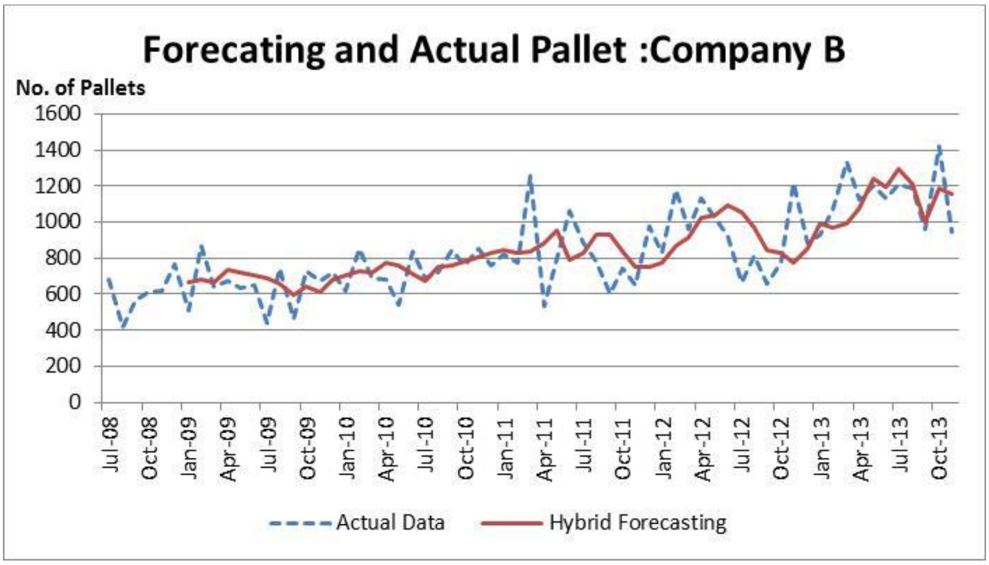


Figure 4.15(b) Actual and Forecasting output of Company B from Hybrid

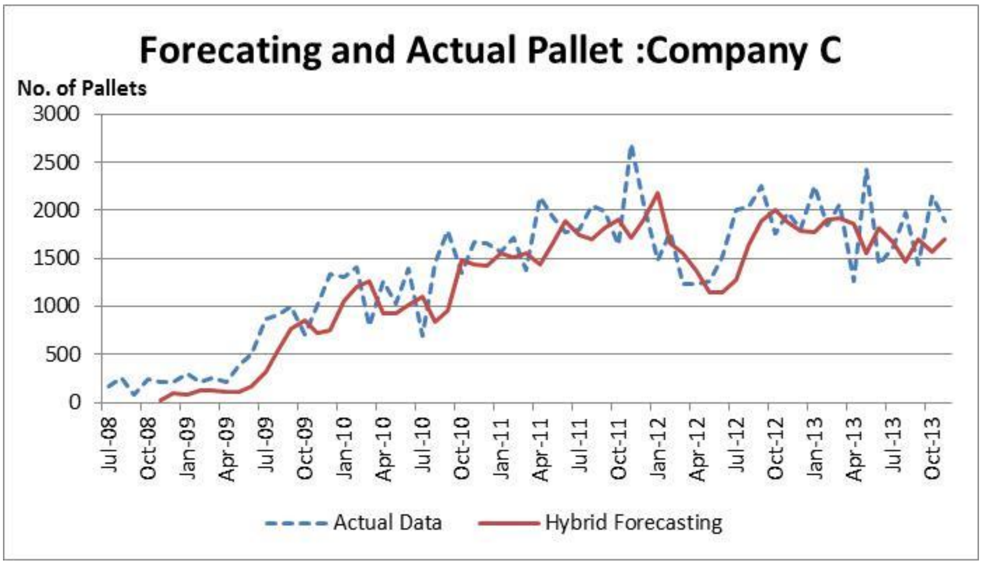


Figure 4.15(c) Actual and Forecasting output of Company C from Hybrid

4.3 Comparison of model performances

In order to compare the efficiency of forecasting model, the comparison could be separated as 2 aspects. The first aspect is to compare the error factor which occurs from each forecasting model. The second aspect is plotting x-y graph between forecasting factor and actual factor, in order to show the linearly relation.

The fitted model for each method is shown as the previous section. In this topic, the model will be compared to each other in accuracy by using RMSE, MAD and MAPE. The result shown that hybrid performs the best model for only company A because the pattern of data is linear with seasonal component. So, the hybrid which combines both ARIMA and TDNN can forecast the output close to actual value. For company B and C, TDNN is the best model for forecasting because the pattern of data for this two company is non-linear as table 4.6. ARIMA and hybrid model cannot forecast in high accuracy because the data is not suitable to the model.

Model	Trend	Seasonal	MAD	RMSE	MAPE
Company A	✓	✓			
ARIMA			65.94	74.93	11.59%
TDNN			127.21	144.23	25.98%
Hybrid			**58.13**	**70.85**	**9.21%**
Company B	✓	✗			
ARIMA			113.15	148.28	12.52%
TDNN			**87.98**	**107.11**	**7.26%**
Hybrid			105.35	135.52	11.10%
Company C	✓	✗			
ARIMA			379.20	419.74	19.93%
TDNN			**77.04**	**97.90**	**6.05%**
Hybrid			290.78	349.75	14.72%

Table 4.6 Comparison of model performances

Moreover, the performances of forecasting can be measure by using linearly relation of actual values and forecasting value. To find the linearly relation, the forecasting values and the actual values will be plot in x-y coordinates, and find the trends line of these coordinates. If the forecasting value equal or close to the actual value, the trend line will be linear. The linearly can be measure by using R-square value (R^2). If r-square value is closer to 1, the coordinate is almost in linear relation. In the other word, the forecasting values and actual value are nearly the same value. From the figure , the forecasting values and actual values of company A were plot in x-y axis separated by model. R^2 value of Hybrid model is the highest values which is 0.8711. From the figure , the forecasting values and actual values of company B were plot in x-y axis separated by model. R^2 value of TDNN model is the highest values which is 0.7211. From the figure , the forecasting values and actual values of company C were plot in x-y axis separated by model. R^2 value of TDNN model is the highest values which is 0.7641. The result show that the Hybrid perform the best model for Company A, while the TDNN is the best for Company B and C.

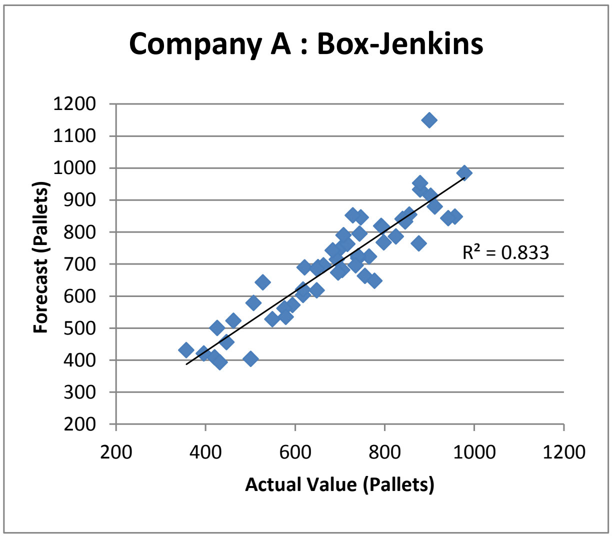


Figure 4.16(a) Actual and Forecasting plot from Company A by Box-Jenkins

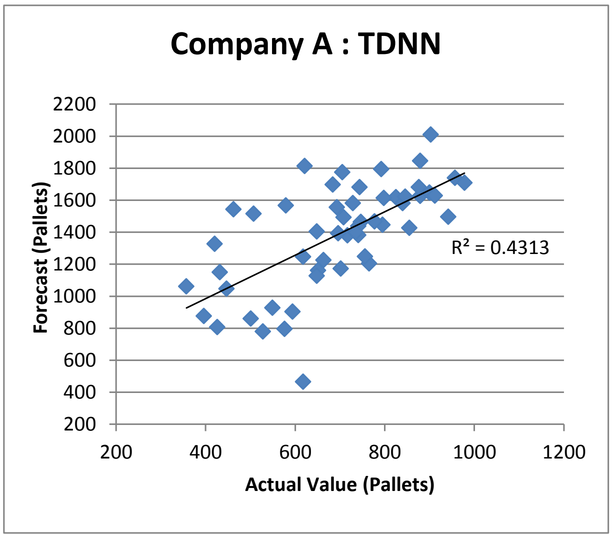


Figure 4.16(b) Actual and Forecasting plot from Company A by TDNN

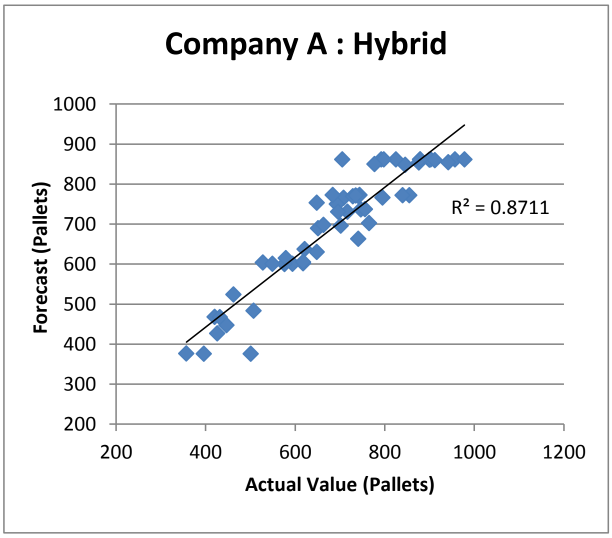


Figure 4.16(c) Actual and Forecasting plot from Company A by Hybrid

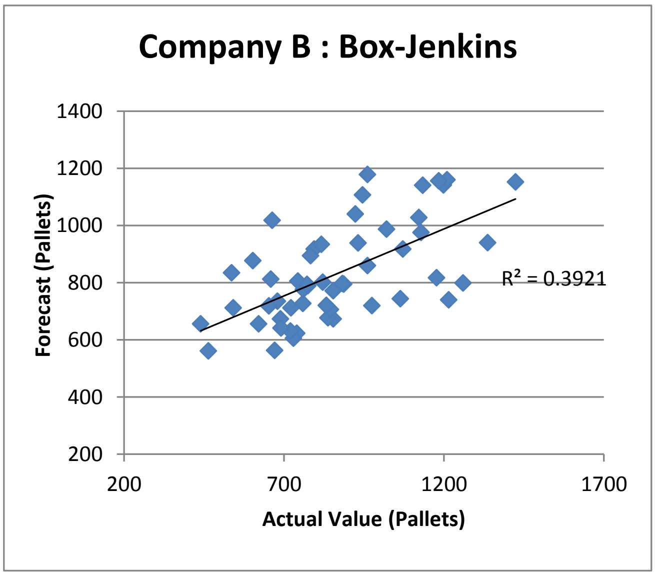


Figure 4.17(a) Actual and Forecasting plot from Company B by Box-Jenkins

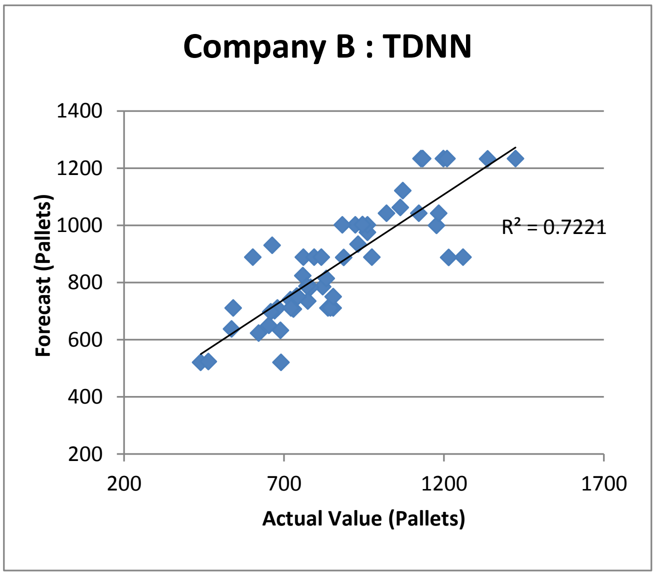


Figure 4.17(b) Actual and Forecasting plot from Company B by TDNN

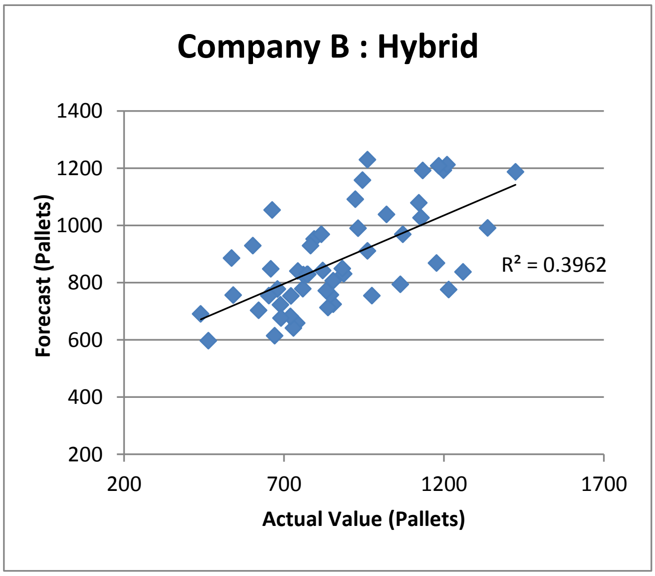


Figure 4.17(c) Actual and Forecasting plot from Company B by Hybrid

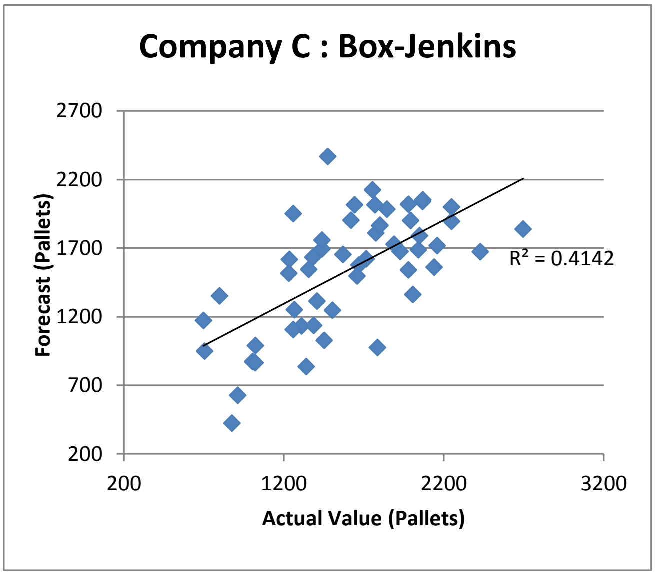


Figure 4.18(a) Actual and Forecasting plot from Company C by Box-Jenkins

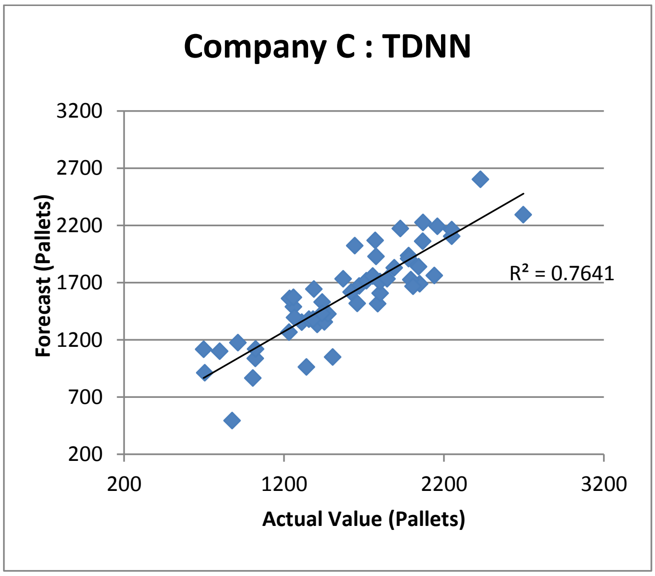


Figure 4.18(b) Actual and Forecasting output of Company C by TDNN

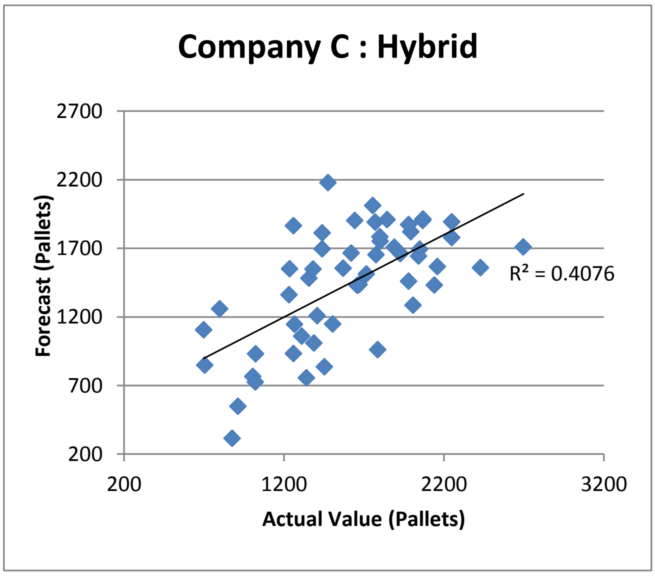


Figure 4.18(c) Actual and Forecasting output of Company C by Hybrid

4.4 The contribution of forecast modeling to warehouse management

4.4.1 Model implementation

TDNN perform the best for both short-term and long term forecasting. The output of short term and long term were compared in this study while the short term forecasting means 1 month forecasting and long term means 3 months forecasting. TDNN also perform the best performance to forecast with lowest error. For ARIMA and Hybrid model, the performance of these two method are related because the hybrid model combine the ARIMA and TDNN. If ARIMA can forecast with high accuracy, the output will feed to hybrid model and the hybrid model will forecast in high accuracy as table 4.7.

Model	MAD			RMSE			MAPE		
	1Month	3Month	Difference	1Month	3Month	Difference	1Month	3Month	Difference
Company A									
ARIMA	21.25	56.74	35.49	62.05	74.38	12.33	14.07%	20.83%	6.76%
TDNN	186.74	237.92	51.18	195.87	237.92	42.05	45.61%	66.64%	21.03%
HYBRID	**27.11**	**49.82**	22.71	**31.87**	**49.82**	17.95	**6.94%**	**13.96%**	7.01%
Company B									
ARIMA	37.93	57.31	19.37	43.82	57.31	13.49	4.78%	7.28%	2.50%
TDNN	**36.12**	**55.02**	18.90	**36.12**	**65.32**	29.20	**3.01%**	**5.18%**	2.17%
HYBRID	43.05	66.38	23.33	43.05	69.71	26.66	3.59%	7.30%	3.71%
Company C									
ARIMA	431.35	757.46	326.11	489.28	757.46	268.18	22.08%	31.17%	9.09%
TDNN	**93.73**	**182.39**	88.66	**119.11**	**182.39**	63.29	**6.05%**	**7.51%**	1.45%
HYBRID	451.07	915.84	464.77	575.43	915.84	340.41	17.03%	37.69%	20.66%

Table 4.7 Comparison of model performances in short and long term

The results of the study show that the model TDNN is appropriate to the non-linear data because of high accuracy. From the case study, almost 70% from total company has the non-linear trend data. Therefore, TDNN model should be the forecasting model for import and export warehouse. For implementation, warehouse user should re-train model every 3 months for accuracy and update the data. User need to adjust the model in two parts which are input and hidden node. The other part such as training algorithm or activation function should not adjust.

1. Input : The new data should be update to the input data set

2. Hidden node : User should adjust the hidden node for the lowest error by trial and error while the hidden node should not be exceed to the input.

4.4.2 Improve operation performance

For the As-is operation, the new shipment product will delivery to warehouse for separate the products by SKUs and lot number before transfer to storage in rack because of GMP. and warehouse policy. The high number of pallet were transfer to the location where use to storage the product. If there is no location, the warehouse staff needs to transfer the other product and prepare the area. This process will take a time and interrupt the put away process, so the average process time is trend to be high. The as is operated time for put away was measure by using average time stamp in system and divided by number of pallet. For the process time after implemented forecasting, the process time is from the put away time and advanced location preparation. In addition, bin to bin product transfer time will be add to the process time if the forecasted values less than the actual income product. The calculation algorithm is summarized by the equation as follow.

$$Total\ time\ saving = (As\ is\ operated\ time - To\ be\ operated\ time) * total\ pallet$$

while

$$As\ is\ operated\ time = (average\ put\ away\ time\ per\ shipment)/(number\ of\ pallet)$$

<u>Case forecasting value exceed the actual value</u>

$$To\ be\ operation\ cost = preparation\ time + average\ put\ away\ time$$

<u>Case forecasting value less than the actual value</u>

$$To\ be\ operation\ cost = preparation\ time + average\ put\ away\ time + bin\ to\ bin\ transfer\ time$$

After forecasting implementation, the process time for put away reduce by 31%,16% and 18% for Company A , B and C respectively as figure 4.16.

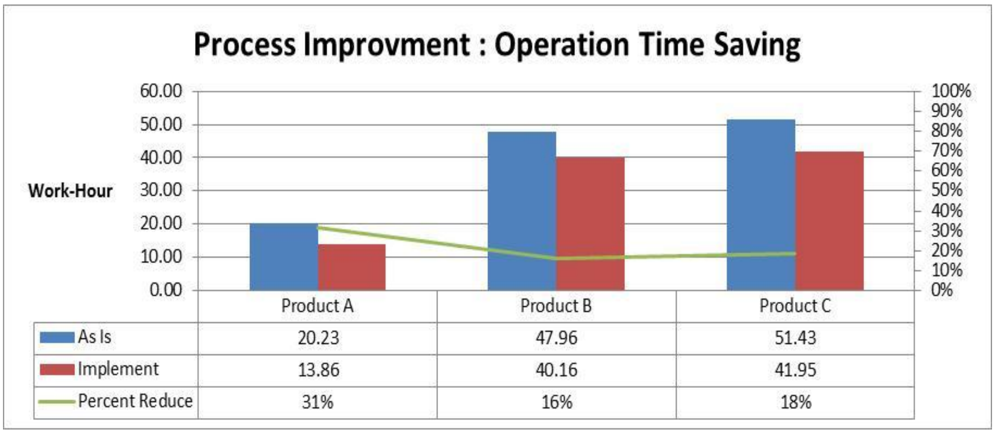


Figure 4.19 As is and implement operation time

For to be operation, the high accuracy can reduce operation time in put away process. The result show that the average process time per shipment reduce average 22% in average as figure 4.17 .

Figure 4.20 As is and To be time line for warehouse process time

4.4.3 Cost Saving

Accuracy forecasting is necessary to the decision making for warehouse management especially in space preparation. Warehouse manager can prepare the bin location for the coming product in advance. From the previous part, the high accuracy forecasting can reduce the operation time by average 22%. The decrease of operation time can convert to the cost of staff and machine as below.

The saving cost calculate by finding the difference between as-is operation cost and to be operation cost. The as-is operation cost is from the product of average as-is process time and average cost per hour. To be operation cost was divided for calculate in two cases. First ,the forecasting value is exceed than actual value. The to be operation cost calculate from the product of total process time and average cost per hour while the total process time is summation of average to be process time and preparing time per pallet. Second, the forecasting value is less the actual value. In this case , warehouse staff need to transfer the other company pallet for the location. The to be operation cost calculate from the product of total process time and average cost per hour while the total process time is summation of average to be process time , preparing time per pallet and the as is process time for remain pallet. For the cost calculation, the equations are shown as follow.

$$Cost\ saving\ per\ pallet = As\ is\ operation\ cost - To\ be\ operation\ cost$$

while

$$As\ is\ operation\ cost = (average\ time\ per\ pallet * operation\ cost\ per\ hour)$$

Case forecasting value exceed the actual value

$$To\ be\ operation\ cost = (preparation\ time + average\ time\ per\ pallet) * operate\ cost$$

Case forecasting value less than the actual value

$$To\ be\ operation\ cost = (preparation\ time + average\ time\ + as\ is\ averagr\ time\ per\ pallet) * operate\ cost$$

The operation cost decease due to the reduce of process time. The results showed that the to be cost reduce from as is costs about 20% or 1.81 THB per pallets. If this forecasting can implement to overall product , the total saving cost will be approximately 500,000 THB per year.

CHAPTER V
CONCLUSION

This chapter is divided into two sections; section 5.1, the conclusion of this research and section 5.2, the recommendations for the hospital and a future research.

5.1 Conclusions

In this research has developed a model in order to forecasting the pallet quantity to be actually stored in warehouse for import and export product by monthly. The Artificial neural network approach compared with 2 different methods which are Time series by Box-Jenkins and Hybrid model, which combines the two models together. The model was constructed by using Time delay that have sigmoid function as activation function and using Bayesian regularization as learning process to forecast the model. However, it was found that TDNN provides the highest accuracy with error less than Box-Jenkins model and hybrid model for both short term (1 month) and long term (3 months) forecasting. However, the development of this predictive model, can improve the operation time reduce by 22% efficiency and reduce cost of operation activities in warehouse average 20% .

According to the data collection, Time series data elements are involved Trend and seasonal components which related to almost every business and can be founded that the developing of forecasting model, Box-Jenkins forecasting model were more accurate whereas TDNN model is likely suitable for the data with no seasonal component is involved. Therefore, when combined forecasting model together with the Hybrid, the results has the most accurate data when data composed with seasonal component. But if consider data with no seasonal component involved, TDNN models can forecast with the higher performance and results give the lowest moving average. In which the result is consistent with related research on the demand forecast for consumer products by Zhang and qi (2005).

From this research studied is to compared the forecasting model between Box-Jenkins TDNN, and Hybrid by TDNN. The resulted from TDNN that the predictive value close to real data as possible which is unlike the research of Zhang (2003). He stated that concluded of Hybrid provides the most accurate due to the nature of the information that Zhang used data similar to Company A , which looks quite linear , which is inclined to the seasonal component and contains different characteristics from that used in this study that very high Non-linear and no matter the season irrelevant. So the results are not consistent with related researches. In order to select the appropriate forecasting model, not only consider the suitable forecasting method but having regard to considering the characteristic of the data is also important because of the characteristic of the data will affect to the efficiency of the forecast models.

Forecasting the quantity of pallet storage in warehouse can be applied to the import and export for businesses that with similar characteristics. Warehouse must support products from many source of companies with vary kind of products in large quantity and consolidate shipment to reduce the cost of transportation. When products delivered to the warehouse, product is sorted into many small pallets by product type and the lot number, so that warehouse requires lots of storage space. However, if the company lack of information system linking between manufacturer, warehouse cannot be able to know the amount of the actual product shipped and prepare for the storage space in advance.

Forecasting pallet for each company is suitable and can be able to prepare for warehouse storage space if divided pallets into separate storage based on each company but not classified into each category of products. And if company divided pallets classified into each category of product, the forecasting product to a product type will be more suitable for the preparation of space.

5.2 Recommendation

Not only Artificial neural network in forecasting with the addition to the input time series forecasting method, but it also be predicted by other factors which involved and the impact on the forecast used in forecasting for example net sale or inventory level as the input data for forecasting. If study the factors that affect the amount for ordering product to store in warehouse, it may be used to study and modify the forecast in the future.

In term of forecasting software, Box-Jenkins model were analyzed in the SPSS software while TDNN were analyzed in Matlab software. Hybrid model were analyzed in both software. So this can make it complicated to develop model and comparing performance especially Hybrid model.

For the problems that occurred in this study, due to the lack of warehouse data interchange with pharmaceutical companies, agents, and manufacturer companies which affect to the warehouse management process. Therefore in the future if it can connect and develop transmission of information through the entire supply chain, it may be used to support IT system and help Warehouse to reduce currently problem and management work flow more efficiency.

REFERENCES

Aggarwal,K.K , Singh,Y.,Chandra,P.,Puri,M. (2005). Bayesian regularization in a neural network model to estimate lines of code using function points. *Computer sciences ,4,505-509.*

Anderson, D., Hines, E.L., Arthur, S.J., and Eiap, E.L., (1993). Application of artificial neural networks to prediction of minor axis steel connections. *Neural Networks and Combinatorial Optimization in Civil and Structural Engineering, 31-37.*

Bo K. Wong, Vincent S. Lai, Jolie Lam,(2000). A bibliography of neural network business applications research: 1994-1998 .*Computers & Operations Research 27 ,1045-1076*

Bonissone, P.P. (1997). Soft computing: the convergence of emerging reasoning technologies. *Soft Computing, 1, 6-18.*

Ching-Wu Chu, Guoqiang Peter Zhang,(2003). A comparative study of linear and nonlinear models for aggregate retail sales forecasting. *Production Economics 86, 217–231*

Dutta, P., Chakraborty, D. & Roy, A.R. (2007). An inventory model for single-period products with reordering opportunities under fuzzy demand. *Computers and Mathematics with applications, 53, 1502-1517.*

Fabio Bellini & Gianna Figa Talamanca, (2005). Runs tests for assessing volatility forecastability in financial time series. *European Journal of Operation Research ,163,102-114.*

Flood,I. and Kartam,N. (1994). Neural Networks in Civil Engineering. Principles and Understanding. *ASCE journal of computing in civil engineering. 8(2), 131-148*

Grayden, D. B., Scordilis, M.S.,(1992) , TDNN vs. Fully Interconnected Multilayer Perceptron: A Comparative Study on Phoneme Recognition, *Proceedings of the Fourth Australian International Conference on Speech Science and Technology, SST-92, Brisbane, Australia, 214-219*

Hamzacebi ,C. ,(2008), Improving artificial neural networks' performance in seasonal time series forecasting , *Information science,178, 4550-4559*

Henry C. Co , Rujirek Boosarawongse, (2007). Forecasting Thailand's rice export: Statistical techniques vs. artificial neural networks .*Computers & Industrial Engineering 53, 610–627*

Hirschen,K. and Schafer,M., (2006). Baysian regularization neural networks for optimizing fluid flow processes, *Computing methods in applied mechanics and engineering,195,481-500* .

Ilan Alon, Min Qi, Robert J. Sadowski,(2001). Forecasting aggregate retail sales : a comparison of arti"cial neural networks and traditional methods. *Journal of Retailing and Consumer Services 8, 147-156.*

Jiawei Han,(2007). Data mining:concepts and techniques.Morgan Kaufmann publishers,2007

Khan, A.I., Topping, B.H.V., and Bahreininejad, A., (1993). Parallel training of neural Networks for finite element mesh generation. *Neural Networks and Combinatorial Optimization in Civil and Structural Engineering: 81-94.*

Ko, M., Tiwari, A. & Mehnen, J. (2010). A review of soft computing applications in supply chain management. *Applied Soft computing, 10, 661-674.*

Kumar, M., Vratb, P. & Shankar, R. (2004). A fuzzy goal programming approach for vendor selection problem in a supply chain. *Computers & Industrial Engineering, 46, 69–85.*

Kuo R.J.,(2001). Sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm. *European Journal of Operational Research 129 , 496-517*

Leo W.G.Strijbosch, ArisA.Syntetos, JohnE.Boylan, EllekeJanssen,(2009).The interaction between forecasting and stock control: The case of non stationary demand . *Production Economics*

Li S.G., X. Kuo, (2008). The inventory management system for automobile spare parts in a central warehouse. *Expert Systems with Applications 34 , 1144–1153*

Lippmann, R.P., (1987). An introduction to computing with neural nets. *IEEE ASSP Magazine,4–22.*

Luis Aburto ,(2007). Richard Weber Improved supply chain management based on Hybrid demand forecasts. *Applied Soft Computing 7,136–144*

Peter Zhang (2003), Time series forecasting using a hybrid ARIMA and neural network model, *Neurocomputing 50 , 159 – 175*

Peter Zhang , Min Qi,(2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research 160 ,501–514*

Rafael S. Gutierreza, Adriano O. Solis, Somnath Mukhopadhyay,(2008). Lumpy demand forecasting using neural networks .*Production Economics 111 ,2409–420*

Reymann, P.J. (1995). Pharmaceutical inventory forecasting at the wright - Patterson medical center. *The Degree of Master of Science in Logistics Management, Department of The Air Force, The Faculty of the School of Logistics and Acquisition Management of the Air Force Institute of Technology Air University.*

Rodriguez, M.A. & Vecchietti, A. (2010). Inventory and delivery optimization under seasonal demand in the supply chain. *Computer and Chemical Engineering, 34, 1705-1718.*

Rogers, J.L., and Lamarsh, W.J., (1992). Application of a neural network to simulate analysis in an optimization process. *Artificial Intelligence in Design, 739-754.*

Samer Saab , Elie Badr , George Nasr,(2001). Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. *Energy 26 ,1–14*

Sibi,P.Jones,S. and Siddarth,P., (2013). Analysis of different activation functions using back propagation neural networks. *Journal of Theoretical and applied information technology ,47,1264-1268*

Simon Heykin (1999). Neural Network: A comprehensive foundation. *Prentice Hall International*

Taesombat (1996), Quantitative forecasting, *Physics center*

Vellidoa, P.J.G. Lisboaa, J. Vaughanb , (1999). Neural networks in business: a survey of applications (1992–1998) .*Expert Systems with Applications 17, 51–70*

Victor R. Prybutok , Junsub Yi , David Mitchell , (2000). A Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations . *European Journal of Operational Research 122, 31-40*

Volkan Ediger, Sertac Akar, Berkin Ugurlu ,(2006) Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model.*Energy Policy 34 , 3836–3846*

William W.S. Wei. (1989). Time Series Analysis. Department of Statistics ,*Temple University.*

Wu,X. and Lim, S.Y., (1993). Prediction of maximum scour depth at spur dikes with adaptive neural networks. *Neural Networks and combinatorial optimization in Civil and structural engineering*

Zadeh, L. A. (1994). Soft computing and fuzzy logic. *IEEE, 48-56.*

APPENDICES

APPENDIX A
CHARACTERISTICS OF ACF AND PACF FOR STATIONARY PROCESSE

1. ARMA(0,0)

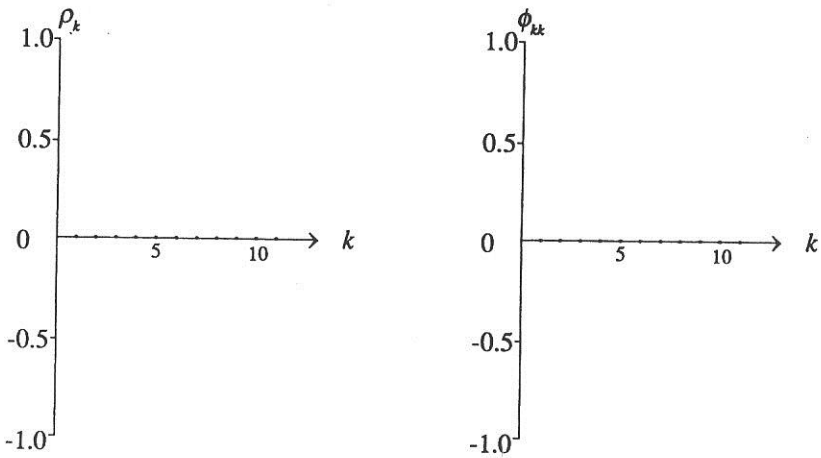


Figure A.1 ACF and PACF correlogram for ARMA(0,0)

2. ARMA(1,0)

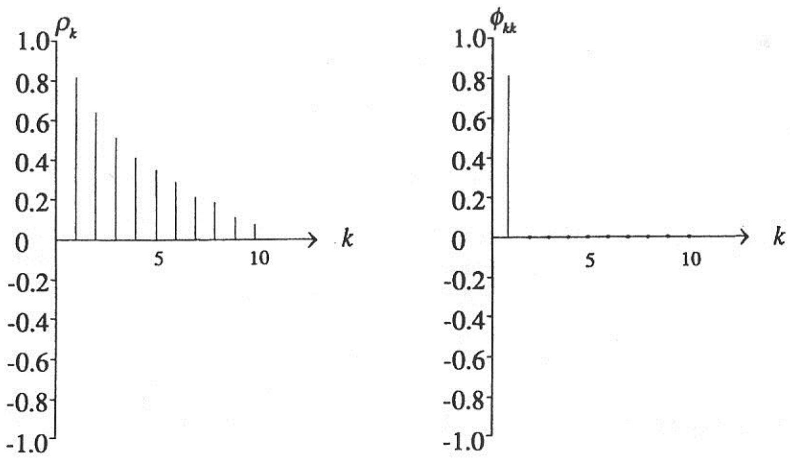


Figure A.2 ACF and PACF correlogram for ARMA(1,0)

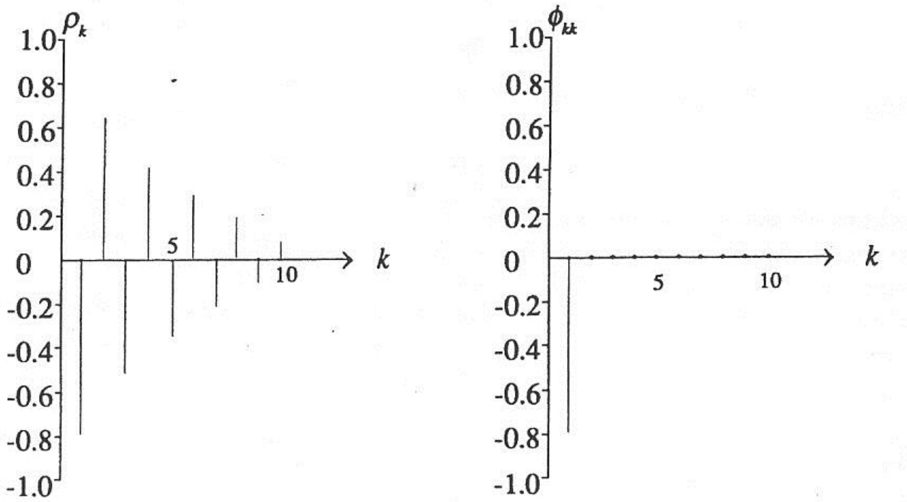


Figure A.2 ACF and PACF correlogram for ARMA(1,0) (cont.)

3. ARMA(2,0)

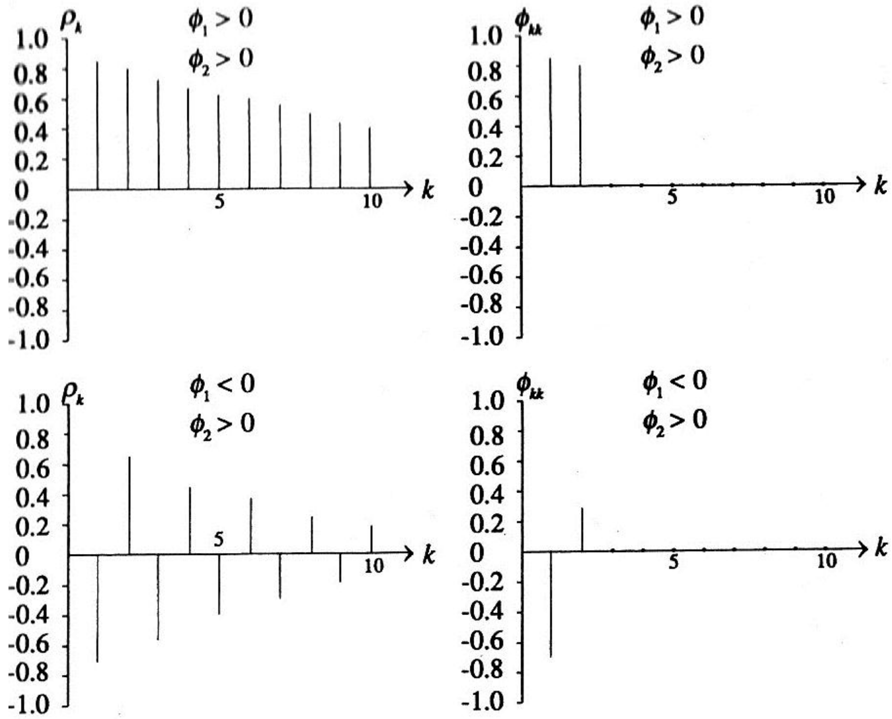


Figure A.3 ACF and PACF correlogram for ARMA(2,0)

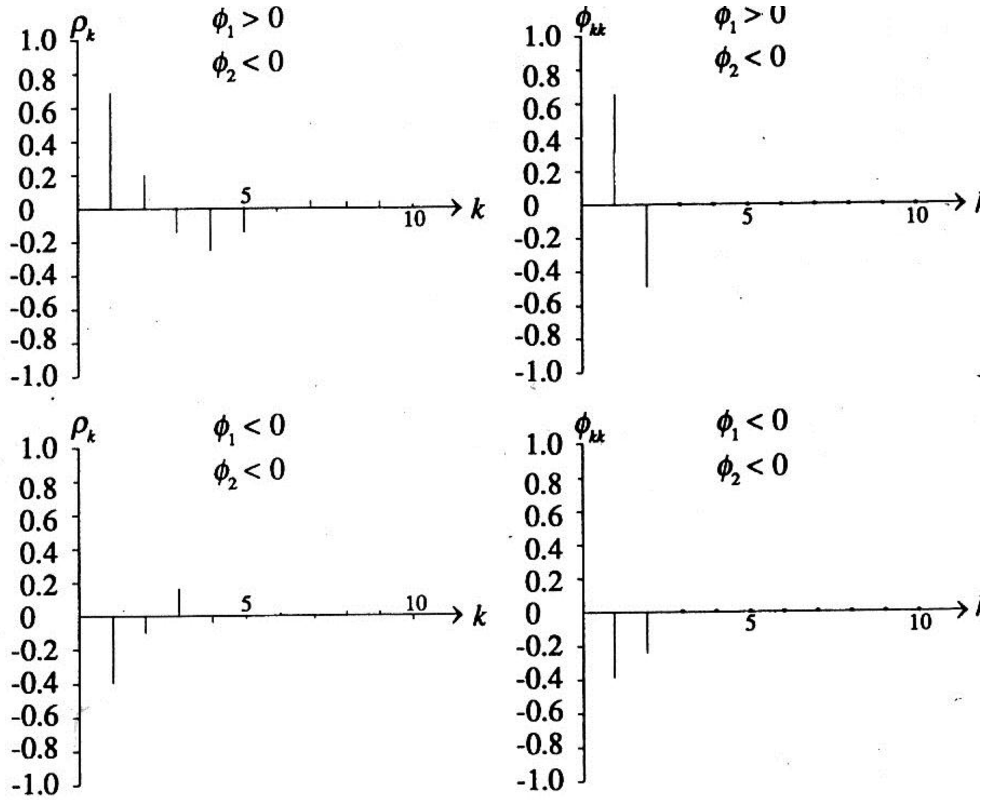


Figure A.3 ACF and PACF correlogram for ARMA(2,0) (cont.)

4. ARMA(0,1)

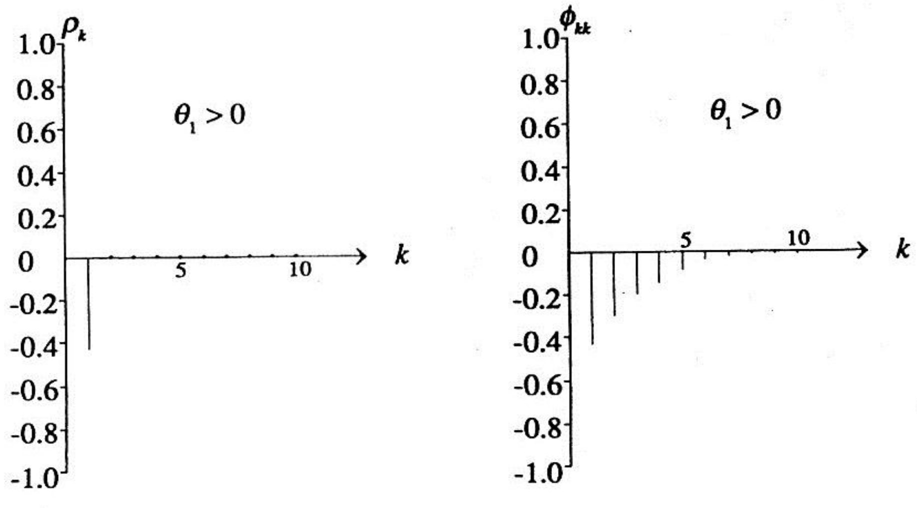


Figure A.4 ACF and PACF correlogram for ARMA(0,1)

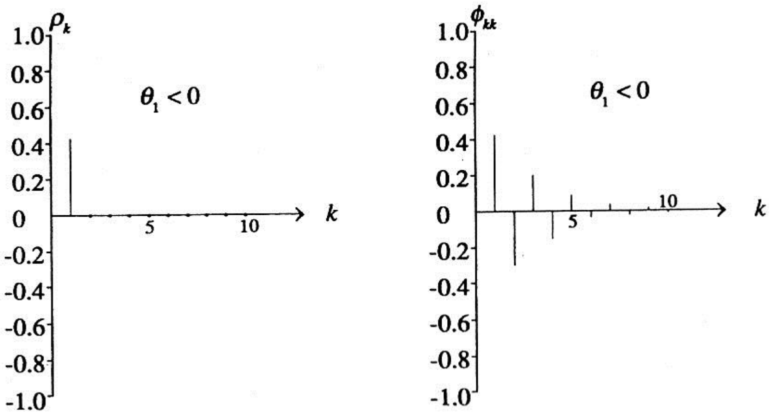


Figure A.4 ACF and PACF correlogram for ARMA(0,1) (cont.)

5. ARMA(0,2)

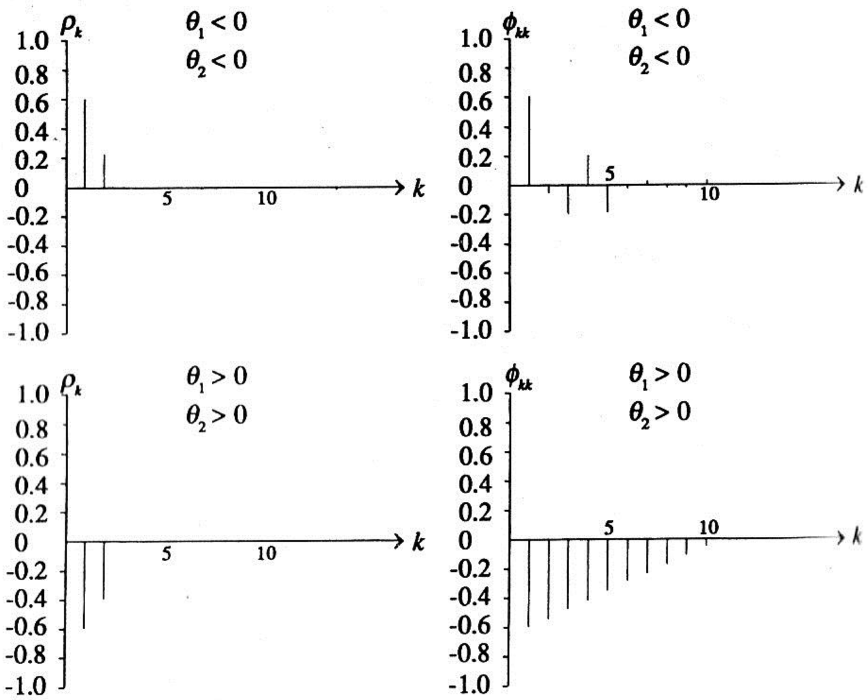


Figure A.5 ACF and PACF correlogram for ARMA(0,2)

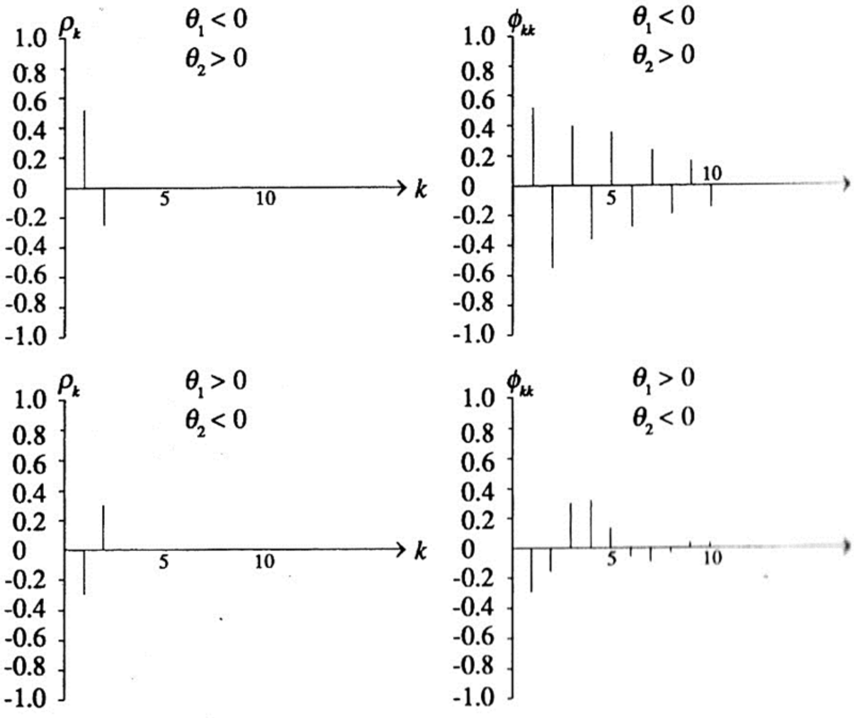


Figure A.5 ACF and PACF correlogram for ARMA(0,2) (cont.)

6. ARMA(1,1)

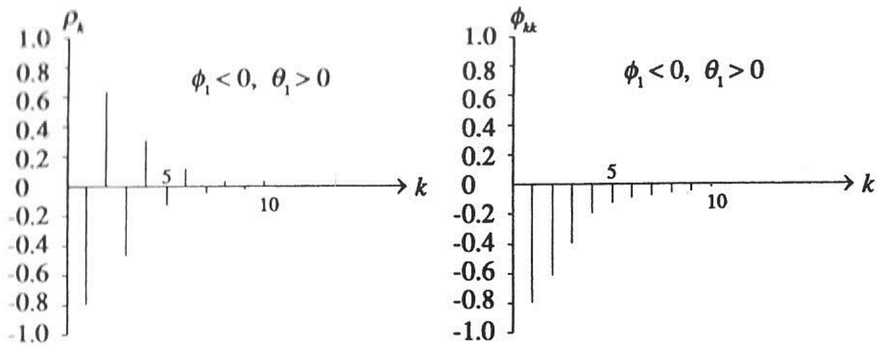


Figure A.6 ACF and PACF correlogram for ARMA(1,1)

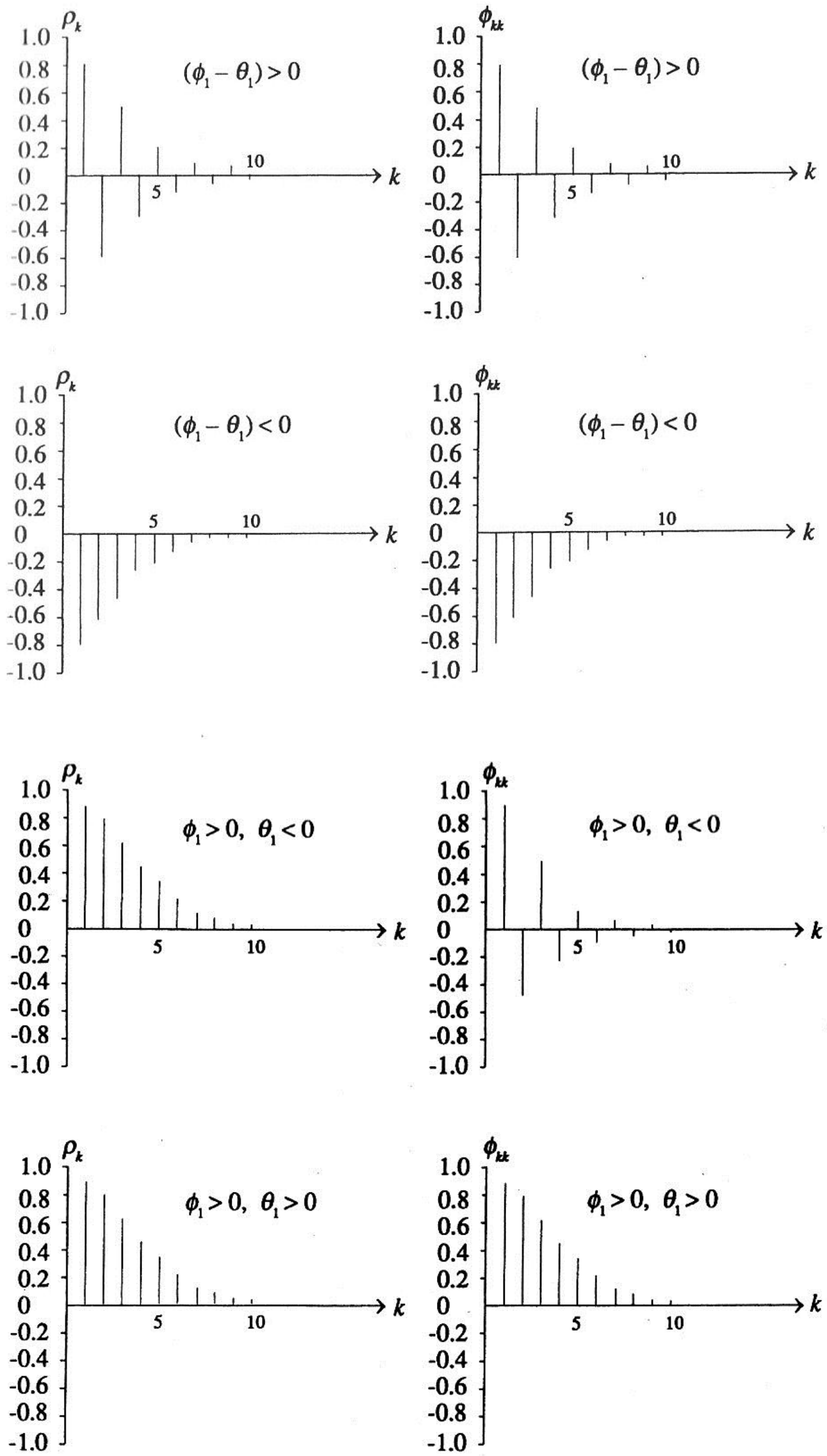


Figure A.6 ACF and PACF correlogram for ARMA(1,1) **(Cont.)**

APPENDIX B
MATLAB CODE OF TIME DELAY NEURAL NETWORK AND HYBRID MODEL

1. MATLAB code for Time delay neural network

Line	Source Code	Definition
1	clear all,close all	Close the other application and clear the work space of matlab
2	%% Load Data	
3	load pallet	Load data set name "Pallet"
4	% Divide data into training set and testing set	
5	training_set = pallet(1:58);	Set the training data
6	testing_set = pallet(59:65);	Set the testing data
8	% Prepare Data for Function	
9	y = con2seq(training_set');	change pattern of input to sequence keep in variable y
10	% Setting	
11	Numbers_Neural = 12;	The Numbers of Neural Node in Hidden Layer
12	Max_Epochs = 1000;	Maximum number of iteration
13	Training_Fcn = 'trainlm';	The training algorithm which is baysian regularization back propagation
14	Perform_Fcn = 'mse';	Use MSE for performance measurement
15	%% Time Delay	
16	for i = 1:4	Set the time delay which is lag 1,2,3, and 12
17	switch i	
18	case 1	
19	% Last Month	
20	delay_vector = 1;	
21	case 2	
22	% Last 2 Months	

Table B.1 Source code for TDNN model

Line	Source Code	Definition
23	`delay_vector = 1:2;`	Set the time delay which is lag 1,2,3, and 12
24	`case 3`	
25	`% Last 3 Months`	
26	`delay_vector = 1:3;`	
27	`case 4`	
28	`% Last Year`	
29	`delay_vector = 1:12;`	
30	`end`	
31	`% Create Time-delay Neural Network`	
32	`net = timedelaynet(delay_vector,Numbers_Neural);`	Define the structue of Network
33	`net.trainParam.epochs = Max_Epochs;`	
34	`net.trainFcn = Training_Fcn;`	
35	`net.performFcn = Perform_Fcn;`	
42	`[p,Pi,Ai,t] = preparets(net,y,y);`	
43	`% Start Training`	
44	`[net,tr{i}] = train(net,p,t,Pi);`	Training command for model
45	`yp = net(p,Pi);`	The output of network will store in variable Yp
46	`e = gsubtract(yp,t);`	Calculate for the residual e
47	`% Root Mean Square Error`	
48	`rmse(i) = sqrt(mse(e));`	Calculate for the RMSE for training
50	`% Testing`	
51	`y_test = sim(net,con2seq(testing_set'),Pi);`	Test the model br testing data set
52	`y_test = cell2mat(y_test)';`	
53	`result_rmse(i) = sqrt(mse(testing_set-y_test));`	Calculate for the RMSE for testing
63	`end`	

Table B.1 Source code for TDNN model (cont.)

2. MATLAB code for Hybrid model

Line	Source Code	Definition
1	clear all,close all	Close the other application and clear the work space of matlab
2	%% Load Data	
3	load pallet_train	Load trainging data set name "Pallet_train"
4	load pallet_test	Load testing data set name "Pallet_test"
5	load residual	Load residual from ARIMA
6	load arima_test	Load output data from ARIMA
8	%% #of Node set	
9	node = [1:12];	The Numbers of Neural Node in Hidden Layer
10	% Prepare Data for Function	
11	%% Input & Target	
12	y = con2seq(pallet_train');	change pattern of input to sequence keep in variable y
13	t = con2seq(residual');	change pattern of input to sequence keep in variable t
14	for i=1:numel(node)	Set the time delay which is lag 1,2,3, and 12
15	%% Setting	
16	delay_vector = 1:13;	
17	Numbers_Neural = node(i);	The Numbers of Neural Node in Hidden Layer
18	Max_Epochs = 1000;	Maximum number of iteration
19	Training_Fcn = 'trainlm';	The training algorithm which is baysian regularization back propagation
20	Perform_Fcn = 'mse';	Use MSE for perforemance measurement
21	%% Time Delay	
22	% Create Time-delay Neural Network	

Table B.2 Source code for Hybrid model

Line	Source Code	Definition
Line	`net = timedelaynet(delay_vector,Numbers_Neural);`	Define the structue of Network
23	`net.trainFcn = Training_Fcn;`	
24	`net.performFcn = Perform_Fcn;`	
25	`net.trainParam.epochs = Max_Epochs;`	
33	`[p,Pi,Ai,tt] = preparets(net,y,t);`	
34	`% Start Training`	
35	`[net,tr] = train(net,p,tt,Pi);`	Training command for model
36	`yp = net(p,Pi);`	The output of network will store in variable Yp
37	`e = gsubtract(yp,tt);`	Calculate for the residual e
38	`% Root Mean Square Error`	
39	`rmse(i) = sqrt(mse(e));`	Calculate for the RMSE for training
40	`% Testing`	
41	`y_test = sim(net,con2seq(pallet_test'),Pi);`	Test the model br testing data set

Table B.2 Source code for Hybrid model (cont.)

BIOGRAPHY

NAME	Mr. Wanlop Fuangfoo
DATE OF BIRTH	22 January 1987
PLACE OF BIRTH	Samutsongkhram, Thailand
INSTITUTIONS ATTENDED	Mahidol University, 2005-2009: Bachelor of Science (Mathematic)
	Mahidol University, 2009-2012: Master of Engineering (Industrial Engineering)
RESEARCH GRANTS	This thesis is partially supported by Graduate Studies of Mahidol University Alumni Association
HOME ADDRESS	112/38 Phraholyothin Road Bang Khen , Bangkok 10220 Email: oshiyoh_f@gmail.com
PUBLICATION / PRESENTATION	The 10th ThaiVCML Conference (ThaiVCML 2010) Krabi Resort Krabi Thailand