Forecasting International Tourist Arrivals from Major Countries to Thailand

Miss Ontheera Hwandee

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering in Industrial Engineering
Department of Industrial Engineering
Faculty of Engineering
Chulalongkorn University
Academic Year 2018

การพยากรณ์จำนวนนักท่องเที่ยวจากกลุ่มประเทศหลักที่เดินทางเข้ามาในประเทศไทย

น.ส.อรธีรา หวานดี

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมอุตสาหการ ภาควิชาวิศวกรรมอุตสาหการ
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561

Thesis Title	Forecasting International Tourist Arrivals from Major Countries to Thailand
By	Miss Ontheera Hwandee
Field of Study	Industrial Engineering
Thesis Advisor	Assistant Professor NARAGAIN PHUMCHUSRI, Ph.D.

Accepted by the Faculty of Engineering, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Engineering

.... Dean of the Faculty of Engineering
(Professor SUPOT TEACHAVORASINSKUN, Ph.D.)

THESIS COMMITTEE

.... Chairman
(Associate Professor PAVEENA CHAOVALITWONGSE, Ph.D.)

.... Thesis Advisor
(Assistant Professor NARAGAIN PHUMCHUSRI, Ph.D.)

.... Examiner
(Amonsiri Vilasdaechanont, Ph.D.)

.... External Examiner
(Assistant Professor Nantachai Kantanantha, Ph.D.)

อรธีรา หวานดี : การพยากรณ์จำนวนนักท่องเที่ยวจากกลุ่มประเทศหลักที่เดินทางเข้ามาในประเทศไทย. (Forecasting International Tourist Arrivals from Major Countries to Thailand) อ.ที่ปรึกษาหลัก : ผศ. ดร.นระเกณฑ์ พุ่มชูศรี

อุตสาหกรรมการท่องเที่ยวเป็นหนึ่งในอุตสาหกรรมหลักที่มีความสำคัญอย่างมากต่อเศรษฐกิจของไทย เพื่อทำให้การตลาดและการวางแผนทรัพยากรมีประสิทธิภาพ การพยากรณ์นักท่องเที่ยวจากประเทศหลักที่เดินทางมายังประเทศไทยอย่างแม่นยำจึงเป็นสิ่งจำเป็นสำหรับอุตสาหกรรมการท่องเที่ยวของไทย ในวิทยานิพนธ์นี้นำเสนอรูปแบบการพยากรณ์ที่หลากหลาย เพื่อคาดการณ์การจำนวนนักท่องเที่ยวรายเดือนจากประเทศหลัก ได้แก่ จีน มาเลเซีย เกาหลี ญี่ปุ่น รัสเซีย สหราชอาณาจักรและสหรัฐอเมริกา โดยตัวแบบพยากรณ์ที่นำเสนอประกอบด้วยทั้งตัวแบบอนุกรมเวลาเช่น Seasonal Autoregressive Integrated Moving Average (SARIMA) และ Holt-Winter และตัวแบบที่ใช้ปัจจัยภายนอกในการพยากรณ์ได้แก่ การถดถอยพหูคูณและโครงข่ายประสาทเทียมแบบไปข้างหน้า (FANN) โดยปัจจัยที่ใช้คือปัจจัยทางเศรษฐกิจเช่น รายได้ ราคาสัมพัทธ์ อัตราแลกเปลี่ยนของเงินตรา และ ตัวแปรเชิงคุณภาพ เช่น ฤดูกาล และ เหตุการณ์ต่างๆที่เกิดขึ้นในประเทศไทย ซึ่งถูกนำมาศึกษาผลกระทบของนักท่องเที่ยวที่จะเดินทางเข้ามาในประเทศไทย สำหรับการเปรียบเทียบตัวแบบพยากรณ์ใช้การประเมินค่าเฉลี่ยร้อยละสัมบูรณ์ (MAPE) โดยผลวิจัยพบว่าตัวแบบพยากรณ์โครงข่ายประสาทเทียมแบบไปข้างหน้าสามารถให้ความแม่นยำมากที่สุดโดยที่ MAPE น้อยกว่า 10% สำหรับทุกประเทศที่ศึกษาและแม่นยำกว่าตัวแบบอย่างง่ายเช่นการถดถอยพหูคูณเชิงเส้น อย่างไรก็ดีบางประเทศเช่น สหรัฐอเมริกาและญี่ปุ่นมีความเหมาะสมกับตัวแบบตัวแบบอนุกรมเวลา Holt-Winter และ SARIMA ตามลำดับเนื่องจากกลุ่มประเทศดังกล่าวมีแนวโน้มและพฤติกรรมตามฤดูกาลอย่างชัดเจน

สาขาวิชา	วิศวกรรมอุตสาหการ	ลายมือชื่อนิสิต	....
ปีการศึกษา	2561	ลายมือชื่อ อ.ที่ปรึกษาหลัก	....

6070373221 : MAJOR INDUSTRIAL ENGINEERING
KEYWORD: Tourist Arrival Forecasting Multiple Regression Seasonal ARIMA Holt and Winter Method Feed Forward Artificial Neural Networks New Shock Effects

Ontheera Hwandee : Forecasting International Tourist Arrivals from Major Countries to Thailand. Advisor: Asst. Prof. NARAGAIN PHUMCHUSRI, Ph.D.

Tourism industry is one of industries that are very important for Thai economy. In order to achieve effective marketing and resource planning, accurate forecasting of tourist arrivals from major countries to Thailand is necessary for Thai tourism industry. In this paper, various forecasting models are explored to forecast monthly tourist arrivals from China, Malaysia, Korea, Japan, Russia, UK and US. The proposed models include both time series models, i.e., SARIMA, Holt-Winter, and explanatory models, i.e., Multiple Regression and Feed Forward Artificial Neural Networks (FANNs). Economic factors such as income, relative price, exchange rates, and dummy variables of seasonality and news shock effect are explored to understand their effects on international tourism demand. Mean absolute percentage error (MAPE) is used for model comparison. It was found that the more advanced model like FANNs can produce high levels of forecasting accuracy (with MAPE $\leq$ 10% for all studied counties) and outperform simpler forms of explanatory model like Multiple Linear Regression. However, there are counties such as US and Japan that are suitable for Holt-Winter and SARIMA, respectively, due to their obvious seasonality and trends.

Field of Study:	Industrial Engineering	Student's Signature ....
Academic Year:	2018	Advisor's Signature ....

ACKNOWLEDGEMENTS

I would like to express my gratitude to my dissertation major advisor who has guided, assisted and encouraged me throughout each stage of my research. Her critical and insightful comments have provided thoughtful ideas that have greatly improved the quality of the thesis. I would also like to thank Assistant Prof. Dr. Nantachai Kantanantha in the Faculty of industrial Engineering, Kasertsat University for his helpful assistance and care given to me. Furthermore, I would like to thank you, my lovely family, particularly my friends, who have encouraged and supported my ambitions over the years.

Ontheera Hwandee

TABLE OF CONTENTS

Page

List of Tables

Page

List of Figure

Page

Chapter 1 Introduction

1. Tourism industry

Travel and tourism industries are some of the world's largest industries with a global economic contribution (direct, indirect and induced) of over 7.6 trillion U.S. dollars in 2016. The direct economic impact of the industry, including accommodation, transportation, entertainment and attractions, was approximately 2.3 trillion U.S. dollars that year. A number of countries, such as France and the United States, are consistently popular tourism destinations, but other, less well-known countries are quickly emerging in order to reap the economic benefits of the industry. Every year, the UNWTO compiles a list of the most visited countries according to the number of international tourist arrivals for every country. European destinations like France, Spain and Italy are still leading the list, while Thailand and Mexico are continuing to climb up the list [1].

Thailand is a country with many natural and cultural tourism products and beautiful beach destinations. All these elements profoundly attract tourists from all over the world especially from Europe, USA, South Asia, the Oceania, the Middle East, and Africa. Moreover, these attributes also combine to draw millions of tourists who come for holiday, leisure, or adventure. In many travel magazines and lifestyle surveys, Thailand is consistently ranked in the top ten for its beaches, entertainment and dining, value of products, recreational facilities, and shopping. A survey by Lonely Planet, which is particularly popular among young visitors, shows that Thailand is a top-rated destination in terms of value and food. It is credited with the most exciting outdoor market [2].

Therefore, tourism is a very important industry to Thailand's economy. It contributes significantly to Thailand's gross domestic product (GDP), affecting employment, investment, and foreign exchange earnings. International tourism is the fastest growing industry in Thailand. The country has continuously experienced growth in the number of tourists and revenues from the industry. The number of international tourists in Thailand increased from 10.87 million in 2002 to 35.38 million in 2017 as shown in figure 1. Direct contribution of travel and tourism to GDP was worth 455.22 billion US dollars in 2017, the 8th largest economy of Asia [3]. The GDP value of Thailand represents 0.73 percent of the world economy [4]. And is predicted to rise by 4.1% in 2018 and to rise by 5.7%, from 2019 – 2028 [5].

Tourism in Thailand has also directly supported 2,337,000 jobs (6.2% of total employment). Besides its direct economic impact, the tourism industry has significant indirect contribution to the economy, especially for tourism services, transport, and the sale of food, drink and souvenirs to the tourists from all over the world. The revenue was announced as 99,114 billion baht in 2017 from Thailand's tourism [6].

Tourism has helped Thailand's economic growth, creating opportunities for Thailand to expand its market presence to high-end buyers, as well as to attract tourists interested in tourism. In 2018, Thailand expects 37.55 million visitors to spend 2.1 trillion baht [6]. Thailand would have more special interest tourism which the National

Tourism Policy Committee has been approved, designed "Amazing Thailand Tourism". This aims to further strengthen the Thai tourism industry by making it more sustainable.

The Amazing Thailand Tourism Year 2018 promotion focused on seven distinct categories. These include Sports Tourism, Gastronomy Tourism, Maritime Tourism, Wedding and Honeymoon Tourism, Medical and Wellness Tourism, Community-based Tourism, and Leisure Destinations. It would be followed the development agenda of Thailand 4.0, based on the government's 12th National Economic and Social Development Plan and 2nd National Tourism Development Plan that maintained a balance between the economy, environment and benefits for Thai society [7].

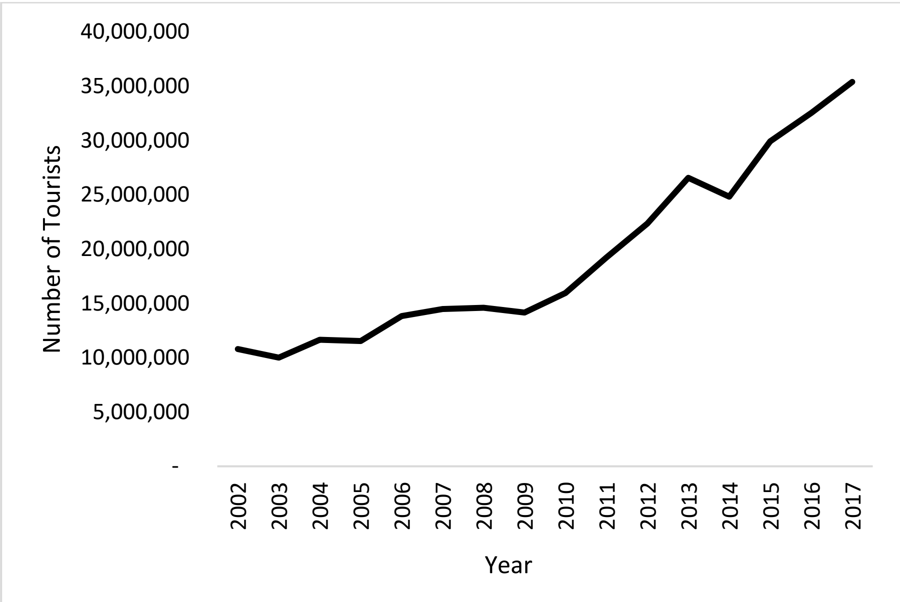

Figure 1: The International Tourist Numbers Visited Thailand from 2002-2017
Source: Ministry of Tourism and Sports

1.2 Number of tourists visiting Thailand in 2017

The number of tourists visiting Thailand has been on a steady rise throughout the years reaching up to 30 million visitors annually, although the numbers sometimes decline due to unrests and security threats. However, Thailand's tourism industry remains an all-time high, especially during the December-January peak season.

The top 10 most popular nationalities to visit Thailand in 2017 shown in Table 1. The major countries from East Asia markets contained almost 80% of all regions were China, Malaysia, Korea, Laos, and Japan as shown in Figure 2.

Table 1: Top 10 arrivals in Thailand by nationality

Rank	Country	Number of Tourists (2017)	Number of Tourists (2016)	% Change
1	China	9,805,753	8,757,646	+10.7
2	Malaysia	3,354,800	3,494,890	-4.0
3	Korea	1,709,070	1,464,200	+14.6
4	Laos	1,612,647	1,388,020	+13.9
5	Japan	1,544,328	1,439,510	+6.8
6	India	1,411,942	1,193,822	+15.4
7	Russia	1,346,219	1,089,992	+19.0
8.	USA	1,056,124	975,643	+7.6
9	Singapore	1,028,077	966,909	+5.9
10	UK	994,464	1,003,386	-0.9

Source: Ministry of Tourism and Sports

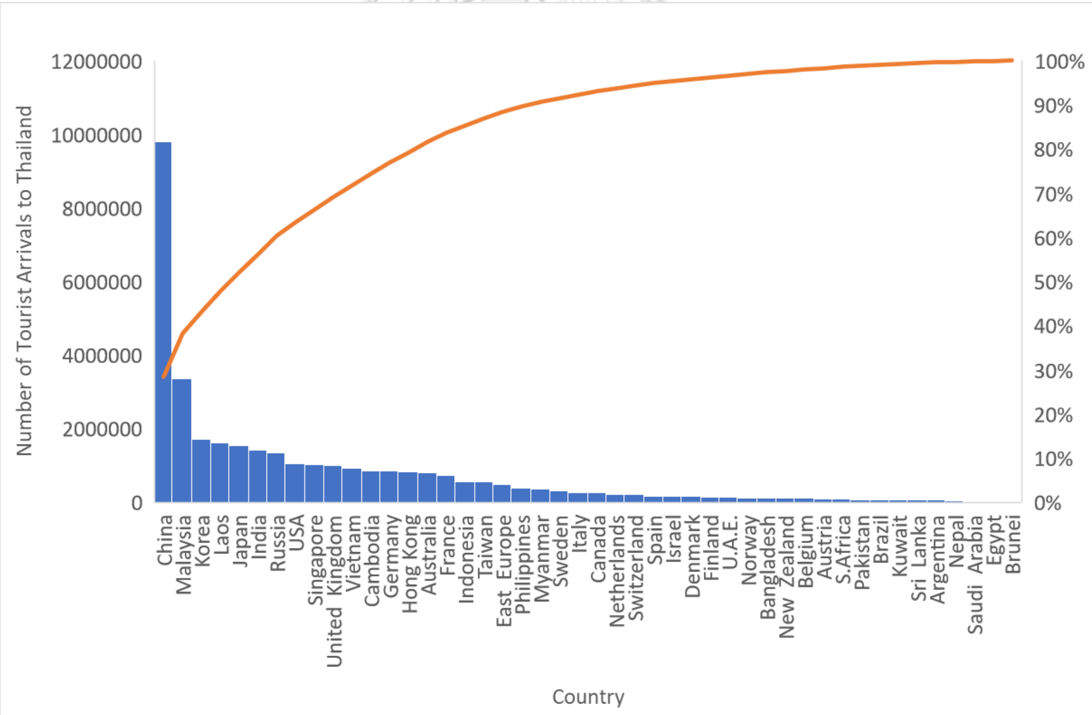

Figure 2: International Tourist Arrivals to Thailand in 2017 by Country
Source: Ministry of Tourism and Sports

Since 2013, by region, East Asia, led by China, has had the largest number of tourists visit Thailand over the decades with the number expected to keep rising. In 2016, about 8,757,466 Chinese tourists visited Thailand making up for more than 25%

of the total number of tourists. In 2017, China recorded a total of 9,805,753 tourists bound for Thailand. A large number of Chinese tourists are linked to the low-price package tours which enable small budget travelers to visit Thailand regularly. However, the Thai government has discouraged these package tours as they limit the amount of spending by the tourists, earning little revenue despite a significant number of visitors.

Malaysia is the second leading country in the world and the leading ASEAN country which sends tourists to Thailand with a total of 3,494,890 visitors in 2016. In 2017, Thailand received a total of 3,354,800 Malaysia tourists and recorded a 4.0 % reduction in tourist receipts. Close relations between the two countries as well as the promotion of Thailand's tourist attractions across ASEAN countries has played a significant role in increasing the number of Malaysian visitors. The shared Buddhist traditions also attract Malaysian tourists who wish to attend pilgrimage and festivals outside their own country.

Korea and Thailand have bilateral relations, with Thailand being considered a top destination for South Korean tourists who visit in large numbers. South Korea makes up about 6% of the foreign tourists in Thailand. About 1,464,200 of South Koreans toured Thailand in 2016 and another 1,709,070 in 2017.

Laos was among the top source market for tourists visiting Thailand from January to December 2016, the number of Lao nationals visiting Thailand was recorded at 1.38 million. In December 2017 alone, the number of Lao visitors to Thailand was recorded at 1,612,647 people, an increase of 13.9 percent compared to the previous year.

Japan would continue to be one of most important markets for Thai tourism as they are high-quality tourists. TAT statistics shown visitors from Japan have increased significantly from 1.43 million in 2016 to 1.54 million in 2017, representing an increase of 6.8 percent over the last year. Japan would continue to be one of most important markets for Thai tourism as they are quality tourists. As for 2018, TAT has set a target for Japanese arrivals at about 1.56 million, an increase of 4%. The number is expected to bring in 70 billion baht in tourism income for Thailand, or an increase of 8%.

India is ranked sixth in terms of tourist arrivals in Thailand and South India contributes around 30 percent of them. As many as 8 lakh Indian tourists visited that country till this year, which has gone up by 16 per cent over 2016. Thailand has registered a 15.4 percent growth in Indian tourist arrivals in 2017. In 2018, the number of India tourists has increased trend and would continue growth in 2019. Indian visitors know Thailand as a hospitable and a year-round tourism destination with a high quality of value-for-money products and services. Indians get visa-on-arrival facilities at Thailand's international checkpoints. In 2017, Thailand won eight Indian tourism awards; such as, Best Wedding Destination, Best Destination, Best International Tourism Board, Best Spa Destination and Best Value Destination [1].

The European market has long played a major role in the development of Thai tourism. According to the Tourism Authority of Thailand (TAT), European visitor arrivals in 2016 totaled 6.17 million, up 6.3% over 2017 [2]. Russia is the largest source market out of Europe, with arrivals of 1.34 million, up by 19.0 percent over 2017. The Russian tourist market following a drop in Russia tourist arrivals in 2014. While Russian in 2016 has continued growth. By this index Russia was ranked the first among European countries and the seventh in the world. This statistic show that Russian tourists willingly choose Thailand as a place for leisure and a travel destination.

1.3 Impact Factors for Tourism Demand

Travel and tourism sector have direct economic impact, significant indirect and induced impacts to the country; hence, it can increase the capital investment and trade of the country, create jobs for the unemployed citizens and protect heritage as well as the cultural values. Moreover, the rapid growth of the tourism sector can be attributed to a number of economic factors. From the previous studies, many researchers investigated factors affecting the number of tourists. A number of one hundred tourism studies conducted on this particular topic over the past twenty years focusing predominantly on economic factors such as income, relative prices and exchange rates shown in Table 2[3].

As a result, the major factors considered to influence tourism demand are incomes, relative prices and exchange rates. These determinants will be described in more depth on the following

Table 2: Classification by Number of Explanatory Variables Used

Explanatory	**Number of Review**
Income	84
Relative Prices	74
Transportation	55
Exchange Rates	25
Trend	25
Competing Destinations	15
Seasonal Factor	14
Qualitative Factors	60

Source: [4]

1.3.1 Frequently used variables

Income

Income in the origin country is the most frequently used explanatory variable in the published tourism studies. Overseas travel (especially recreational travel) is expensive and is generally regarded as a luxury good, in which case the discretionary income (or income remaining after expenditure on necessities) of origin country would be an appropriate income variable. However, discretionary income is a subjective variable and is not precisely measurable. Hence, most researchers have relied on nominal or real (per capita) personal, disposable, or national income, and GDP or GNP as measures (or proxies) for income in the origin.

It is expected that tourism demand will not only be influenced by current, but also by lagged income in the origin, since changes in income may take some time to affect tourism demand. When both current and lagged income are used in a study, the latter would be classified as reflecting dynamics. However, if only lagged income is used (for example, lagged real GDP per capita), it will be regarded as a measure (or proxy) for income. The same reasoning applies to all other variables when current and/or lagged explanatory variables are used [4].

Relative Price

Relative or tourism prices, which are the second most frequently used explanatory variables in the literature, are costs of goods and services that tourists are likely to pay while at the destination (such as accommodation, local transportation, food, and entertainment). In measuring relative price movements in the origin and destination, it is desirable to have indices constructed using a basket of goods purchased by tourists. Since tourist price indices (TPI) are typically unavailable, the consumer price index (CPI) of the origin and destination are used as proxies to reflect the relative prices of foreign tourism goods and services. adjusted CPI ratio formula can be used:

$$\text{Relative Prices} = \frac{\text{CPI(Destination)}}{\text{CPI (Origin)}} \tag{1}$$

Exchange Rates

Exchange rates are often introduced into tourism demand models in addition to, and separately from, the relative price variable [5]. Such studies specifically examine the influence of nominal exchange rates on international tourism demand. Data on exchange rates are readily available because they are widely published and are reasonably accurate. Some researchers argue that tourists respond to exchange rate movements but not to changes in relative inflation rates when they make their decision to travel, because of limited knowledge. Tourists are well-informed of changes in exchange rates, whereas information on price changes in destinations is generally not known in advance [5].

Transportation costs

This refers to the cost of round-trip travel between the origin and destination countries. Unlike other goods, the consumer (tourist) has to be transported to the product (destination) rather than the reverse. Hence, the demand for transportation in international travel is a derived demand, namely to purchase tourism services. However, only 55% of the published papers included this important explanatory variable in their studies. Transportation costs are usually measured by the price of air travel. However, the problem of measuring the effective transportation cost arises, namely, the actual costs borne by tourists. This problem is caused mainly by the pricing practices of airlines, which have resulted in two main categories of passenger fares on scheduled airline services, namely, “normal” (unrestricted) fares and “special” fares. Normal fares are available for first, business and economy class, whereas special fares are only available for economy class, which includes excursion and promotional fares. The most widely offered type of special fares is the excursion fare.

Dummy variables of Seasonality

Dummy variables may be used to capture seasonal variations in tourism demand. Seasonal patterns in tourist flows and expenditures are well-known characteristics of international tourism demand, but only 15 studies have tried to account for seasonality in modeling tourism demand. Although marketing expenditures by private or national agencies are vital for promoting the country as a destination, especially where tourism makes significant contributions to the economy, it is somewhat strange that very few studies (only 7 of the 100 studies) have included this variable in the demand models [6].

Specific time of the year, like a season or a period of school holidays, can have a significant effect on tourism demand. Seasonality has been dealt with by many authors but has been avoided by some due to modelling tourism demand based on annual data. Typically, if using monthly data, twelve seasonal dummy variables are included in the model and similarly four seasonal dummy variables are incorporated regarding the quarterly data [7].

Dummy Variables of news shock effects

The impact of events on the tourism can be incorporated as dummy variables to understand on tourism demand. In terms of tourism demand, news shocks are likely to influence directly or indirectly the demand on tourism. Some news may have a positive impact on some potential tourists and some tourist destinations, whereas other news may have an adverse effect. Thus, this study will analyse the news shocks effects on the tourism demand in Thailand, as follows.

Thai political crisis

The 2013–2014 Thai political crisis was a period of political instability in Thailand. Anti-government protests took place between November 2013 and May 2014, organized by the People's Democratic Reform Committee (PDRC), a political pressure group led by former Democrat Party parliamentary representative (MP) Suthep Thaugsuwan. The protests eventually resulted in the removal of incumbent Prime Minister Yingluck Shinawatra, a coup d'état, and the establishment of the military junta. Thai anti-government protests that have shut down parts of Bangkok may cost the nation's tourism industry as Chinese visitors cancel trips during the lunar new year holiday that starts this week. Other countries also have issued travel alerts about Thailand because of the political demonstrations, including the United States, Malaysia, Hong Kong, Australia, and the Philippines[8].

Bomb in Bangkok

On 17 August 2015, a bombing took place inside the Erawan Shrine at Ratchaprasong intersection in Pathum Wan District, Bangkok, Thailand, killing 20 people and injuring 125 [9]. Thailand's tourist industry has been dealt a huge blow after the country's worst-ever bombing caused a 17 per cent fall in visitor numbers, tourist arrivals into country have fallen from 85,000 a day to 70,000 [10].

Floods in South of Thailand

The 2017 Southern Thailand floods occurred from December 2016 throughout early months of 2017 and are the biggest floods in over 30 years in the southern part of the country during the regional annual monsoon season, which is distinct from other parts of the nation and mirrors that the Malay Peninsula. hey have also caused damages in Kelantan and Terengganu states in Malaysia the weeks prior. The floods began in early January. Some 120 billion baht ($4 billion USD) in damages are foreseen as of mid-January, much of this due to lost production in agriculture tourism and infrastructure[11]. However, the flooding had not affected the overall tourism situation in the South. Major tourist destinations, especially Phuket and Krabi, had escaped lightly.

Thai King Rama9 Passed Away

King Bhumibol Adulyadej of Thailand died at the age of 88 on 13 October 2017, after a long illness. A year-long period of mourning was subsequently announced. A royal cremation ceremony took place over five days at the end of October 2017. The actual cremation, which was not broadcast on television, was held in the late evening of 26 October 2017. A number of events have been canceled during the 30-day mourning period, this includes the Full Moon Party, a monthly all-night beach party on Koh Pha Ngan island that's a huge draw for international tourists. Canceled Bangkok events include the upcoming sold-out Morrissey concert and the Scorpions' 50th Anniversary World Tour [12]. This news might be influenced on tourism demand in Thailand, some country has canceled flight due to the canceled events.

Boat Accident in Phuket

On Thursday 5 July 2018, the double-decker cruise boat Phoenix PC Diving sailed from Phuket, one of the most visited resorts in Thailand, for Koh Racha, a popular snorkeling island off the coast of Phuket. It carried 101 people, including 89 tourists, all but 2 of whom were Chinese nationals, 11 crew members and a tour guide [13]. Phuket was already experiencing a huge drop in the number of Chinese tourists coming to the island despite the number of Chinese tourists coming to Thailand still growing – highlighting a major shift that was already underway before the Phoenix tour boat disaster.

1.4 The Models for Forecasting Tourism Demand

There are basically two kinds of methods: time-series and causal methods. Each method has different strengths and weaknesses. The causal models require the need to identify and forecast independent variables in the defined model in order to use the model. This represents a significant challenge as the incorrect prediction of these independent variables will result in an incorrect forecast from the causal model. Moreover, the causal based methodologies have become very complex in recent years with the use of cointegration modelling and hence very expensive for use by business. Time-series methods do not have the above causal requirements and can be more practical for business use, primarily because they cost less to execute within a business workplace. Furthermore, purely from the point of view of accuracy, time-series models can be at least as accurate as causal models.

In this study, the explanatory methods and time series methods are used to compare with machine learning, namely artificial neural network for the international tourism demands in Thailand.

Time Series Methods

Holt-Winters' Exponential Smoothing

Holt-Winters is a model of time series behavior. Forecasting always requires a model, and Holt-Winters is a way to model three aspects of the time series: a typical value (average), a slope (trend) over time, and a cyclical repeating pattern (seasonality). Holt-Winters uses exponential smoothing to encode lots of values from the past and use them to predict "typical" values for the present and future. The model predicts a current or future value by computing the combined effects of these three influences. The model requires several parameters: one for each smoothing (α, β, γ), the length of a season, and the number of periods in a season. As with the trend, the seasonality may be modeled as either an additive or multiplicative process for a linear or exponential change in the seasonality [14].

Seasonal Autoregressive Integrated Moving Average Model (SARIMA)

The ARIMA model is time series model which explains a variable with regard to its own past and a random disturbance term. Particular attention is paid to exploring the historic trends and patterns (such as seasonality) of the time series involved, and to predict the future of this series based on the trends and patterns identified in the model. Since time series models only require historical observations of a variable, it is less costly in data collection and model estimation. ARIMA models have been widely used for tourism demand forecasting in the past four decades proposed by [15].

Time series forecasting techniques is depending on the frequency of the time series, either simple ARIMA or seasonal ARIMA (i.e. SARIMA) models could be used with the latter gaining an increasing popularity over the last few years, as seasonality is such a dominant feature of the tourism industry that decision makers are very much interested in the seasonal variation in tourism demand. With regard to the forecasting performance of the ARIMA and SARIMA models, empirical studies present contradictory evidence [16]. This model will show the response to shocks in the short run used to forecast tourism demand.

Causal Method

The Multiple Regression (MR)

Multiple regression is an extension of simple linear regression. It is used to predict the value of a variable based on the value of two or more other variables. This model is commonly used to study the tourism demand in many countries. The economic variables in this model are income, price, exchange rates, and qualitative variables affecting tourist arrivals to destination country. It is also used to identify the important factors of tourism demand and to forecast their impact. In addition, the dummy variable of seasonality and new shock effect such as natural disasters can be included in the model to explain the seasonal pattern.

Artificial Neural Networks (ANNs)

Beside the time series forecasting methods, technique such as artificial intelligent (AI) have emerged in the tourism demand forecasting. In recent years, the study of artificial neural networks (ANNs) has aroused great interest in fields as diverse as biology, psychology, medicine, economics, mathematics, statistics and computers. The reason behind this interest is that ANN are universal function approximators capable of mapping any linear or non-linear function [17]. Due to their flexibility in function approximation, ANN are powerful methods in tasks involving pattern classification, estimating continuous variables and forecasting. In the last case, neural networks offer several potential advantages over alternative methods— mainly the ARIMA time series models—when handling problems with non-normal and non-linear data.

The first advantage is that ANN are very versatile and do not require formal specification of the model nor acceptance of a determined probability distribution for data. As for the second advantage, ANN are capable of tolerating the presence of chaotic components better than most alternative methods. This capacity is particularly important, as many relevant time series possess significant chaotic components.

ANN have been applied in the many fields mentioned above and have been a pioneer in the field of tourism data analysis. Thus, neural network models have recently been used as a statistical technique in the main fields of tourism research, such as demand and consumer behavior forecasting.

1.5 Contribution of This Study

Since international tourism plays an important role to strengthen the Thai economy, forecasts of tourism demand are of great economic value both for the public and private sector. In the last few decades, numerous researchers have studied international tourism demand and a wide range of the available forecasting techniques have been tested [7]. In Thailand, most research focused on international forecasting with time series using ARIMA. Moreover, there are not many studies using explanatory method with economic factors in Thailand. Although there are a few comparisons between time series forecasting and explanatory forecasting, was not used explanatory as an economic factor.

In addition, the tourism market in the form of Thai culture to attract foreign tourists to travel in Thailand increased in 2018. Tourism marketing also focused on other activities such as food tourism (Gastronomy Tourism and Medical Tourism). For these reasons, it is important to thoroughly examine all aspects of the tourist demand on each country which impact on special interest tourism. It is tremendously significant for the government. Determining the level of tourist demand helps planners reduce the risk of decisions regarding the future. It also provides information for tourism for suppliers to prepare suitable products for each group of tourists.

From the perspective of practical forecasting in industry the greatest need is for short term tourist arrival forecasts. Industry needs in the tourism sectors have become more short-term focused, and designed to change rapidly with changing market demand. Partly in consequence of this, the longest terms econometric modeling has become less relevant to industry.

The most related research is [2] where they studied the factors that influence the behavior of international tourists in Thailand from an economic perspective. A number of important economic factors such as income, price, and exchange rate are studied regarding international tourism demand in the long-run. However, seasonal effect in this research has not yet been included in the model and they focused on forecasting model from 6 regions (Americas, Europe, Oceania, ASEAN, East Asia, and East Asia). Moreover, SARIMA model was examined in country forecasting term, so the models could not able to be compared accuracy.

In order to comprehend at the best tourism demands, it is important to forecast tourism demand for each country and understand the factors affecting them. Thus, this research focuses on capturing seasonality factors and unexpected news by comparing the multiple regression (focusing on economic factors and dummy variable of seasonal effect and news shock effects), the SARIMA model, Holt-Winters, and artificial neural network. This can help to explain the change in tourism demand in the short-run and monthly arrivals can be predicted.

The results of the study will provide information to increase the effectiveness of the strategic plan to develop the tourism industry and action plans to promote Thailand as a tourist destination. Major users of the results would be the Tourism Authority of Thailand (TAT) and the Ministry of Tourism and Sports. In essence, an area of this research that would contribute to the programs of TAT is the analysis of tourism from the demand perspective. An analysis of the economic factors that influence international tourist arrivals to Thailand would increase the knowledge base needed to stimulate tourism growth, as well as provide a basis for future research on tourism forecasting. Assessing the effect of news shocks on tourism demand is particularly important regarding the effort to promote Thailand as a tourist destination in various markets.

1.6 Objectives of The Study

1. To develop and compare forecasting models to forecast monthly international tourist arrivals from major countries to Thailand.

2. To examine factors that influence the behavior of international tourists in Thailand from an economic perspective and effects of news shocks so that the results can be useful for Thai policy makers in boosting tourism demand.

1.7 Scope of The Study

The scope of this study focuses on the following:

1. The international arrivals from major countries to Thailand have been chosen for this study because of their importance on the special interest tourism. The major countries are considered from 80% of all countries. They are China, Malaysia, Korea, Japan, Russia, UK and US.

2. The criteria for country selection are the first seven countries with a high proportion of tourists and tourist country’s spending in Thailand using Pareto analysis.

3. The data of international tourist arrivals to Thailand are divided into 2 intervals as follow.
    - The monthly tourist arrivals in 2013-2017 from secondary sources are employed from Ministry of Tourism and Sports used for constructing models.
    - The monthly tourist arrivals in 2018 are used for testing models.

4. A review of the literature was undertaken to find factors that are useful to explain to international tourists.

5. The models focused on this thesis are time series models, i.e., SARIMA, Holt and Winter, and causal models, i.e., Multiple Regression and Feed Forward Artificial Neural Networks (FANNs)
6. The performance of forecasting models is evaluated in term of mean absolute percentage (MAPE). To understand the performance of all methods, the randomized complete block designs (RCBD) and Tukey method were tested at 95% confident interval.

1.8 Outcomes of this Thesis

The expected outcomes of this research are as follow:

(1) Forecasting model which is suitable for the international tourism arrivals from each major country to Thailand.

(2) Factors contributing to number of international tourists from major countries.

1.9 Benefits of this Thesis

(1) Accurate forecasting model can provide useful guidelines for development of Thailand tourism marketing strategy for each major country. It can also be extended to other countries in the future

(2) The results from this study can be useful for other companies in this tourism industry to apply the analysis to their data for their own insights as well.

(3) Understanding of factors contributing to number of international tourists from each country can give insights on different behavior of tourists from different nations.

Chapter 2 Literature Reviews

This chapter reviews the various quantitative forecasting methods that have been applied to tourism demand forecasting. The quantitative methods reviewed include the Holt and winters exponential smoothing, Box-Jenkins models, multiple regression, and Neural Network models.

2.1 Summary of The Time Series Model for Tourism Demand

A time series refers to observations on a variable that occurs in a time sequence. Most time series are stochastic in that the future is only partly determined by past values. A time-series forecasting model predicts future values of a time series solely on the basis of the past values of the time series.

Time series models have been widely used for tourism demand forecasting in the past four decades with the dominance of the integrated autoregressive moving-average models (ARIMAs) proposed by [15]. Different versions of the ARIMA models have been applied in over two-thirds of the post-2000 studies that utilized the time series forecasting techniques. Depending on the frequency of the time series, either simple ARIMA or seasonal ARIMA (i.e. SARIMA) models could be used with the latter gaining an increasing popularity over the last few years, as seasonality is such a dominant feature of the tourism industry that decision makers are very much interested in the seasonal variation in tourism demand. to the forecasting performance of the ARIMA and SARIMA models, empirical studies present contradictory evidence.

For example, [18] showed that the ARIMA model outperformed two other time series models in all cases. [19] suggested that the SARIMA models outperformed eight other time series methods while the non-seasonal (simple) ARIMA model's performance was above the average of all forecasting models considered. However, [20] found that the ARIMA or SARIMA model could not even outperform the Naïve 1 (no-change) model.

[21], monthly number of international arrivals in Bumthang district, Bhutan is modelled by using Box-Jenkins seasonal Autoregressive Integrated Moving Average (SARIMA) model. The forecasted result can be used as a tool to handle future challenges and to bring further development in tourism industry.

Another extension of the univariate time series analysis of tourism demand has been the application of the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model. GARCH models have been widely used in the financial modelling context to investigate the volatility of the time series. [50] applied three multivariate GARCH models to examine the volatility of tourism demand and the effects of various shocks in the tourism demand models. They found that tourism demand was affected by the conditional variances of the models that underline the demand for Australian tourism by the four leading tourism source markets. However, the forecasting performance of these multivariate GARCH models was not assessed.

[22] applied Naïve 1 (or no-change), naïve 2 (or constant-growth-rate), exponential smoothing models, and simple autoregressive models for tourism forecasting studies, they are usually used as benchmarks for forecasting accuracy evaluation.

[23] developed triple exponential smoothing for a seasonal time series with trend while [24] applied the Holt-Winters, Naive 1 and Naive 2 models to generate domestic forecasts of visitor arrivals in Las Vegas. The study shows that the Holt-Winters exponential smoothing method generates forecasts with lower error magnitudes (MAPE) than both the naive models for domestic tourism demand forecasting in Las Vegas. This supported the suggestion that the Holt-Winters model is particularly appropriate for data series where trend and seasonality components are present.

[25]tested the accuracy of selected forecasting models (naive 1, naive 2, exponential smoothing, trend curve analysis, autoregression, and econometric) for outbound tourism from France, Germany, the United Kingdom, and the United States to six of their main destinations. The results reveal that the simple naive no-change extrapolation model generally generates the most accurate one-year ahead tourism demand forecasts with lower mean absolute percentage errors (MAPEs).

[26] used the time-series forecasting techniques to generate forecasts of international tourist arrivals from Thailand to Hong Kong: 1. naive, 2. moving average, 3. single exponential smoothing, 4. linear moving average,5. Brown's one-parameter linear exponential smoothing, 6. Holt's 2-parameter linear exponential smoothing, and 7. Winter's 3-parameter exponential smoothing. Their results confirm that simple techniques may be just as accurate and often more time and cost-effective than more complex ones.

[27] estimated various exponential smoothing models over the period 1975 to 1999 to forecast quarterly tourist arrivals to Australia from Hong Kong, Malaysia, and Singapore. Using the root mean squared error criterion as a measure of forecast accuracy, they found that the Holt-Winters Additive and Multiplicative Seasonal models outperform the Single, Double, and the Holt-Winters Non-Seasonal exponential smoothing models for the one-quarter-ahead forecasts for the period 1988 to 2000. They also concluded there is a need to be concerned about seasonality in forecasting of international tourism demand for Australia.

2.2 Summary of Explanatory for Tourism Demand

The quantity of tourism demand was most frequently measured in terms of tourist arrivals [30]. One of the major advantages of the econometric approaches over the time series models lies in its ability to analyze the causal relationships between the tourism demand variable and its influencing factors (explanatory variables). Tourism demand can be affected by determinants such as income, exchange rates and relative price.

Income

Income plays a vital role in travelling decision. Many researchers find that income is the most common used independent variable for tourism demand. This statement is supported by [4] he discovered that out of 118 surveys of tourism modelling, about 83 studies on tourism demand has applied income as one of the explanatory variables in their research.

According to the [2] studied the international tourist arrivals to Thailand, he suggested that international demand of luxurious tourist countries for Thailand came from the region of America, Europe, South East Asia, East Asia, South Asia and Oceania. American region and European region have the highest income elasticity among these regions which are 3.40 and 4.31 respectively. These results showed that the gross domestic product (GDP) growth have positive effect on international tourist arrivals to Thailand. This result is corresponding to the [31] which shown that the GDP growth is positively related with the international visitor arrivals to Thailand in the long run. While in short-run, China, England, German, France, America and Canada have a positive relationship with visitors travelling to Thailand except Malaysia that has an inverse impact. In the study of [32] the result shown that real income per capita of Singapore has significant relationship on inbound tourism from Singapore to Australia and states that income and price is inelastic.

[33] investigated the impact on tourism industry when Chinese tourists were allowed in Taiwan in 2001. In long-run, the result shows that the income of Chinese tourists is positively related with the number of Chinese tourists in Taiwan. Chinese tourists first concerned about their income before travelling. Other than that, Chinese tourists consider travel to Thailand is luxury good as its income elasticity is 1.37 [34] and [35] found that tourists' incomes mainly affect tourist arrivals in Hong Kong.

Relative Price

The effect of relative prices has also been suggested as an important variable in explaining international tourism flows. This suggests that tourists are likely to react when there is a change in the ratio between prices in the exporting country relative to prices in the receiving country or prices in alternative tourist destinations. Hence, as relative prices decline, an increase in the quantity of international tourism services imported by a tourist-generating country should be anticipated, assuming other things being equal.

The destination price, the cost of goods and services bought at the destination, or living costs at the tourist destination would normally account for the major portion of the total price. In basic economic demand theory, the price of a good or service is negatively related to the quantity demanded by consumers. In the context of foreign travel demand, as the prices of goods and services in a destination country increase, tourists are expected to make fewer visits to the destination, holding other things being constant.

[36] examined the impact of relative prices on tourism flows in Mauritius using time series analysis, autoregressive model is employed. The results show that relative price measures have a long run impact on international tourism flows, indicating that tourists are sensitive to price levels. The relative average cost in the different competing destinations is also reported to be positive and significant, indicating that the impact of relative price changes in foreign destinations competing with Mauritius tourism matters.

Tourism price refers to the price of all goods and services consumed by tourists at the destination, but most countries do not have a price index for goods and services consumed by tourists. The best alternative is the general Consumer Price Index (CPI) of that particular country. Hence, a consumer price index (CPI) was used instead as a proxy measure in most studies

[37]; [38]; [39]. They used the calculation of tourism price which based on the consumer price index (CPI) of the destination divided by the CPI of the country of origin

Exchange Rate

The rapid changes in exchange rates are noticed more readily by potential foreign travelers than changes in the country's price levels. Research from Tourism Australia and Tourism Research Australia (referenced in the report) shown that exchange rates were shown to have some influence on destination choice and travel purchases, but that influence is modest and only one part of the consumer's decision-making process. However, exchange rates were found to have more bearing on tourism expenditure once in the destination. The destination country is expected to enjoy an increasing demand, all other factors remaining constant. In the past, the effect of exchange rate changes often was accounted for indirectly by converting the destination price variable. However, the recent practice is to include exchange rate as an independent variable [30]. [40] found that the exchange rate variable is an equivalent measure of the change in destination price.

In the study of [41], the researchers also claimed that a successful tourism in a country will become an important factor in terms of economic that many businesses and government will be more concentrate and interest in the progressively favorable influence such as the production of foreign exchange rate. According to the [42] a high exchange rate fluctuation will affect the travel performance of a country and thus will lead to a change in their business from a country with high exchange rate fluctuation into a country with low exchange rate fluctuation.

[43] examined regarding the tourism demand based on a few of economic determinants included exchange rate as one of their key determinants and the focused country was in Malaysia. The study has found that exchange rate was as hypothesized negatively correlated with the tourist arrivals in Malaysia.

However, the researchers also claimed that the outcome from the indices of the exchange rate can be both positive and negative. This This all relied whether the relative value of the based country has comparatively raised or drop. In other words, travelers from the richer countries will be more likely to choose to travel to a country which is less wealthy due to the lower exchange rate of the money currency of the country the travelers decided to travel [44].

Transportation costs

International tourism is also dependent upon transportation costs. It may be anticipated that an increase in relative transportation costs would result in a decline in international tourism, all things being equal. A number of researchers have used the cost of travel in their models [45]. However, others used air distances rather than monetary costs of travel, while still others did not include any form of travel cost in their studies.

In this study, transportation cost was not considered. Several reasons contributed to this decision:

- It is difficult to obtain accurate data on transportation costs; the data would have had to be aggregate in nature and the quality of such data were suspect;
- it creates the problem of multicollinearity, income and air fares are highly correlated [46].

Since 1987, evaluated the usefulness of multivariable regression analysis in identifying factors which influence tourists' decisions to visit Kenya. The choice of independent variable varies, and it includes incomes in origin countries, exchange rates, transport costs, population of the origin countries and price variables (such as hotel rates or consumer-price indices [22].

Dummy variables of seasonality and new shock effect

Many studies on international tourism demand mostly used quantitative independent variables such as GDP, relative price, and exchange rates. However, considering an accurate forecasting result by understanding and capturing the effects of news shocks and seasonality on the tourism demand is necessary [28]. In studies that forecast international tourism demand, variables relating to tourists' safety and health were use as dummy variables and found to be statistically significant [29].

The dummy variables represent the binary independent variable. Thus, it takes two values: '1' if an event was built and '0' if it was not built. For captured seasonality in the pattern of time series, the monthly dummy variables are put into the multiple regression model [30].

[31] explored the inclusion of marketing variables in international tourism demand models, as National Tourist Organizations often spend considerable sums in foreign countries on promoting the particular country as a tourist destination. However, marketing activity is seldom incorporated as an explanatory variable in models of the demand for international tourism. It was highlighted where marketing is used to explain international tourism demand, caution must be exercised in interpreting the empirical results; poor results cannot be used to reach sensible conclusions, and when good results are obtained the full implications of the estimated coefficients need to be explored.

[32] attempted to identify those economic variables that are most important in influencing business trips to Australia from four of Australia's most important travel and trade partners. They found that the importance of the economic variables varies from country to country, although overall openness to trade and origin country real income are important variables explaining business travel to Australia from these origin countries.

As a way looking into the behavior of tourism demand, the Artificial Neural Networks (ANN) was used to forecast international tourists demand. This method had been considered interesting in economics and business areas since it was viewed as a valid alternative to the classical forecast approaches event in complex situations. [33] highlighted how effectively a Neural Network technique was compared with other forecasting methods and how well the Neural Network technique was implemented, was a key area to consider before concluding on the effectiveness of Neural Networks for forecasting and prediction.[34] had reviewed the main ANN modelling issues which include: (1) the network architecture, i.e., determining the number of input nodes, the number of hidden layers and hidden nodes, and the number of output nodes; (2) the activation function, (3) the training algorithm, (4) the training sample and test sample, and (5) performance measures. In summary, there is no formal systematic ANN model building approach [35] For instance, there was no standard formula for calculating the number of layer and nodes needed in the hidden layer [36] and [37]. In particular, the problem of network "tuning" such parameters of the backpropagation algorithm as well as the Neural Network design needed to be adjusted for optimal performance. Therefore, there was no one single ANN configuration that is suitable for all applications; and a process of trial and error is needed to find the right parameters for an ANN forecasting model.

[27] used ANNs method to predict the tourism demand for two regions (North and Centre of Portugal). The comparing performance of different artificial neural network techniques for tourist demand forecasting was developed by [38] and [6]. [[39], [40]] developed several models based on artificial neural networks, linear regression models, Box- Jenkins methodology and ARIMA models to predict the time series of tourism.

2.3 Review of international tourist arrivals models in terms comparison

Tourism demand modelling and forecasting research relies heavily on secondary data in terms of model construction and estimation. Although the explanatory variables are included in the tourism demand models vary enormously with research objectives and researchers' backgrounds, the employment of certain indicators as the measurement of tourism demand variables in modelling and forecasting tourism demand have been less controversial as suggested in [41]. The tourist arrivals variable is still the most popular measure of tourism demand over the past few years. Specifically, this variable was measured by total tourist arrivals from an origin to a destination.

[31] compared the accuracy of seven forecasting methods for simulating visitor flows among twenty-four origin-destination pairs. They found that exponential smoothing was the second most accurate model in terms of mean absolute percentage error (MAPE). Also, advance forecasting models have been widely used for tourism demand forecasting in the past four decades with 12 the dominance of the integrated autoregressive moving-average models (ARIMAs).

[42] compared five time series forecasting methods in international tourism demand forecasting; naïve trend, double moving average with linear trend, double exponential smoothing, linear trend and autoregressive method. The results showed that all used models have good forecasting performances, but the double moving average method performed the best forecasting performance due to the smallest mean absolute percentage error.

[43] developed tourism demands econometric models based on the monthly data of tourist arrivals to Taiwan and applied Multivariate Adaptive Regression Splines (MARS), Artificial Neural Network (ANN) and Support Vector Regression (SVR).

Furthermore, the Artificial Neural Network combined with the Genetic Algorithm (GA) and Fuzzy were studied to establish a prediction model of international tourists.

[20] investigated the combination of individual forecasting models and their roles in improving forecasting accuracy and proposed two non-linear combination forecasting models using Radial Basis Function and Support Vector Regression neural networks for UK inbound tourism.

[21] examined the forecasting accuracy of different forecasting techniques in modelling and forecasting international tourism demand in Croatia, German visitor arrivals is taken as a measure of tourism demand in Croatia using four techniques: seasonal naïve model, Holt-Winters triple exponential smoothing, seasonal autoregressive integrated moving average model (SARIMA) and multiple regression model. All models are compared considering the in sample and the out of sample mean absolute percentage error (MAPE). The results showed that all used models have good forecasting performances.

[23] estimated and predicted international tourist arrivals to Cambodia. Factors affecting tourist arrivals, simple regression are applied. On the other hand, several time series models of ARIMA (p, d, q), GARCH (s, r) and the hybrid of ARIMA (p, d, q)-GARCH (s, r) are employed to forecast tourist arrivals in line with AIC and BIC in selecting the best modified models. The empirical results primarily reveal that tourist arrivals are affected by exchange rate, dummy factors such as the AEC, global finical crisis, national election and Cambodia's e-Visa. With regard to forecasting stage, the result indicates that tourist arrivals are shocked by time trend in the past period.

2.4 Review of forecasting demand for Thai tourism

The existing literature on forecasting tourism demand is wide ranging both in terms of the different techniques employed and in terms of the different countries covered. The selection of the most accurate forecasting model for a particular destination is often based on the out-of-sample forecasting performance. The mean absolute percentage error (MAPE) or root mean squared percentage error (RMSPE) are computed and compared. There are also several forecasting models for the international tourist arrivals to Thailand.

[44] examined the demand for Thai tourism by seven major origin countries – Australia, Japan, Korea, Singapore, Malaysia, the UK and the USA using classical regression analysis, with ordinary least square method (OLS) as the main estimation procedure. Economic theory is used to recommend what variables should be included in the demand model. The empirical results shown that the income, own price, cross price and trade volume variables were significant in the demand model.

[45] applied multiple regression analysis with independent variables, such as income level in the country of origin, prices of tourism goods in the destination country, currency exchange rate between the origin and destination country, and rooms supply in destination. Qualitative factors, represented by dummy variables, namely special promotional program and political unrest, show slight impact on demand. A model was developed for each country in the East Asia and Pacific Region to forecast tourism demand from that market to Thailand.

[46] examined the factors that influence the behavior of international tourists in Thailand from an economic perspective so that the results can be used as a guideline for Thai policy makers in boosting tourism demand. This paper also estimated the international tourism demand for Thailand during the period 2008 to 2010 by using the time-series seasonal model. Moreover, the impact of news shocks on the international tourism was analyzed

The nonlinear forecasting models for Thailand tourism has already propose by [47]. The seasonal unit root test (HEGY-test extent version) was carried out to test this data, both the MS-VAR model and AR model are employed to predict this data for future of Thailand. The empirical results from this research was concluded that in high seasonal period can be use AR (2)-MLE, AR (2)-MLE bootstrapping, and AR (1)-ME-bootstrapping to predict the number of international tourist arrivals to Thailand for future years. However, in low seasonal period only AR (1)-ME-bootstrapping can be used to predict the number of international tourist arrival to Thailand for future years.

[7] studied forecasting the tourist arrivals from East Asia, namely China, Korea, and Japan, to Thailand. Two forecast models are applied: the AR(m)-GARCH (p, q), and the Kink AR-GARCH model (Kink AR(m)-GARCH (p, q)) that combined the classical GARCH model with the Kink model. The accuracy of the forecast models is evaluated in terms of the RMSE, the MAE and the MSPE.

[48] employed two hybridization models, i.e. hybrid ARIMA-RBFNN model and hybrid RBFNN-ARIMA model to examine the Chiangmai's tourist time series data. Moreover, ANN model used to forecast foreign tourist in Thailand for the variety of scenario including the normal, positive and negative condition by set up the model. The estimated ANN with quarterly data during 2003 to 2015 can be employed for ex pose forecast with only 5.5 percent of forecast error. Tourist income and government budget to promote tourism, are the most important factors for forecast foreign tourists in Thailand.

In recent year, there are also several forecasting models for the international tourist arrivals to Thailand in term of technical comparison including those by [49] Box-Jenkins, regression analysis, and Brown's double exponential smoothing. [50] comparing with the Box–Jenkins and Winter's methods. Moreover, the neural network was developed to forecast international tourists in Thailand, Tourist income and government budget to promote tourism, are the most important factors for forecast foreign tourists in Thailand [51].

For number of tourist's demand, most research used time series forecasting, some used explanatory factors, and some used comparison between time series and explanatory factors. Table 3 presented summary of previous literature on Tourism demand forecasting models. The key contribution of this research is that we focus on finding the most suitable model to forecast monthly tourism demand from main countries to Thailand, while most previous work for Thai tourism had focused on annual demand. In addition, this paper compares performance of time series and causal models and recommend which model is the best for each particular country. Insights are also provided for future model selection for this industry.

Table 3: presented summary of previous literature on Tourism demand forecasting models

Study	Data frequency	Destination focused	Modelling	
			Causal Models	Time series Models
Hao, Var, and Chon (2003) [32]	A	Thailand	MR	
Song and Witt (2003) [31]	A	Thailand	MR	ADLM
Lin, and Lee (2013) [52, 53]	M	United Kingdom	Radial Basis SVR w ANN	
Vu (2006) [57]	Q	Japan		BSM Naïve T ES
Sookmark (2011) [2]	M	Thailand	MR	SARIMA
Baldigara (2013) [24]	M	Croatia	LR	DEST Naïve T AR
C. Lin and T. Lee (2013) [25]	M	Taiwan	SVR MARS ANN ANN-GA	
Chaitipa, Chaiboonsri (2014) [33]	M	Thailand		MS-VAR AR
Saothayanun et al. (2014) [35]	M	Thailand		ARIMA HW
Maja Mamula (2015) [27]	Q	Croatia	MR	Naïve S HW SARRIMA
Wongsathan (2016)[48]	Q	Thailand	ANN	RBFNN-ARIMA
Nattana Boonaom (2016) [34]	A	Thailand	LR	ARIMA HW
Chinnakum and Boonyasana (2016) [19]	M	Thailand		AR-GARCH
Chaivichayachat, (2018) [54]	A	Thailand	ANN	
This research	**M**	**Thailand**	**MR ANN**	**SARIMA HW**

Legend
Data frequency
A: Annual   M: Monthly   Q: Quarterly
Modelling and forecasting methods
AR: Autoregressive method
ADLM: Autoregressive distributed lag model
ARIMA: Autoregressive integrated moving average
ANN: Artificial Neural Network
AR(m)- GARCH (p, q): Autoregressive with generalized autoregressive conditionally heteroskedastic
BSM: non-causal basic structural model
DEST: Double exponential smoothing with trend
ES: Exponential smoothing
Naïve T: naïve with trend, Naïve S: seasonal naïve model
LR: linear regression, MR: multiple regression model
MARS: Multivariate Adaptive Regression Splines
SVR: Support Vector Regression
SARIMA: Seasonal autoregressive integrated moving average
MS-VAR: Markov-switching vector autoregressive
HW: Holt-Winter method

Chapter 3 Methodology

The objective of this chapter is to estimate the international tourism demand for Thailand. To handle the increasing variety and complexity of managerial forecasting problems, many forecasting techniques have been developed in recent years. Each has its special use, and care must be taken to select the correct technique for a particular application. This section provides details about how we selected main countries, data categories and division, data exploration about trends and seasonality, and selected models in this research.

The selection of a method depends on many factors, the context of the forecast, the relevance and availability of historical data, the degree of accuracy desirable, the time period to be forecast, the cost/ benefit (or value) of the forecast to the company, and the time available for making the analysis.

3.1 Research methodology

3.1.1 Country Selection

This study chooses the international arrivals from major countries to Thailand because of their importance on the special interest tourism. The major countries are considered from 80% of all countries.

The criteria for country selection are the first seven countries with a high proportion of tourists and tourist country’s spending in Thailand using Pareto analysis which consider the scores of each country from the proportion of tourists and the major list of tourist expenditure for 2017 which shows in Table 4. Since the tourist expenditure on each country directly impact on Thai economic, so 2.0 is weight for tourist expenditure and 1.0 is weight for proportion of tourists.

Table 4: The major list of tourist expenditure for 2017

Country	Expenditure (billion)
1. China	5.31
2. Russia	1.05
3. Malaysia	0.86
4. USA	0.78
5. United Kingdom	0.77
6. Korea	0.76
7. Japan	0.69
8. Australia	0.65
9. India	0.62
10. Germany	0.57

The scoring of each criteria between 0-10 which score 10 means the greatest number of tourists and spending and score 0 means the least number of tourists and spending. Each country scores with weight which shows in Table 5.

Table 5: The total scores of each county

Criteria	Proportion of Tourist	Tourist Spending	Scores
Weight	**1.0**	**2.0**	
China	10	10	30
Malaysia	7	5	17
Korea	5	3	11
Laos	4	1	6
Japan	4	3	10
India	3	2	7
Russia	2	6	14
USA	1	4	9
Singapore	1	1	3
UK	1	4	9
Vietnam	1	1	3
Cambodia	1	1	3
Germany	1	2	5

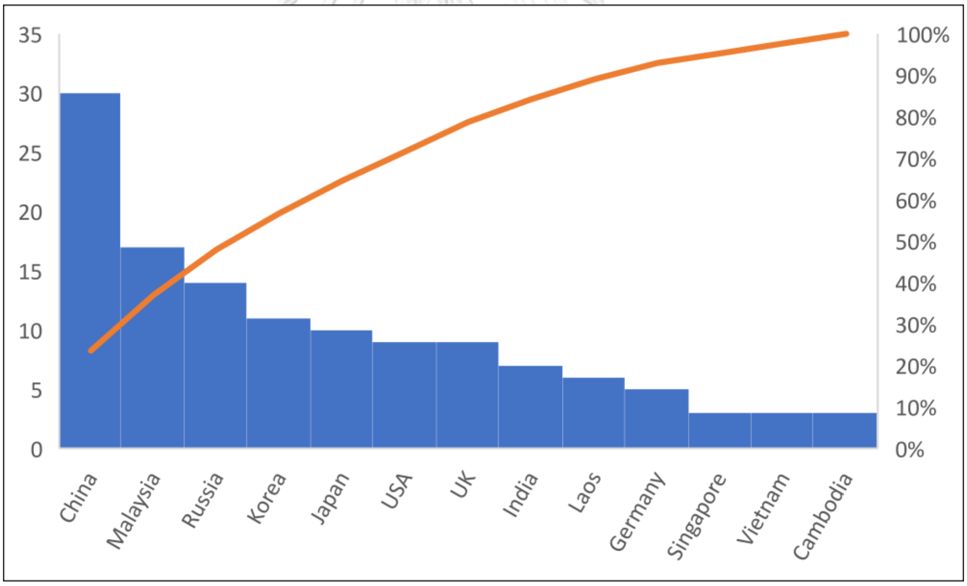

Figure 3: Pareto chart of major countries

From Figure 3, the 80% of all countries are China, Malaysia, Russia, Korea, Japan, USA and United Kingdom. Therefore, these countries are the main focus to identify their suitable demand forecasting models.

3.1.2 The Variables

This study uses the important variables suggested by the literature on tourism demand. The details for each variable are discussed as follows.

- This study selects the international tourist data from major countries using high proportion of tourists and country's spending in Thailand, namely China, Malaysia, Russia, Korea, Japan, USA and United Kingdom. The data set corresponds with the monthly tourist arrivals during the 6 years period from Jan 2013 to Dec 2018. The total data sets consist of 72 observations. The data on the tourist arrivals to Thailand are from the Tourism Statistic Report organized by Tourism Authority of Thailand (TAT).

- The independent variables in this study use the suggested by literature on tourism demand, the used variables are income, price, exchange rates, and dummy variables of seasonality and new shock effects. Details are as follows.

Income Variable

This variable refers to gross domestic product (GDP) per capita of each country in real terms, using constant prices. It is used as a proxy variable for income and is used to explain whether tourists' income plays an important role in explaining tourist arrivals to Thailand. The source of data in each country are from the Economic Trading Statistics [55].

Tourism Price Variable

This variable represents the cost of goods and services purchased by tourists in Thailand. It is measured by relative prices. The relative price variable is given by the ratio of the consumer price indices (CPI) of the destination, which is Thailand, and the origin countries. The relative consumer price can be calculated as follow:

$$\text{Relative Prices} = \frac{\text{CPI(Destination)}}{\text{CPI (Origin)}} \tag{1}$$

The data on CPI of each country are from the Economic Trading Statistics[56].

Exchange Rate Variable

The nominal exchange rate measures the effective prices of goods and services in Thailand. It is expressed in terms of the number of units of origin currency needed to purchase Thai baht. The exchange rate refers to the ratio of the baht and the currency of each origin country. The data on the exchange rate of each country are from the IMF International Financial Statistics, the World Bank, and the Bank of Thailand[57].

Dummy Variables Seasonality and News Shock Effects

Dummy variables can also be used to accommodate the effects of seasonality when quarterly data are used in tourism model estimation. Additionally, for multi destination models, dummy variables are used to represent destination ‘attractiveness’ with the assumption that the relative attractiveness of the destinations remains the same over the time period considered. These are included in the regression model to allow for the effect of 'one-off events.

These variables take the value 1 after 2 months until 3 months from the time the event occurs and 0 otherwise. Thus, the impact of events such as the antigovernment and flood in Bangkok can be incorporated as dummy variables.

3.1.3 Data Division

The data of international arrivals to Thailand divide into 2 intervals as

- The monthly data of Jan 2013 to Dec 2017 use for constructing models.
- The monthly data of Jan 2018 to Dec 2018 use for testing models.

This study also aims to analyze historical data on the international tourist arrivals in Thailand in order to identify any pattern or trends for insights for future planning. To investigate of trend and seasonality of international tourist arrivals to Thailand with monthly data, the time series plots of international tourist arrivals of each country are explored. In this and following sections, the example of time series of number of tourist arrivals to Thailand from main country between Jan 2013 and Dec 2018 are of my interest.

- **China**

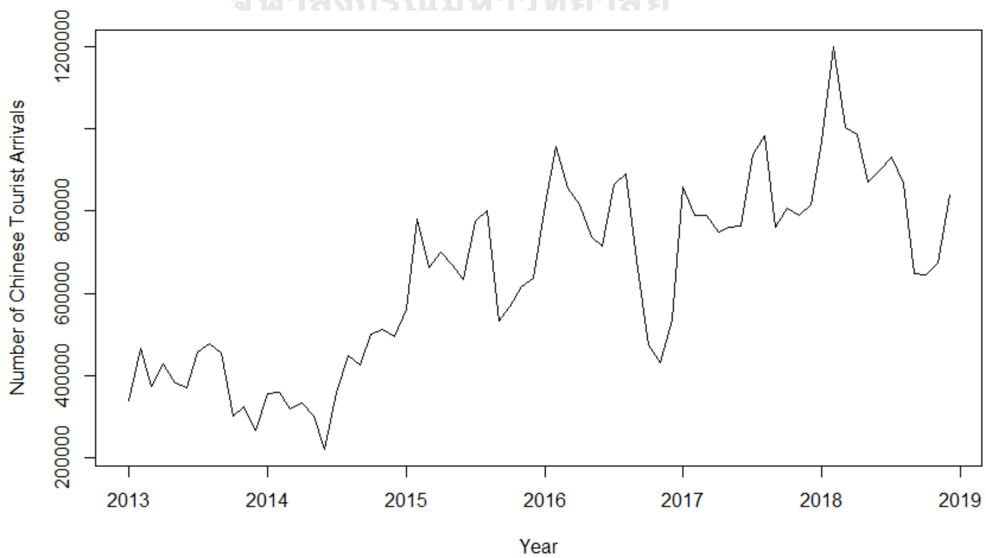

Figure 4: Number of Chinese tourist arrivals to Thailand in 2013 to 2018

From Figure 4, it shows that typically, the pattern of Chinese tourist data is likely to increase at the beginning of the year and continuous fell down until June. After that it rose again between July and August. Then, it decreased to constant at the end of the year. But in 2018, the pattern trend was gradually reduced. China's tourist arrivals in August 2018 were down by 11.8%, the trigger for the drop was a tour boat accident in Phuket in July that killed dozens of Chinese holidaymakers, sparking safety concerns. Thailand's image in China has also been hurt by a dengue outbreak, the strength of the baht and, most recently, a viral video of an airport guard apparently punching a Chinese tourist.

- **Malaysia**

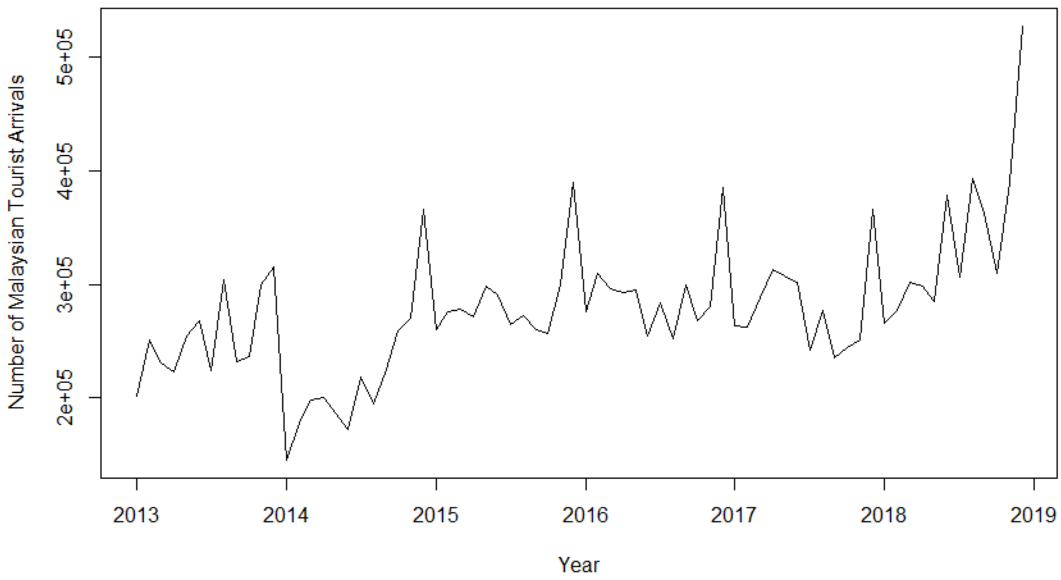

Figure 5: Number of Malaysian tourist arrivals to Thailand in 2013 to 2018

From Figure 5, it shown that at Malaysian tourist came to Thailand at the end of the year, after that, it fell down until October. In 2018, pattern has changed since in June and July, it contained more Malaysian tourist than the last year. It might result from Malaysia's national day, Malaysian tourists celebrated the long holiday in Thailand and after the election in Malaysia, confidence in the economy revived. The number of Malaysian tourists visiting Thailand would return to increase again

➢ **Korea**

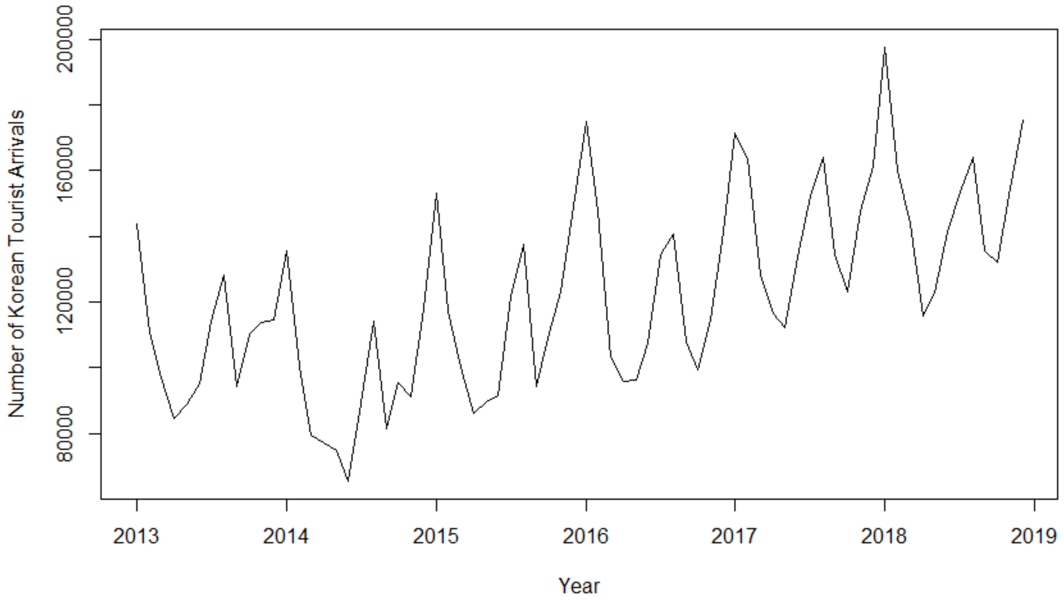

Figure 6: Number of Korean tourist arrivals to Thailand in 2013 to 2018

From Figure 6, it shows that there are more Korean tourists at the first of the year. After that, it decreased gradually in June. Then, it increased again in August. The pattern in every year found that there are trend and seasonality in the pattern.

➢ **Japan**

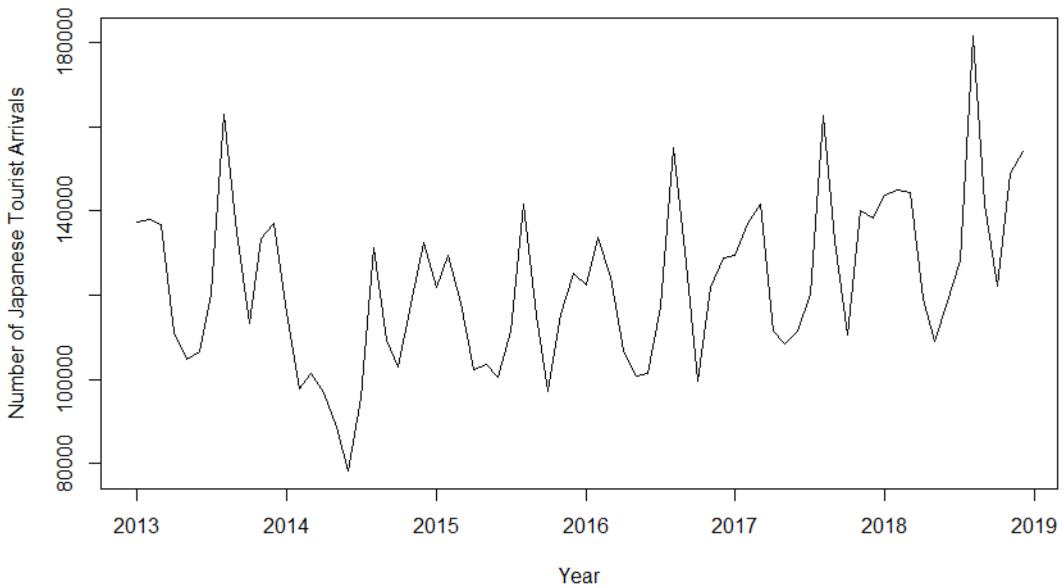

Figure 7: Number of Japanese tourist arrivals to Thailand in 2013 to 2018

From Figure 7, the highest number of Japanese tourists exist in August of each year. Japanese tourists decreased in September and October and it increased again in December to February. These show that there is clearly seasonal effect in the pattern. The Japanese tourist pattern is likely to remain the same since Jan 2013-Sep 2018. The pattern also shows increasing trend of Japanese tourists.

- **Russia**

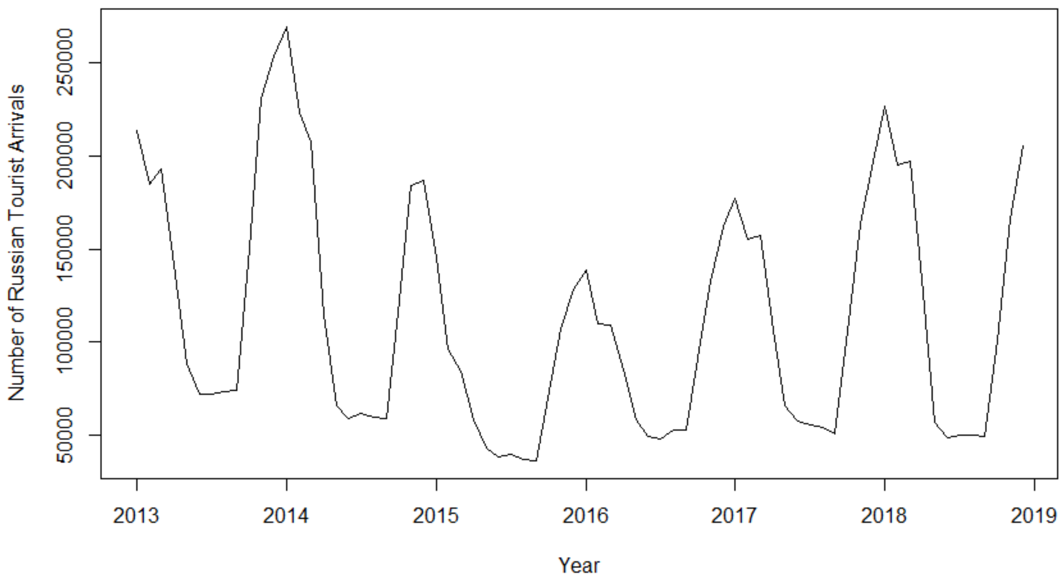

Figure 8: Number of Russian tourist arrivals to Thailand in 2013 to 2018

Russia is a large tourist market with high purchasing power. The program is available throughout the year. From winter (from September to March) This is the season for the European market as a whole. After that, the rainy season (between June and September), Russian tourists have decreased. Then, it rose again at the end of the year. In 2013, the number of Russian tourists in Thailand rose to 1.7 million, before continuing to decline from 2014 to less than 1 million last year. The pattern is likely to decrease from 2013-2016. However, Russian tourists are expected to return to positive after a negative quarter in the third quarter. According to preliminary data, Russian travel agencies have increased their frequency of flights to Thailand since October.

- **UK**

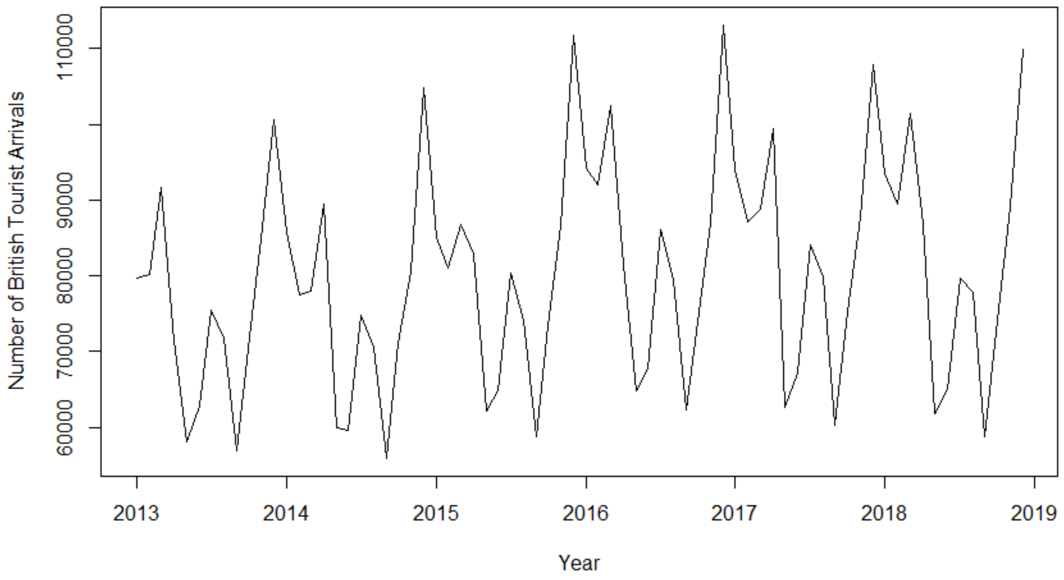

Figure 9: Number of British tourist arrivals to Thailand in 2013 to 2018

From Figure 9, it shows that there is clearly seasonal exist. December is peak season while summer is low season for British tourists. This pattern of British tourist arrivals to Thailand has been remaining the same every year.

- **United State**

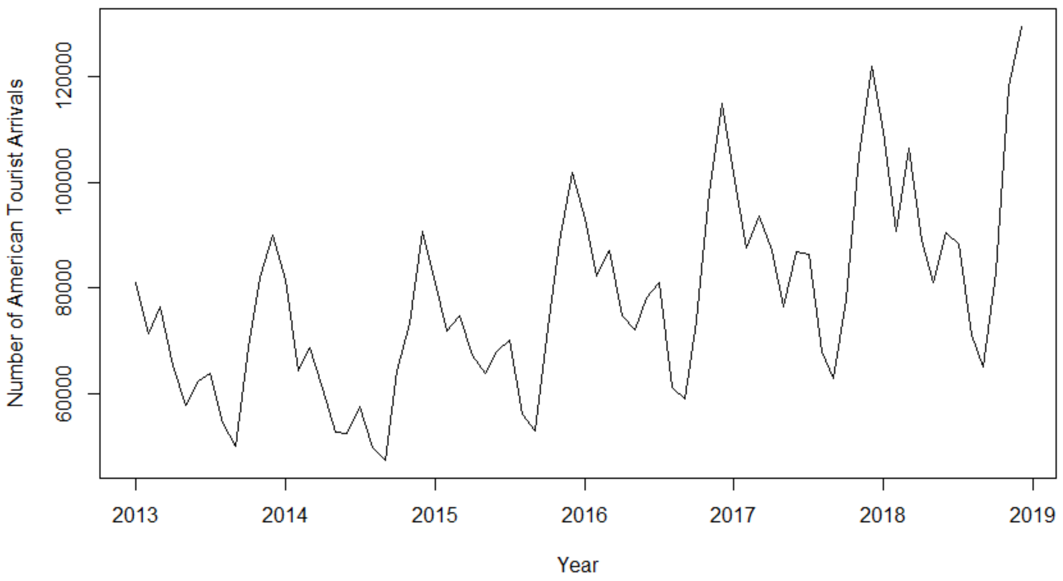

Figure 10: Number of American tourist arrivals to Thailand in 2013 to 2018

The time series pattern in Figure 10, it shows that there are trend and seasonal exist. American tourist pattern has been increased every year. As the same British tourists, December is peak season but number of American tourists decreased in rainy season in Thailand.

3.2 Forecasting Models

Since the time series plots of major country are analyzed, it is found that there are trend and seasonality existing in each country. Therefore, the forecasting methods are considered from patterns notification in each country's data. The example of time series plot of American tourist arrivals is demonstrated in Figure 11.

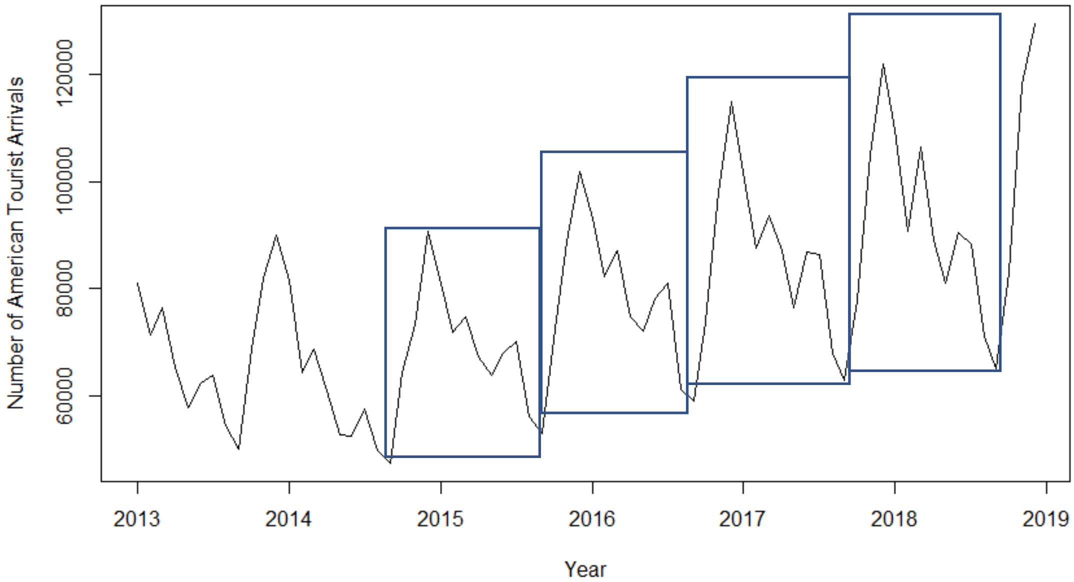

Figure 11: Time series plot of American tourist arrivals which show seasonality and trend

There are two main categories for forecasting methods: qualitative and quantitative methods [56].
Qualitative methods are used typically when there's no information for mathematically models. Although qualitative methods offer often high-class information about future demand [57]. It can be divided into time series methods and causal methods. Time series methods are used for a short period forecasting and there's used historical data about demand. Although causal methods are used to forecast medium and long time period demand [58]

The forecasting methods explored in this research are both time series models, i.e., SARIMA and Holt Winters because these methods can be captured trend and seasonality. For explanatory models, i.e., Multiple Regression and Feed Forward Artificial Neural Networks (FANNs) are employed, these methods can be used to analyze the linear and nonlinear relationships between factors and number of tourist arrivals in each country. In addition, explanatory models are also used to capture the patterns of time series that is not clearly seasonality, the example of time series plot for

Malaysian tourist arrivals is shown in Figure 12. It can be noticed that seasonality is not very obvious in these data.

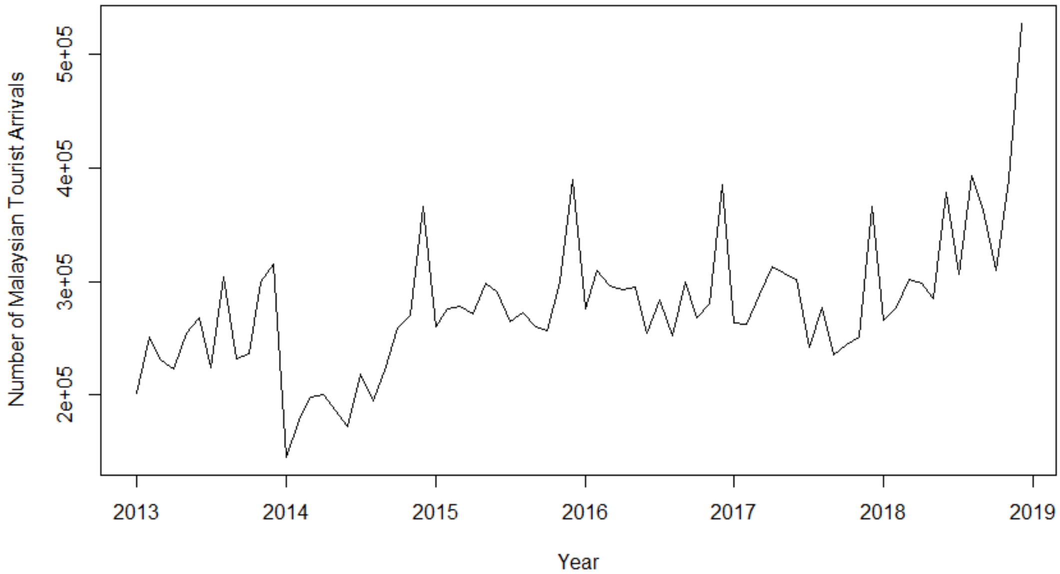

Figure 12: Time series plot for Malaysian tourist arrivals

This study present two quantitative methods as follows; explanatory methods and time series methods. To build accurate models for each country, this study focuses on methods screening from simple methods to advanced and complex methods. The time series models are screened and the explanatory models are analysed if the time series models still have high error for forecasting result. Details of each models are as follows.

3.2.1 Time series forecasting method

Holt and Winters exponential method

The Holt-Winters model was frequently used for the short-term forecasting of tourist arrivals so it is selected as baseline model. The Holt-Winters multiplier model involves triple exponential smoothing (overall smoothing, trend smoothing and seasonal smoothing). This method is a suitable seasonal model that could reasonably be applied (from past studies) to offer an acceptable level of forecasting accuracy. After the values of level, trend and seasonality were calculated, the following equation was used for forecasting [58]:

$$L_t = \alpha\,(Y_t / S_{t-p}) + (1-\alpha)\,[L_{t-1} + T_{t-1}] \tag{2}$$

$$T_t = \beta\,[L_t - L_{t-1}] + (1-\beta)\,T_{t-1} \tag{3}$$

$$S_t = \gamma\,(Y_t / L_t) + (1-\gamma)\,S_{t-p} \tag{4}$$

$$\hat{Y}_t = (L_{t-1} + T_{t-1})\,S_{t-p} \tag{5}$$

Where, L_t = level at time t,

α = the weight for the level

T_t = trend at time t,

β = weight for the trend

S_t = seasonal component at time t

γ = weight for the seasonal component

S = seasonal period

Y_t = data value at time t

$\hat{Y}_t$= fitted value, or one-period-ahead forecast, at time t

In this study calculate the values of the three smoothing constants using R studio program based on the minimum absolute percentage error (MAPE). Furthermore, an analyzed variation used patterns of time series to identify weather time series is additive variation or multiplicative variation.

Seasonal ARIMA method

Time-series methods make forecasts based solely on historical patterns in the data. Time-series methods use time as independent variable to produce demand. In a timeseries, measurements are taken at successive points or over successive periods. The measurements may be taken every hour, day, week, month, or year, or at any other regular (or irregular) interval. A first step in using time-series approach is to gather historical data. The historical data is representative of the conditions expected in the future. Time-series models are adequate forecasting tools if demand has shown a consistent pattern in the past that is expected to recur in the future.

ARIMA models are commonly used in tourism forecasting, The Box-Jenkins methodology is based on the assumption that the underlying time series is stationary or can be made stationary by differencing it one or more times. This is known as the ARIMA (p, d, q) model, where d denotes the number of times a time series has to be differenced to make it stationary. For seasonal time series literature suggests seasonal autoregressive integrated moving average models, also called SARIMA (p, d, q) (P, D, Q) $_s$ models. The ARIMA model can be expressed as:

$$\Phi(L^s)\ \phi(L)\ \Delta^d \Delta_s{}^d y_t = \theta_0 \Theta(L^s)\ \theta(L)\ \varepsilon_t \qquad (6)$$

where s is the seasonal length, for example s = 12 for monthly and s = 4 for quarterly data, L is the lag operator and Δ_t is assumed to be a Gaussian white-noise process with mean zero and variance σ^2. The difference operator is Δ_d where d specifies the order of differencing and the seasonal difference operator is $\Delta^D{}_s$ where D is the order of seasonal differencing. The difference operators are applied to transform the observed non-stationary time series y_t to the stationary [15]. Fitting Box–Jenkins models for a seasonal model is demonstrated in Figure 13.

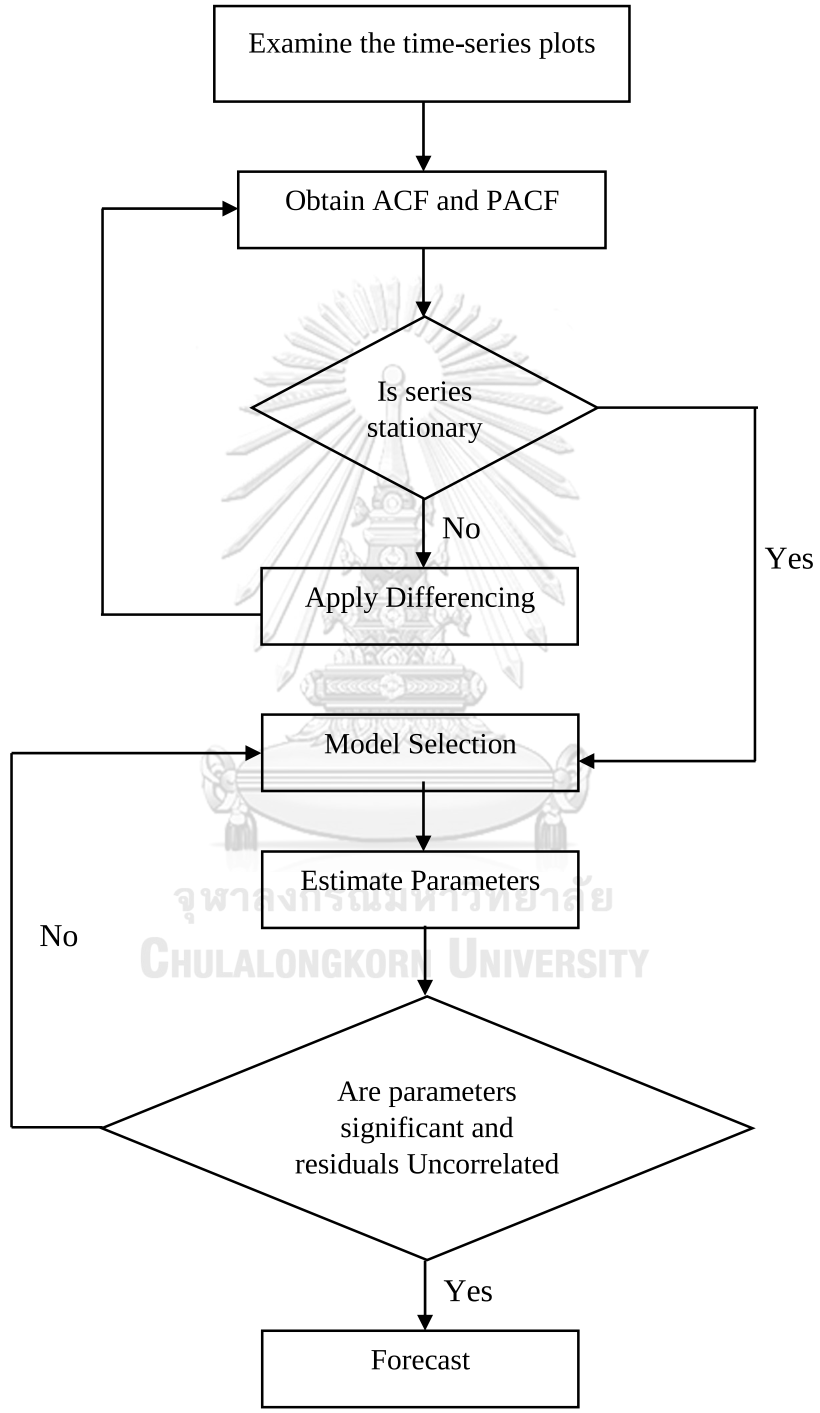

Figure 13: Fitting Box–Jenkins models for a seasonal model

From Figure 13, the time-series plots for seasonality and trend (i.e., check for stationarity using Augmented Dickey Fuller Test (ADF) are examined. If the time series is stationary, the parameters of AR and MA terms are identified from considering ACF and PACF. When the model is selected, its parameters can be estimated using statistical techniques, such as the maximum likelihood, least-squares, or Yule–Walker method and then tests on the residuals are performed in order to determine whether the model is adequate for the data. It is necessary to check the assumptions of normality and homoscedasticity, and also to check for autocorrelations (using the Ljung–Box test). It is sensible to use a p-value threshold of 0.05 (and equivalently a confidence level of 95%), since this is the most widely used value and allows comparison to other studies. The model selecting used AIC values to select the accurate models for each country. After the model for each country is selected, the forecasting errors are calculated.

A refinement of the Box-Pierce test proposed by Ljung and Box appears to provide a better approximation to the chi-square distribution of the Q statistic under the null hypothesis. The Ljung-Box Q* statistic is calculated in Equation

$$Q^* = n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{n-k} \tag{7}$$

Where; n = number of observations in the time series,

h = maximum number (about 20) of first set of residual lags being tested, and

n = the kth residual autocorrelation; k = 1, 2, 3, …, h

A p-value below .05 leads to reject the null hypothesis that the residuals are random at the a = 0.05 significance level. In other words, p-values below 0.05 signals model inadequacy and p-values greater than 0.05 are desired. If these estimated residuals appear satisfactorily to imitate a white noise process, and then the chosen model can proceed to the forecasting phase. Otherwise, a new tentative Box-Jenkins model should be considered.

3.2. 2 Explanatory forecasting method

Explanatory forecasting models usually consider several independent variables that are related to the dependent variable being predicted. Once these related variables have been found, a statistical model is built and used to forecast the variable of interest. This approach can be more powerful than the time series methods that use only the historic values of the variable to be forecasted.

The most common quantitative causal forecasting method is linear regression analysis. The linear regression model involves identifying variables for defining the model and specifying the residual variable, the context in which the regression model is used. The goal of using the regression model is to obtain the parameters corresponding to the set of variables formulated by analyzing dependence between variables, where data series are recorded at the level of population statistics for a period or a moment, and to highlight the dependence between variables in a given time horizon [59].

Multiple Regression

The estimation of international tourism demand to Thailand from the 7 major countries of origin is presented in the following formula:

$$Q_{ij} = f(G_i , R_i ,Ex_{ij} , D_j, S_{jt}, \varepsilon_{ij}) \quad (8)$$

where:
Q_{ij} = quantity of tourism demanded in destination i by tourists from country j,
Ex_{ij} = Exchange rate from origin country j to destination country i. It expects a positive sign because an increase in exchange rate implies that the currency of the source country will have greater purchasing power in Thailand.
R_{ij} = tourism price which based on the consumer price index (CPI) of the destination i divided by the CPI of the country of origin j. The calculation is shown in this equation:

$$Relative\ \mathrm{Price} = \frac{\mathrm{CPI\ (Destination)}}{\mathrm{CPI\ (Origin)}} \quad (1)$$

G_{it} = GDP from origin country j in the t period, with a positive expected sign
S_{it} = Seasonal effect, set of k-1 monthly seasonal dummy variables where k is an indicator of the month (k=1~12).
N_{it} = News shock effects, these variables being expected to have affect to individual decision to participate in news shock.

Events selecting are considered from time series plot in each country, the time series of Chinese tourist arrivals is the example as shown in Figure 14. From Figure 14, it shows that the number of Chinese tourists will be affected by each news shock effects after 1-2 months and the impact continues for the next 1-2 months.

This study focuses on the news shock that negatively impacted tourist arrivals, the timeline of news shock effects shown in Figure 14 are as follows.

- Anti-Government happened in November 2013 and May 2014., Since the Chinese tourist decreased in first of the 2014. It means that the news shock affected to Chinese tourists from January to March in 2014.
- Bomb in Bangkok happened in August 2015, this news shock affected the Chinese tourists from October to December 2015.
- Floods in the South of Thailand happened in December 2016, this news shock suddenly affected the Chinese tourists in December to February.
- Thai King Rama9 Passed Away in October 2017, this news shock affected the Chinese tourists in December 2017.
- Boat accident in Phuket happened in July 2018, this news shock affected the Chinese tourists from September to December 2018.

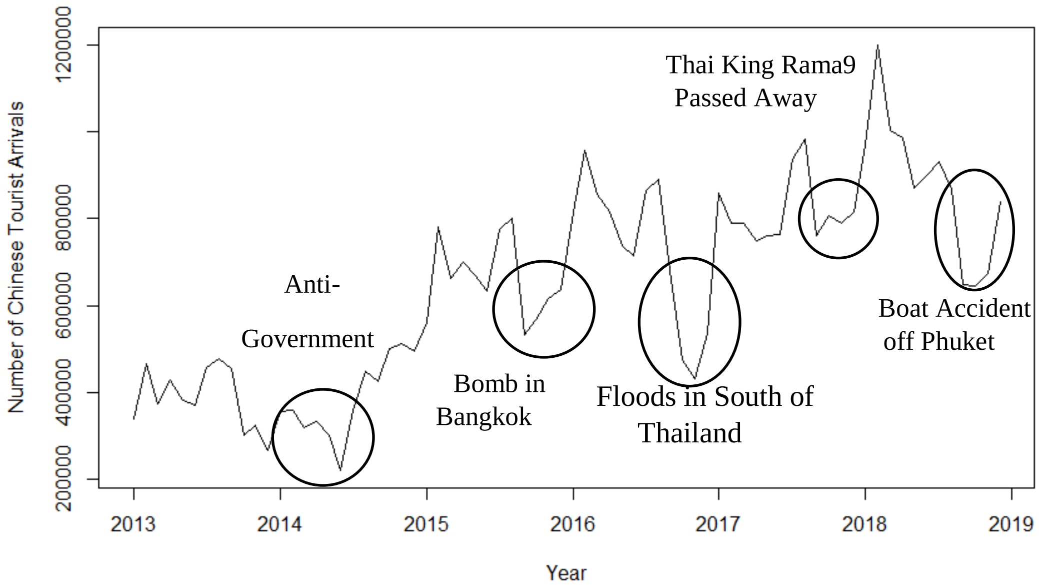

Figure 14: Timeline of news shock that impacted Chinese tourist

The dummy variables of news shock effects used to define in the models are as follows.

1) Anti-Government

The news shocks of anti-government was expressed as dummy variables. This variable is expected to be negatively associated with tourist demand in Thailand. Dummy variable was defined as Anti and was equal to "1" for shocks of anti-government and "0" otherwise. In the model training, this dummy variable is 1 from January to March.

2) Bomb in Bangkok

The news shocks from Bangkok, Thailand was expressed as dummy variables. This dummy variable on news shocks regarding the bombing in Thailand. The variable was expected to be negatively associated with tourist demand. Dummy variables were used to define as Bomb equal to "1" for news shocks of the boom in Thailand and "0" otherwise. In the model training, this dummy variable is 1 from October to December.

3) Floods in South of Thailand

The news shocks from domestic national disasters were expressed as dummy variables. This dummy variable on news shocks regarding the domestic natural disasters in Thailand in this study was the flood in south of Thailand. The variable was expected to be negatively associated with tourist demand. Dummy variables were used to define as Flood equal to "1" for news shocks of the Flood disaster in Thailand and "0" otherwise. In the model training, this dummy variable is 1 from December to February.

4) Thai King Rama9 Passed Away

The news shocks of Thai king rama9 passed away was expressed as dummy variable. This dummy variable on news shocks of Thai king death in Thailand is presented as Bangkok, Thailand. This variable is expected to be negatively associated with tourist demand. Dummy variables were defined as King rama9 passed away and were equal to "1" for shocks of king death in Bangkok, Thailand and "0" otherwise. In the model training, this dummy variable is 1 from December to February.

5) Boat accident in Phuket

The news shocks of Boat accident in Phuket was expressed as dummy variable. This variable is expected to be negatively associated with tourist demand. Dummy variable was defined as Boat accident in Phuket and were equal to "1" for shocks of Boat accident in Phuket and "0" otherwise. In the model training, this dummy variable is 1 from September to December.

In this study, economic explanatory factors are used as independent variables. The general form of a multiple regression model is:

$$Y_t = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \varepsilon \quad (9)$$

Where;
X_1 is income in term of GDP per capita
X_2 is relative price
X_3 is exchange rate
X_4 is dummy variable of seasonal effect
X_5 is dummy variable of news shock effect

β_0 is constant variable

β_1 is coefficient of the first control variable, X_1

β_2 is coefficient of the second control variable, X_2

β_3 is coefficient of the third control variable, X_3

β_4 is coefficient of the fourth control variable, X_4

β_5 is coefficient of the fifth control variable, X_5

ε is error term
The dependent variable is number of tourists visited Thailand of each country.

This study applies backward elimination regression to reduce the set of predictor variables to those that are necessary and significant. In addition, backward elimination regression also helps to determine the level of importance of each predictor variable. It also assists in assessing the effects once the other predictor variables are statistically eliminated.

Artificial Neural Network

ANN construction is generated by the inspiration of biological neural networks. ANN can have human-like simple determination capability and judgment, which is advantageous to formal logic reasoning. A common multilayer feedforward network consists of three parts as shown in Figure 15.

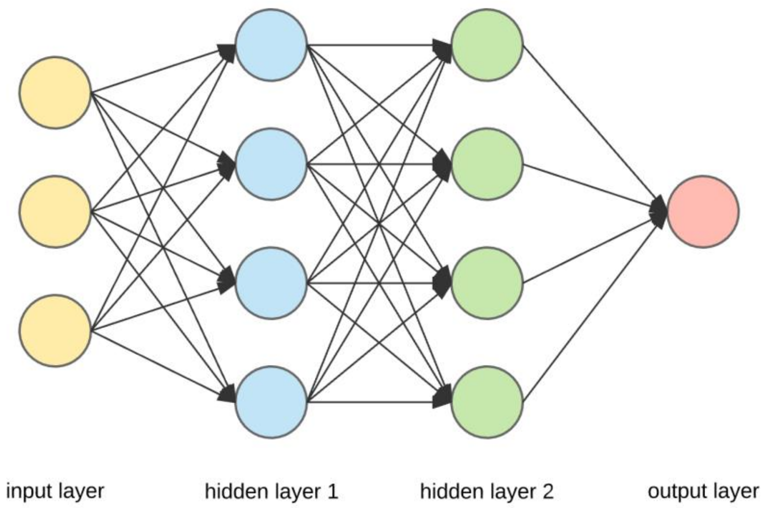

Figure 15: A multilayer feedforward neural network [59]

There are multiple methods in neural network literature. This study focuses on feedforward neural network, since the result of MAPE when trying to build back propagation neural network in China is higher than that of feedforward neural network. The feedforward neural network is mostly used for supervised machine learning tasks where we already know the target function. The multilayer feedforward network is the flow of information takes place in the forward direction, as x is used to calculate some intermediate function in the hidden layer which in turn is used to calculate y. The main goal of a feedforward network is to approximate some function f* in hidden layers. For example, a regression function $y = f^*(x)$ maps an input x to a value y. A feedforward network defines a mapping $y = f(x; \theta)$ and learns the value of the parameters θ that result in the best function approximation. The function of this network was described as follows:

$$Y_j = f(\Sigma W_{ij} X_{ij}) \tag{10}$$

where Y_j = the output of node j,

$f(.)$=the transfer function,

W_{ij} = connection weight between node j and node i in the lower layer

X_{ij} =the input signal from the node i in the lower layer to node j [60].

Architecture Design and Model Training

Since there is no recognized number of neurons in the hidden layer, when the number of neurons is larger, the nonlinearity will be more significant and the robustness of the neural network will be more significant while too few hidden units underfit the model, and is not sufficiently accurate. Therefore, this study focuses on optimizing the number of hidden nodes in hidden layers to improve the performance of ANNs by minimizing the mean absolute percentage error over noisy training data.

In this work, the number of nodes in hidden single and double layer is optimized with comparing single layer and double layers by additional coding in R program. The range of number of nodes in each layer is 1 to 70. The maximum is 70 because after 70 nodes are too high error training. The combination of single layer is 70 values while combination of double layers is 4900 values (1-70, 1-70).

For each hidden layer, the log Sigmoid function which is the most widely used activation function is used in this research. It is defined as [58]:

$$S(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}. \tag{11}$$

The supervised feedforward neural network learns from training data to discover patterns representing input and output variables. Usually, the process of learning involves the following stages.

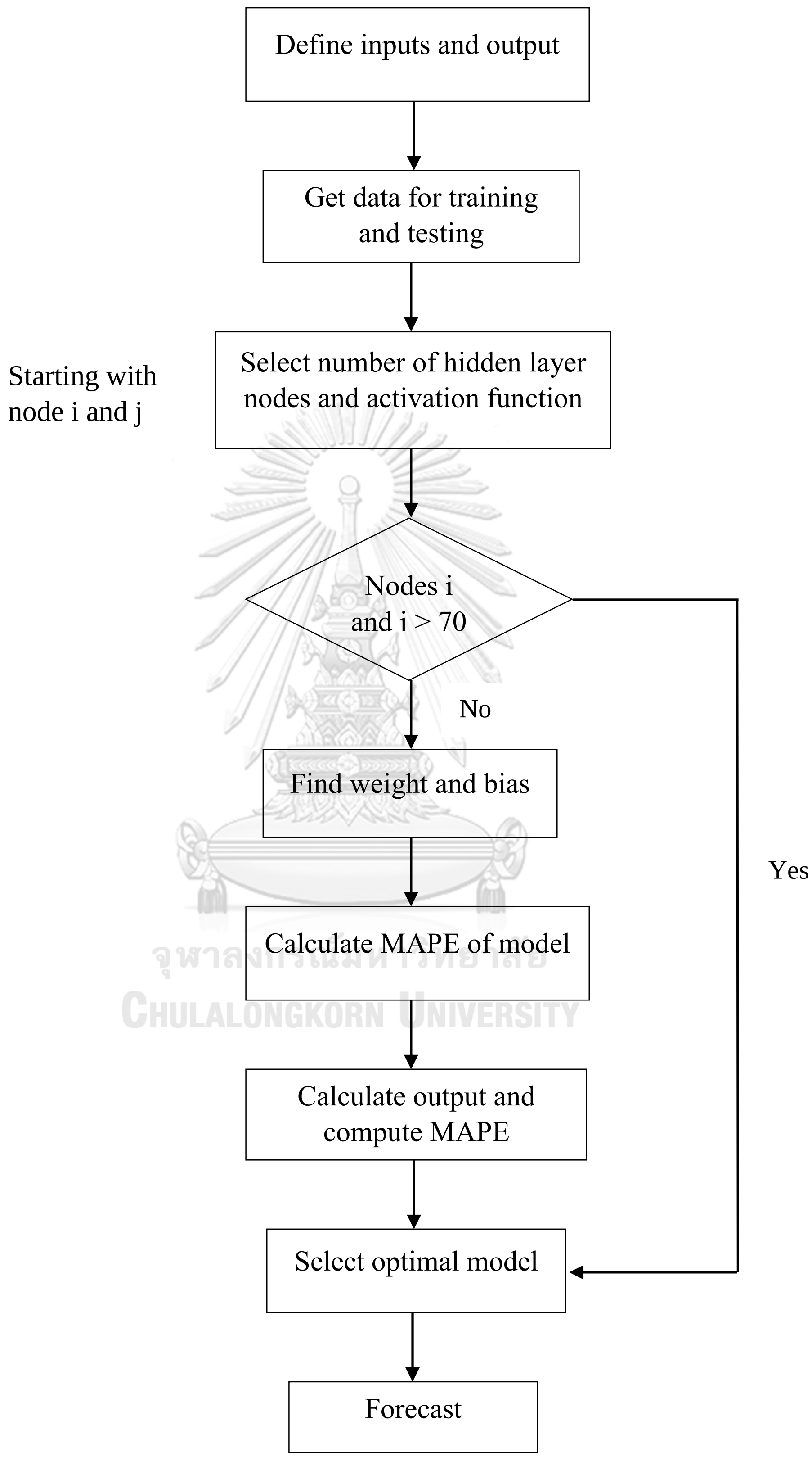

Figure 16: The steps of building feedforward neural network

From Figure 16, the numbers of hidden nodes in each hidden layer are defined in i and j nodes; i and j are 1-70. Since the program start running, the weight in each input is assigned randomly. MAPE of models in each node are calculated, and then the optimal number of nodes is selected to forecast number of tourist arrivals to Thailand for each country.

3.3 Evaluation of forecasting performance

Since some companies relies on a single forecasting method for a specific given data it is important to identify the most accurate method. For generalizing about forecasting Methods: Empirical Comparisons" five different error measures in terms of reliability, construct validity, sensitivity to small changes, protection against outliers, and their relationship to decision making.

Five different error measures are present to assess the performance of the different forecasting method.

Where; A_t = the actual value in the time series at time t.

F_t = the forecasted value in the time series at time t.

n = the length of the time series.

The following error measures are calculated:

• Mean Absolute Error (MAE): Measures average absolute deviation of forecast from the actual.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t| \tag{12}$$

• Mean Absolute Percentage Error (MAPE): The average absolute percentage of errors to the actual values. Accuracy is expressed as a percentage.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{13}$$

• Mean Square Error (MSE): The average of the square of the difference between the actual and the forecast.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2 \tag{14}$$

• Root Mean Square Error (RMSE): Expresses the variance plus the bias of the estimator Standard Squared Error:

Where;

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(A_t\text{-}F_t)^2} \tag{15}$$

To find out how accurate the predicted numbers really are, it will be taken a closer look at the error values itself. Since it is more difficult to interpret the RMSE, this method is also not reliable, in the field of comparing accuracy across series, the main focus in interpreting the accuracy of the forecasting models lies on the MAPE. The difficulty in the interpretation lies in the format of the result, because in contrast to the MAPE, which provides a percentage, the RMSE delivers only an absolute value. There is no limit of the value, the level of the result normally depends on the height of the actual and the forecast values.

3.4 Comparing of Forecasting Methods

It is important to identify the most accurate method for each country using suitable measures. In this study, Mean Absolute Percentage Error (MAPE) is selected for forecasting accuracy measurement since it is insensitive to means of data and appropriate for comparison of different counties. The interpreting the MAPE are shown in Table 5.

Table 5: Accuracy of the forecasts

MAPE value	Interpretation
<10%	Highly accuarate
10-20%	Good
20-50%	Reasonable
>50%	Bad

Source: [61]

In addition, to test of different forecasting methods provide significantly different accuracy, this research examines Randomized Complete Block Design (RCBD). The p-value that is greater than 0.05 indicated that there are no statistically significant differences between group means as determined by RCBD. In addition, pairs of methods are analyzed using the Tukey test to identify pairs that are significantly different.

Chapter 4 Results and Discussion

This chapter presents the empirical results of the time-series forecasting performance analysis compared on the causal forecasting performance measures. Firstly, the various forecasting performance measures are discussed. Next is comparing forecasting results.

4.1 Time Series Method

4.1.1 Holt and Winters Method

The R Studio software program automatically provides optimized estimates of model parameters (alpha, beta, and gamma) that minimize forecast errors.

The forecast values and the Mean Average Percentage Errors (MAPEs) are then calculated for each of the China, Malaysia, Korea, Japan, UK, and USA tourist flows. A summary of the parameters and MAPEs for the China, Malaysia, Korea, Japan, Russia, UK, and US tourist arrivals are shown in Table 6.

Table 6: Summarized results of parameters and MAPE from Holt-Winters Method

Origin Country	Type of variation	Parameters			MAPE	
		Alpha	Beta	Gamma	Training	Testing
China	M	0.629	0	1	14.95	12.13
Malaysia	M	0.562	0	0.906	12.49	39.74
Korea	M	0.635	0.002	1	7.95	19.60
Japan	M	0.765	0	1	7.07	26.25
Russia	M	0.299	0.109	1	8.02	13.43
UK	A	0.136	0.009	0.441	3.70	16.23
US	A	0.548	0.040	1	4.46	2.88

M: multiplicative Variation A: Additive Variation

The MAPE results in Table 6 show that Holt-Winters Method is the most accurate for forecasting American tourist arrivals to Thailand in 2018 while all other counties still have high MAPE value. A possible reason is the time series pattern of American tourist arrivals (shown in Figure 10) has not been changed in 2018 and clearly shows seasonality and trend.

4.1.2 Seasonal ARIMA Method

Monthly tourist arrivals data from the seven countries are used to estimate ARIMA parameters using least square method. The adequate models selected for major countries' tourist arrivals are based on significant t-statistics at the 95% confidence level of significance. In addition, modes are selected using Akaike Information Criterion (AIC), whereby smaller values are preferred. The results of estimating the various ARIMA processes of tourist arrivals from seven major countries to Thailand are presented in Table 7.

Table 7: Parameter optimization for SARIMA for each country

Origin Country	**Modelling**	**AIC**
	(1,1,0)(1,0,0)$_{12}$	1522.47
	(0,1,1)(0,0,1)$_{12}$	1525.44
China	(1,1,0)(1,0,1)$_{12}$	1523.39
	(0,1,0)(1,0,0)$_{12}$	**1518.46**
	(0,1,0)(1,0,1)$_{12}$	1519.15
	(1,1,0)(1,0,0)$_{12}$	1520.39
	(2,0,2)(1,1,1)$_{12}$	1149.51
	(2,0,0)(1,1,0)$_{12}$	1138.37
Malaysia	(2,0,0)(1,1,1)$_{12}$	1143.67
	(2,0,1)(1,1,0)$_{12}$	**1141.62**
	(2,0,0)(0,1,0)$_{12}$	1158.94
	(1,0,0)(1,1,0)$_{12}$	1142.11
	(0,1,0)(0,1,0)$_{12}$	**1005.69**
	(0,1,0)(0,1,0)$_{12}$	1007.02
	(0,1,1)(0,1,1)$_{12}$	1006.79
Korea	(0,1,0)(1,1,0)$_{12}$	1006.20
	(1,1,0)(0,1,0)$_{12}$	1006.85
	(1,1,1)(0,1,0)$_{12}$	1007.71
	(0,1,0)(0,1,0)$_{12}$	993.64
	(0,1,1)(0,1,1)$_{12}$	991.23
Japan	(0,1,1)(0,1,0)$_{12}$	994.71
	(0,1,0)(1,1,0)$_{12}$	991.83
	(0,1,0)(0,1,1)$_{12}$	**989.20**
	(1,1,1)(0,1,1)$_{12}$	991.94
	(0,1,0)(0,1,0)$_{12}$	1067.97
	(1,1,0)(1,1,0)$_{12}$	1057.08
Russia	**(0,1,2)(0,1,1)$_{12}$**	**1050.69**
	(0,1,1)(0,1,0)$_{12}$	1058.56
	(1,1,2)(0,1,1)$_{12}$	1052.39
	(0,1,3)(0,1,1)$_{12}$	1051.786
	(0,0,0)(1,1,0)$_{12}$	**972.99**
	(0,0,1)(1,1,0)$_{12}$	975.21
UK	(1,0,1)(1,1,0)$_{12}$	977.76
	(1,0,0)(0,1,0)$_{12}$	981.59
	(1,0,0)(1,1,0)$_{12}$	975.22
	(0,0,0)(0,1,0)$_{12}$	979.54
	(0,1,0)(0,1,0)$_{12}$	917.35
	(1,1,0)(1,1,0)$_{12}$	917.891
US	(0,1,1)(0,1,1)$_{12}$	916.18
	(0,1,1)(0,1,0)$_{12}$	**914.53**
	(0,1,1)(1,1,1)$_{12}$	918.54
	(1,1,0)(0,1,0)$_{12}$	916.43

The results from Table 7 show parameters optimization for SARIMA for each country, the parameters selecting base on AIC. The optimum model of China is $(0,1,0)(1,0,0)_{12}$ with 1518.46, Malaysia is $(2,0,1)(1,1,0)_{12}$ with 1141.62, Korea is $(0,1,0)(0,1,0)_{12}$ with 1005.69, Japan is $(0,1,0)(0,1,1)_{12}$ with 989.20, Russia is $(0,1,2)(0,1,1)_{12}$ with 1050.69, UK is $(0,0,0)(1,1,0)_{12}$ with 972.99, and US is $(0,1,1)(0,1,0)_{12}$ with 914.53.

The Box-Jenkins methodology is required to diagnose the selected models. This diagnostic checking conducts through examining residuals from the fitted model. The China model represents an example of residual analysis to check the assumption of SARIMA model as shown in Figure 17-20.

From Figure 17-18, it shows that residuals after fitting the model was white noise, there exists no significant autocorrelation and partial autocorrelation in the residual series.

From Figure 19-20, it shows that the estimated autocorrelations of these residuals are approximately normal distribution and uncorrelated from the fitted model. The residual plot of the rest countries shown in appendix 2. Furthermore, the Ljung-Box test is not statistically significant (p-value = 0.05144) confirming the fact that residues are randomly distributed. The values of the Ljung-Box statistic in other countries are presented in appendix 1.

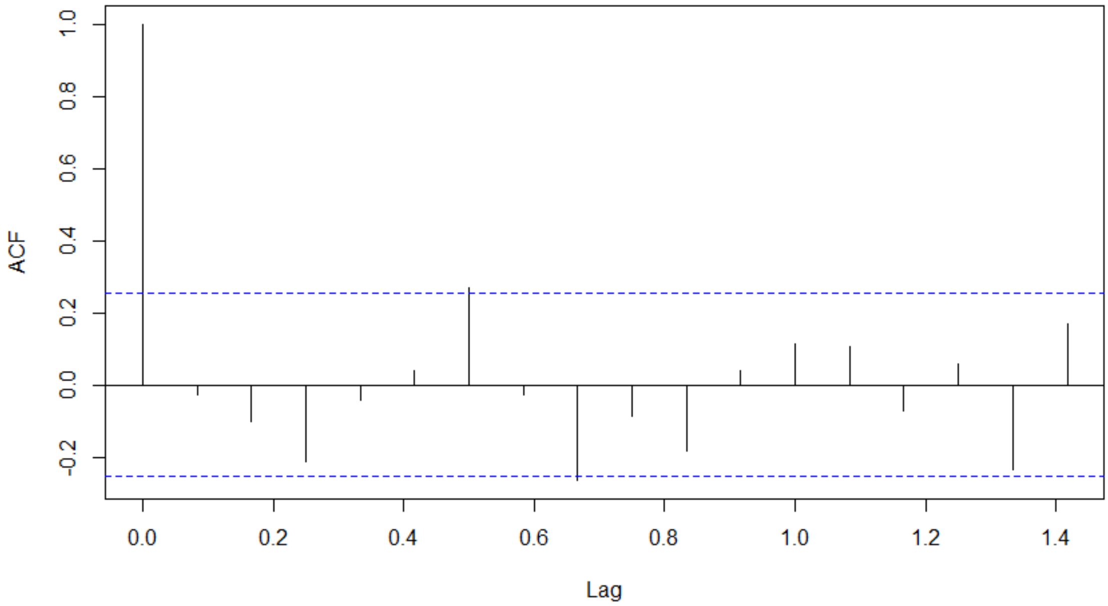

Figure 17: Autocorrelation function of residual for China

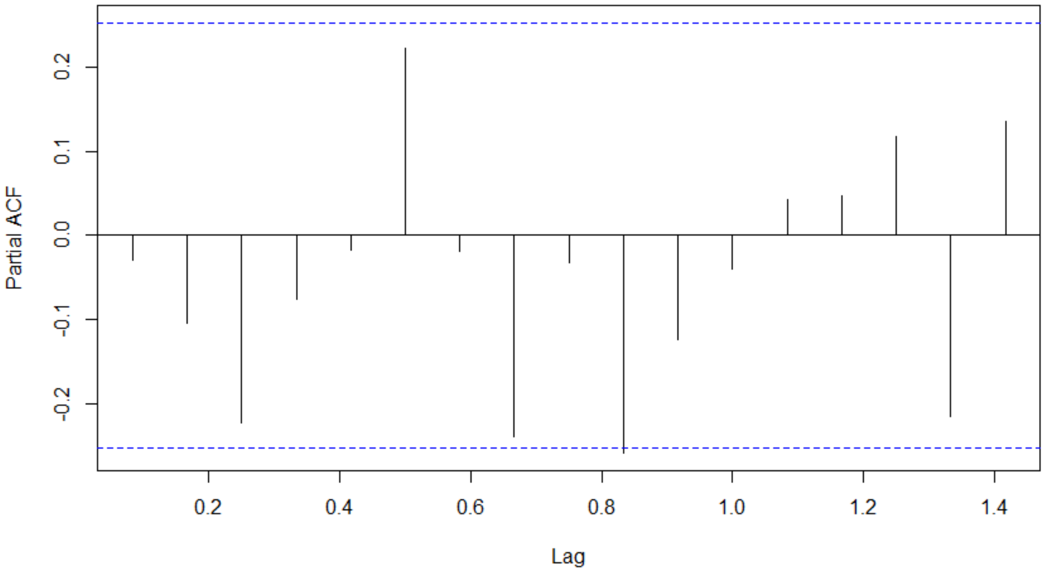

Figure 18: Partial autocorrelation function of residual for China

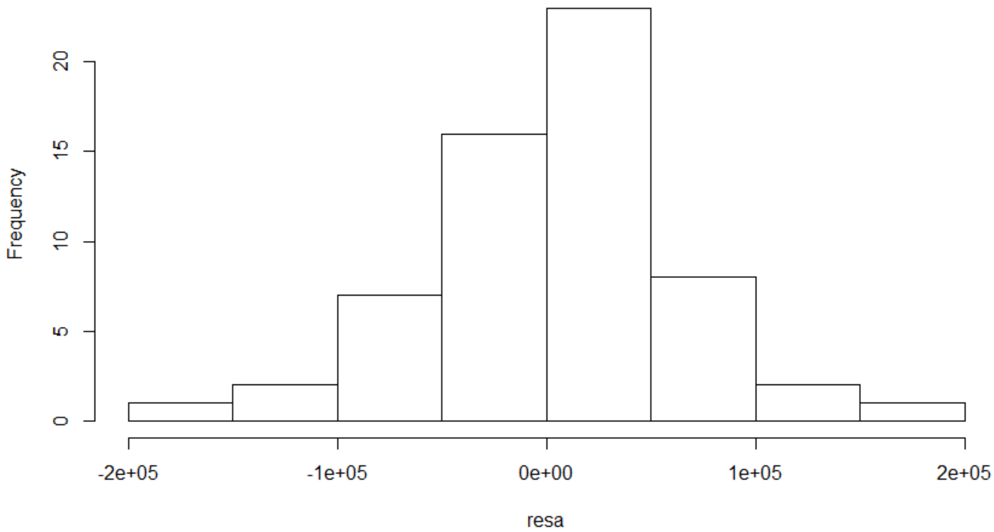

Figure 19: Normality test of residual for China

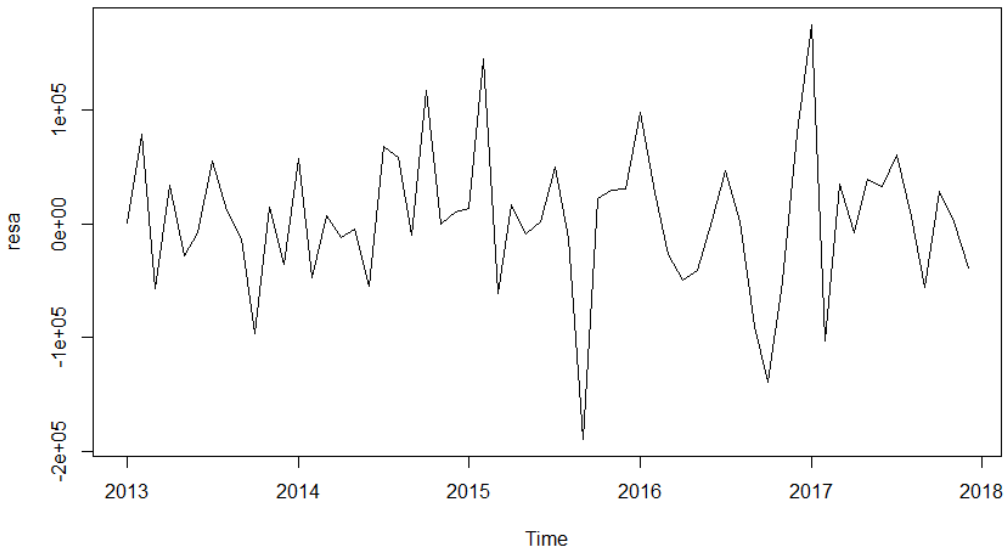

Figure 20: The residual plot over time for China

After the empirical examination the most appropriate models for tourist arrivals from seven major countries were determined. The best fitting model for each tourist arrival series, and the accuracy of the twelve months ahead ex post forecast from January 2018 to December 2018, were examined for their forecast performance. The MAPE of the best fitting ARIMA models for the twelve months ahead ex post forecast, 2013 to 2017, are presented in Table 8 for the forecasting period January 2018-December 2018.

Table 8: The MAPE of Forecasting in each country 2018

Origin Country	Modelling	MAPE	
		Training	Testing
China	$(0,1,0)\ (1,0,0)_{12}$	11.56	19.99
Malaysia	$(2,0,1)\ (1,1,0)_{12}$	7.90	17.34
Korea	$(0,1,0)\ (0,1,0)_{12}$	4.50	9.43
Japan	$(0,1,0)\ (0,1,1)_{12}$	4.50	6.63
Russia	$(0,1,2)\ (0,1,1)_{12}$	8.02	32.02
UK	$(0,0,0)\ (1,1,0)_{12}$	4.32	3.51
US	$(0,1,1)\ (0,1,0)_{12}$	3.34	3.91

The results from Table 8 show that only Korea, Japan, UK, and US provided small values (MAPE <10%). A possible reason is the pattern of time series of these countries are obvious and remain the same every period. On the other hand, the patterns of time series have been changed in 2018 for arrivals from China, Malaysia, and Russia. The example of time series of Chinese and Malaysian tourists represented the pattern in 2018 as shown in Figure 21 and 22.

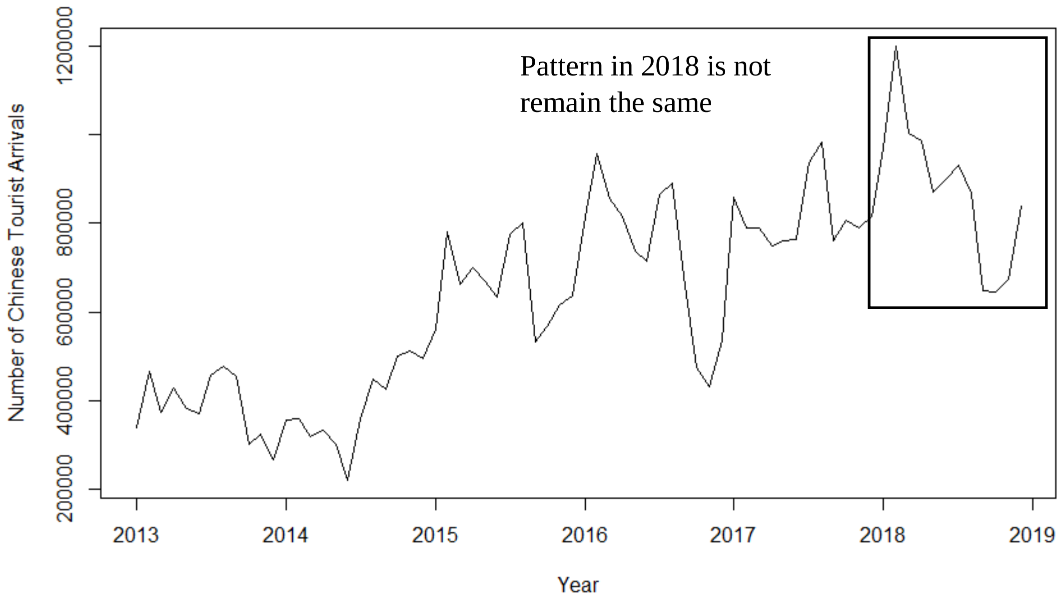

Figure 21: The time series pattern of Chinese tourist arrivals to Thailand

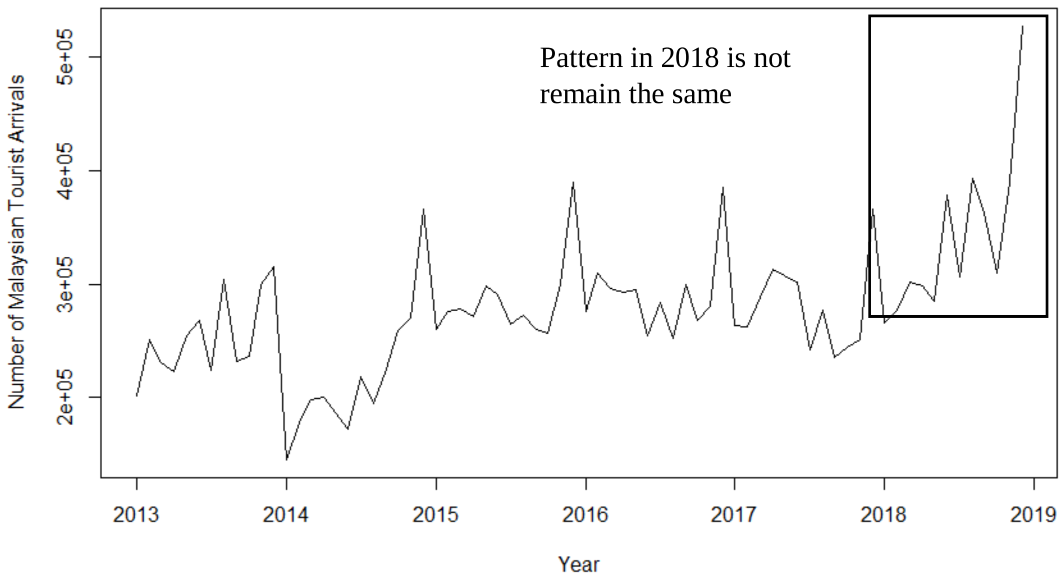

Figure 22: The time series pattern of Malaysian tourist arrivals to Thailand

4.2 Casual Method

4.2.1 Multiple Regression

In this study, multiple regression divides in to two parts. The first part focuses on multiple regression with dummy variables of seasonality while the second part focuses on multiple regression with dummy variables of seasonality and news shock effects. Both parts are compared to find the one with lower MAPE.

a. Multiple Regression with Dummy Variables of Seasonality

The factors influencing international tourist flows to Thailand from major countries were analyzed using the ordinary least squares multiple regression technique. The backward elimination is used to remove factors that are not significant in each country at 95% confidence. The results for all countries are shown in Table 9. Only variables with a significant p-value less than 0.05 are included in the final equations reported.

Table 9 divides into four parts. The top part shows constant term, and macroeconomic variables affecting number of tourist arrivals. The second and third part show yearly seasonal effects and monthly effect of tourists from each source country that 2014, September is used as the base line. And the fourth part shows how much the estimated equation can be explained by the relation.

From Table 9, only Russia has a significant exchange rate term indicating that the exchange rate impact on the number of Russian tourist arrivals to Thailand. As the weak exchange rate between the ruble and other currencies, Russian tourists face to higher costs in terms of the ruble. Relative CPI are significant determinants for only 4 countries—China, Malaysia, Japan, and US. While GDP are important only for Korea, Japan, UK, and US.

Seasonal factors largely reflect the holiday pattern for each source country. For most countries, December is the peak travel season for tourists except China.

The r-squared statistical results indicate that estimated equations explained 73 to 93 percent of the monthly variation in tourist arrivals.

Table 9: Estimated models without news shock effects for top seven countries arrivals to Thailand 2013-2017

Equation	**China**	**Malaysia**	**Korea**	**Japan**	**Russia**	**UK**	**US**
Constant	8771675	997045	-991763	-1828972	-1633	-10690	296915
GDP			38.57	60.41		2.443	3.86
Relative CPI	-8284103	-858913		-368169			-1053321
Exchange rate					95394		
Yearly Effect							
2013		45320	88505	-116250			7222
2014			Base				
2015	88822		48115	217353	-34085		
2016		35594	39129	-42838	-13588	3509	2765
2017							
Seasonal Effect							
Jan		-24764	51763		132342	27615	32660
Feb	103425		23972		98663	23411	20552
Mar	98824				94858	29389	25124
Apr	116713		-12007	-20462	44244	24775	16723
May	124764		-11632		-25478		10975
Jun	130985			-25611		4280	15268
Jul	255326		18308	-15277		20015	16559
Aug	218201		32866	24471		14982	
Sep			Base				
Oct				-21777	51459	12984	14995
Nov			14106		111078	25155	32510
Dec		94756	34000	5235	135263	47457	46127
R-square	89.01	73.82	90.84	92.5	92.16	93.22	97.03
R-square adj	86.41	70.85	88.5	90.58	90.56	91.49	96.10

b. Multiple Regression with Dummy of seasonal effect and News Shock Effects

The news shock effects are put in regression equation to see how these news shock can explain the variation of number of tourist arrivals. The factors influencing international tourist flows to Thailand from major countries are analyzed by using the ordinary least squares multiple regression technique. The backward elimination is used to remove factors that are not significant in each country at 95% confidence. The results for all countries are shown in Table 10.

Only variables with a significant p-value less than 0.05 are included in the final equations reported. The estimated parameters shown in Table 10 indicate that the model performs satisfactorily. The results from Table 10 show Korean, Japanese, and British tourist are sensitive to income, it means that when incomes of these countries increase, the number of these tourist arrivals to Thailand will increase. The impact of price affects to Chinese, Malaysian, Korean, Japanese, and American tourists, it means that when prices of goods and services in Thailand increase, the number of these tourists will decrease. In addition, only Chinese tourists is sensitive to exchange rates. However, all effects from economic factors do not affect Russian tourists.

When considering the seasonality, December is the peak season for all tourist arrivals except Chinese. From the result, Chinese tourists is peak in July and August. In addition, the numbers of tourists from Russia, UK, and US are high in winter (December-January).

Table 10: Estimated models with news shock effects for top seven countries arrivals to Thailand 2013-2017

Equation	**China**	**Malaysia**	**Korea**	**Japan**	**Russia**	**UK**	**US**
Constant	8352557	531115	349319	-1706429	59812	-10690	659538
GDP			16.75	58.41		2.443	
Relative CPI	-8599482	-332068	-725736	-415289			-1438331
Exchange rate	162259						
Yearly Effect							
2013		24937	42082	-111169	26809		
2014			Base				
2015		38101	20943	211522	-39341		
2016	51204	43886	19311	-40400	-22179	3509	2148
2017							

Seasonal Effect							
Jan		-19825	53767		129204	27615	33718
Feb	46139		25210		94140	23411	21649
Mar	54921				90366	29389	26208
Apr	98598		-8305	-21217	41637	24775	18064
May	97246		-6627	-25275			12582
Jun	74035			-26234		4280	16620
Jul	173596		22599	-13182		20015	17561
Aug	169823		34590	25346		14982	
Sep			Base				
Oct			11311	-19304	53595	12984	15030
Nov		31793	23210		103734	25155	32893
Dec	-65296	109355	35504	5193	125807	47457	46178
News Shock Effect							
Anti	-134100	-32805			34314		-3311
Bomb	-111969			-8163			
Flood	-105096	-51622					
King Death	-246611	-29883	-21609	-8114			
R-squared	95.49	83.31	91.99	93.62	92.85	93.22	96.94
R-squared adj	93.95	79.91	89.49	91.64	91.21	91.49	96.08

From Table 9 and10, the R squared adjust in multiple regression with seasonality and news shock effects are concluded in Table 11.

Table 11: Comparing R square of multiple regression with and without dummy variable of news shock effects

Origin Country	R squared adj	
	MR with dummy of seasonal effect	MR with dummy of seasonal effect and new shock effect
China	86.41	93.95
Malaysia	70.85	79.91
Korea	88.5	89.49
Japan	90.58	91.64
Russia	90.56	91.21
UK	91.49	91.49
US	96.10	96.08

From Table 11, It is found that adding the dummy variables of news shock effects to the models will increase R-squared adj values for Chinese, Malaysian, Korean, Japanese, and Russian tourists. However, adding news shock effects does not affect R- squared adj for the number of British and American tourist arrivals. It shows that these countries are not sensitive for the news shock effects.

Therefore, the multiple regression with dummy variables of seasonality and news shock effect is selected to analyze impact of economic factors and seasonality.

The news shock mostly affects the tourist from all major countries except British tourists. It indicates that this tourist is not sensitive to the news shock effects. From Table 11, it shows that all news shock are negative effects for all tourist arrivals except Russian tourists. It is possible that Russian tourists are affected by other news shock that is not in this study.

- *Antigovernment on Thailand*

The results from Table 10 show news shocks that the antigovernment on Thailand had a negative and significant effect on the demand of international tourist arrivals from China, Malaysia, and US. This suggests that this news shock was important in tourism demand in Thailand and implies that the antigovernment dramatically discouraged Thai tourism growth because tourists from these countries were shocked by the unexpected terrible damage seen in the media. Moreover, China, which is the highest-spending country, declined rapidly because tourists felt unsafe for tourism. On the other hand, Russia has a positive significant effect, it indicating that this news shock was not directly affected on Russia tourists because this news occurred in the high season.

- *Bomb in Bangkok*

According to these results, the negative coefficient for news shocks on Bomb in Bangkok variable had a negative impact on international tourist arrivals to Thailand from China and Japan. It may result from the announcement of travel advice to Thailand and flight cancelation (August to September 2014) from origin countries such as China, Japan, Hongkong, and Singapore. Tourist has a shock on this news; they are responsive to this news shock. Thus, the news shock on Thailand Bombing should be given due attention by the Thai tourism industry. The result is a slight impact of this factor on the Thai tourism industry.

- *Flood*

These results show that news shocks of the flood disaster on Thailand had a negative and significant effect on the demand of international tourist arrivals from China. This news shock occurred in peak season for these tourists who flock to the kingdom's island resorts. But the deluge disrupted beach holidays in several traveler hot spots, including the popular islands of Samui and Phagan.

- *Thai king rama9 passed away*

These results show that news shocks of Thai king rama9 passing away on Thailand had a negative and significant effect on the demand of international tourist arrivals from China, Malaysia, Korea, and Japan. The effect of the Thai king rama9 passing away can mostly be seen in the cities and islands like Koh Phi Phi, Koh Lanta, Koh Tao etc. Any public gathering is not being recognized except the mourning of the Thai people for the king and gathering in a stadium. So, currently, no bars and pubs are allowing parties in open. That means loud music, open music, live concerts etc. are not allowed there at the moment.

After approaching the training model, residual analysis diagnostic was checked. The example residuals analysis of China shown in Figure 23 and 24.

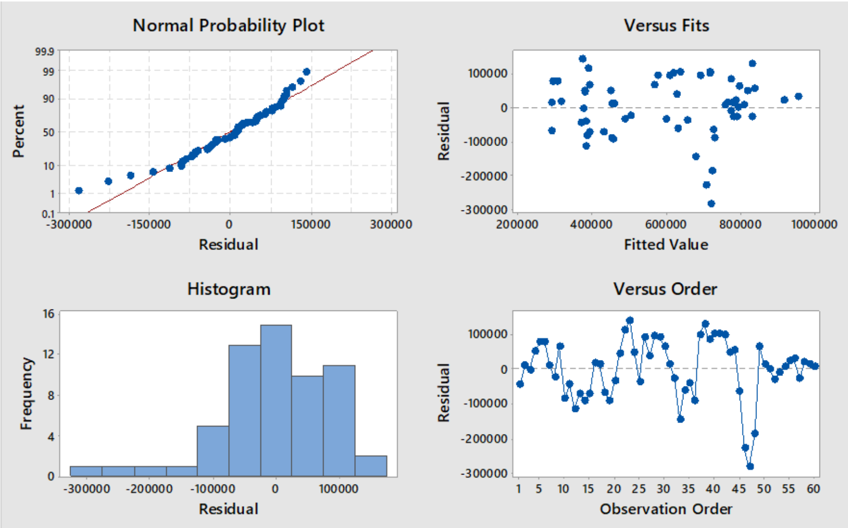

Figure 23: Residual Plot for China Tourists using multiple regression without dummy variable of news shock effects.

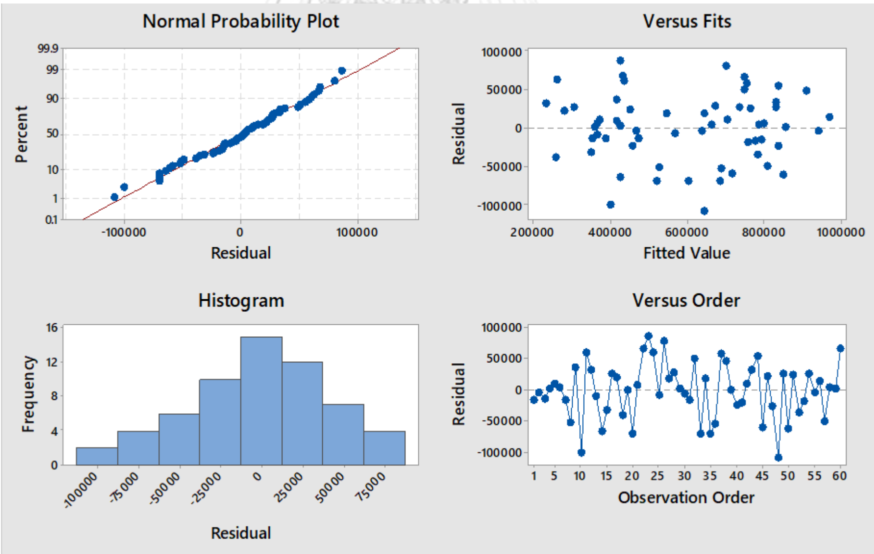

Figure 24: Residual Plot for China Tourists using multiple regression with dummy variable of news shock effects

The comparison on Figure 23 and 24, multiple regression with dummy variable of news shock effects is more adequate because the error has a mean of zero, a constant covariance, normality, and be uncorrelated with each other over time. The residual plot of each country is in appendix 3.

Comparison the multiple regression models using MAPE

The multiple regression models from two parts are used to forecast the number of tourist arrivals to Thailand in each country for 2018. The MAPE values of each country are calculated and the comparing multiple regression with dummy variables of seasonality and news shock effects shown in Table 12.

Table 12: Comparing the multiple regression with dummy of seasonality and news shocks effect using MAPE

Origin Country	MAPE			
	MR with dummy of seasonal effect		MR with dummy of seasonal effect and new shock effect	
	Training	Testing	Training	Testing
China	12.07	17.96	6.67	14.61
Malaysia	8.69	16.93	6.39	24.60
Korea	4.81	9.10	4.14	5.14
Japan	3.52	6.86	3.44	8.22
Russia	17.26	13.07	16.07	16.07
UK	3.10	6.79	3.10	6.79
US	3.09	6.87	3.05	8.62

The results from Table 12 show that MAPE of both regression models for testing periods for China, Malaysia, and Russia are still higher than 10 percent, while others are less. It is supposed that relationships between dependent variable and independent variables are not simply linear. More advanced casual methods should be applied next.

4.2.2 Artificial Neural Network

Artificial Neural Network model is created in R studio Program. The input variables of ANN in this study are similar to those described in Multiple Regression model. Data for output layer in the testing periods are monthly arrivals in 2018. The parameters for ANN are largely focused upon the number of nodes in each hidden layer and the choice of the number of layers. In this study I find the most suitable number of nodes from 1 to 70 and the number of layers of one or two layers. The parameter results for each main country are shown in Table 13.

A summary of selected the number of hidden node and MAPEs for China, Malaysia, Korea, Japan, Russia, UK, and US shown in Table 14.

Table 13: The optimal number of hidden layer nodes and MAPEs value for each country

Country	One Layer		Two Layers	
	No. of nodes	MAPE Testing	No. of nodes	MAPE Testing
China	45	13.17	(15,63)	4.70
Malaysia	39	9.41	(44,31)	8.60
Korea	23	5.60	(55,12)	4.64
Japan	42	20.03	(56,52)	12.77
Russia	6	13.97	(19,50)	9.57
UK	2	3.19	(2,58)	2.55
US	7	4.45	(22,41)	3.98

Table 14: Summarized results of selected hidden node and MAPE for each country

Origin Country	Number of Hidden Node (i,j)	MAPE	
		Training	Testing
China	(15,63)	2.01	4.70
Malaysia	(44,31)	10.04	8.60
Korea	(55,12)	8.10	4.64
Japan	(56,52)	4.89	12.77
Russia	(19,50)	7.20	9.57
UK	(2,58)	4.31	2.55
US	(22,41)	6.8	3.98

i = Number of hidden nodes in first layer

j = Number of hidden nodes in second layer

The results from Table 13 show ANN models with two layers are more accurate than one layer due to lower MAPE. Therefore, the ANN models with two layers are then selected to forecast international tourist arrivals to Thailand from each major country. The results in Table 14 show that ANNs model performs very well for all countries except Japan. When MAPE in training period of Japan is low but MAPE in testing period is high, it may result from overfitting issue.

4.4 Models Comparison and Selection

In this section, a comparison among four proposed models, namely, Seasonal ARIMA, Holt-Winters, multiple regression and ANN is discussed. China is used as an example and forecasting results in the testing period is shown in Figure 25. It can be seen that predicted values by ANN are relatively close to the actual values. The low MAPE indicated that ANN model provided the most accurate forecasting value, followed by Holt-Winter, multiple regression and Seasonal ARIMA model.

To understand the performance of all methods, the randomized complete block designs (RCBD) and Tukey method are tested at 95% confidence interval. Only absolute percentage forecasting error values (APE) of different methods are of interest while month factors are blocked. The absolute percentage error of China shown in Table 15. And RCBD using Tukey test of China shown in Table 16 and 17.

Table 15: Absolute Forecasting Error value of China tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	19.21	10.72	3.57	3.49	1.54
Feb	16.49	13.81	19.69	19.20	12.63
Mar	8.84	5.16	4.02	11.44	5.70
Apr	10.56	4.00	4.78	15.11	1.33
May	4.48	2.46	9.04	8.79	9.22
Jun	6.33	2.68	5.25	24.56	1.38
Jul	7.00	6.77	12.97	12.18	4.19
Aug	19.29	19.23	24.14	3.72	6.44
Sep	26.59	20.81	46.23	15.58	2.13
Oct	26.42	21.16	50.61	14.74	0.60
Nov	30.18	24.88	42.86	9.94	6.17
Dec	40.21	43.65	16.73	6.80	5.17

Table 16: Grouping Information Using the Tukey Method at 95% Confidence

Method	N	Mean	Grouping	
SARIMA	12	19.99	A	
MR+S	12	17.96	A	
MR+S+E	12	14.61	A	B
HW	12	12.13	A	B
ANN (15,63)	12	4.71		B

Table 17: Tukey Simultaneous Tests for Differences of Means

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	7.42	4.57	(-5.46, 20.30)	1.63	0.488
MR+S - ANN	13.26	4.57	(0.38, 26.14)	2.90	0.041
MR+S+E - ANN	9.90	4.57	(-2.98, 22.78)	2.17	0.207
SARIMA - ANN	15.28	4.57	(2.40, 28.16)	3.35	0.012
MR+S - HW	5.84	4.57	(-7.04, 18.72)	1.28	0.705
MR+S+E - HW	2.48	4.57	(-10.40, 15.36)	0.54	0.982
SARIMA - HW	7.86	4.57	(-5.02, 20.74)	1.72	0.429
MR+S+E - MR+S	-3.36	4.57	(-16.24, 9.52)	-0.74	0.947
SARIMA - MR+S	2.02	4.57	(-10.86, 14.90)	0.44	0.992
SARIMA - MR+S+E	5.38	4.57	(-7.50, 18.26)	1.18	0.764

From the results in Table 16 and 17 show the comparison of the paired of method, the paired of method which provided p-value below 0.05, are MR+S vs ANN and SARIMA vs ANN. It indicated that the methods are different at 95% confidence intervals while others paired of method are not different at 95% confidence intervals. The absolute percentage error and Tukey tests of all countries are present in appendix 4.

Therefore, to select the performance of method, the MAPE value and Tukey test are considered. The recommended method for each country is the on providing the lowest MAPE. The forecasting performance accuracy comparisons of all methods shown in Table 18.

Table 18 Forecasting performance accuracy comparisons of each model

Country	Method (MAPE Testing)				
	MR+S	MR+S+E	SARIMA	HW	ANN
China	17.97	14.61*	19.99	12.13*	4.70**
Malaysia	16.93*	24.60	17.34*	39.74	8.60**
Korea	9.10*	5.15*	9.43*	19.60	4.64**
Japan	6.86*	8.22*	6.63**	26.25	12.77
Russia	13.07*	16.07*	32.02	13.43*	9.57**
UK	6.79*	6.79*	3.51*	16.23	2.55**
US	6.87*	8.66*	3.91*	2.88**	3.98*

** the method that provided the lowest MAPE value.
*the methods that are not significantly different from the method that provided the lowest
MR+S: Multiple Regression with Monthly Dummy Variables of Seasonal Effects
MR+S+E: Multiple Regression with Dummy Variables of Seasonal Effect and News shock Effects
SARIMA: Seasonal Autoregressive integrated moving average
HW: Holt – Winters Exponential Smoothing Method
ANN: Artificial Neural Network

The results from Table 18 show ANN model provides the lowest MAPE for China, Malaysia, Korea, Russia, and UK. However, there are many models that are comparable to ANN denoted with * in the table. For example, the ANN model did not appear to be significantly better than Holt-Winters and the multiple regression with dummy variables of seasonality and news shock effects. The SARIMA provided the lowest MAPE for Japan, a possible reason is the pattern of time series for Japan has not been changed in 2018. Although ANN is most accurate for Russia but the result of Turkey test shows all methods are not significantly different except SARIMA model. Holt-Winters model is more accurate for American tourist arrivals. However, this model is not significantly different when compared with other models.

Therefore, ANN should adequate for time series that are not obvious and have changed pattern in forecast period while time series model like Holt-Winters and SARIMA are accurate and adequate for time series that remain the same in each year. However, time series pattern of Russian tourist arrivals show obvious pattern and it remained the same in each year as shown in Figure 25.

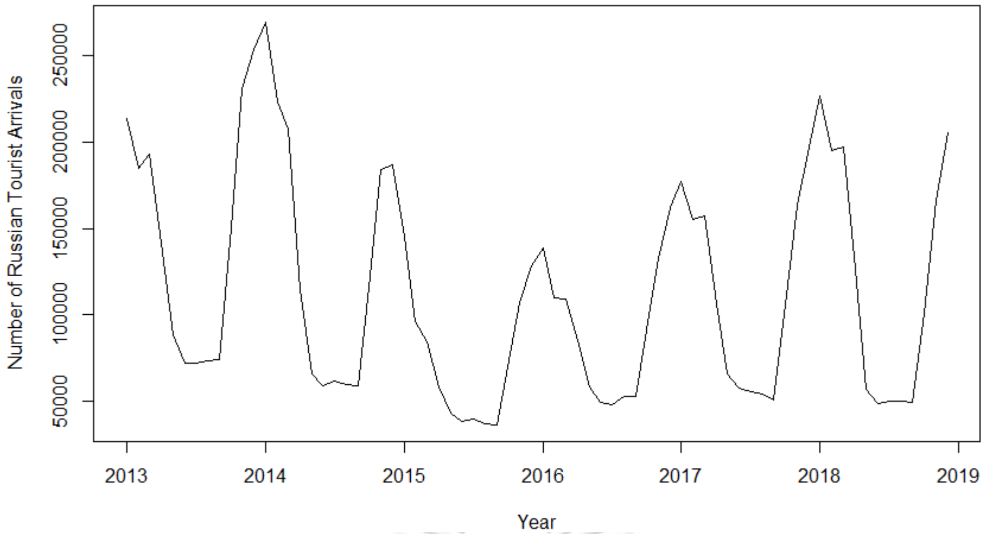

Figure 25: Time series pattern of Russian tourist arrivals to Thailand

From time series pattern in Figure 25, it shows that the seasonality pattern exists. However, time series methods like SARIMA provided relatively high MAPE (32.02%), is means that there are other factors affecting these data. Therefore, although the time series pattern can be used to screen the forecasting methods, the screening methods should analyze the factors that will affect to those countries.

The forecasting results for tourist arrivals from all models are compared with actual Chinese Tourist as shown in Figure 25-31.

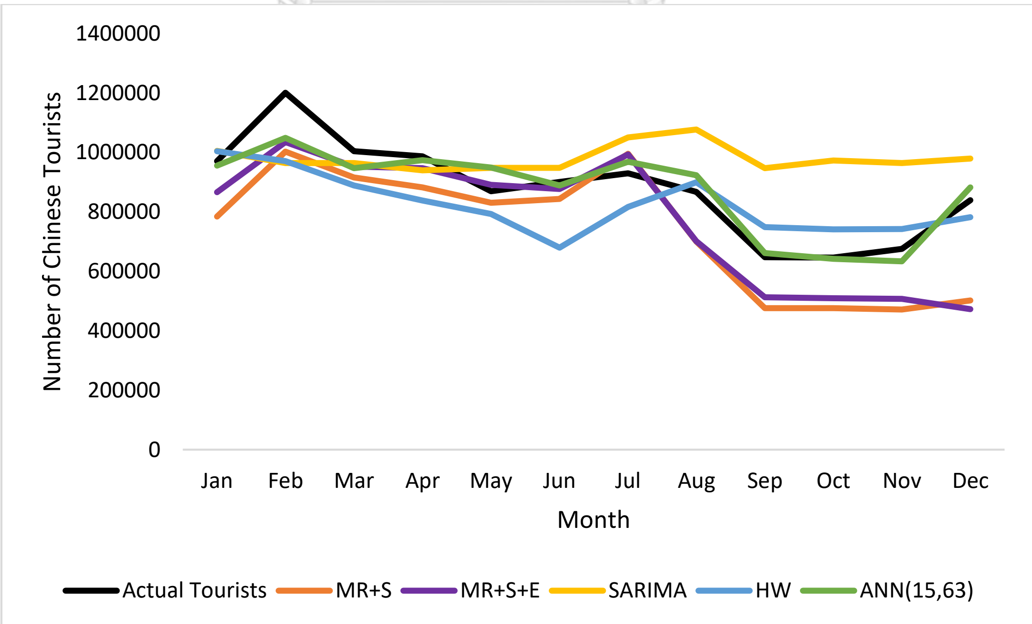

Figure 26: Graphical presentation of different forecasting methods of China Tourists in 2018

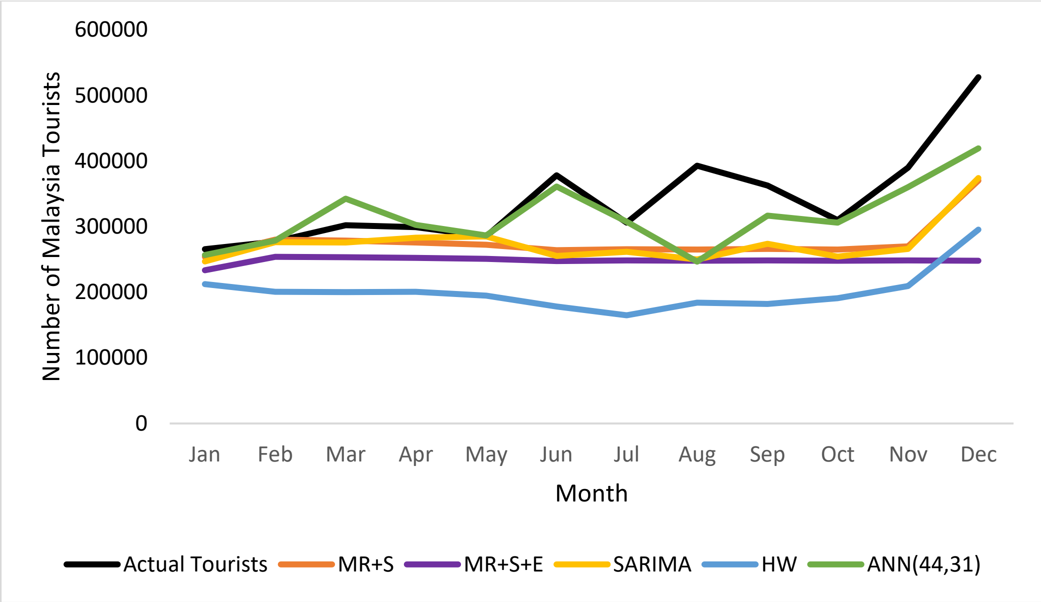

Figure 27: Graphical presentation of different forecasting methods of Malaysia Tourists in 2018

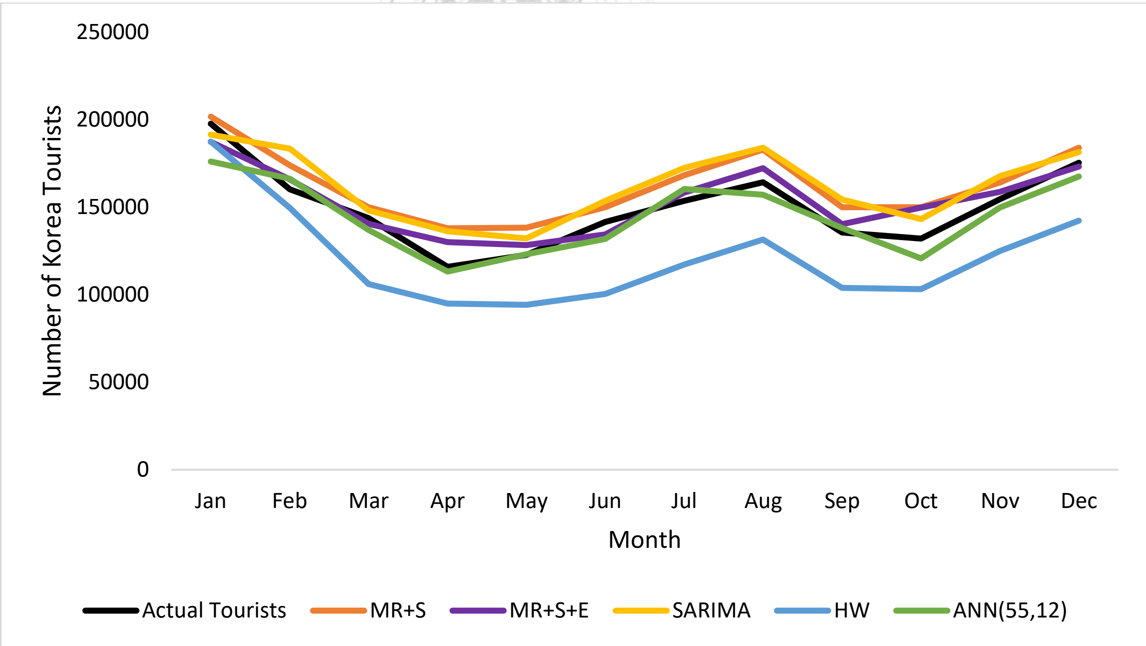

Figure 28: Graphical presentation of different forecasting methods of Korea Tourists in 2018

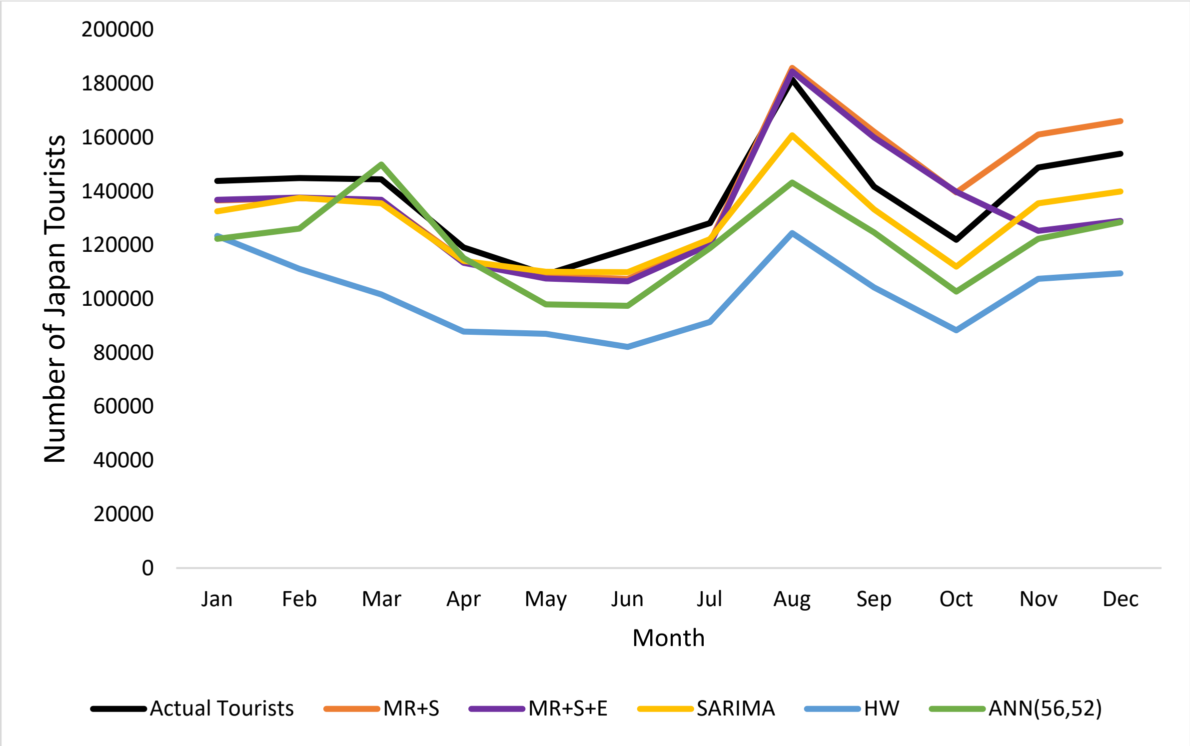

Figure 29: Graphical presentation of different forecasting methods of Japan Tourists in 2018

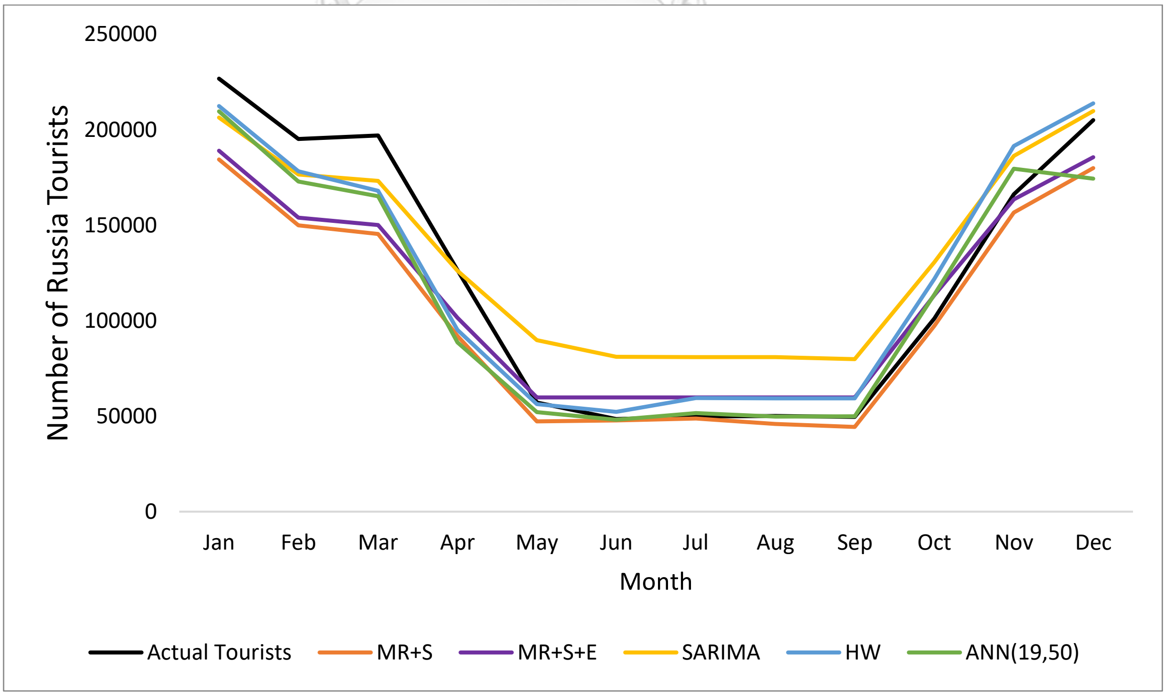

Figure 30: Graphical presentation of different forecasting methods of Russia Tourists in 2018

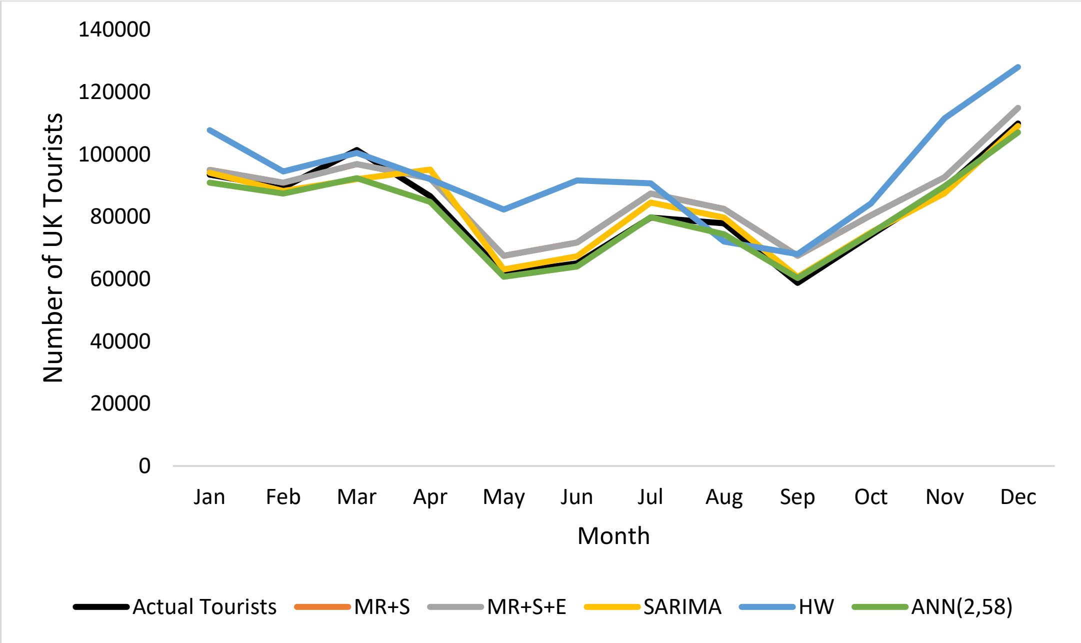

Figure 31: Graphical presentation of different forecasting methods of UK Tourists in 2018

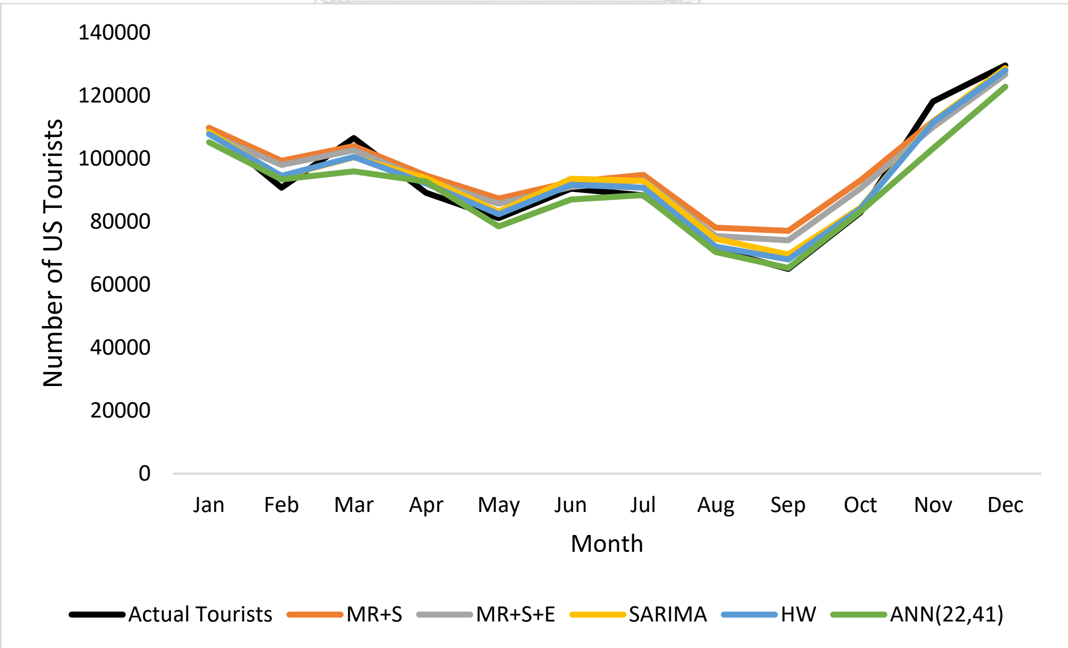

Figure 32: Graphical presentation of different forecasting methods of US Tourists in 2018

Chapter 5 Conclusion and Future Work

5.1 Conclusions

This thesis compared forecasting model and examined the factors that influence the demand for international tourism in Thailand. Empirical results in this study show that the relative performance of a particular forecasting model varies for each of the tourist arrivals. This thesis was divided into two parts.

The first part was to identify and estimate the important of economic factors and dummy variables of seasonal and news shock affecting tourism demand for Thailand. The Multiple Regression model (MR) was used to study the tourism demand in the short-run. The result found that putting dummy variables of news shock effects would increase the adjusted R-square for China, Malaysia, Korea, and Japan because these countries were sensitive for unexpected news. However, news shock effects were not important for Russia, UK, and US. In terms of tourism price, the results showed that tourism demand from China, Malaysia, Japan, and US were very sensitive to price variations while Korea, Japan, and UK were sensitive to GDP. In addition, only Russia had a significant exchange rate term, indicating that the exchange rate impacted the number of Russian tourist arrivals to Thailand. The MAPE from regression model for testing periods for China, Malaysia, and Russia are still higher than 10 percent, while others are less. It is supposed that relationships between dependent variable and independent variables are not simply linear for those counties.

The second part was to compare time series methods namely, Holt and Winters exponential smoothing (HW) and SARIMA with explanatory methods namely, multiple regression (MR) and Artificial Neural Network (ANN). From empirical results, it was found that There was no single model that is 'best' for all the origin-Thailand pairs. The more sophisticated Neural Network models are more accurate for China, Malaysia, Korea, Russia, and UK tourist arrivals to Thailand in all the short-term forecasting horizons used in this study. On the other hand, SARIMA outperformed other methods with the lowest MAPE value for Japan and Holt and winters provided the lowest MAPE for US. To understand the performance of all methods, the randomized complete block designs (RCBD) and Tukey method were tested at 95% confident interval. The overall results support the conclusion that Neural Network models are useful for forecasting monthly tourist arrivals to Thailand. In addition, the appropriate selection of the associated functions and number of hidden nodes is important and will improve the forecasting accuracy of a Neural Network.

Even though ANN was considered to be the over-all most accurate model, there were many models that are comparable to ANN for some countries. So, my research identified not only the best model for each country but also the comparable models for optional choices.

5.2 Practical Contributions

This study makes significant and practical contributions to tourism demand forecasting in the following aspects.

(1) This thesis focused on finding the most suitable model to forecast monthly tourism demand from main countries to Thailand, while most previous work for Thai tourism had focused on annual demand.

(2) This thesis compared performance of time series and causal models and recommend which model is the best for each particular country. Insights are also provided for future model selection for this industry. It was shown that the neural techniques are capable of producing high levels of forecasting accuracy (with MAPE $\leq$ 10%) and outperform simpler forms of explanatory model.

5.3 Limitation for this research

This research focused on forecasting models that used to forecast international tourist arrivals to Thailand. Building the model in Jan 2013 to Dec 2017 to forecast 2018. A limitation of this is if the planner wants to plan for next year around December of this year, there will be no data in December. The effect of deleting December data of the current year to forecast all 12 months of this year has not yet been analyzed in this study. Also, in this study, only linear relationship was considered in screening factors. If there is non-linear relationship in some variables, they may be selected to be kept in the model.

5.4 Recommendations for Future Research

1) The idea can be extended to cover all other countries that are important to Thai tourism
2) Future researchers could consider expanding tourism demand forecasting with the use of other deep learning techniques that were not yet explored in this thesis, e.g. hybrid ANN-Time series, etc.
3) Future work could consider nonlinear relationship between factors and number of tourists from each country in multiple regression part to see if there are other factors should be included in the non-linear model.

Appendix 1

Table 19: Ljung-Box tests of each country

Country	Ljung-Box Q		
	Statistics	DF	Sig.
China	30.896	19	0.0514
Malaysia	20.35	19	0.3738
Korea	24.143	19	0.1907
Japan	16.32	19	0.6359
Russia	26.124	19	0.1268
UK	16.358	19	0.6332
US	31.465	19	0.0359

Appendix 2

Diagnosis Checking of SARIMA Model

➢ **China**

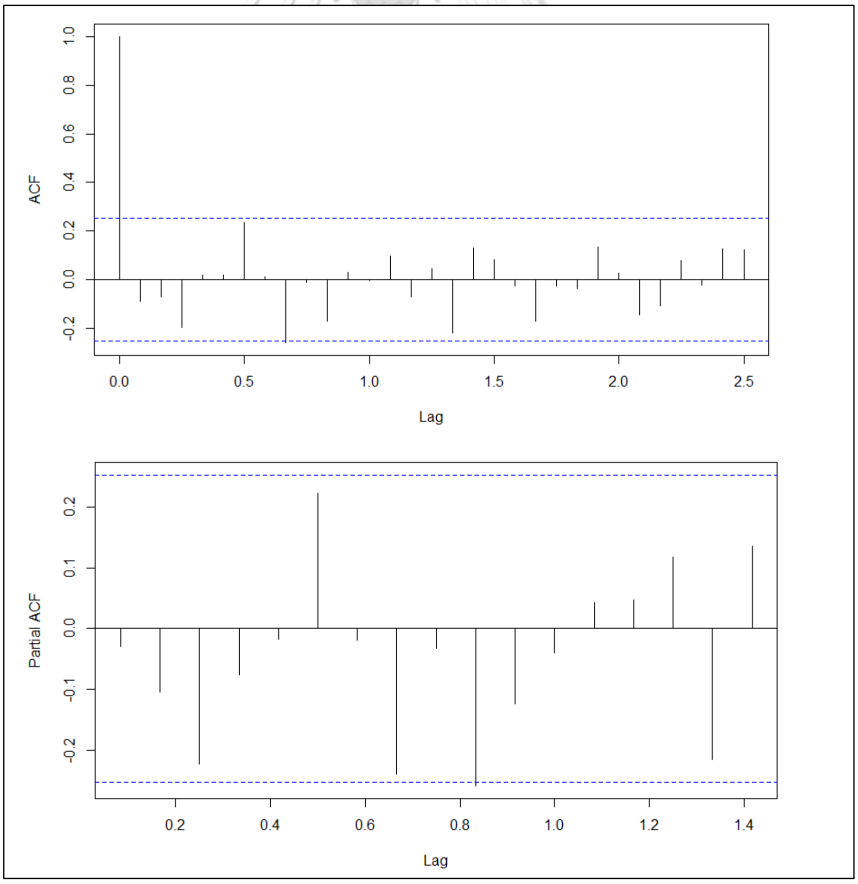

Figure 33: ACF and PACF of residuals for China

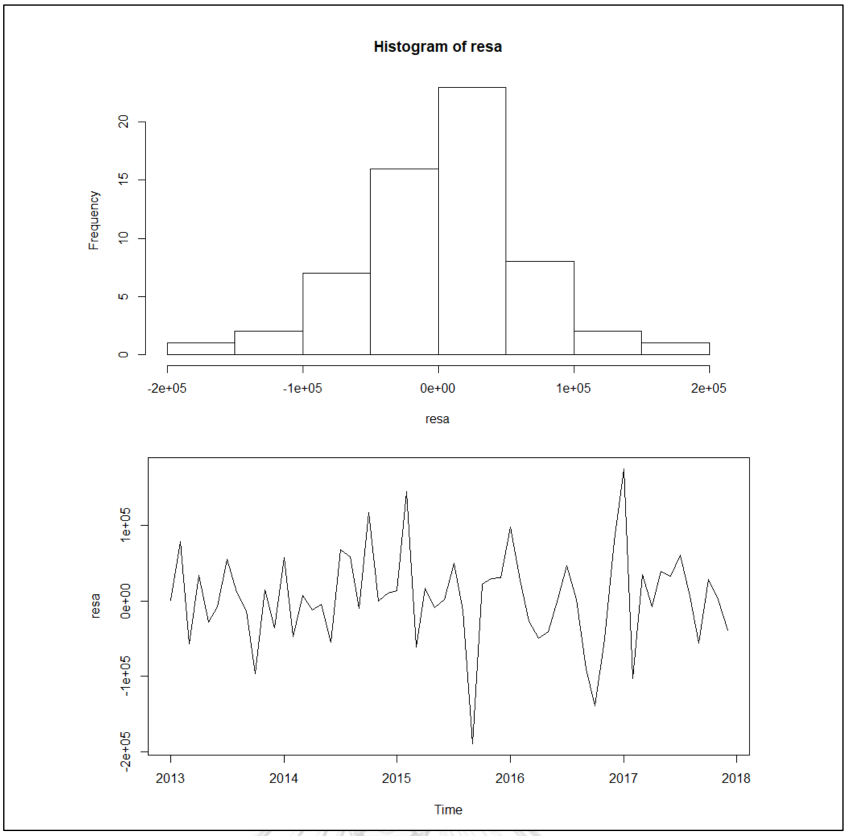

Figure 34: Residual normality and residual plot over time of China

- **Malaysia**

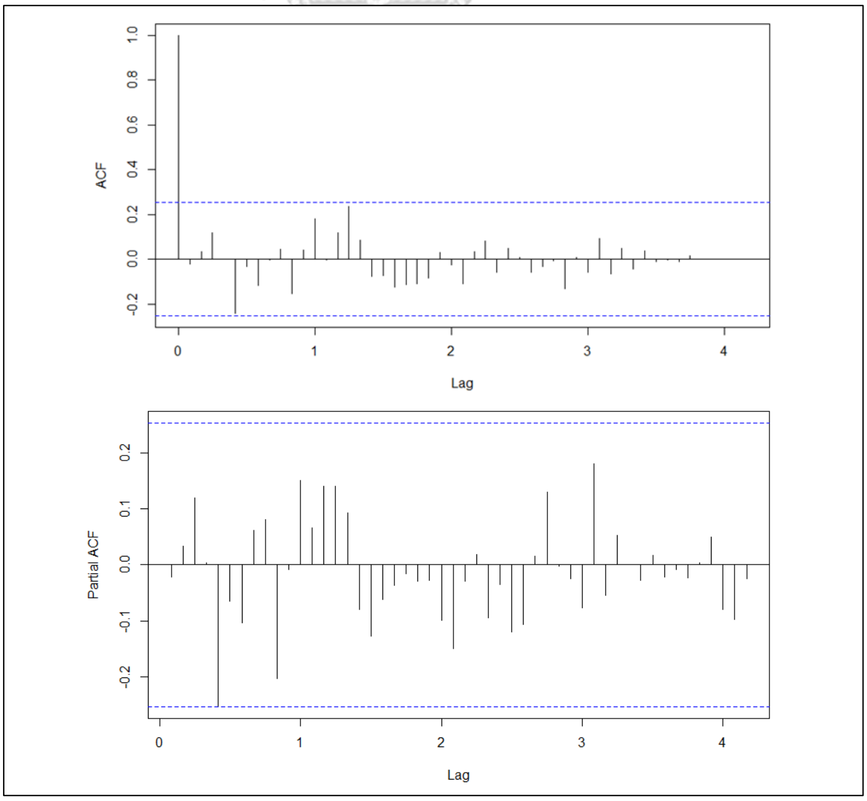

Figure 35: ACF and PACF of residuals for Malaysia

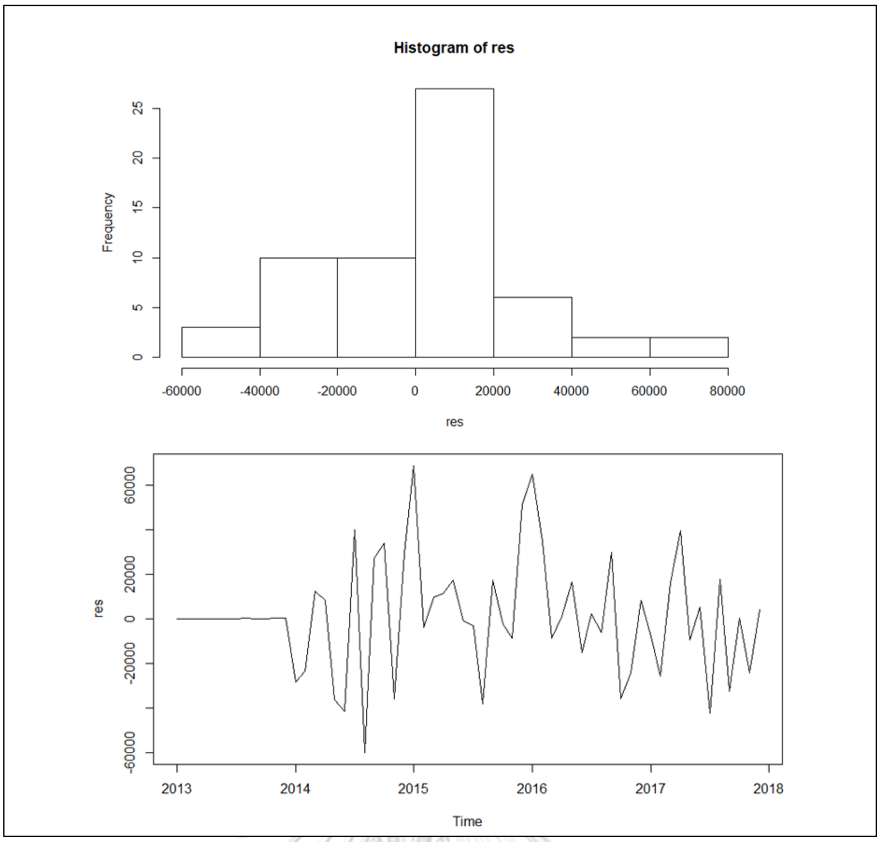

Figure 36: Residual normality and residual plot over time of Malaysia

- **Korea**

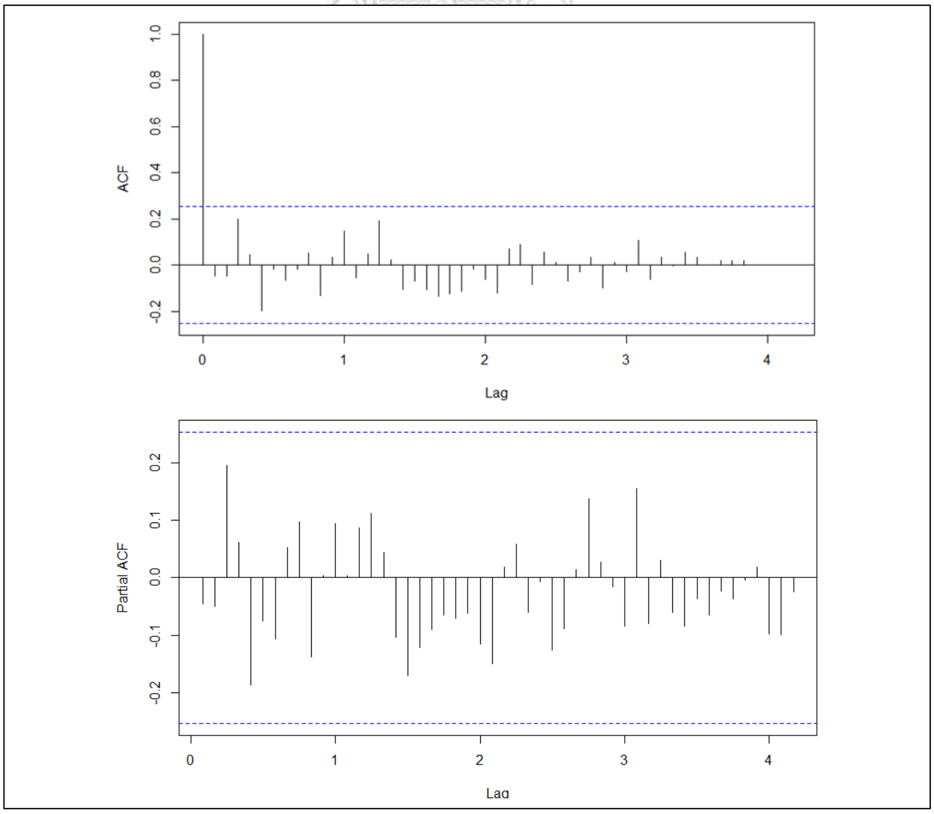

Figure 37: ACF and PACF of residuals for Korea

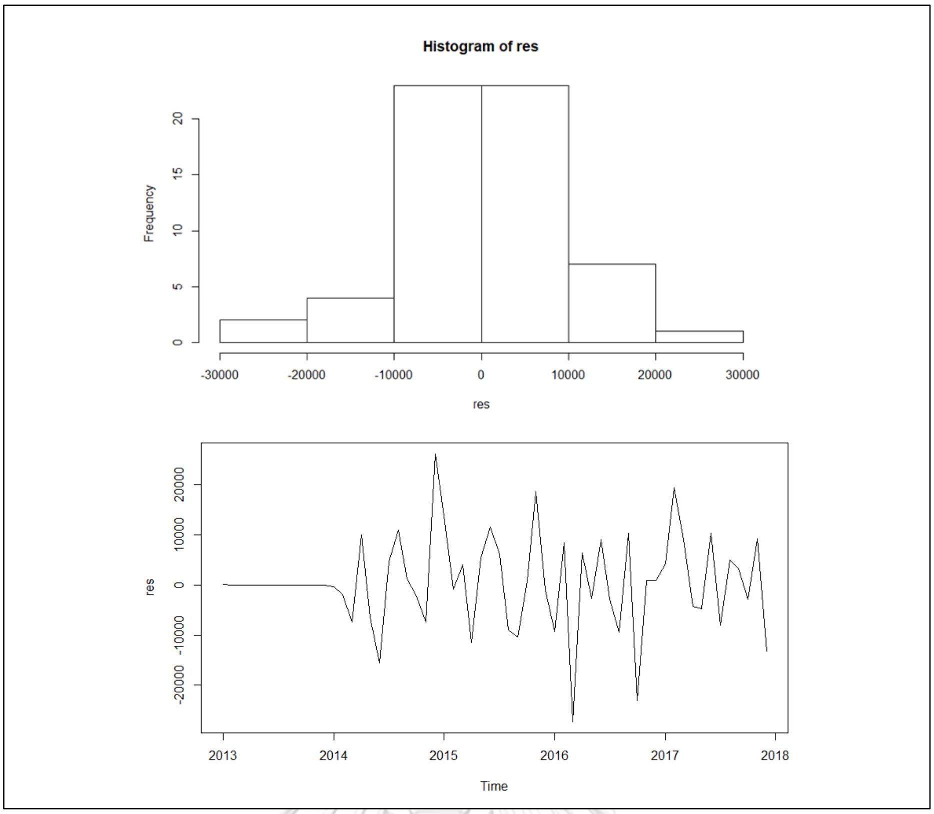

Figure 38: Residual normality and residual plot over time of Korea

- **Japan**

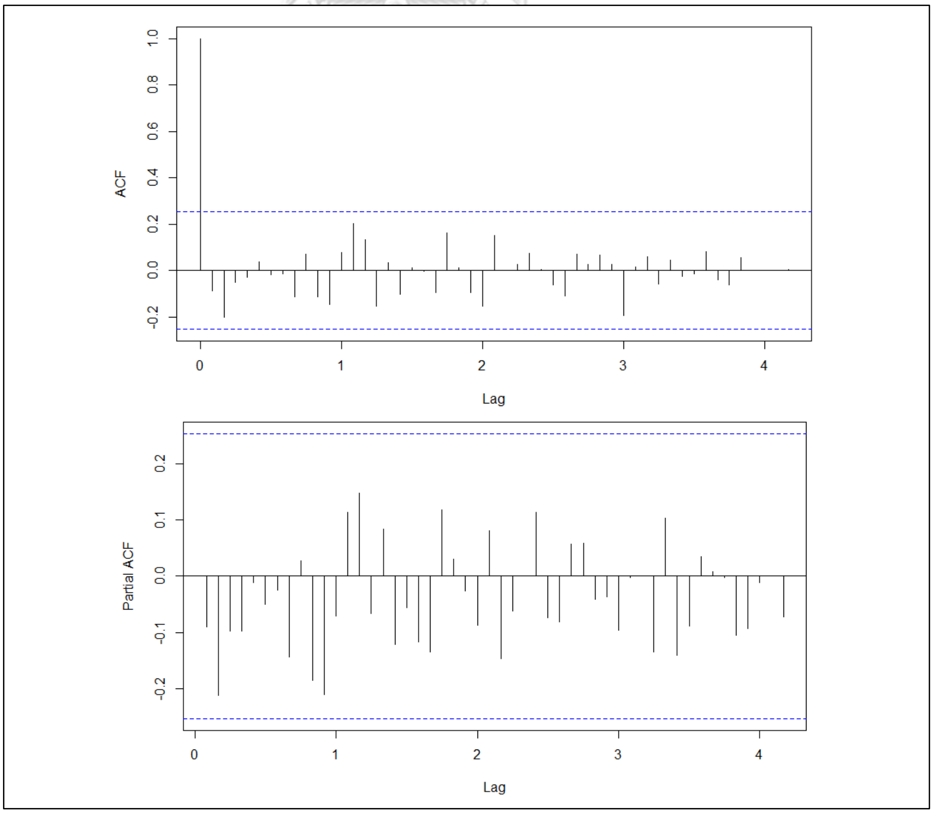

Figure 40: ACF and PACF of residuals for Japan

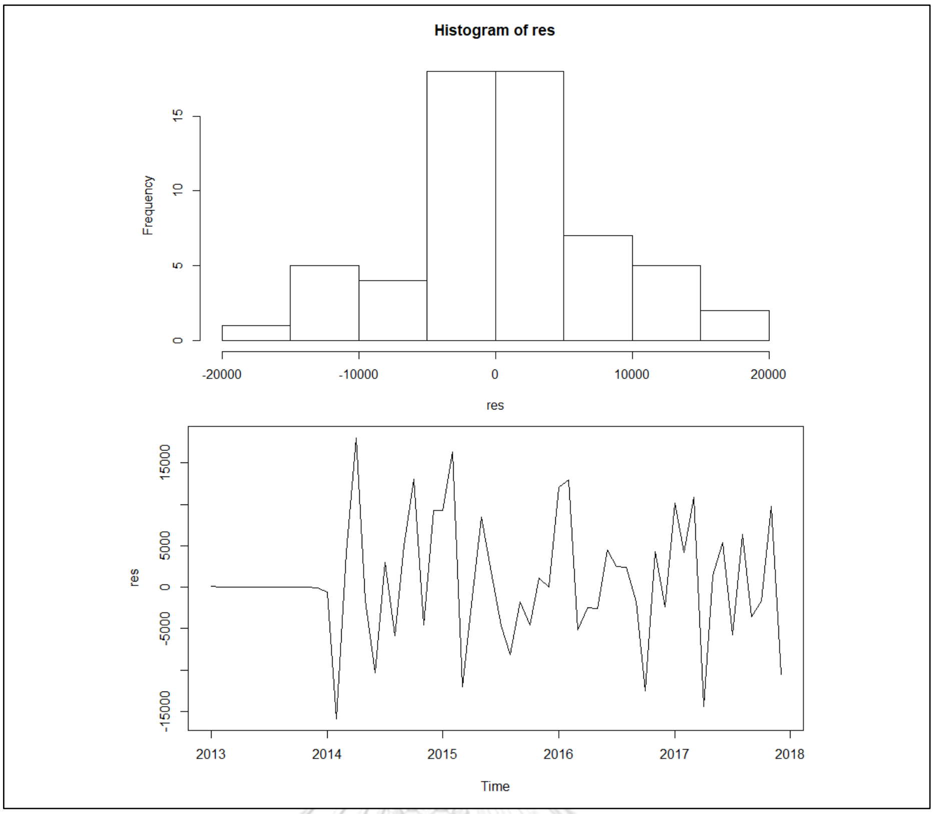

Figure 41: Residual normality and residual plot over time of Japan

- **Russia**

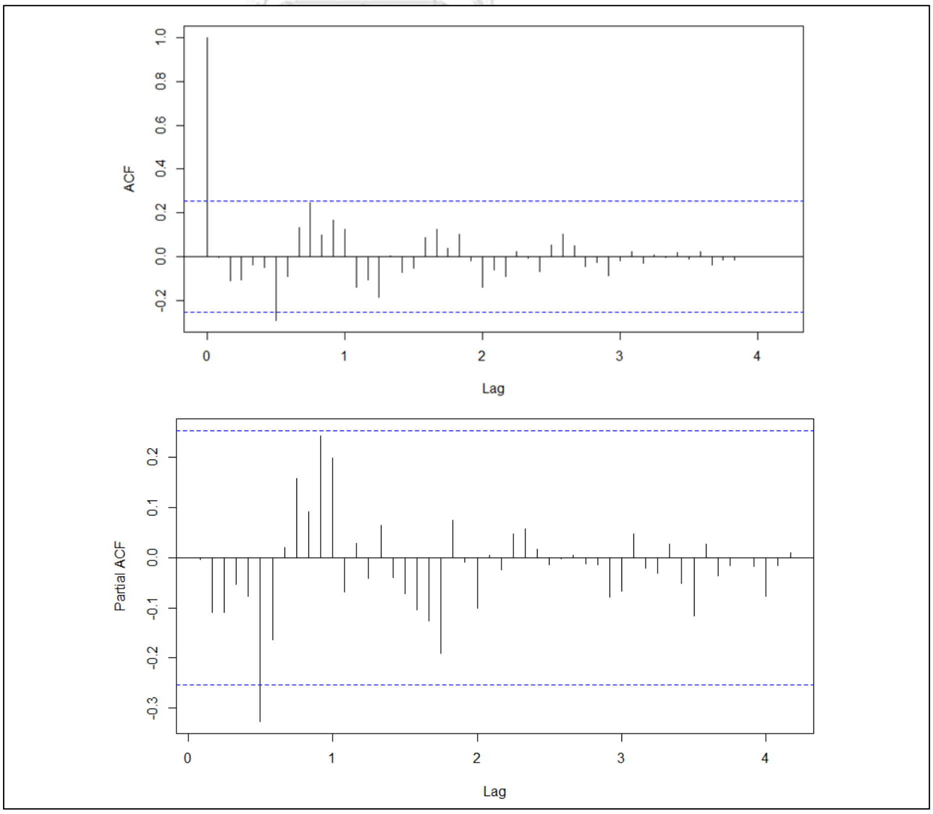

Figure 42: ACF and PACF of residuals for Russia

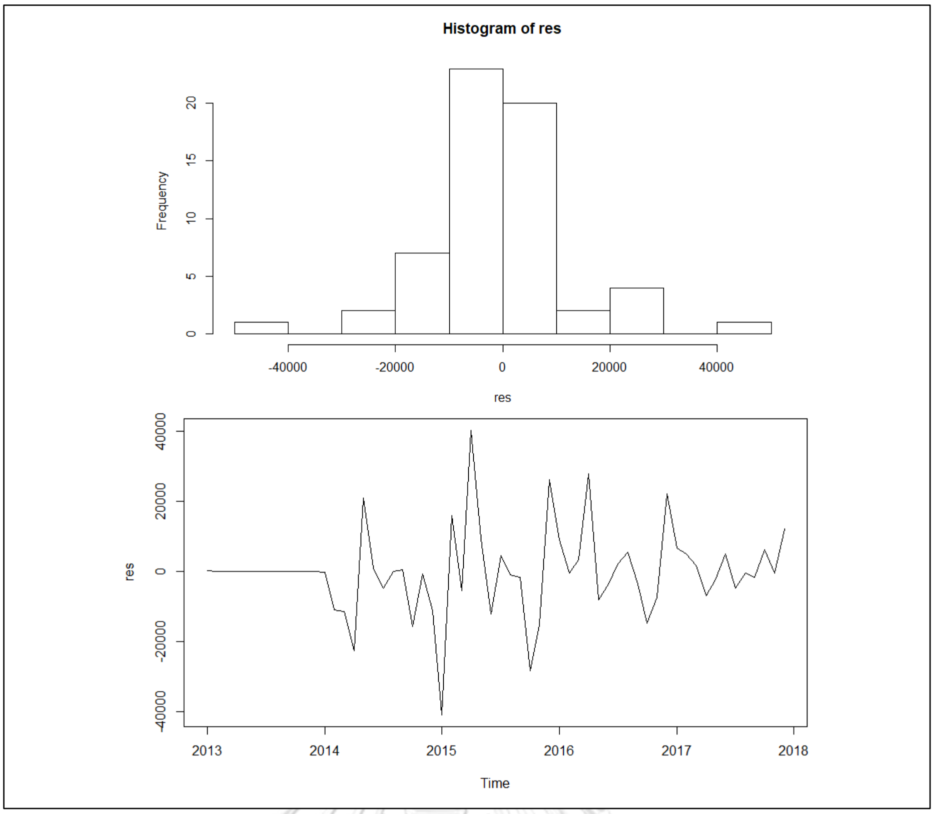

Figure 43: Residual normality and residual plot over time of Russia

- **UK**

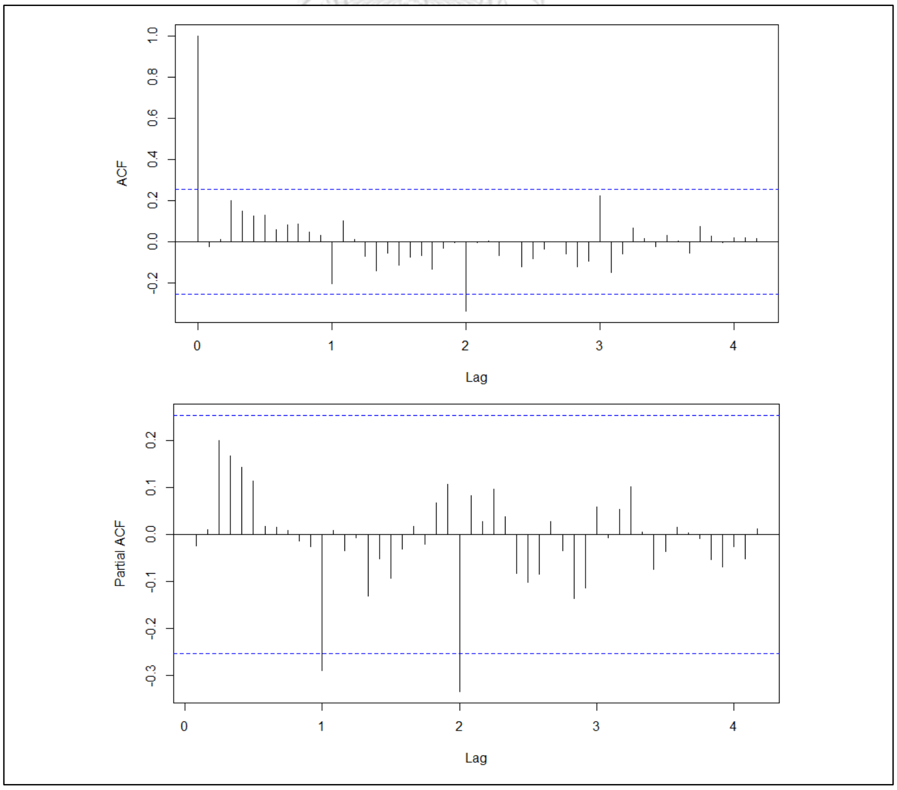

Figure 44: ACF and PACF of residuals for UK

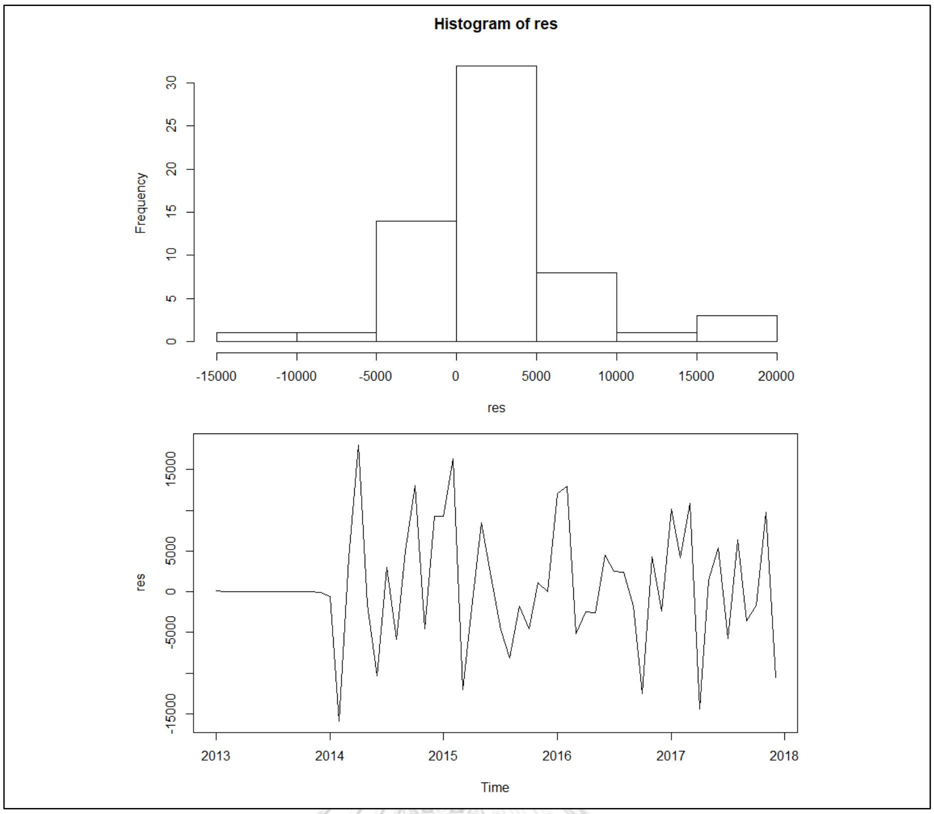

Figure 45: Residual normality and residual plot over time of UK

- **US**

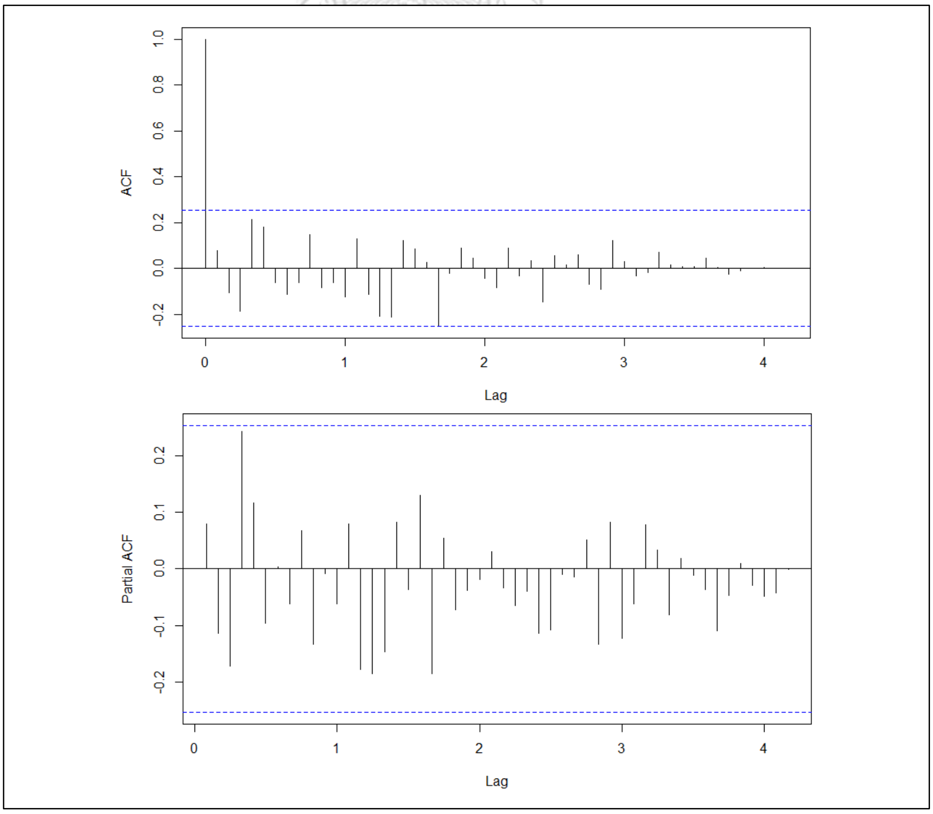

Figure 47: ACF and PACF of residuals for US

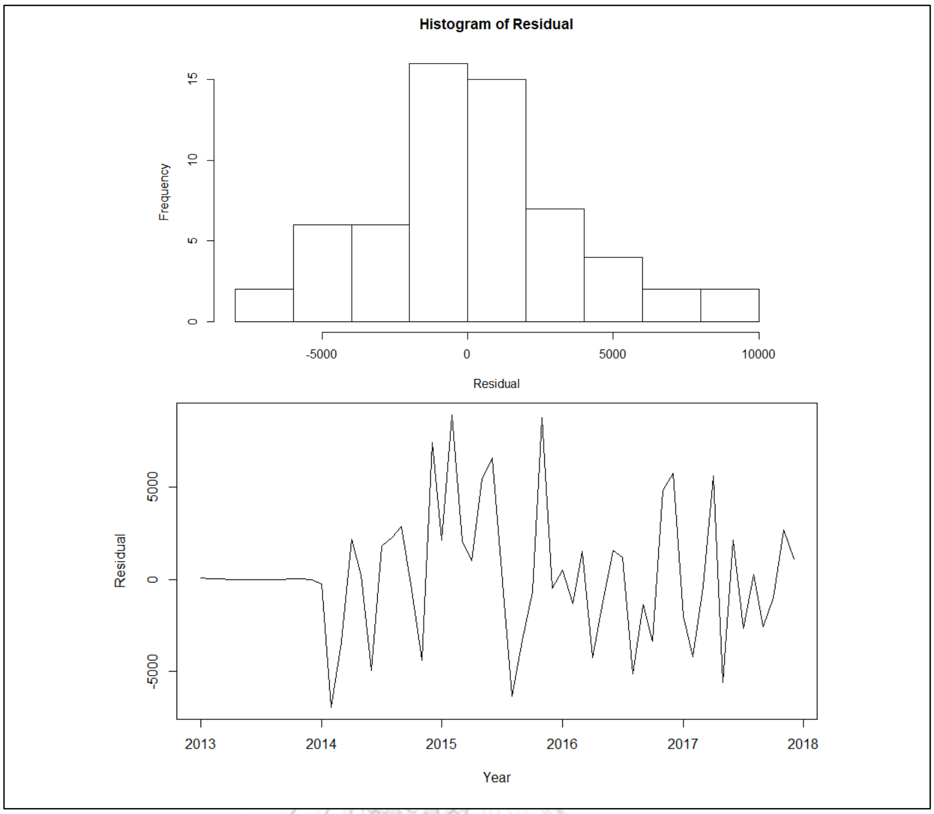

Figure 48: Residual normality and residual plot over time of US

Appendix 3

Diagnosis Checking of SARIMA Model

➢ **China**

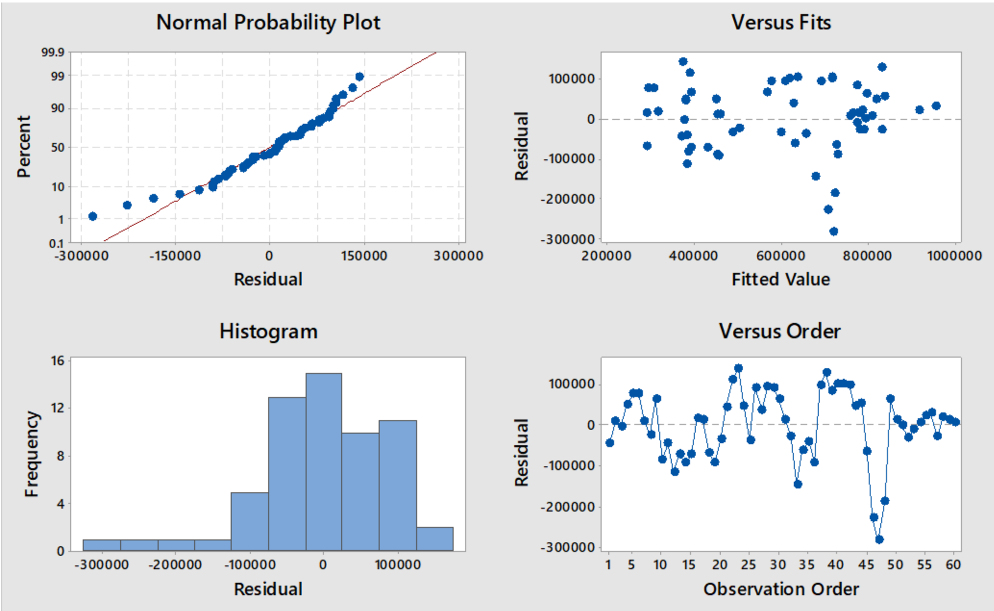

Figure 49: Residual plot for China tourists with dummy variable of seasonality

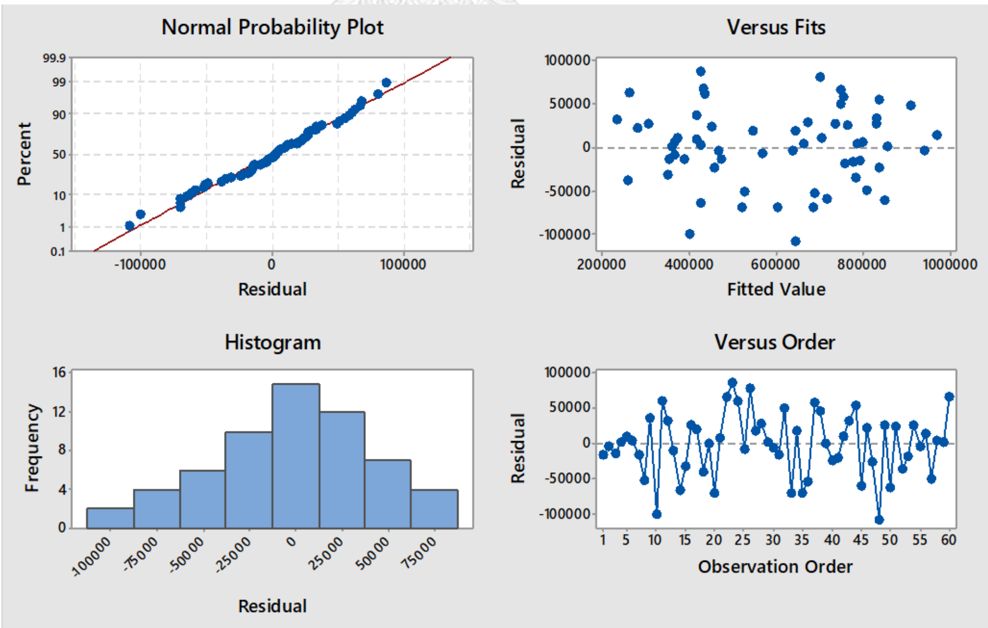

Figure 50: Residual plot for China tourists with dummy variable of seasonal and news shock effects

➢ Malaysia

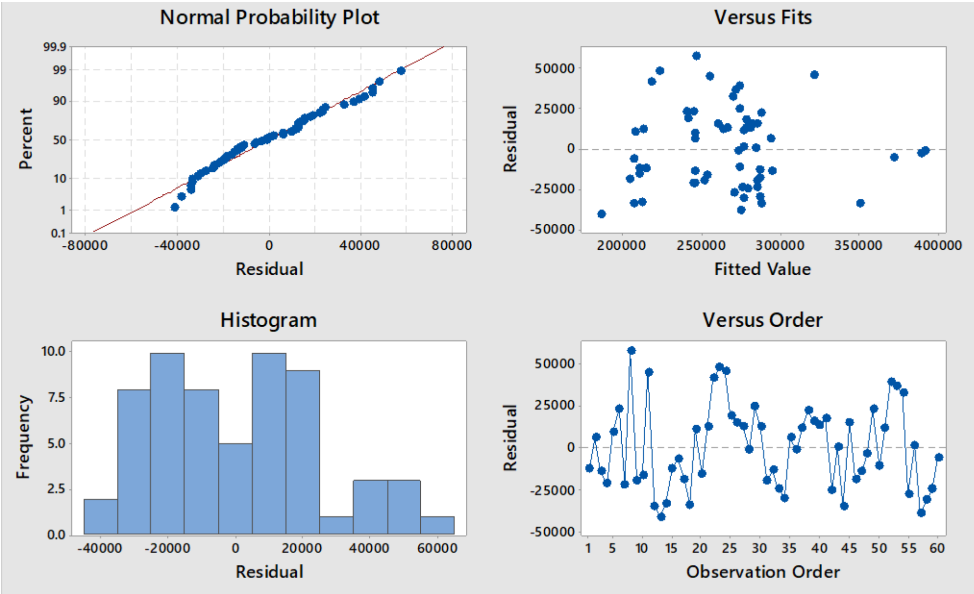

Figure 51: Residual plot for Malaysian tourists with dummy variable of seasonality

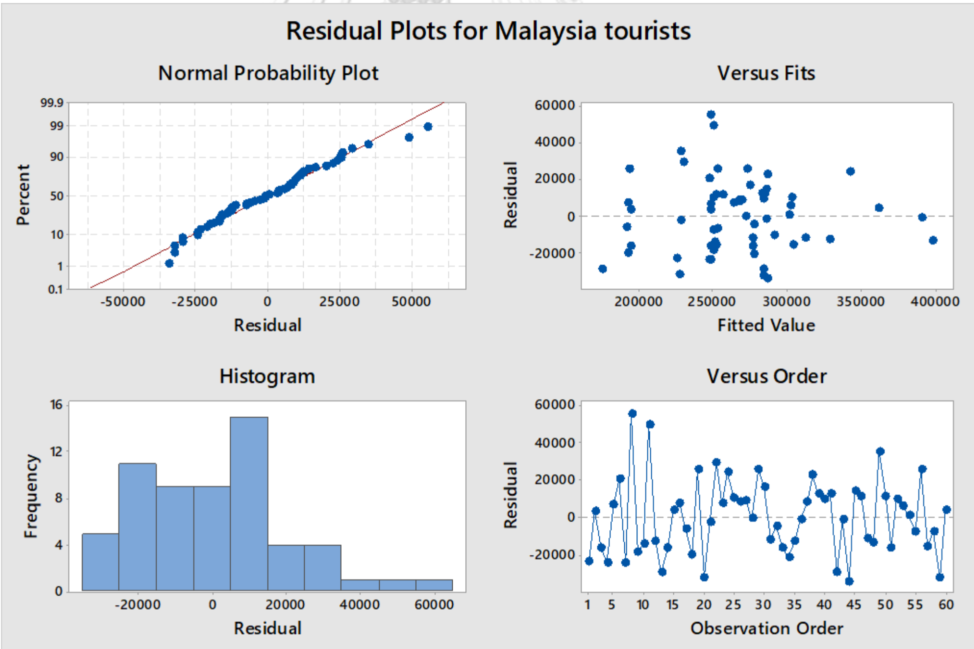

Figure 52: Residual plot for Malaysian tourists with dummy variable of seasonal and news shock effects

➢ Korea

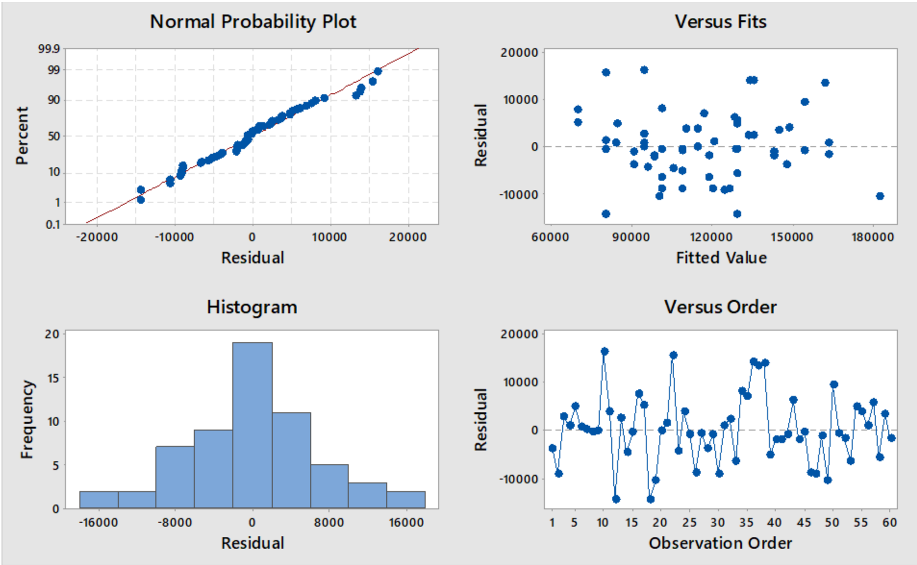

Figure 53: Residual plot for Korean tourists with dummy variable of seasonality

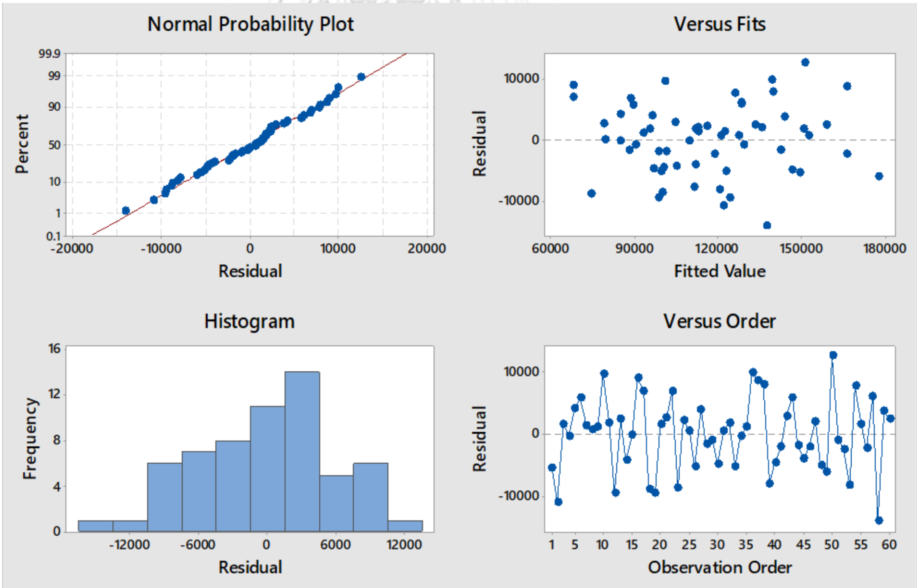

Figure 54: Residual plot for Korean tourists with dummy variable of seasonality and news shock effects

➢ **Japan**

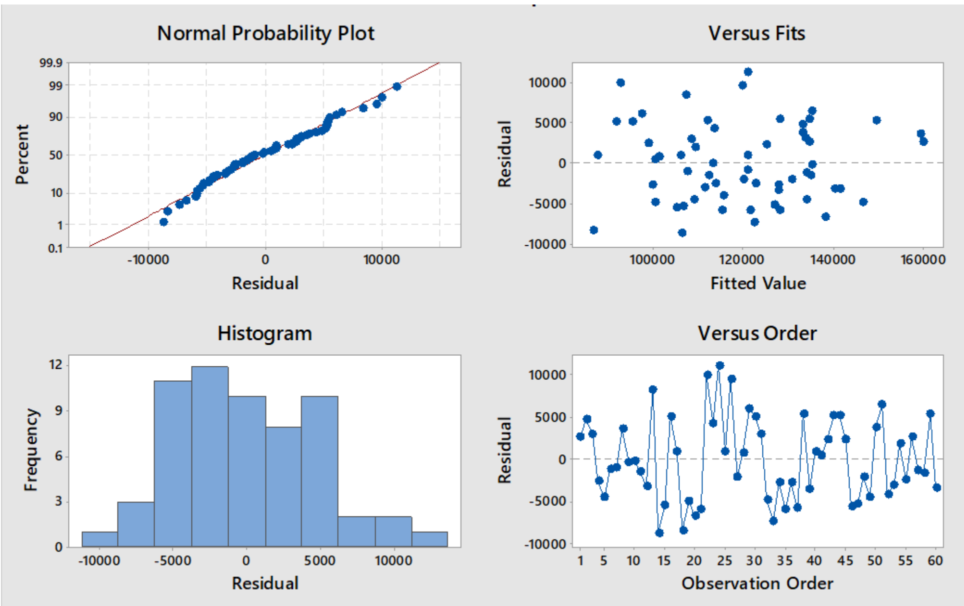

Figure 55: Residual plot for Japanese tourists with dummy variable of seasonality

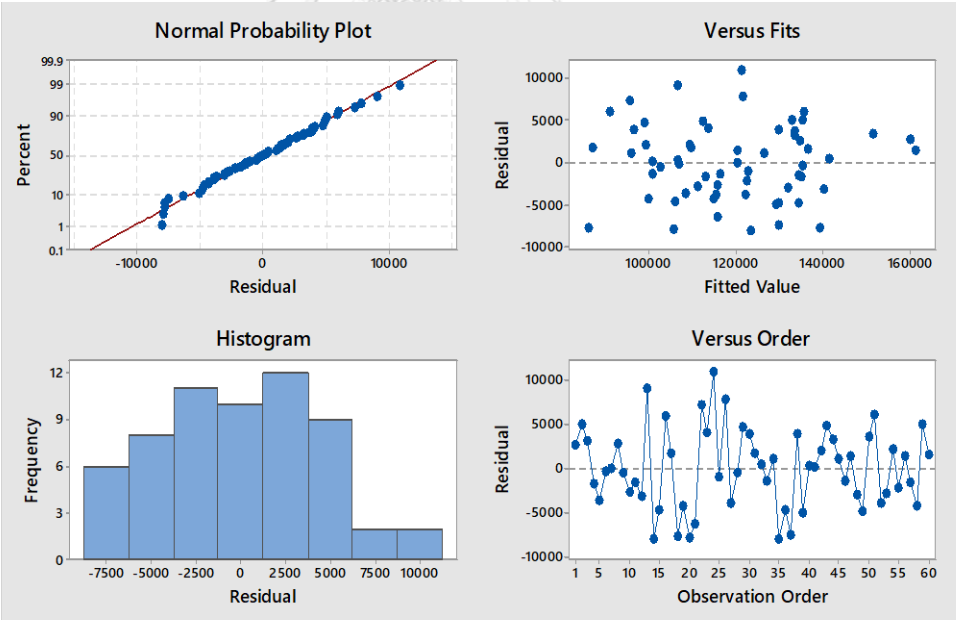

Figure 56: Residual plot for Japanese tourists with dummy variable of seasonality and news shock effects

- **Russia**

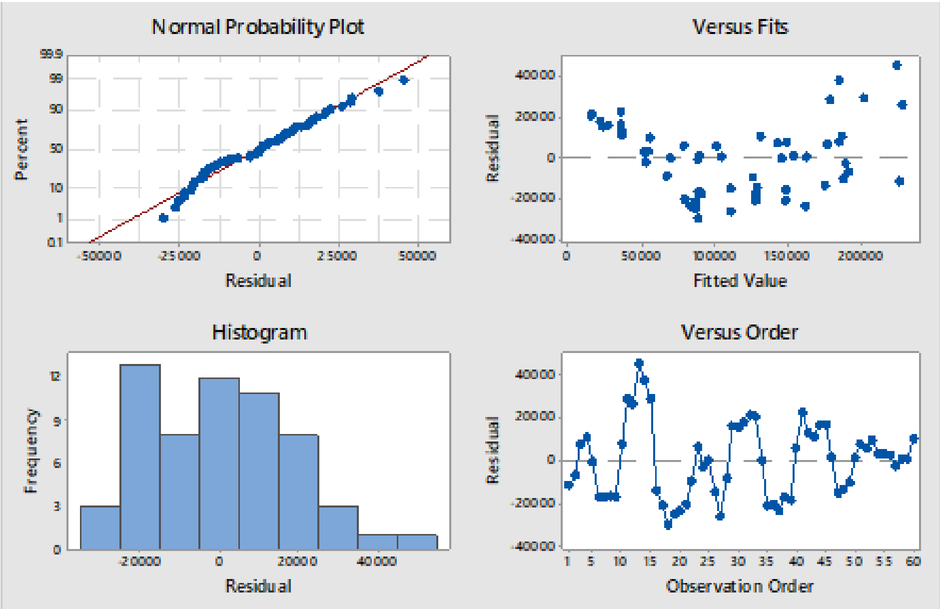

Figure 57: Residual plot for Russian tourists with dummy variable of seasonality

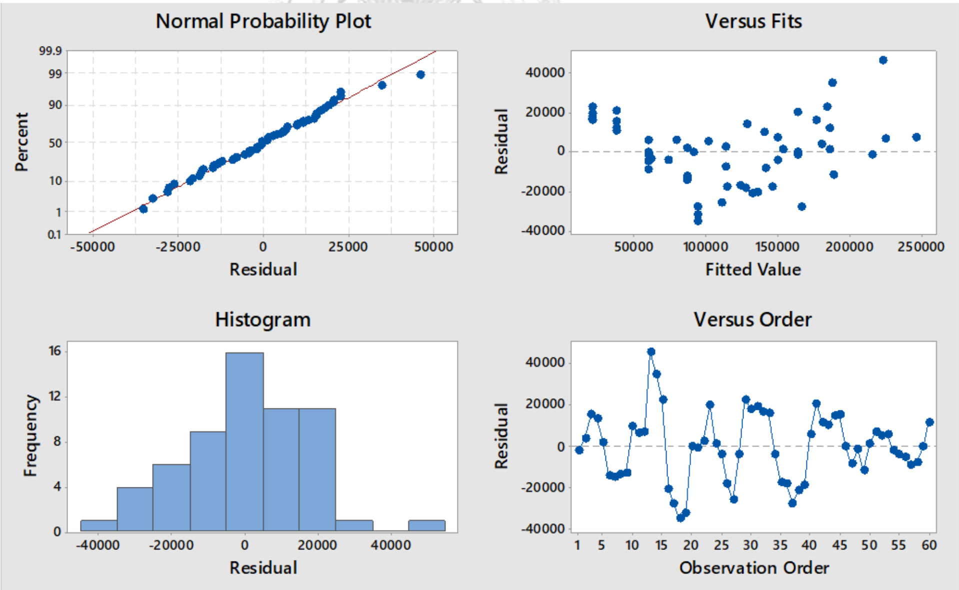

Figure 58: Residual plot for Russian tourists with dummy variable of seasonality and news shock effects

- **UK**

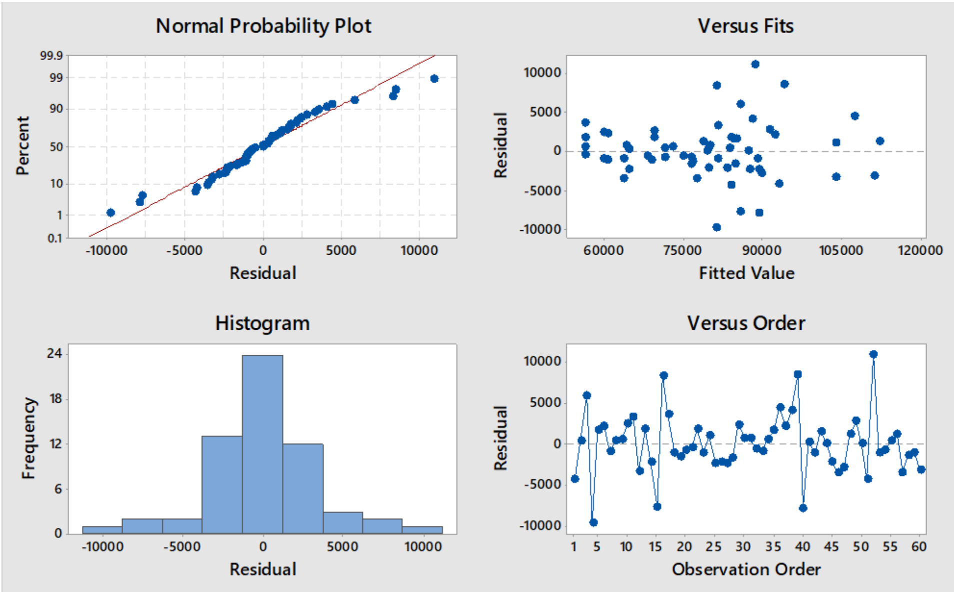

Figure 59: Residual plot for UK tourists with dummy variable of seasonality

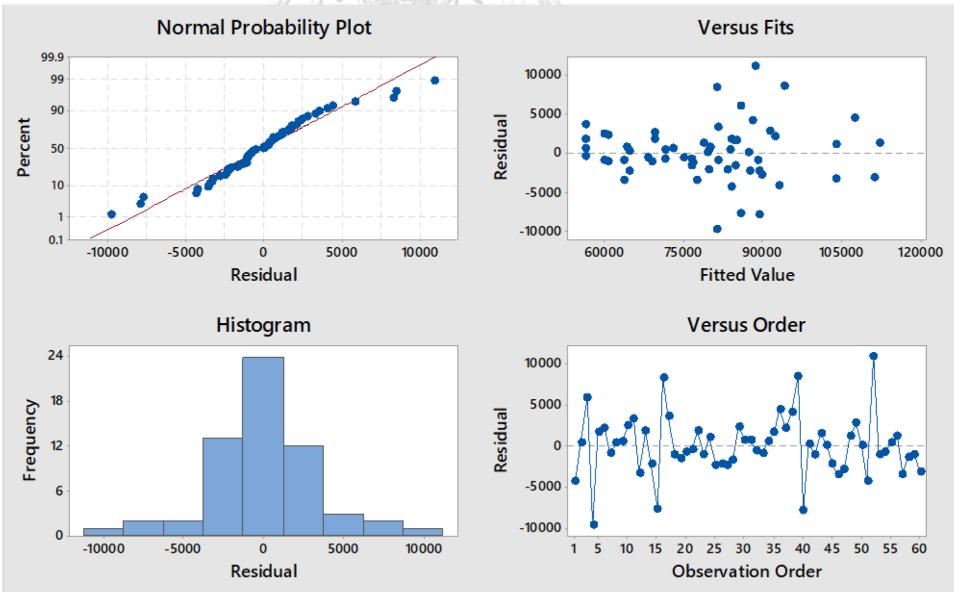

Figure 60: Residual plot for UK tourists with dummy variable of seasonality and news shock effects

➢ **US**

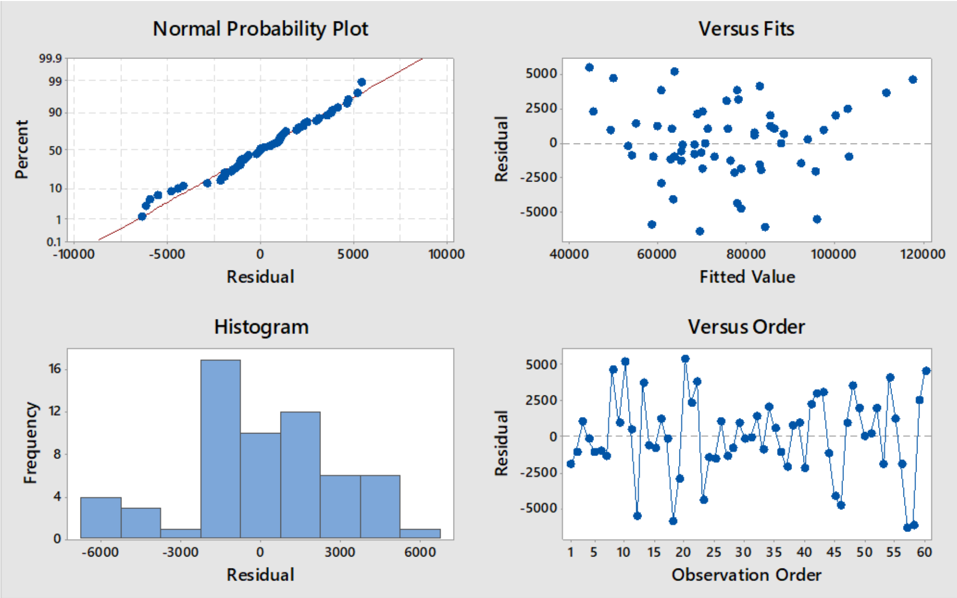

Figure 61: Residual plot for US tourists with dummy variable of seasonality

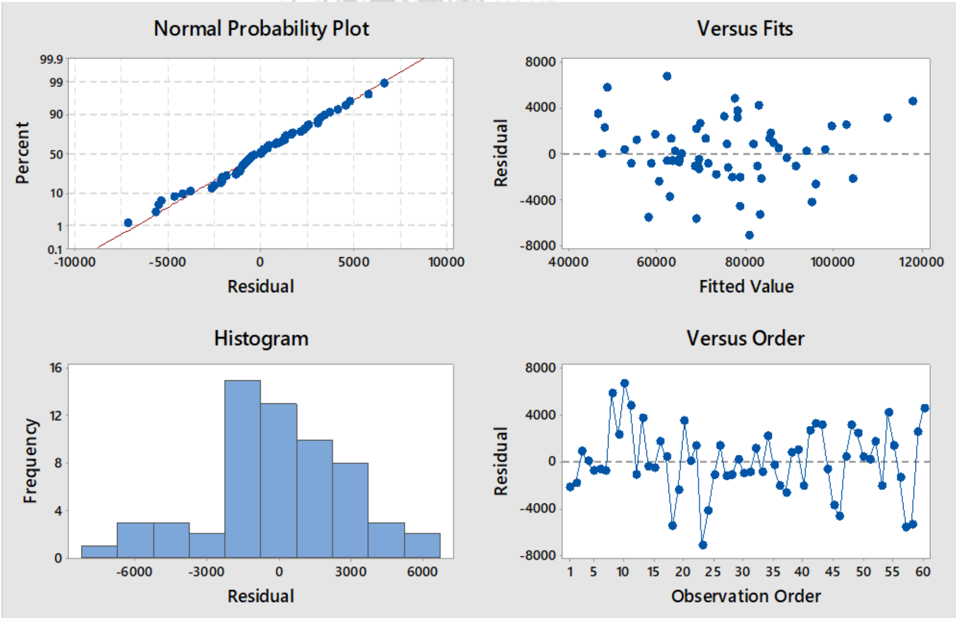

Figure 62: Residual plot of China tourists with dummy variable of seasonal and news shock effects

Appendix 4

Absolute Forecasting Error value of international tourist arrivals to Thailand in 2018 from major countries

➢ **China**

Table 20: Absolute Percentage Forecasting Error value of China tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	19.21	10.72	3.57	3.49	1.54
Feb	16.49	13.81	19.69	19.20	12.63
Mar	8.84	5.16	4.02	11.44	5.70
Apr	10.56	4.00	4.78	15.11	1.33
May	4.48	2.46	9.04	8.79	9.22
Jun	6.33	2.68	5.25	24.56	1.38
Jul	7.00	6.77	12.97	12.18	4.19
Aug	19.29	19.23	24.14	3.72	6.44
Sep	26.59	20.81	46.23	15.58	2.13
Oct	26.42	21.16	50.61	14.74	0.60
Nov	30.18	24.88	42.86	9.94	6.17
Dec	40.21	43.65	16.73	6.80	5.17

Table 21: Grouping Information Using the Tukey Method for China

Method	N	Mean	Grouping	
SARIMA	12	19.99	A	
MR+S	12	17.96	A	
MR+S+E	12	14.61	A	B
HW	12	12.13	A	B
ANN (15,63)	12	4.71		B

Table 22: Tukey Simultaneous Tests for Differences of Means for China

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	7.42	4.57	(-5.46, 20.30)	1.63	0.488
MR+S - ANN	13.26	4.57	(0.38, 26.14)	2.90	0.041
MR+S+E - ANN	9.90	4.57	(-2.98, 22.78)	2.17	0.207
SARIMA - ANN	15.28	4.57	(2.40, 28.16)	3.35	0.012
MR+S - HW	5.84	4.57	(-7.04, 18.72)	1.28	0.705
MR+S+E - HW	2.48	4.57	(-10.40, 15.36)	0.54	0.982
SARIMA - HW	7.86	4.57	(-5.02, 20.74)	1.72	0.429
MR+S+E - MR+S	-3.36	4.57	(-16.24, 9.52)	-0.74	0.947
SARIMA - MR+S	2.02	4.57	(-10.86, 14.90)	0.44	0.992
SARIMA - MR+S+E	5.38	4.57	(-7.50, 18.26)	1.18	0.764

- **Malaysia**

Table 23: Absolute Percentage Forecasting Error value of Malaysian tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	4.42	12.09	7.05	20.06	3.31
Feb	1.17	8.33	0.30	27.64	0.66
Mar	7.86	16.22	8.72	33.82	13.33
Apr	7.73	15.63	5.34	32.95	1.13
May	4.46	12.02	0.03	31.78	0.50
Jun	30.14	34.48	32.45	52.81	4.47
Jul	13.14	18.82	14.47	46.10	0.47
Aug	32.55	36.86	36.31	53.19	37.22
Sep	26.69	31.51	24.48	49.78	12.64
Oct	14.45	19.97	18.03	38.43	1.33
Nov	30.72	36.29	31.77	46.34	7.62
Dec	29.86	52.97	29.12	44.02	20.55

Table 24 : Grouping Information Using the Tukey Method for Malaysia

Methods	N	Mean	Grouping		
HW	12	39.7431	A		
MR+S+E	12	24.6001		B	
SARIMA	12	17.3393		B	C
MR+S	12	16.9329		B	C
ANN	12	8.6013			C

Table 25: Tukey Simultaneous Tests for Differences of Means for Malaysia

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	31.14	4.97	(17.13, 45.15)	6.27	0.000
MR+S - ANN	8.33	4.97	(-5.68, 22.34)	1.68	0.456
MR+S+E - ANN	16.00	4.97	(1.99, 30.01)	3.22	0.018
SARIMA - ANN	8.74	4.97	(-5.27, 22.75)	1.76	0.407
MR+S - HW	-22.81	4.97	(-36.82, -8.80)	-4.59	0.000
MR+S+E - HW	-15.14	4.97	(-29.15, -1.13)	-3.05	0.028
SARIMA - HW	-22.40	4.97	(-36.41, -8.40)	-4.51	0.000
MR+S+E - MR+S	7.67	4.97	(-6.34, 21.68)	1.54	0.539
SARIMA - MR+S	0.41	4.97	(-13.60, 14.41)	0.08	1.000
SARIMA - MR+S+E	-7.26	4.97	(-21.27, 6.75)	-1.46	0.591

- **Korea**

Table 26: Absolute Percentage Forecasting Error value of Korean tourists with different method

Korea	Absolute Forecasting Error				
	MR+S	MR+S+E	SARIMA	HW	ANN(55,12)
Jan	4047	10361	6259	10228	21606
Feb	13733	5620	23292	10549	5998
Mar	6265	3135	4471	37517	6776
Apr	22032	14149	20432	20996	2700
May	15593	5591	9417	28517	446
Jun	8347	7260	12053	41203	9665
Jul	14693	4883	18872	36254	6819
Aug	18640	8046	19747	32710	7015
Sep	14439	4681	18869	31628	2551
Oct	17895	17694	11101	28808	11391
Nov	9454	4114	13082	29682	4759
Dec	8550	2306	5969	33102	7814

Table 27: Grouping Information Using the Tukey Method for Korea

Methods	N	Mean	Grouping	
HW	12	19.5953	A	
SARIMA	12	9.4314		B
MR+S	12	9.0615		B
MR+S+E	12	5.1457		B
ANN	12	4.6399		B

Table 28: Tukey Simultaneous Tests for Differences of Means for Korea

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	14.96	1.99	(9.33, 20.58)	7.50	0.000
MR+S - ANN	4.42	1.99	(-1.20, 10.05)	2.22	0.189
MR+S+E - ANN	0.51	1.99	(-5.12, 6.13)	0.25	0.999
SARIMA - ANN	4.79	1.99	(-0.84, 10.42)	2.40	0.130
MR+S - HW	-10.53	1.99	(-16.16, -4.91)	-5.28	0.000
MR+S+E - HW	-14.45	1.99	(-20.08, -8.82)	-7.25	0.000

SARIMA - HW	-10.16	1.99	(-15.79, -4.54)	-5.10	0.000
MR+S+E - MR+S	-3.92	1.99	(-9.54, 1.71)	-1.96	0.297
SARIMA - MR+S	0.37	1.99	(-5.26, 6.00)	0.19	1.000
SARIMA - MR+S+E	4.29	1.99	(-1.34, 9.91)	2.15	0.215

- **Japan**

Table 29: Absolute Percentage Forecasting Error value of Japanese tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	5.02	4.87	7.80	14.19	14.89
Feb	5.20	4.98	5.10	23.27	12.92
Mar	5.41	5.26	6.17	29.66	3.79
Apr	4.06	4.72	4.15	26.26	3.26
May	0.54	1.40	0.92	20.18	10.23
Jun	9.35	10.09	7.23	30.66	17.76
Jul	5.53	6.19	4.59	28.68	7.18
Aug	2.39	1.66	11.40	31.42	21.08
Sep	14.42	12.93	5.97	26.41	12.01
Oct	14.34	14.55	8.23	27.61	15.77
Nov	8.29	15.76	8.95	27.74	17.75
Dec	7.83	16.19	9.11	28.88	16.57

Table 30: Grouping Information Using the Tukey Method for Japan

Methods	N	Mean	Grouping		
HW	12	26.2462	A		
ANN	12	12.7682		B	
MR+S+E	12	8.2154		B	C
MR+S	12	6.8645			C
SARIMA	12	6.6337			C

Table 31: Tukey Simultaneous Tests for Differences of Means for Japan

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	13.48	1.93	(8.03, 18.92)	6.98	0.000
MR+S - ANN	-5.90	1.93	(-11.35, -0.46)	-3.06	0.027
MR+S+E - ANN	-4.55	1.93	(-10.00, 0.89)	-2.36	0.142
SARIMA - ANN	-6.13	1.93	(-11.58, -0.69)	-3.18	0.020
MR+S - HW	-19.38	1.93	(-24.83, -13.94)	-10.04	0.000
MR+S+E - HW	-18.03	1.93	(-23.48, -12.59)	-9.34	0.000
SARIMA - HW	-19.61	1.93	(-25.06, -14.17)	-10.16	0.000
MR+S+E - MR+S	1.35	1.93	(-4.09, 6.80)	0.70	0.956
SARIMA - MR+S	-0.23	1.93	(-5.68, 5.21)	-0.12	1.000
SARIMA - MR+S+E	-1.58	1.93	(-7.03, 3.86)	-0.82	0.924

- **Russia**

Table 32: Absolute Percentage Forecasting Error value of Russian tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	18.66	16.64	8.98	6.33	7.56
Feb	23.24	21.14	9.59	8.69	11.46
Mar	26.16	23.78	12.12	14.70	16.16
Apr	27.30	19.74	0.25	24.72	29.83
May	17.21	4.75	57.33	1.35	8.64
Jun	1.62	23.34	67.26	7.77	0.79
Jul	1.77	20.29	62.62	19.54	3.67
Aug	8.23	19.50	61.74	18.51	0.62
Sep	10.64	20.50	60.89	19.60	0.75
Oct	3.91	12.04	29.12	20.53	12.33
Nov	5.75	1.63	12.05	15.17	7.97
Dec	12.31	9.49	2.30	4.26	14.99

Table 33: Grouping Information Using the Tukey Method for Russia

Methods	N	Mean	Grouping	
SARIMA	12	32.0214	A	
MR+S+E	12	16.0697	A	B
HW	12	13.4326		B
MR+S	12	13.0653		B
ANN	12	9.5651		B

Table 34: Tukey Simultaneous Tests for Differences of Means for Russia

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	3.87	5.83	(-12.57, 20.31)	0.66	0.963
MR+S - ANN	3.50	5.83	(-12.94, 19.94)	0.60	0.974
MR+S+E - ANN	6.50	5.83	(-9.94, 22.94)	1.12	0.797
SARIMA - ANN	22.46	5.83	(6.02, 38.90)	3.85	0.003
MR+S - HW	-0.37	5.83	(-16.81, 16.07)	-0.06	1.000
MR+S+E - HW	2.64	5.83	(-13.80, 19.08)	0.45	0.991
SARIMA - HW	18.59	5.83	(2.15, 35.03)	3.19	0.019
MR+S+E - MR+S	3.00	5.83	(-13.44, 19.44)	0.52	0.985
SARIMA - MR+S	18.96	5.83	(2.52, 35.40)	3.25	0.016
SARIMA - MR+S+E	15.95	5.83	(-0.49, 32.39)	2.74	0.061

- **UK**

Table 35: Absolute Percentage Forecasting Error value of British tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	1.67	1.67	0.64	15.21	2.75
Feb	1.70	1.70	1.26	5.79	2.18
Mar	4.52	4.52	9.24	0.93	8.93
Apr	6.44	6.44	9.75	6.28	2.14
May	9.32	9.32	2.30	33.37	1.64
Jun	10.28	10.28	3.51	40.91	1.64

Jul	9.72	9.72	6.01	13.80	0.12
Aug	5.90	5.90	2.41	7.47	4.43
Sep	14.83	14.83	3.27	15.73	2.67
Oct	8.63	8.63	1.27	13.71	0.83
Nov	3.89	3.89	1.78	25.11	0.70
Dec	4.54	4.54	0.74	16.46	2.60

Table 36: Grouping Information Using the Tukey Method for UK

Methods	N	Mean	Grouping	
HW	12	16.2316	A	
MR+S	12	6.7871		B
MR+S+E	12	6.7871		B
SARIMA	12	3.5146		B
ANN	12	2.5532		B

Table 37: Tukey Simultaneous Tests for Differences of Means for UK

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	13.68	2.47	(6.70, 20.65)	5.53	0.000
MR+S - ANN	4.23	2.47	(-2.74, 11.21)	1.71	0.435
MR+S+E - ANN	4.23	2.47	(-2.74, 11.21)	1.71	0.435
SARIMA - ANN	0.96	2.47	(-6.01, 7.94)	0.39	0.995
MR+S - HW	-9.44	2.47	(-16.42, -2.47)	-3.82	0.003
MR+S+E - HW	-9.44	2.47	(-16.42, -2.47)	-3.82	0.003
SARIMA - HW	-12.72	2.47	(-19.69, -5.74)	-5.14	0.000
MR+S+E - MR+S	-0.00	2.47	(-6.97, 6.97)	-0.00	1.000
SARIMA - MR+S	-3.27	2.47	(-10.25, 3.70)	-1.32	0.678
SARIMA - MR+S+E	-3.27	2.47	(-10.25, 3.70)	-1.32	0.678

➢ **US**

Table 38: Absolute Forecasting Error value of American tourists with different method

Month	Absolute Percentage Error				
	MR+S	MR+S+E	SARIMA	HW	ANN
Jan	0.04	1.72	1.20	1.82	4.15
Feb	9.30	7.87	3.84	4.08	2.82
Mar	2.41	3.61	5.76	5.68	9.90
Apr	6.14	4.61	5.21	3.32	4.03
May	7.64	5.60	2.44	1.43	3.28
Jun	2.35	0.62	3.52	1.36	3.78
Jul	7.34	5.56	5.23	2.72	0.11
Aug	9.74	6.04	4.79	1.26	1.09
Sep	18.59	13.97	7.06	4.62	0.45
Oct	12.27	9.16	1.78	1.50	0.32
Nov	5.37	6.90	5.37	5.58	12.66
Dec	1.27	37.83	0.66	1.20	5.22

Table 39: Grouping Information Using the Tukey Method for US

Methods	N	Mean	Grouping
MR+S+E	12	8.62465	A
MR+S	12	6.87021	A
ANN	12	3.98387	A
SARIMA	12	3.90555	A
HW	12	2.88066	A

Table 40: Tukey Simultaneous Tests for Differences of Means for US

Difference of Methods Levels	Difference of Means	SE of Difference	Simultaneous 95% CI	T-Value	Adjusted P-Value
HW - ANN	-1.10	2.21	(-7.33, 5.12)	-0.50	0.987
MR+S - ANN	2.89	2.21	(-3.34, 9.11)	1.31	0.687
MR+S+E - ANN	4.64	2.21	(-1.58, 10.87)	2.10	0.233
SARIMA - ANN	-0.08	2.21	(-6.30, 6.15)	-0.04	1.000
MR+S - HW	3.99	2.21	(-2.24, 10.21)	1.81	0.379
MR+S+E - HW	5.74	2.21	(-0.48, 11.97)	2.60	0.084

SARIMA - HW	1.02	2.21	(-5.20, 7.25)	0.46	0.990
MR+S+E - MR+S	1.75	2.21	(-4.47, 7.98)	0.80	0.931
SARIMA - MR+S	-2.96	2.21	(-9.19, 3.26)	-1.34	0.666
SARIMA - MR+S+E	-4.72	2.21	(-10.94, 1.51)	-2.14	0.219

Appendix 5

Monthly predicted number of tourist arrival to Thailand in 2018

Table 41: Predicted number of China tourist with different methods

Month	Actual Tourists	Predicted Number of China Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN (15,63)
Jan	970015	783658	866059	1004680	1003851	955105
Feb	1200479	1002480	1034698	964067	970032	1048825
Mar	1004030	915243	952258	963689	889139	946846
Apr	986729	882524	947306	939569	837592	973575
May	869235	830297	890619	947794	792794	949365
Jun	900665	843626	876549	947981	679472	888237
Jul	929771	994886	992739	1050385	816514	968753
Aug	867461	700116	700638	1076837	899764	923360
Sep	647664	475449	512876	947089	748545	661479
Oct	646143	475449	509413	973128	741366	642271
Nov	675129	471360	507140	964474	742250	633505
Dec	838634	501392	472549	978919	781567	881967

Table 42: Predicted number of Malaysia tourist with different methods

Month	Actual Tourists	Predicted Number of Malaysia Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN (44,31)
Jan	265727	253995	233591	247006	212435	256924
Feb	277140	280387	254045	276299	200548	278963
Mar	302437	278656	253376	276050	200160	342753
Apr	299104	275995	252347	283135	200545	302474
May	285344	272608	251037	285248	194648	286769
Jun	378163	264168	247775	255467	178454	361264
Jul	305960	265750	248386	261692	164899	307387
Aug	392946	265036	248110	250257	183936	246683
Sep	362775	265935	248458	273981	182174	316930
Oct	310172	265365	248237	254252	190965	306055
Nov	389968	270163	248458	266074	209276	360261
Dec	527868	370272	248237	374135	295523	419398

Table 43: Predicted number of Korea tourist with different methods

Month	Actual Tourists	Predicted Number of Korea Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN (55,12)
Jan	197625	201672	187264	191366	187397	176019
Feb	160148	173881	165768	183440	149599	166146
Mar	143644	149909	140509	148115	106127	136868
Apr	115870	137902	130019	136302	94874	113170
May	122684	138277	128275	132101	94167	123130
Jun	141562	149909	134302	153615	100359	131897
Jul	153524	168217	158407	172396	117270	160343
Aug	164135	182775	172181	183882	131425	157120
Sep	135470	149909	140151	154339	103842	138021
Oct	132014	149909	149708	143115	103206	120623
Nov	154561	164015	158675	167643	124879	149802
Dec	175359	183909	173053	181328	142257	167545

Table 44: Predicted number of Japan tourist with different methods

Month	Actual Tourists	Predicted Number of Japan Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN(56,52)
Jan	143810	136597	136805	132599	123410	122398
Feb	144937	137404	137715	137545	111216	126206
Mar	144482	136666	136883	135574	101632	149963
Apr	119101	114267	113481	114162	87819	115214
May	109086	108499	107560	110091	87072	97928
Jun	118509	107427	106555	109946	82170	97457
Jul	128134	121043	120208	122251	91387	118934
Aug	181530	185866	184535	160830	124489	143256
Sep	141678	162106	159992	133227	104266	124668
Oct	122034	139537	139794	111989	88342	102789
Nov	148810	161140	125361	135487	107536	122401
Dec	153989	166050	129060	139967	109511	128478

Table 45: Predicted number of Russia tourist with different methods

Month	Actual Tourists	Predicted Number of Russia Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN(19,50)
Jan	226754	184439	189016	206390	212396	209607
Feb	195219	149849	153952	176489	178249	172851
Mar	197023	145491	150178	173138	168061	165184
Apr	126402	91899	101449	126087	95155	88692
May	57098	47273	59812	89833	56325	52162
Jun	48493	47710	59812	81111	52261	48108
Jul	49723	48844	59812	80860	59440	51548
Aug	50053	45934	59812	80954	59319	49742
Sep	49637	44354	59812	79861	59367	50009
Oct	101217	97262	113407	130692	122001	113692
Nov	166249	156690	163545	186282	191476	179503
Dec	205081	179839	185619	209791	213812	174338

Table 46: Predicted number of UK tourist with different methods

Month	Actual Tourists	Predicted Number of UK Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN(2,58)
Jan	93535	95101	95101	94130	107764	90962
Feb	89379	90897	90897	88252	94554	87426
Mar	101461	96875	96875	92082	100513	92404
Apr	86679	92261	92261	95131	92120	84823
May	61732	67486	67486	63152	82334	60723
Jun	65076	71766	71766	67360	91701	64006
Jul	79751	87501	87501	84544	90758	79843
Aug	77877	82468	82468	79754	72062	74427
Sep	58768	67486	67486	60689	68012	60339
Oct	74075	80470	80470	75014	84231	74692
Nov	89169	92641	92641	87581	111555	89794
Dec	109954	114943	114943	109143	128055	107092

Table 47: Predicted number of US tourist with different methods

Month	Actual Tourists	Predicted Number of US Tourist				
		MR+S	MR+S+E	SARIMA	HW	ANN(22,41)
Jan	109757	109798	107870	108441	107764	105202
Feb	90852	99301	98001	94343	94554	93410
Mar	106562	103989	102719	100424	100513	96015
Apr	89159	94629	93266	93803	92120	92756
May	81177	87375	85725	83156	82334	78517
Jun	90468	92595	91030	93653	91701	87051
Jul	88352	94834	93266	92975	90758	88449
Aug	71163	78096	75460	74573	72062	70389
Sep	65009	77095	74094	69598	68012	65299
Oct	82986	93164	90591	84466	84231	83255
Nov	118146	111807	109993	111805	111555	103183
Dec	129617	127973	126759	128759	128055	122847

REFERENCES

1. Embassy, R.T. 2018.
2. SamuiTime, *Russia Tops European Visitor Arrivals to Thailand; Insights into German Visitors*. 2018.
3. Vencovsk, J., *The Determinants of International Tourism Demand,* in *Faculty of Social Sciences*. 2014, Charles University in Prague.
4. Lim, C., *Rview of International Tourism Demand Models*. 1997, University of Western Australia, Australia.
5. Artus, J.R., *The Effect of Revaluation on the Foreign Travel Balance of Germany.* nternational Monetary Fund Staff Papers 1970. **27**: p. 602-617.
6. Haiyan Song, G.L., *Tourism demand modelling and forecasting - A review of Recent research.* Tourism Management, 2008. **7**.
7. Chinnakum, W. and P. Boonyasana, *Forecasting International Tourism Demand in Thailand.* Thai Journal of Mathematics, 2016: p. 231-244.
8. Wikipedia. *2013–2014 Thai political crisis*. 2013; Available from: https://en.wikipedia.org/wiki/2013%E2%80%932014_Thai_political_crisis.
9. Wikipedia. *2015 Bangkok bombing*. 2015; Available from: https://en.wikipedia.org/wiki/2015_Bangkok_bombing.
10. Telegraph, T. *Bangkok bombing hits Thailand tourism*. 2015.
11. Wikipedia. *2017 Southern Thailand floods*. 2017; Available from: https://en.wikipedia.org/wiki/2017_Southern_Thailand_floods.
12. travel, C. *What visitors to Thailand need to know after King Bhumibol's death*. 2016; Available from: https://edition.cnn.com/travel/article/thailand-king-travel/index.html.
13. Wikipedia. *2018 Phuket boat capsizing*. 2018; Available from: https://en.wikipedia.org/wiki/2018_Phuket_boat_capsizing.
14. Brownlee, J. *A Gentle Introduction to Exponential Smoothing for Time Series Forecasting in Python*. 2018.
15. Box, G.E.P. and G.M. Jenkins, *Time Series Analysis.* Forecasting and Control, 1970.
16. Song, H., *Tourism Demand Modelling and Forecasting—A Review of Recent Research*. 2008, University of Surrey.
17. Khan, N. *ACTIVATION FUNCTION IN DEEP LEARNING*. 2017; Available from: https://medium.com/@najeebnik21/activation-function-in-deep-learning-587e83d5a681.
18. Cho, V., *Tourism forecasting and its relationship with leading economic indicators.* Journal of Hospitality and Tourism Research, 2001. **25**: p. 399-420.
19. Goh, C. and R. Law, *Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention.* Tourism Management, 2002. **23**: p. 499-510.
20. Smeral, E. and M. Wüger, *Does complexity matter? Methods for improving forecasting accuracy in tourism: The case of Australia.* Journal of Travel Research, 2005. **44**: p. 100-110.
21. Choden and S. Unhapipat, *ARIMA model to forecast international tourist visit in Bumthang, Bhutan.* Journal of Physics, 2018.
22. Witt, S.F. and C. Witt, *Modelling and Forecasting Demand in Tourism, Academic Press, London.* 1992.

23. Holt, C., C., *Forecasting Seasonal and Trends by Exponentially Weighted Moving Averages.* Office of Naval Research, Research Memorandum, 1957.
24. Witt, S.F., G.D. Newbould, and A.J. Watidns, *Forecasting Domestic Tourism Demand: Application to Las Vegas Arrivals Data.* Journal of Travel Research, 1992. **31**(1): p. 36-41.
25. Martin, C.A. and Witt, *Forecasting performance.* Tourism Management, 1989. **9** (4): p. 326-329.
26. Athiyaman, A. and R.W. Robertson, *Time Series Forecasting Techniques: Short-Term Planning in Tourism.* International Journal of Contemporary Hospitality Management,, 1992.
27. Lim, C. and M. McAleer, *Forecasting tourist arrivals.* Journals of Tourism Research, 2001. **28**(4): p. 965-977.
28. Crouch, G.I., *The Study of International Tourism Demand: a Survey of Practice.* Journal of Travel Research, 1994. **32**: p. 41-54.
29. Hiemstra, S.a.W., K. K. F., *Factors Affecting Demand for Tourism in Hong Kong.* Journal of Travel & Tourism Marketing, 2002: p. 43-62.
30. NWANKWO, C.H., *dummy variable multiple regression forecasting model.* International Journal of Engineering Science Invention, 2013. **2**(3): p. 42-51.
31. Martin, C.A. and S.F. Witt, *Forecasting tourism demand: A comparison of the accuracy of several quantitative methods.* International Journal of Forecasting, 1989. **5**: p. 7-10.
32. Kulendran, N. and K. M., *Forecasting intemational quarterly tourist flows using error-correction and time-series models.* International Journal of Forecasting, 1997. **13**: p. 319-327.
33. Adya, M. and C. F, *How effective are Neural Networks at forecasting and prediction? A review and evaluation.* Journal of Forecasting, 1998. **17**: p. 451-461.
34. Zhang, G., B.E. Putuwo, and M.Y. Hu, *Forecasting with artificial NeuralNetworks: The state of the ait.* International Journal of Forecasting, 1998. **14**: p. 35-62
35. Qi, M. and G.P. Zhang, *An investigation of model selection criteria for Neural Network time series forecasting.* European Journal of Operational Research, 2001.
36. Lippman, R.P., *An introduction to Computing witt Neural Nets*, in *lEEEASSP Magazine*. 1987. p. 4-22.
37. Gorr, W.L., D. Nagin, and J. Szczypula, *Comparative Study of Artificial Neural Network and Statistical Methods for Predicting Student Grade Pomt Averages.* International Journal of Forecasting, 1994. **10**: p. 17-34.
38. Untong, A., et al., *Tourism demand analysis of chinese arrivals in Thailand.* Tourism Economics, 2015. **6**: p. 1221-1234.
39. Salman, A.K., *Estimating Tourism Demand Through Cointegration Analysis: Swedish Data.* Current Issues in Tourism, 2003. **6**: p. 323-338.
40. Lim, C., *The Major Determinants of Korean Outbound Travel to Australia.* Mathematics and Computers in Simulation, 2004. **64**: p. 477–485.
41. Song and Witt, *Tourism demand modelling and forecasting: modern econometric approaches.* Journal of Retailing and Consumer Services, 2000. **1**: p. 54-55.

42. Baldigara, T., *Forecasting Tourism Demand in Croatia: A Comparison of Different Extrapolative Methods.* Journal of Business Administration Research, 2013. **2**.
43. Huang, N. *10 most visited countries in the world*. 2017; Available from: https://www.wildjunket.com/most-visited-countries/.
44. Song, H., et al., *Modelling and forecasting the demand for Thai tourism.* Tourism Economics, 2003. **4**: p. 363–387.
45. Hao, J., T. Var, and J. Chon, *A forecasting model of tourist arrivals from major markets to Thailand.* Tourism Analysis, 2003. **8**: p. 33-45.
46. Sookmark, S., *A Analysis of International Tourism Demand in Thailand*, in *School of Development Economics*. 2011. p. 154.
47. Chaitip, P. and C. Chaiboonsri, International Conference on Applied Economics, 2014: p. 100-109.
48. Wongsathan1, R. and W. Jaroenwiriyapap, *A Hybrid ARIMA and RBF Neural Network Model for Tourist Quantity Forecasting : A Case Study for Chiangmai Province.* KhonKan University journal, 2016. **21(1)**.
49. Boonaom, N., *Forecasting the Number of Chinese Tourists in Thailand.* Thammasat journal, 2018. **26**.
50. Saothayanun, L., et al., *A Forecasting Methods for the Number of International Tourists in Thailand: Box-Jenkins Method and Winter's Method.* University of the Thai Chamber of Commerce, 2014.
51. Chaivichayachat and Bundit, *Forecasting Foreign Tourist in Thailand by Artificial Neural Network.* Advanced Science Letters, 2018. **28**.
52. Lin, C.J. and T.S. Lee, *Tourism Demand Forecasting: Econometric Model based on Multivariate Adaptive Regression Splines, Artificial Neural Network and Support Vector Regression.* Advances in Management & Applied Economics, 2013. **3**: p. 1-18.
53. Law, R. and R. Pine, *Tourism Demand Forecasting for the Tourism Industry: A Neural Network Approach.* Neural Networks in Business Forecasting, 2004.
54. Chaivichayachat and Bundit, *Forecasting Foreign Tourist in Thailand by Artificial Neural Network.* American Scientific Publishers, 2018. **24**(4): p. 9251-9254.
55. *GDP per capita*. 2018; Available from: https://tradingeconomics.com/china/gdp.
56. *Consumer price index* 2018; Available from: https://tradingeconomics.com/china/consumer-price-index-cpi.
57. Xrate. 2018; Available from: https://www.x-rates.com/.
58. Handbook, E.S.
59. sciences, T.D. *Applied Deep Learning* 2017.
60. Ali, R. and A. Shabri, *Modelling Singapore Tourist Arrivals to Malaysia by Using SVM and ANN.* Journal of Mathematics 2016. **1**(2).
61. Klimberg, R.K., et al., *Forecasting performance measures- What are their practical meaning?* 2010.

VITA

NAME	อรธีรา หวานดี
DATE OF BIRTH	18 มกราคม 2537
PLACE OF BIRTH	กรุงเทพมหานคร ประเทศไทย
INSTITUTIONS ATTENDED	ปริญญาตรี วิทยาศาสตร์บัณฑิต สาขาเคมี เกียรตินิยมอันดับสอง มหาวิทยาลัยธรรมศาสตร์
HOME ADDRESS	2-4 ถ. บัวรอง ต. ในเมือง อ. เมือง จ. นครราชสีมา 30000
PUBLICATION	Proceeding for Forecasting Iternational Tourist Arrivals to Thailand from Major country