A Novel Computing Approach to Mid-Term Load Forecasting for Thailand

Pituk Bunnoon

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Electrical Engineering

Prince of Songkla University

2013

A Novel Computing Approach to Mid-Term Load Forecasting for Thailand

Pituk Bunnoon

A Thesis Submitted in Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Electrical Engineering

Prince of Songkla University

2013

Thesis Title A Novel Computing Approach to Mid-Term Load Forecasting for Thailand

Author Mr. Pituk Bunnoon

Major Program Electrical Engineering

Major Advisor :

....

(Assist. Prof. Dr. Kusumal Chalermyanont)

Co-advisor :

....

(Assoc. Prof. Dr. Chusak Limsakul)

Examining Committee :

....Chairperson

(Assist. Prof. Dr. Pornchai Phukpattaranont)

....

(Assoc. Prof. Dr. Montri Karnjanadecha)

....

(Dr. Titiporn Sangpetch)

....

(Assist. Prof. Dr. Kusumal Chalermyanont)

....

(Assoc. Prof. Dr. Chusak Limsakul)

The Graduate School, Prince of Songkla University, has approved this thesis as fulfillment of the requirements for the Doctor of Philosophy Degree in Electrical Engineering.

....

(Assoc. Prof. Dr. Teerapol Srichana)

Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

..........................Signature

(Assist. Prof. Dr. Kusumal Chalermyanont)

Major advisor

..........................Signature

(Mr. Pituk Bunnoon)

Candidate

I hereby certify that this work has not already been accepted in substance for any degree, and is not being concurrently submitted in candidature for any degree.

..........................Signature

(Mr. Pituk Bunnoon)

Candidate

Thesis Title A Novel Computing Approach to Mid-Term Load Forecasting for Thailand

Author Mr. Pituk Bunnoon

Major Program Electrical Engineering

Academic Year 2012

ABSTRACT

The electricity consumption load demand in kilowatt-hour (kWh) is one of the most interesting indexes in power system because it can identify the customer behavior and trend of electricity consumption in the future. For this reason, the forecasting of the electricity consumption load demand is very important for power system operation. Mainly, the electricity consumption load demand is forecasted in Mid-term load forecasting (timing between 3 months - 3 years) for fuel reserve planning and unit commitment in power system. In the past, many forecasting methodologies were used in the load demand forecasting including statistical methods and artificial intelligence methods. Moreover, the external factors that affect to the forecasted values are extendedly used. In general, the factors can be divided into 2 main groups: weather factors such as temperature, humidity, rainfall, wind speed, and economic factors like industrial index, consumer price index and gross domestic product, etc. The error of electricity load demand forecasting results depend on the size of historical data and forecasting area, the forecasting algorithms and selected factors.

In this research, five forecasting models of energy consumption load demand forecasting are performed for comparison study. The first two statistical methods are Multiple Linear Regression (MLR) and Seasonal Autoregressive Integrated Moving Average (SARIMA). MLR is proposed to analyze the linear part with the kWh time series model. It used all factors to build a model for load forecasting. On the other hand, SARIMA is proposed to investigate the forecast problem by using seasonal method in ARIMA. This method can predict kWh by using and learning the historical demand signals of kWh without using any factor in the model. Since the energy load demand signals are non-linear or non-stationary signals in time series, the forecasting results from MLR and SARIMA may not yet good. Next, the artificial intelligent methodology of neural network algorithm (ANN) is proposed to solve the forecasting problem of

non-linear time series. The forecasting accuracy from this method is better than that of statistical methods. Although, the results of ANN are high accuracy compared to traditional method being statistical methodologies, but it is not yet satisfied the requirement from the Electricity Generating Authority of Thailand (EGAT). The last two forecasting models are included preprocessing stages. They are used for decomposing the electricity load demand signal into sub-level signals before sending to the forecasting stage with ANN. First, by using wavelet transform (WT) in the preprocessing stage, the results show that the 2-level wavelet transform gives the Mean Absolute Percent Errors (MAPE) better than that other level. However, this method is complicate and some of sub-signals are useless because they are not related to any factors. In the last method, the Hodrick Prescott (HP) filter is used for decomposing load demand signal into trend and cyclical components. This model leads good correlations between weather factors and cyclical component as well as economic factors and trend component. As a result, the combination of HP-filter and ANN is high reliability and more accuracy than other approaches.

Since the accuracy depends not only on a model algorithm but also a size of forecasting area. In this research, all of the mentioned forecasting methods above are performed with 2-types of area division: whole country area and substation-control-central area. Most of the results show that the whole area forecasting have slightly better accuracy than the substation control central area forecasting.

Thesis Title A Novel Computing Approach to Mid-Term Load Forecasting for Thailand

Author Mr. Pituk Bunnoon

Major Program Electrical Engineering

Academic Year 2012

ABSTRACT

The electricity consumption load demand in kilowatt-hour (kWh) is one of the most interesting indexes in power system because it can identify the customer behavior and trend of electricity consumption in the future. For this reason, the forecasting of the electricity consumption load demand is very important for power system operation. Mainly, the electricity consumption load demand is forecasted in Mid-term load forecasting (timing between 3 months - 3 years) for fuel reserve planning and unit commitment in power system. In the past, many forecasting methodologies were used in the load demand forecasting including statistical methods and artificial intelligence methods. Moreover, the external factors that affect to the forecasted values are extendedly used. In general, the factors can be divided into 2 main groups: weather factors such as temperature, humidity, rainfall, wind speed, and economic factors like industrial index, consumer price index and gross domestic product, etc. The error of electricity load demand forecasting results depend on the size of historical data and forecasting area, the forecasting algorithms and selected factors.

In this research, five forecasting models of energy consumption load demand forecasting are performed for comparison study. The first two statistical methods are Multiple Linear Regression (MLR) and Seasonal Autoregressive Integrated Moving Average (SARIMA). MLR is proposed to analyze the linear part with the kWh time series model. It used all factors to build a model for load forecasting. On the other hand, SARIMA is proposed to investigate the forecast problem by using seasonal method in ARIMA. This method can predict kWh by using and learning the historical demand signals of kWh without using any factor in the model. Since the energy load demand signals are non-linear or non-stationary signals in time series, the forecasting results from MLR and SARIMA may not yet good. Next, the artificial intelligent methodology of neural network algorithm (ANN) is proposed to solve the forecasting problem of non-linear time series. The forecasting accuracy from this method is better than that of statistical methods. Although, the results of ANN are high accuracy compared to traditional method being statistical methodologies, but it is not yet satisfied the requirement from the Electricity Generating Authority of Thailand (EGAT). The last two forecasting models are included preprocessing stages. They are used for decomposing the electricity load demand signal into sub-level signals before sending to the forecasting stage with ANN. First, by using wavelet transform (WT) in the preprocessing stage, the results show that the 2-level wavelet transform gives the Mean Absolute Percent Errors (MAPE) better than that other level. However, this method is complicate and some of sub-signals are useless because they are not related to any factors. In the last method, the Hodrick Prescott (HP) filter is used for decomposing load demand signal into trend and cyclical components. This model leads good correlations between weather factors and cyclical component as well as economic factors and trend component. As a result, the combination of HP-filter and ANN is high reliability and more accuracy than other approaches.

Since the accuracy depends not only on a model algorithm but also a size of forecasting area. In this research, all of the mentioned forecasting methods above are performed with 2-types of area division: whole country area and substation-control-central area. Most of the results show that the whole area forecasting have slightly better accuracy than the substation control central area forecasting.

ACKNOWLEDGEMENTS

This work was funded by the Office of the Higher Education Commission. I have been supported by CHE510382 Ph.D. Scholarship, Thailand. I would also like to thank the Thai Meteorological Department, Ministry of Transport and Communications, the Ministry of Commerce, the office of the National Economic and Social Development Board, and the Electricity Generating Authority of Thailand (EGAT) in providing the valuable data.

I would like to express my heartfelt gratitude and appreciations to all supporters in the Electrical Engineering Department, Faculty of Engineering, Prince of Songkla University (PSU). Also, I would like to express sincere gratitude to my advisors who are Assistant Professor Dr.Kusumal Chalermyanont and Associate Professor Dr.Chusak Limsakul. They have given me anything about the research steps. Sometimes if I have some problems, they will give the answer and suggest it clearly.

I am grateful to the President of Rajamangala University of Technology Srivijaya (RMUTSV) for providing full time scholarship during the study at the Price of Songkla University and also thank to associate professor Sompan Ampawan, Mr.Konto Pantongkum, Mr.Somkid Leelachanachipong, Mr.Paradorn Ruangkool, and other people for advising during study.

Finally, I would like to thank my family for their love and encouragements, my parents who raised me with the love of science and technology, and supported me in all my pursuits. I would like to thank all my brothers and sister who are care me since I was studied in the university and most of all for my loving, supporting, encouraging, and other during the final stages of this Ph.D. is so appreciated. Thank you very much.

Pituk Bunnoon

CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

Chapter 1
Introduction

1.1 Background

Thailand's power sectors have been dominated by three government-owned enterprises since 1970's. The first is the Electricity Generating Authority of Thailand (EGAT), responsible for generation and transmission. The other two are Metropolitan Electricity Authority (MEA) and Provincial Electricity Authority (PEA), responsible for distribution and retail services for the Bangkok Metropolitan Region and the rest of the country respectively. All electricity organizations especially EGAT will monitor both peak load and electricity consumption load demand that because they are essential in the daily life and the main driving factors for country economic. In order to provide sufficient electricity and make the economic grown continuously, the load forecastings are required for the related electricity producers. Since the construction of a power plant must take 5-10 years from planning, designing, environmental admitting to constructing step and there are few electric networks of Thailand and neighbor countries, the midterm load forecasting (MTLF) and the long term load forecasting (LTFL) are very important for building up the energy stability in Thailand .

Electricity load forecasting is not only significant for investment planning of three electricity authorities (EGAT, MEA and PEA), but also it is useful for estimating the financial statement of three electricity institutes. The accuracy forecasted values make the proper investment for three electricity authorities. If the forecasted values are too high, the exceeding investment is obtained and it will push this expense to consumers. However, if the forecasted values are too low, the inadequate investment is occurred and it will cause electricity deficiency in the country [1-2]. Most of the forecasting flow will put in the historical data to various forecasting techniques to get the forecast results as seen in Figure 1.1

The electricity forecasted demands are defined as two significant values: the peak value (Maximum load demand) and energy value (Electric energy demand). The peak value is used in planning for new electricity plants while the energy value is used in planning for fuel providing. In the past, each electricity authority used the different methods to find these values.

Figure 1.1: Forecasting flow.

Forecasting of the energy value

- The MEA and PEA use the econometric model with Error Correction Model of Engle-Granger method or the econometric method with auto-Regressive Distributed Lag (ARDL) for monthly forecasting. The forecasting variables affected to electric load demands are electricity bill rate, GDP and temperature. Since, there are no monthly GDP data; they used the money quantity that circulates in people hand and the deposit reserve call of people in bank system instead adding with the specific losses in the distribution systems. The loss values in the distribution systems of the MEA and PEA are respectively 3.64% and 5.20% of the electricity demand taking from EGAT.

- Direct customers of EGAT use the direct inquire method from electricity consumers.

- The EGAT uses combining energy values from the MEA, the PEA and direct customers adding with the loss values in the generation system and transmission system to be the energy value of the system. The specific loss is about 5.10 % of energy value of the system.

Forecasting of the peak value

- The MEA and PEA use the character of load profile of each customer to calculate energy value and adjust this value to be equal to the forecasted energy value of each type of customers. The overall peak value can be obtained by adding every load profiles of all customers.

- The direct customers of the EGAT use the principle of load profile by adjusting the load profile of each customer to be equal to the forecasted value and adding all load profiles to get the overall peak value.

- The EGAT adds all the load profile from The MEA and PEA as well as the direct customers to find the peak value. At the same time, this method can determine the peak value of the MEA, PEA and the direct customers in the same system (Coincident Peak). In all, the country's system peak generation requirement in 2011 stood at 23,900.21 megawatts (MW), occurring on May 24, 2011 at 14.00 hours. This marked a 0.46 percent decline compared with the 2010 record high peak demand level. Total energy generation requirements throughout the year decreased slightly from the prior year by 0.77 percent to 158,963.30 million kilowatt-hour (kWh). The country's power system had a total installed capacity of 31,446.72 MW, 1.70 percent higher than the previous year (2010). A greater number of electrical appliances are thus widely used, for example, refrigerators, electric fans, rice cookers, televisions, light bulbs, including air-conditioners, computers, and so on. These appliances have become necessities in Thai households. As electricity becomes a major factor in the country's economic development and people's living, energy development plans have so far focused on expansion of electrification to ensure adequate and extensive service coverage [1-2].

In Figure 1.2, fuel consumption, in term of oil equivalent, for electric generation in 2011 amounted to 32,389 ktoe, 5.3% decreased from the previous year. Natural gas consumption for electric generation in 2011 totaled 865,561 MMscf, or 2,371 MMscfd in average, decreased 15.5% over the previous year, and accounted for 65.1% of the total fuel consumption of national grid generation, Coal and lignite consumption for electric generation in 2011 totaled 24,044 thousand tons, or 66 thousand tons per day, increased 17.1% from the previous year, and accounted for 26.4% of the total fuel consumption of national grid generation, Fuel oil consumption for electric generation in 2011 was 447 million liters, increased 85.5% from the previous year, and accounted for 1.3% of the total fuel consumption of national grid generation, Renewable fuel consumption for electric generation in 2011 was 9,530 thousand tons with an average rate of 26 thousand tons per day, increased 32.9% over the previous year, and accounted for 6.1% of the total fuel consumption of national grid generation. All renewable fuel (paddy husk, bagasse, garbage, and agricultural waste) was consumed by private power producer and very small renewable energy power producers. In addition, in 2011 biogas, residual gas from production processes and black liquor consumed by cogeneration and gas engine amounted to 197,130,034 cubic meters, 709,662 gigajoules and 8,233,596 gigajoules respectively, with accounted for 1.0% of the total fuel consumption of national grid generation. Consumption, the electric consumption of end-users in national grid in 2011 was 148,700 GWh, 0.4% decreased from 2010. The fuel used in Thai electricity in 2007 and 2011 shows in Figure 1.2. Details of energy consumption demand are as follows:

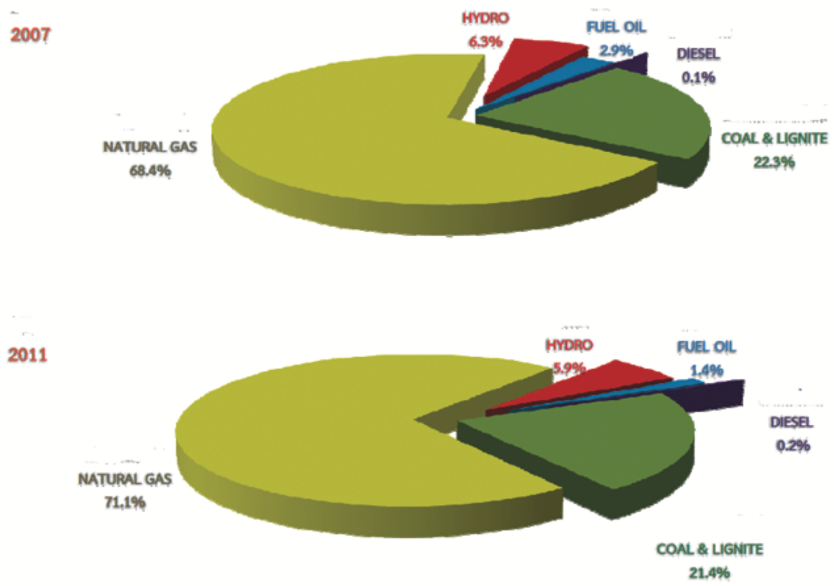

Figure 1.2: Fuel used in Thailand electricity.

1. By economic sectors

1.1 Industrial sector, among economic sectors industry was the largest electric consumer. Electric consumption in industrial sector was 63,418 GWh, decreased 0.3% from the previous year, and accounted for 43.7% of the total electric consumption for the whole country.

1.2 Commercial section (including government sector and non-profit organizations), the commercial sector consumed electricity 51,019 GWh, decreased 0.3% from the previous year, and accounted for 34.3% of the total electric consumption.

1.3 Residential sectors, the residential sector consumed electricity 32,920 GWh, decreased 1.3% from the previous year, and accounted for 22.1% of the total electric consumption.

1.4 Agricultural sector, the agricultural sector consumed electricity 304 GWh, decreased 9.5% from the previous year and accounted for 0.2% of the total of the total electric consumption.

1.5 Others, electric consumption in others (temporary consumers) in 2011 amounted to 933 GWh, increased 18.4% from the previous year, and accounted for 0.6% of the total electric consumption.

1.6 Transportation section, a small volume of electricity 106 GWh was consumed by sky train, subway and Airport Rail Link in Bangkok Metropolitan Region, increased 43.3% over the previous year.

2. By area

2.1 Bangkok Metropolitan Region (BMR) was 44,191 GWh, or 29.7% of the total consumption for the whole country, a decrease of 1.9% from the previous year.

2.2 Outside Bangkok Metropolitan Region consumed 104,509 GWh, or 70.3% of the total electric consumption, an increase of 0.2% over the previous year.

The electricity consumption by each sector of the country from 1986 to 2009 has been shown in Figure 1.3 [1-2].

Power system expansion planning starts with a forecast of anticipate future load requirements. Estimates of both demand and energy requirement are crucial to effective system planning. Demand forecasts are used to determine the capacity of generation, transmission, and distribution system additions, and energy forecasts determine the type of facilities required. Load forecasts are also used to establish procurement policies for construction capital, where, for sound operation, a balance must be maintained in the use of debt and equity capital. Further, energy forecasts are needed to determine future fuel requirement and, if necessary when fuel prices soar, rate relief to maintain an adequate rate of return. In summary, a good forecast reflecting current and future trends, tempered with good judgment, is the key to all planning, indeed to financial success. The accuracy of a forecast is crucial to any electric utility, since it dictates the timing and characteristics of major system addition. A forecast that is too low can easily result in lost revenue from sales to neighboring utilities or even in load curtailment. On the other hand, forecasts that are too high can result in severe financial problems due to excessive investment in an electric plant that is not fully utilized or, equivalently, is operated at low capacity factors. Capacity factor is defined as the ratio of average energy supplied to maximum energy capability. Unfortunately, an accurate forecast depends on the judgment of the forecaster, and it is impossible to rely strictly on analytical procedures to obtain an accurate forecast. Good judgment cannot be emphasized enough in forecasting future requirements.

For forecasting approaches, the expert systems such as the artificial intelligent (AI) are frequently used for the system that required training and making decision based on the massive data. In the past, many artificial intelligence (AI) and Expert system (Es) methods such as artificial neural network (ANN), Fuzzy logic (Fs) and Genetic algorithm (GA) are proposed for the electricity load forecasting in short-term, mid-term

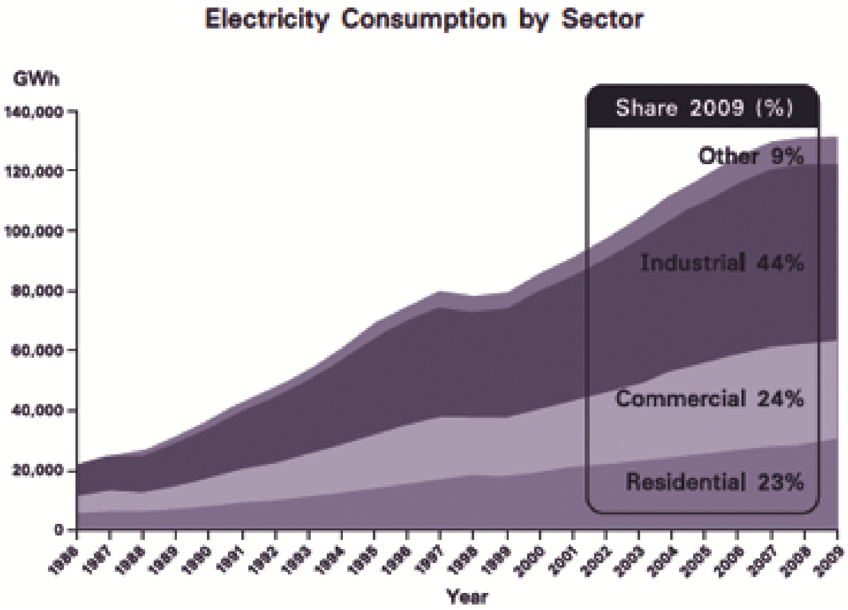

Figure 1.3: For example of electricity consumption by sector in year 2009.

and even long-term forecasting. This research focuses on Mid-term load forecasting unit in kWh that it is the most important for fuel reserve planning in power system interval month to year. The result of electricity demand forecasting can assist the operators on important decision making in fuel reserve planning and electric power generation preparation. However, for this research, there are two reasons that they are firstly to solve the demand forecasting problem of the country by using the artificial neural network and secondly to propose a new model for mid-term load forecasting in Thailand.

1.2 Objectives

1. To develop a new model approach for mid-term electricity consumption load demand forecasts by improving reliability, accuracy and also using affecting variables for prediction.

2. To study the affecting factors and analyze the dependence of factors to accuracy an accuracy of the load demand of the country.

3. To study and compare the results of the prediction by using the actual data and forecasted value based on the other related predictions, especially, from EGAT.

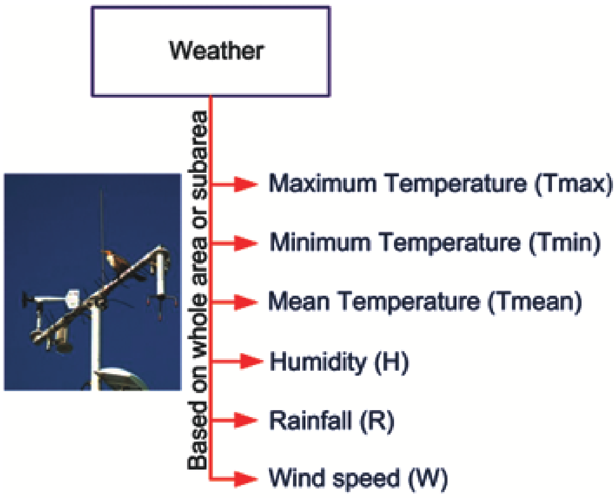

Figure 1.4: Weather factors.

1.3 Frameworks

The research proposes a mid-term load forecasting in the year of 2007 by using the historical data in ten years ago. The results of this work will be compared with the EGAT forecasting based on mean absolute percent error (MAPE) at +/- 3 percent.

1.3.1 Factor Correlations

This part will be explained about the factors that have an influence to load demand. The appropriate factors and the use of those factors for the energy consumption load demand forecasting are presented in Chapter 3. In this research, a case study employs the data information collected from whole area and substation control areas of the EGAT that the substation control areas are covered the whole area of the country [3-7]. It is necessary to analyze the appropriate factors in Figure 1.4 and Figure 1.5 before and after using them for feature inputs for forecasting by finding the significant values (sig.) between load demand and all factors. The study will rely on relative analysis of factors by using Pearson correlation for finding the appropriate relationship between factors and electricity load demand. The results obtain a coefficient of relative factors running from -1 to 1. The definition of this method shows that if there is a negative value, the correlation of variables will be in the reverse direction. On the other hand, if there is positive value, the correlation of variables will be in the accordingly identical direction. Additionally, if there is a zero value, it means that is do not have any correlation. The correlation of factors in each area may have the different values fol-

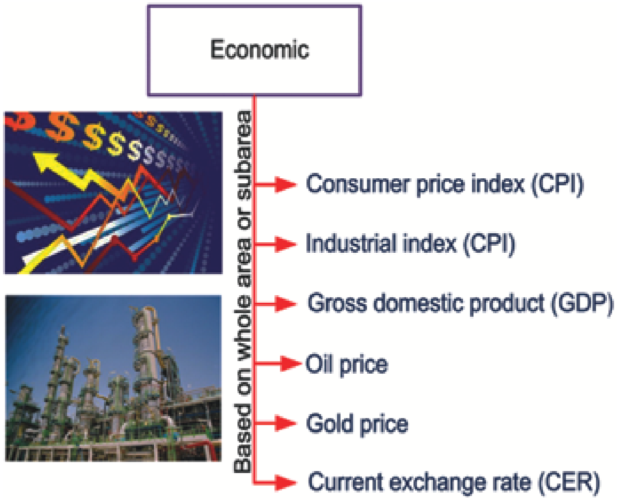

Figure 1.5: Economic factors.

lowing the vital of the economic and weather factors. Thus, the factor correlated values are important for finding the appropriate factor for load forecasting.

1.3.2 Algorithms

In this research, five forecasting models of energy consumption load demand forecasting are performed for comparison study shown in Figure 1.6. The first two statistical methods are Multiple Linear Regression (MLR) and Seasonal Autoregressive Integrated Moving Average (SARIMA). MLR is proposed to analyze the linear part with the kWh time series model. It used all factors to build a model for load forecasting. On the other hand, SARIMA is proposed to investigate the forecast problem by using seasonal method in ARIMA. This method can predict kWh by using and learning the historical demand signals of kWh without using any factor in the model. Since the electricity load demand signals are non-linear or non-stationary signals in time series, the forecasting results from MLR and SARIMA may not yet good. Next, the artificial intelligent methodology of neural network algorithm (ANN) is proposed to solve the forecasting problem of non-linear time series. The forecasting accuracy from this method is better than that of statistical methods. Although, the results of ANN are high accuracy compared to traditional method being statistical methodologies, but it is not yet satisfied the requirement from the Electricity Generating Authority of Thailand (EGAT). The last two forecasting models are included preprocessing stages. They are used for decomposing the electricity load demand signal into sub-level signals before sending to the forecasting

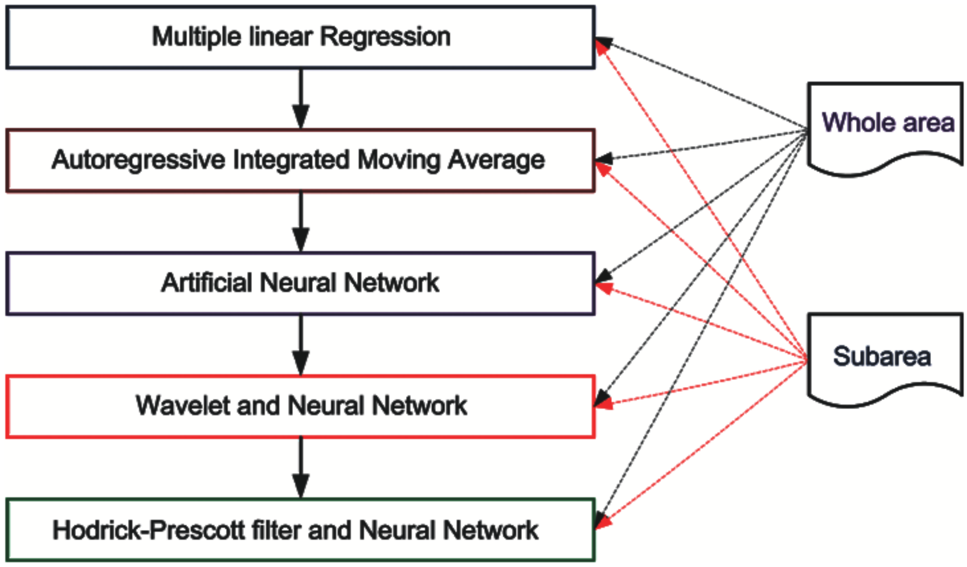

Figure 1.6: Algorithms flow in the research.

stage with ANN. First, by using wavelet transform (WT) in the preprocessing stage, the results show that the 2-level wavelet transform gives the Mean Absolute Percent Errors (MAPE) better than that other level. However, this method is complicate and some of sub-signals are useless because they are not related to any factors. In the last method, the Hodrick Prescott (HP) filter is used for decomposing electricity demand signal into trend and cyclical components. This model leads good correlations between weather factors and cyclical component as well as economic factors and trend component. As a result, the combination of HP-filter and ANN is high reliability and more accuracy than other approaches.

1.3.3 Area Separations

In Thailand, there are several factors that affect electricity appliances, income, temperature, and consumer load pattern which differ by regions and consumer groups. A reliable load forecast methodology must "correctly" gauge the effects of the key factors on electricity demand [1]. This research proposes both whole area forecasting and substation central area forecasting which composes of four regions, EGAT's power system operation has, since 1995, been divided into five areas. Metropolitan Area Operation covers Bangkok and two nearby provinces namely Nonthaburi and Samut Prakan. Central Region Operation is responsible for 27 provinces in the central part of the country; Northeastern Region Operation for 19 provinces in the Northeastern;

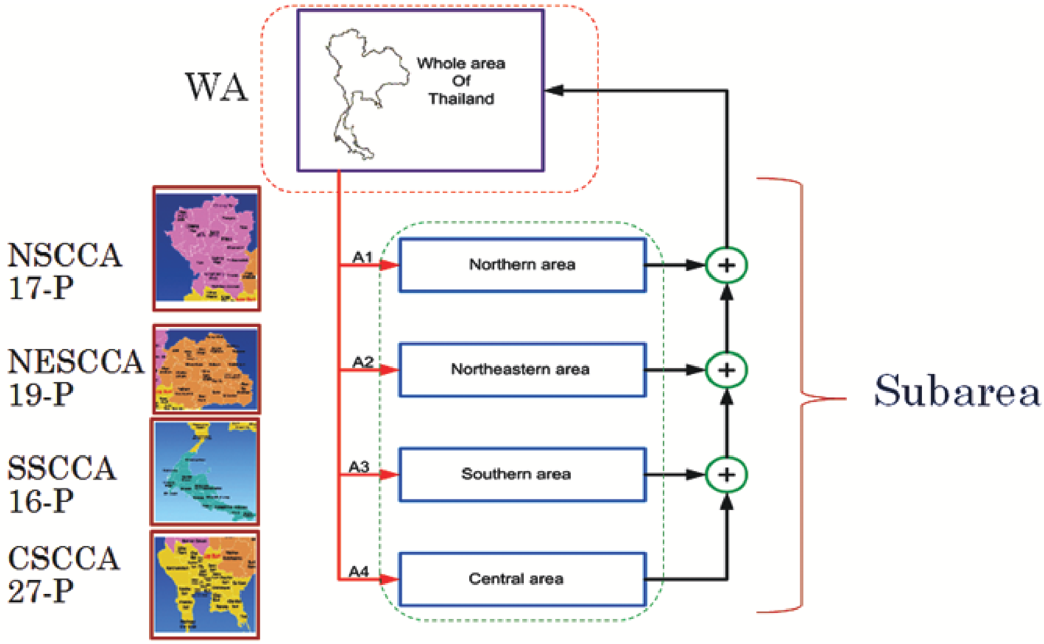

Figure 1.7: Whole and subarea of Thailand.

Southern Region Operation for 16 southern provinces and Northern Region Operation for 17 provinces in the North and the upper Central area. Each of these areas is interlinked with the transmission network and operated under the command of the National Control Centre based at EGAT Headquarters and four regional control centers to keep a fine balance between generation and consumption [3]. The subarea to whole area has been shown in Figure 1.7.

Since the accuracy depends not only on a model algorithm but also a size of forecasting area. In this research, all of the mentioned forecasting methods above are performed with 2-types of area division: whole country area and substation-control-central area. Most of the results show that the whole area forecasting have slightly better accuracy than the substation control central area forecasting.

1.3.4 Preprocessing

Preprocessing is one of the approaches which is important for this research. It uses for separating the time series of electricity load demand before finding the appropriate factors. In the research, two types of preprocessing methods are presented as shown in Figure 1.8. One is the wavelet transform and another one is HP-filter approach.

Several preprocessing approaches were used for many problems, for ex-

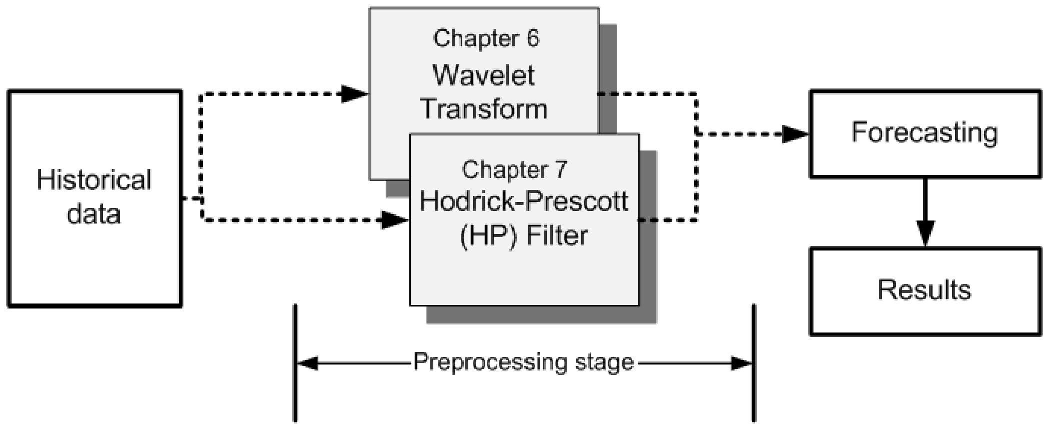

Figure 1.8: Preprocessing block diagram.

ample image or signal processing. In this research, we use preprocessing approach for finding the appropriate factors before using for feature inputs for forecasting stage. It shows that there are the differences between these algorithms such as the non-complexity and accuracy of methods. Finally, HP-filter shows both non-complexity and good accuracy of method more than wavelet transform and other approaches in this research which the preprocessing approaches are presented in chapter 6 and 7.

1.4 Contributions

Research advantages for power system are

1. A new model that can be applied to mid-term or medium term forecasting in power system.

2. The advantages of the mid-term load forecasting of the country are improved.

1.5 Research Methodology

This work studies the application of neural network (NN) with non-linear time series models to improve mid-term load demand forecasting model. Recently, ANN have widely been used and applied to load demand prediction. The main advantage of neural network is its flexibility for nonlinear time series built the models. Using hybrid model or combination of several methods has greatly become a common practice to improve and develop the forecast model which the combination of forecasts from one or more model often leads to improve the forecasting performance of models. This investigation, an approach for mid-term ECLD forecasting problems based on hybrid correct methods, which is a combination of an artificial neural network (ANN) and preprocessing approach that is a wavelet transform (WT) or HP-filter methods [10]. Wavelet transforms [9], [12], [13] and HP-filter methods are used to analyze and find the relationship between factors and demand before choosing factors for feature for neural network. For the WT-ANN approach, in preprocessing stage, is proposed to implement to the times series decomposed the data into a number of wavelet coefficient signals. Then, the decomposed signals are fed into neural network model for training. In wavelet transform decomposition, the load demand is decomposed into one to four levels for the model comparison. Also, for HP-filter approach, in preprocessing stage, is led to two major signals that are trend and detail components. Each signal is associated with some factors that will be used for the models. This approach, HP-filter in preprocessing stage, can reduce the complexity of factor separation before going to forecasting stage.

1.6 Outlines of Dissertation

This dissertation is divided into nine chapters, including "Introduction chapter and conclusion".

Chapter 2, the article reviews the historical papers and analyzes the methods, feature input, models, and the accuracy [15]. This research work done in the chapter is published in "International journal of engineering and technology, vol.2, no.1, pp.94-100 (2010)".

Chapter 3, factors and electricity consumption load demand are basically analyzed. Several factors are employed in the research because their effects to the load demand such as temperature and industrial index. In the chapter, we analyze the total factors and compute the correlated value of them. This research work done in this chapter is published in all our articles [8-15].

Chapter 4, time series in statistical methods being the tradition methodology for forecasting is proposed, for instance multiple linear regression (MLR) and seasonal autoregressive integrated moving average (SARIMA). These approaches will be used to ECLD prediction by building many models for comparison. The best model will be employ for ECLD forecasts [11]. This research work done in the chapter is published in "International conference of signal processing, pp.924-928 (2010)".

Chapter 5, the chapter presenting ANN approach is used for load forecasting in WA and SCCA patterns. Artificial neural network approach is an appropriate method for nonlinear time series as electricity load demand. ANN has widely been used and applied to load demand prediction. The main advantage of neural network is its flexibility for nonlinear time series built the models [8], [9], [10], [12], [13], [14]. This research work done in the chapter is published in "International journal of computer and electrical engineering, vol.2, no.2, pp.338-343 (2010)".

Chapter 6, WT-ANN approach to mid-term load forecasts is offered in preprocessing stage. WT is used to decompose the signal of load demand. Decomposing the data into a number of wavelet coefficient signals is used. Then, the decomposed signals are fed into neural network model for training. In wavelet transform decomposition, the load demand is decomposed into one to four levels for comparing and finding the best model [9], [12], [13]. This research work done in this chapter is published in "International journal of Soft Computing and Engineering, vol.1, no.6, pp.81-86 (2012)", / "International journal on energy procedia (International conference of advance energy engineering), vol.14, pp.438-44 (2012)", /"International conference of digital image processing, vol.7546, pp.75460B-1-9 (2009)".

Chapter 7, presenting the HP-filter and ANN is proposed to implement to the time series data. HP-filter method, in preprocessing stage, is led to two groups of the data separation, the trend and detail components. Each group is associated with factors that will be used for the models. This approach, HP-filter in preprocessing stage, can reduce the complexity of factor analysis before forecasting stage [10]. This research work done in this chapter is published in "IEEJ Transactions on Power and Energy, vol.132, issue.3, pp.235-243(2012)", /"International journal of electrical power and energy engineering, vol.44, issue 1, pp.561-570 (2013)" [16].

Finally, Chapter 8, research synthesis, all of chapter will be analyzed, discussed, and concluded in this chapter.

References

[1] http://thailand business-news.com/lifestyle/28830-electricity-generation-what-are-the-options-for-thailand

[2] http://electricitygovernance.wri.org/files/egi/egi-thailand-report-0.pdf

[3] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th

[4] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[5] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[6] Ministry of Industry Thailand, http://www.industry.go.th.

[7] Ministry of Energy, Thailand, http://www.energy.go.th.

[8] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Mid-tem load forecasting: Level suitably of wavelet and neural network based on factor selection," Energy procedia 14, pp.438-444, 2012.

[9] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Wavelet and neural network approach to demand forecasting based on whole and electric sub-control center area," IJ Soft Computing and Engineering, vol.1, issue. 6, pp.81-86, 2012.

[10] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Improving the model for energy consumption load demand forecasting," IEEJ Transactions on Power and Energy, vol.132, issue. 3, pp.235-243, 2012.

[11] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid term load forecasting of the country using statistical methodology: case study in Thailand," 2009 International conference on signal processing systems, pp.924-928, 2009.

[12] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," proceedings of SPIE second International conference on digital image processing, vol. 7546, pp.75460B-1-75460B-9, 2010.

[13] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," International conference on advance energy engineering, pp.438-444, 2011.

[14] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "The comparision of mid term load forecasting between multi-regional and whole country area using artificial neural network," International journal of computer and electrical engineering, vol. 2, no. 2, pp.338-343, 2010.

[15] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "A computing model of artificial intelligent approaches to mid-term load forecasting: a state-of-the-art-survey for the research," IACSIT International journal of engineering and technology, vol.2, no.1, pp.94-100, Feb.2010.

[16] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Multi-substation control central load area forecasting by HP-filter and double neural networks (HP-DNNs)," International journal of Electrical Power and Energy Engineering, vol.44, no.1, pp.561-570, 2013.

CHAPTER 2

Reviews of Literature

2.1 Load Forecasting

2.1.1 Definition

Load forecasting minimizes utility risk by predicting future consumption of commodities transmitted or delivered by the utility. Accurate models for electricity power load forecasting are essential to the operation and planning of a utility company. Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development. Load forecasts are extremely important for energy suppliers, ISOs, financial institutions, and other participants in electric energy generation, transmission, distribution, and markets. If the forecasted values are too high, the exceeding investment is obtained and it will push this expense to consumers. However, if the forecasted values are too low, the inadequate investment is occurred and it will cause electricity deficiency in the country. The forecasted electricity demands are defined as two values:

- The peak value (Maximum load demand: kilowatts)
- The energy value (Electric energy demand: kilowatthours)

2.1.2 Types of Load Forecasting

Load forecasts can be divided according to the forecast period into three categories:

- Short-term load forecasting
- Mid-term load forecasting
- Long-term load forecasting

In each load forecasting, period of time, forecasted values and aims of forecasting are noticeably different and they are comparably described. The forecasts for different time horizons are important for different operations within a utility company. Because of the difference of time period, forecasted values and aims of each load forecasting type, researchers in the past proposed many different algorithms and methods in order to obtain the precise load forecasting values. Next, relative papers and research topics of each load forecasting type are briefly concluded.

2.2 Important Factors for Forecasts

For short-term load forecasting several factors should be considered, such as time factors, weather data, and possible customers' classes. The medium- and long-term forecasts take into account the historical load and weather data, the number of customers in different categories, the appliances in the area and their characteristics including age, the economic and demographic data and their forecasts, the appliance sales data, and other factors.

The time factors include the time of the year, the day of the week, and the hour of the day. There are important differences in load between weekdays and weekends. The load on different weekdays also can behave differently. For example, Mondays and Fridays being adjacent to weekends, may have structurally different loads than Tuesday through Thursday. This is particularly true during the summer time. Holidays are more difficult to forecast than non-holidays because of their relative infrequent occurrence.

Weather conditions influence the load. In fact, forecasted weather parameters are the most important factors in short-term load forecasts. Various weather variables could be considered for load forecasting. Temperature and humidity are the most commonly used load predictors. Most electric utilities serve customers of different types such as residential, commercial, and industrial. The electric usage pattern is different for customers that belong to different classes but is somewhat alike for customers within each class. Therefore, most utilities distinguish load behavior on a class-by-class basis.

The summary of the important factors for forecast related to the forecasting methods will be shown in a next chapters.

2.3 Forecasting Methods

Most forecasting methods use statistical techniques or artificial intelligence algorithms such as regression, neural networks, fuzzy logic, and expert systems. Two of the methods, so-called end-use and econometric approach are broadly used for medium- and long-term forecasting. A variety of methods, which include the so-called similar day approach, various regression models, time series, neural networks, statistical learning algorithms, fuzzy logic, and expert systems, have been developed for short-term forecasting. The short review about various forecasting methods recently used for short term load forecasting will be mentioned in the first paragraph.

In [10-16], neural network for short-term load forecasting is used based on historical load and temperature input data. Moreover, some papers use additional input data from day types, humidity, wind speeds and seasons. These methods are performed in comparison with conventional methods. Training network is achieved by supervise learning and back propagation algorithm. Another technique for short-term load forecasting is using fuzzy logic and neural network [17]. This paper presents that the neuro-fuzzy method gives more accuracy results compared to one of the neural network method. In [18-19], types of input data used in fuzzy logic and neural network algorithms are historical load and weather with the case study of Electric Company in China (Hang Zhou Electric Power Company). In this paper, the principle of fuzzy rough sets is used to help neural network in forecasting. In [20], fuzzy logic with back propagation algorithm (BP) is used for short-term load forecasting in the uncertainty of the data input case. In this paper, the network composes of 51 inputs and 24 outputs and it is simulated by MATLAB. Paper [21] proposes short-term load forecasting by combining neural network and genetic algorithm with the case study in Taiwan while [22] presents the implementation of genetic algorithm method for fastening computation and increasing forecasting accuracy. The time period of this load forecast value is in 24 hours. In 2001, some of load forecasting models used the principle of wavelet decompositions to bring the more accuracy in electric load forecasting [23]. The article [24] presents a short-term load forecasting by using fuzzy logic algorithm with input data of time and temperature. The input variable liked "time" has been divided into eight triangular membership functions. The membership functions are Mid Night, Dawn, Morning, Fore Noon, After Noon, Evening, Dusk and Night. Another input variable temperature has been divided into four triangle membership functions. They are Below Normal, Normal, Above Normal and High. The forecasted load's outputs have been divided into eight triangular membership functions such as Very Low, Low, Sub Normal, Moderate Normal, Normal, Above Normal, High and Very High. In this paper, the case study has been carried out for the Neyveli Thermal Power Station Unit-II (NTPS-II) in India. The article [25], in 2005, proposes a short term load forecasting using ARIMA and ANN method based on non-linear load. It is concluded that using both methods can help each other in short-term load forecasting of the system. In 2007, the article [26] proposes a novel method approach of load forecasting by using regressive model and ANN model with the case study for Turkey. In this research, two methods are separately performed and compared. It shows that both methods give high accuracy results. In [27-29], combination of ANN, Genetic algorithm and Fuzzy logic (Fs) methods are proposed for adjusting short-term load forecasting of electric system. Genetic algorithm is used for selecting better rules and back propagation algorithm is also for this network. The paper shows that they give more accuracy results and faster processing than other forecasting methods. The article [30] proposed short-term load forecasting for holiday by using fuzzy linear regression method. The proposed algorithm shows good accuracy and the average maximum percentage error of 3.57 percent in the load forecasting of the holidays. The article [31] proposes a novel hybrid load forecasting algorithm, which combines the fuzzy linear regression method and the general exponential smoothing method with the analysis of temperature sensitivities. The article [32] proposes the development of load forecasting which combines the fuzzy logic, neural network and chaos and another algorithm. The proposed algorithm shows good accuracy or better than conventional method. The article [33] proposes an approach based on combined regression method and fuzzy inference system that developed for load forecasting. In addition, the fuzzy inference system makes a load correction inference from historical information and past forecast load errors from a multi linear regression model to infer a forecast load error. The effectiveness of the proposed approach to the short term load forecasting problem is demonstrated by practical data from the Taiwan Power Company. Paper [34] presents the development of a neuro-expert system for medium term load forecasting. Back propagation algorithm is slightly modified and is used to train the artificial neural network. The proposed algorithm is tested on the practical 66/11 kV primary distribution system of Mysore, Karnataka State, and the South India.

Table 2.1: Classified by algorithm.

Reference	Algorithm
10-16, 23, 60	Artificial neural network
17-20	Artificial neural network and Fuzzy logic
21-22	Artificial neural network and Genetic
24	Fuzzy logic
25, 29	ARIMA and Artificial neural network
26	Regression and Artificial neural network
27-29	Artificial neural network, GAs, and Fuzzy logic
30, 33	Fuzzy logic and Regression
31, 33	Hybrid approach
55	Support vector machine

Table 2.2: Classified by input (short term load forecasting).

Input	References
Historical load	10-11, 13, 17-19, 20, 23-25
Temperature	10-20, 23-25, 27-31
humidity	12, 14-16, 27-31
Rainfall	12, 14-16, 27-31
Wind speed	12, 14-16, 27-31
Seasons	12, 14-16, 27-31
Weekday-monday-friday	12, 14-16, 27-31
Weekend-saturdeay-sunday	12, 14-16, 27-31
Special day	12, 14-16, 27-31

Table 2.1 shows the summary of the methods of the load forecasting mainly classified by algorithms. The article interests do not give absolutely the importance with algorithm but they study in the part of input variables for learning algorithms to forecast in short term load forecasting such as temperature and historical load data from the past. Other input variables can be added to perform short term load forecasting as shown in Table 2.2. Data inputs have many and difficult predictions which must choose the data before the forecasting [17], [55], [42] shown in Table 2.3.

Table 2.3: Input selection (short term load forecasting).

References	Preprocessing
17, 55	Self-organizing map
13	Graphical modeling for selection input variable
57	Input dimension reduction
42	Fuzzy clustering

2.4 Mid-term Load Forecasting Explained in Details

The article preceding summarize that is the topic that bring to composed of mid term load forecasting. There are three principles as follows:

2.4.1 Data Input

The data input of this research will be shown as follows.

2.4.1.1 Input Classification

Paper [35, 36, 37, 39, and 40] presented about historical inputs and meteorologies such as maximum temperature, minimum temperature in monthly or economic variables in some articles for feature inputs for mid-term load forecasting [38]. In this research, historical load data from EGAT, temperature data from meteorological department and economic data bring from the government of Thailand are used.

2.4.1.2 Input Improvement

Paper [13, 17, and 42] proposed the improvement or input selection for preprocessing of load forecasting in conventional method, non interested input. Many researchers must lead arrangement group data technique to be suitable a problem for decreasing the complexity and over time of the network, and by this way has various the techniques such as Self-Organizing Map, Data mining etc., for classifying inputs before taking theirs to the process of MTLF forecasting.

2.4.2 Artificial Intelligence Technology Method

2.4.2.1 Expert Systems (ES)

Paper [34] proposed neuro-expert to electric mid term load forecasting for planning system by S. Chandrashekara, T. Ananthapadamanabha and A.D. Kulkarni , by using the principle of Heuristic Rules for sorting out the data not good go out and back propagation algorithm for forecasting for neural network can decrease the time and memory in training stage.

2.4.2.2 Artificial Neural Network (ANN)

Paper [34], [35] proposed artificial neural network (ANN) approach to electrical demand load, and case studies have been carried out such as The Island [34], data training for neural network learning. The interest of this research is forecasting during yearly (to 15 years), weekly (to 3 years) and hourly (to 24 hours). There are many groups of researchers for electricity mid-term load forecasting, several techniques. This method has receives many attentions from general researcher, which might bring to electricity peak load forecasting of distribution system [37]. The relationship of learning approach among the data in the past, the present and the future in the temperature, about the neural network in prediction can separate to daily peak load, total load of day and monthly electricity load consumption. The article [38] showed a development of demand forecasting using historical load, economic and population variable and using neural network, time lagged feed forward network (TLFN) to forecasting, and the forecasting is satisfy.

2.4.2.3 Fuzzy Logic (FS)

Paper [42] presented the principle of fuzzy logic coming to integrate with other approaches, for taking its merit coming to manage the data-tolerance to imprecision. Several articles had separate the data input [42] before and some articles led to forecasting again for getting the best answer [17-19], [28-29].

2.4.2.4 Genetic Algorithms

Papers [48-49] proposed the genetic algorithms for electricity load demand forecasting. It is the science that call hereditary, at pertaining to new born and mutate

Table 2.4: Accuracy for each type of inputs.

References	Types of input	Accuracy
24	Historical load, temperature	MAPE<2 percent
38	Historical load, temperature, GDP, CPI, HIS, population	Can decrease error
40	Historical load , temperature, humidity, wind speed, rainfall	MAPE<2 percent

with another algorithm. Paper [48] presented the the approach integrated with neural network for load forecasting, by Ronaldo R. B . de Aquino, Otomi Nobrega, Middl M. S . Lira A. Ferreira and Manoel. The research was hourly load forecasting based on monthly load, forecasting time in 45 days and 49 days ahead, historical data in 2005. The objective of method used the genetic algorithm to seeking the initial weight of the neural network, does not random initial weight.

2.4.2.5 Support Vector Machine (SVM)

Paper [50] proposed SVM to electricicy load forecasting, relationship the historical data in the past, the present and the future in the weather and load, in 2001, forecasting in 31 days ahead. The approach resembled with neural network, unless, the SVM method be completed more than neural network method.

2.5 Experimental and Result Analysis

Many articles have proceeding to try on a summarization in tabular brief for comparison which is no meant the best accuracy, be good way. Because the forecasting research techniques are the different objectives of the researches. The forecasting will give a significance with error which comes from the comparison between actual and forecasted demand, the definition following in equation (2.1).

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\left|\frac{X_k^R - X_k^F}{X_k^R}\right| \times 100 \tag{2.1}$$

Where X_k^R is actual load of monthly load in k-th year , is parameter of forecasted in same year, and N is number of test set (See in Table 2.5, 2.6, 2.7, and 2.8).

Table 2.5: Preprocessing.

References	Preprocessing	Algorithm	Accuracy
41	Knowledge based	ANN and Knowledge based	MAPE=2.29 percent
48	Fuzzy clustering	Fuzzy and ANN	MAPE=1.569 percent
55	Self organizing map	SVM	MAPE(w)=1.65 percent and MAPE(s)=2.42 percent

Table 2.6: Algorithms.

References	Input	Algorithm	Accuracy
24	Historical load, temperature, day types	Hybrid	MAPE=1.60 percent
37	Historical load, temperature	ANN and adjust hidden layer	two hidden layers (best result)
46	Historical load, day types	Hybrid and Genetic	MAPE=2.80 percent

Table 2.7: Filtering input.

Reference	Input	Algorithm	Output	Accuracy
	Historical load			decreased form 3.24 percent to 1.26 percent
	Weather	ANN and MA filtering	Peak Monthly load	decreased form 9.55 percent to 4.81 percent
56	Economic			
	Polution			

2.6 Summary

Mid-term load forecasting (MTLF) becomes an essential tool for power systems today, and also uses in all countries where power system operates in a deregulated environment. This objectives of load forecast have many applications such as for scheduling and developing the fuel purchasing strategies. This chapter proposes several researches using many approaches that use in demand forecasting by applying from the articles and improving the new approach to load prediction. Lastly, it has summarized appropriate approach which will be used for mid-term load forecasting.

References

[1] Energy to Energy Policy and Planning Office, Ministry of Energy, Royal Thai Government, Thailand.

[2] Energy to Energy Policy and Planning Office, "Load forecasting reports 2547," Ministry of Energy, Royal Thai Government, Thailand.

[3] Peter Vas.,"Artificial-Intelligence-Based Electrical Machines and Drives," Oxford science publication, 1999.

[4] Lefteri H.Tsoukalas, Robert E.Uhrig, "Fuzzy and Neural Approaches in Engineering," A wiley-Interscience publication, JohnWilley and Sons INC, 1997.

[5] Peter M.ills, Albert Y.Zomaya, and Moses O.Tade, "Neuro-Adaptive Process Control A Practical Approach," John Willey and Sons Ltd, 1996.

[6] R.A. Aliev, R.R.Aliev,"Soft Computing and its Application," World scientific publication, 2001.

[7] Stamatios V.Kartalopoulos.Understanding, "Neural Network and Fuzzy logic Basic Concepts and Applications," AT and T, 1996.

[8] K. Metaxiotis, A. Kagiannas, D. Askounis, J. Psarras, "Artificial intelligence in short term electric load forecasting :a state-of-the-art survey for the research," Energy Conversion and Management 44, pp.1525-1534, 2003.

[9] Georgr gross,franscisco D.galiana, "Short-term load forecasting," Preceeding of the IEEE, vol.75,no.12, pp.1558-1573, 1987.

[10] Tomonoobu senjyu, hitoshi takara,katsumi ueezato,toshihisa funabashi, "One-hour ahead load forecasting using neural network," IEEE Trans. power syst., vol.17, no.1, pp.113-118, 2002.

[11] Zhiyong Wang,Yijia Cao, "Mutual Information and Non-fixed ANNs for Daily Peak load forecasting," Power Systems Conference and Exposition, IEEE PES, pp.1523 - 1527, 2006.

[12] wenjin Dai, ping Wang, "Application of pattern recognition and artificial neural network to load forecasting in electric power system," Third international conference on natural computation, ICNC 2007.

[13] Hiroyuki mori,eitaro kurata, "Graphical Modeling for Selecting Input Variables of Short-term Load Forecasting," Power Tech, 2007 IEEE Lausanne, pp.1084 - 1089, 2007.

[14] G.A.Adepoju, S.O.A.ogaunjuyigbe, and K.O.Alawode, "Application of neural network to load forecasting in nigerian electrical power system," The pacific journal of science and technology, vol.8, no.1, pp.68-72, 2007.

[15] Mohsen Hayati, Yazdan Shirvany, "Artificial neural network approach for short term load forecasting for Illam region," International journal of electrical, computer, and system engineering, vol.1, no.2, pp.121-125.

[16] H.A.Salama, A.F.AbdElGawad, H.M.Mahmond, E.A.Mohamed, S.M.Saker, "Short term load forecasting investigations of eqyptian electrical network using ANNs," Universities Power Engineering Conference, pp.550 - 555, 2007.

[17] Bhavesh kumar chauha, amit sharma, and m.hanmandlu, "Neuro-fuzzy approach based short term electric load forecasting," 2005 IEEE/PES Transmission and Distribution Conference and Exhibition, Asia and Pacific Dalian, Chaina.

[18] Z.Y.Wang, C.X.Guo , Y.J.Cao, "A New method for short term load forecasting integrating fuzzy tough sets with artificial neural network," Power Engineering Conference, IPEC 2005, pp.1 - 173, 2005.

[19] Cuiru wang, Zhikun cui , Qi chen, "Short term load forecasting based on fuzzy neural network," Intelligent Information Technology Application, pp.335 - 338, 2007.

[20] Hari seetha and R.saravanan, "Short term electric load prediction using fuzzy BP," Journal of computing and information technology-CIT 15, pp.267-283, 2007.

[21] Chih-hsien kung, michael J.Devaney, chung-ming huang, chih-ming kung, "An adaptive power system load forecasting scheme using a genetic algorithm embedded neural network," IEEE Intrumentation and Measurement technology conference St.Paul, Minnesota, USA, May 18-20.

[22] L.L.Lai,H.Subasinghe, N.Rajkumar, E.Vaseekar, B.J.Gwyn, V.K.Sood, "Object oriented genetic algorithm based artificial neural network for load forecasting," Springer-verlag berlin, pp.462-469, 1999.

[23] Zhao-yang dong, Bai-ling zhang, Qian huang, "A daptive neural network short term load forecasting with wavelet decompositions," IEEE Porto power tech conference, 10-13 september, porto, Portugal.

[24] S.Chenthur Pandian, K.Duraiswamy, C.Christer Asir Rajan, N.Kanagaraj, "Fuzzy approach for short term load forcasting," Electric Power Systems Research 76 pp.541-548, 2006.

[25] Jian-Chang Lu, Dong-xiao niu, zheng-yuan jia, "A study of short term load forecasting based on arima-ann," Machine Learning and Cybernetics, vol.5, pp.3183 - 3187, 2004.

[26] Ummuhan basaran filik, Mehmet kurban, "A new approach for the short term load forecasting with autoregressive and artificial neural network models," International journal of computational intelligence research , vol.3, no.1, pp.66-71.

[27] P.K.Dash, S.Mishra, S.Dash, A.C.Liew, "Genetic optimization of a self organizing fuzzy-neural network for load forecasting," IEEE 2000.

[28] Gwo-ching liao, Ta-peng tsao, "Integrated genetic algorithm/Tabu search and neural fuzzy networks for short-term load forecasting," Power Engineering Society General Meeting, pp.1082 - 1087, 2004.

[29] Gwo-ching liao, Ta-peng tsao, "Novel GA-Based Approach and Neural fuzzy networks application in short-term load forecasting," Power Engineering Society General Meeting, pp.589-594, 2004.

[30] Kyung-bin song, young-sik baek, dug hun hong and gilsoo jang, "Short-term load forecasting for the holidays using fuzzy linear regression method" IEEE Trans. power syst., vol.20, no.1, pp.96-101, 2005.

[31] Kyung-bin song, seong-kwan ha, jung-wook park, dong-jin kweon, kyu-ho kim, "Hybrid load forecasting method with analysis of temperature sensitivities," IEEE Trans. power syst., vol.21, no.2, pp.869-876, 2006.

[32] Gwo-ching liao, Ta-peng tsao, "Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short term load forecasting," Evolutionary Computation, vol.10, issue 3, 2006, pp:330 - 340.

[33] R.H.Liang and C.-C.Cheng, "Combined regression-fuzzy approach for short-term load forecasiting," Generation, Transmission and Distribution, vol.147, issue 4, pp.261 - 266, 2000.

[34] Adiga S. Chandrashekara, T. Ananthapadmanabha, A. D. Kulkarni. A neuro-expert system for planning and load forecasting of distribution systems. International Journal of Electrical Power and Energy Systems, vol.2, issue 5, pp.309-314, 1999.

[35] J.V.Ringwood, D.Bofelli and F.T.Murray, "Forecasting electricity demand on short, medium and long time scales using neural networks," Journal of Intelligent and Robotic System 31, pp.129-147, 2001.

[36] Ioan Borlea, Adrian Buts, "Some aspects concerning mid term monthly load forecasting Using ANN," EUROCON , 22-24 Nov.2005.

[37] T.Yalcinoz, U.Eminoglu, "Short term and medium term power distribution load forecasting by neural networks," Energy Conversion and Management 46, pp.1393-1405, 2005.

[38] Danilo Bassi, Oscar Olivares, "Medium term electric load forecasting using TLFN neural networks," IJCCC, vol.I, no.2, pp.23-32, 2006.

[39] M.Ghiassi, David K.Zimbra, H.Saidane, "Medium term system load forecasting with a dynamic artificial neural network model," Electrical Power system Research 76, pp.302-316, 2006.

[40] M.C.Falvo, R.Lamedica, "Meteorological parameters influence for medium term load forecasting," IEEE PES Transmission and distribution conference and exhibition, pp.1296 - 1301, 2006.

[41] M.C.Falvo, R.Lamedica, "A Knowledge based system for medium term load forecasting," IEEE PES Transmission and distribution conference and exhibition,, pp.1291 - 1295, 2006.

[42] Lu Yue, Yao Zhang, Huifan Xie, Qing Zhong, "The fuzzy logic clustering neural network approach for middle and long term load forecasting, Grey Systems and Intelligent Services(GSIS 2007), pp.963 - 967, 2007.

[43] Nima Amjady, Farshid Keynia, "Mid-term load forecasting of power system by a new prediction method," Energy conversion and management (science direct), pp.1-10, 2008.

[44] E.Gonzalez-Romera,M.A.Jaramillo-Moran,and D.Carmona-Fernandez, "Monthly electric energy demand forecasting with neural networks and fourier series," Energy conversion and management (science direct), pp.1-8, 2008.

[45] R.E.Abdel-Aal, "Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks," Computer and Industrial Engineering 54, pp.903-917, 2008.

[46] Ronaldo R.B. de Aquino, Aida A.Ferrira, Manoel A.Carvalho Jr, "Development of a hybrid intelligent system for electrical load forecasting," IBERAMIA-SBIA, pp.228-237, 2006.

[47] Ronaldo R.B. de Aquino, Aida A.Ferrira, Manoel A.Carvalho Jr., "Combining Artificial neural networks and heuristic rules in a hybrid intelligent load forecast system," ICANN, pp.757-766, 2006.

[48] Ronaldo R.B.de Aquino, Otomi Nobrega Neto, Milde M.S.Lira,Aida A.Ferreira,Manoel A, "Development of an artificial neural network by genetic algorithm to mid term load forecasting," Neural Networks 2007(IJCNN 2007). International Joint Conference, pp.1726 - 1731, 2007.

[49] Ronaldo R.B.de Aquino, Otomi Nobrega Neto, Milde M.S.Lira, Aida A.Ferreira, Manoel A., "Using Genetic Algorithm to Develop a Neural network based load forecasting," ICANN, pp.738-747, 2007.

[50] [50] M Gavrilas, I Ciutea and C Tanasa, "Load Forecasting Using support vector machines," IEE 2001, pp.1-7, 2001.

[51] S.Mirasgedis, Y.Sarafidis, E.Georgopoulou, D.P. Lalas, "Models for mid-term electricity demand forecasting incorporating weather influences," energy 31, pp.208-227, 2006.

[52] Xie Da, Yu Jiangyan, Yu Jilai, "The physical series algorithm of mid-long term load forecasting," Electric power research , pp.31-37, 2000.

[53] Alexander Bruhns, Gilles Deurveilher, Jean-sebastien roy, "A non-linear regression model for mid-term load forecasting and improvements in seasonality," 15th PSCC Liege, August 2005.

[54] G.J.Tsekouras, P.B.Kotoulas, C.D.Tsirekis, E.N.Dialynas, N.D.Hatziargyriou, "A pattern recognition methodology for evaluation of load profiles and typical days of large electricity customers," Electric power systems research78, pp.1494-1510.

[55] Shu fan, Chengxiong Mao, Luonan Chen, "Electricity Peak Load Forecasting with Self-Organizing Map and Support Vector Regression," Transactions on Electrical and Electronic Engineering, pp.330-336, 2006.

[56] L.D. Voss , M.M.A. Salama , J.Reeve, "A Practical Approach to Electric Load Forecasting Using Artificial Neural Networks with Corrective Filtering," IEEE, 1995.

[57] Xu Tao, He Renmu, Wang Peng and Xu Dobgjie, "Input dimension reduction for load forecasting based on support vector machines," IEEE International Conference on Electric Utility Deregulation ,April 2004.

[58] A.A. El Desovley and M.M. ElKateb, "Hybrid adaptive techniques for electric load forecasting using ANN and ARIMA," IEE proceeding, vol.147, no.4, July 2000.

[59] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid Term Load Forecasting of the Country Using Statistical Methodology: Case study in Thailand," 2009 International Conference on Signal Processing System, pp.924-928, May 2009.

[60] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," Proceeding of SPIE, second international conference on digital image processing, Feb.26-27, 2010.

Chapter 3
Electricity Demand and Relative Factors

3.1 Electricity Load Demand

Electricity load demand is the amount of electricity being consumed at any given time. It rises and falls throughout the period of time in response to a number of things, including the time and environmental factors. For example, Figure 3.1 shows monthly electricity energy load demand of Thailand from 1997 to 2006 and the subarea shows in Figure 3.2. In the mid-term load forecasting, the electricity load demand forecasting is required. As we know the growth rate of energy consumption load demand (kWh) follows both seasonal and economic factors. However, load demands in some areas of the country are mainly affected by the internal and external economic while, in some areas, the seasonal factors obviously impact in electricity consumption. Therefore, in this research, we separate the whole country area into four subareas: northern, northeastern, southeastern, and central areas followed the substation division of the Electricity Generating Authority of Thailand (EGAT)[1]. Figure 3.3(a-d) shows the record of electricity consumption demand in each subarea from 1997 to 2006.

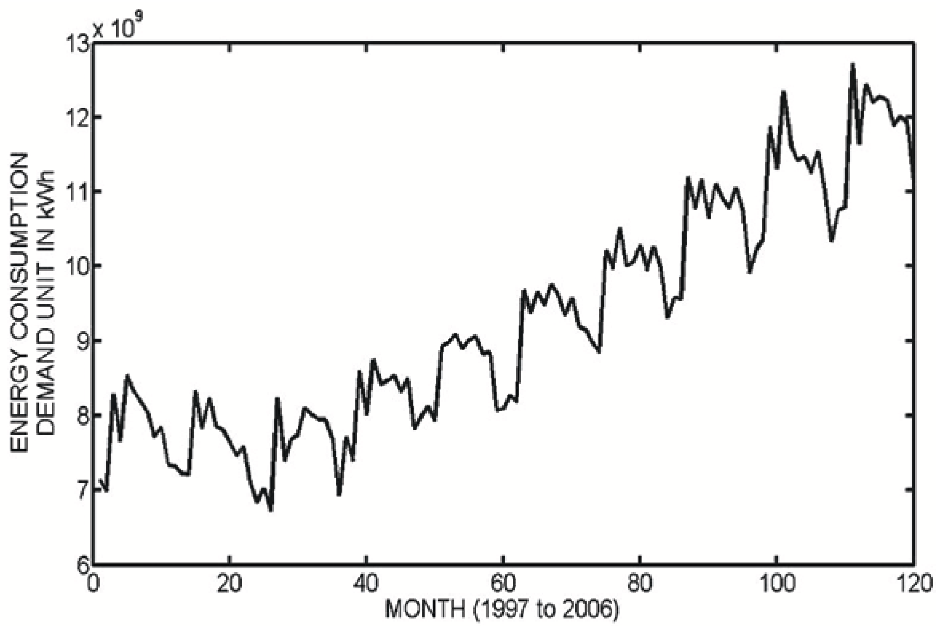

Figure 3.1: Electricity consumption load demand of the country.

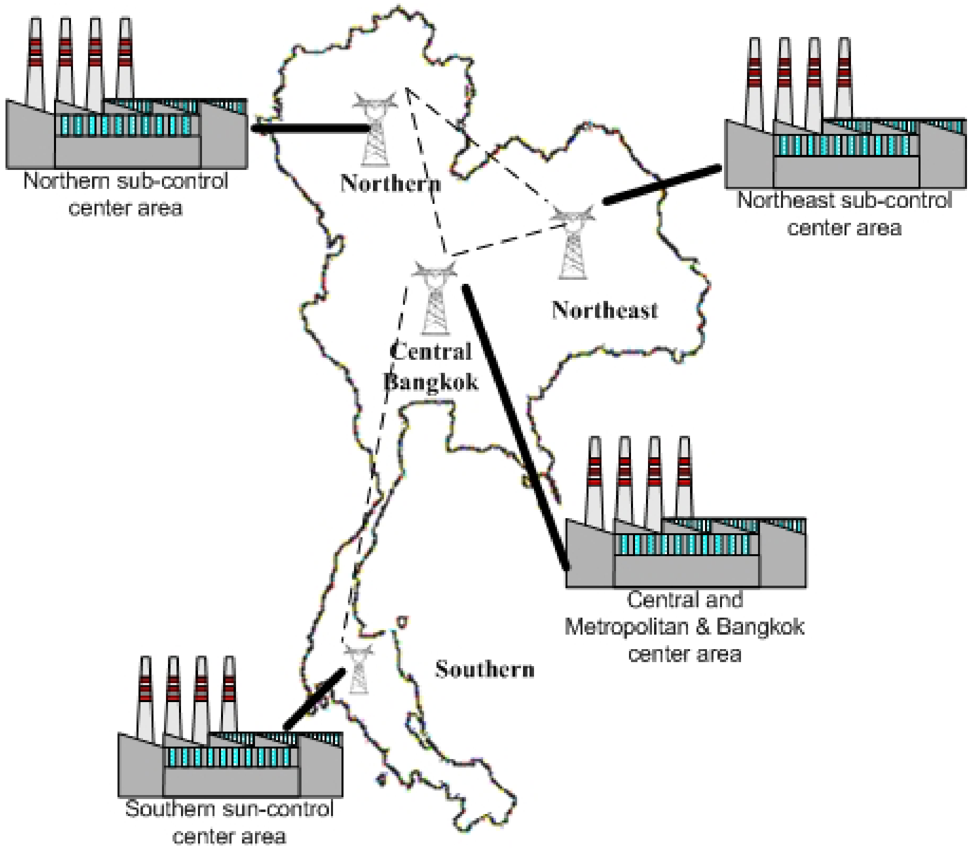

Figure 3.2: Substation control central area (SCCA) of EGAT-Thailand.

Figure 3.3(a) illustrates the behavior of the electricity demand in central area including a capital city of Bangkok. Since this area are clouded with industrial sectors and population, the load demand obviously goes after economic factors. In April 2000, a maximum demand hits about 9×10^9 kWh because the usage of air conditions increases during summer time.

Figure 3.3(b) shows the behavior of electricity demand in the northern area. In this part, the seasonal factors fairly influence on the electricity consumption demand especially in the winter time of November- February. Moreover, several economic factors may have an effect on the agriculture industrials such as flowers and fruits. The maximum energy demand is about 9×10^8 kWh in April 2006.

Figure 3.3(c) shows the behavior of the electricity demand in the north-eastern area of the country. Since this area is quite similar to the northern area in both seasonal and economic factors, the electricity consumption demand is not significantly different from the previous one. The maximum energy demand is about 12×10^8 kWh in April 2006.

Figure 3.3(d) shows the behavior of the electricity demand in the southern area of the country. The load demand behavior in this area is rather different than other areas because the weather in this region has almost rainy all year long. However, just like other areas, the weather will heat up in same period of the year. The economic factors may have an influence in important agricultures and industrials like rubber, palm and seafood. The maximum demand is about 9.9×10^8 kWh in April 2006. The electric demand hardly changes followed the weather factors.

The total electric demand of the country illustrates in Figure 3.3(e). The demand is integrated from load consumptions (kWh) in four subareas. It grows at higher rate followed the internal and external economic aspects all year. The peak of the demand is about to 12.2×10^9 kWh in 2005 and 12.8×10^9 kWh in April 2006 and it occurred in a similar period. Obviously, the behavior of electricity consumption load demand depended on two major factors; weather and economic factors. These will be discussed in detail in the next section.

3.2 Relative Factors

As we mentioned previously, the growth rate of electricity consumption load demand (kWh) follows both seasonal and economic factors. This section will be explained about the factors that have an influence to the load demand and summarized methods that can be used to identify a relationship that describes weather's and economy's impact on demand such as time factors, weather data, and possible customers' classes.

The medium-and long-term forecasts take into account the historical load and weather data, the number of customers in different categories, the appliances in the area and their characteristics including age, the economic and demographic data and their forecasts, the appliance sales data, etc. For medium or mid-term load forecasting, several factors impacting to the load demand forecasting which can be classified into two major groups; weather group [2], and economic group [3], [4]. In fact, the weather factors are the most important factors to the very short-term, the short-term, the mid-term or the long-term load demand forecasting. Weather factors consist of temperature, humidity, rainfall, and wind speed. On the other hand, the economic factors are important to the mid-term or the long-term load demand forecasting.

- Weather effect

  Weather factors play a major role in the load demand forecasting, especially mean temperature. It has been found that other weather factors are not effectively influent to the load demand. In Thailand, maximum temperatures are observed the summer time of March to June. In this research, we only focus on the monthly maximum, minimum, and mean temperatures.

  The load demand level in a power system is sensitive to weather. Hot summer conditions give rise to high peak demands. Moreover, cold winter conditions increase the use of electric heating, also creating high demand levels. An important aspect of forecasting is to study this relationship and adjust all forecasts to a similar level of "assumed weather". Generally, it is required extensive analysis of weather, and demand, and adjustment of all load histories used in forecasting to a standard set of weather conditions.

- Economic effect

  In load demand forecasting, economic factors are very important because they identify the growth of electricity consumption. For example, in 2011, Thai economy growth rate is in consistency with higher demand of energy consumption. They rise and increase in tandem. The economic influence in load demand forecasting is markedly important. Because it shows very good relationship with the load demand by using the Pearson's coefficient correlation [8], which is used to evaluate the correlation between variables. This method is one of the most familiar measures of dependence between two quantities to show how good in a linear relationship among variables or factors. In the case that the Pearson correlation is +1, it will signify a perfect linearly positive increasing correlation trend. If the Pearson correlation is -1, it will indicate a perfect linearly negative declining correlation trend. If the Pearson correlation is between 1 and -1, it will indicate a degree of linear dependence between the two variables. Lastly, if the Pearson correlation is zero, it shows that there is no relationship between the variables. Before obtaining load forecasting model: all variables must be correlated with the electricity consumption load demand in order to choose the appropriate variables to be the best feature inputs for the research model.

Table 3.1: Whole area relative factors.

Order	Factors (samples)
1	Electricity consumption load demand (kWh) in WA (ref.1)
2	Maximum temperature (T_{max}) (ref.2)
3	Minimum temperature (T_{min}) (ref.2)
4	Mean temperature (T_{mean}) (ref.2)
5	Consumer price index (CPI) (ref.3)
6	Industrial index (IDI) (ref.4)

$$r_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 . \sum (Y_i - \bar{Y})^2}} \tag{3.1}$$

$$r_{XY} = \frac{\sum (X_i Y_i) - \frac{(\sum X_i)(\sum Y_i)}{n}}{\sqrt{\left[\sum X_i^2 - \left(\frac{\sum X_i}{n}\right)^2\right]\left[\sum Y_i^2 - \left(\frac{\sum Y_i}{n}\right)^2\right]}} \tag{3.2}$$

Equation 3.1 and 3.2 are used to calculate the relationship between two variables of X and Y. r is the Pearson product moment of correlation. In this research, we compute the correlation value by using SPSS program V.11, V.12 based on Window XP and Window 7.

3.2.1 Whole Area (WA) Relative Factors

In the whole area load forecasting models, several relative factors impact to an electricity consumption load demand (ECLD) of the country are analyzed. Maximum temperatures (T_{max}), minimum temperature (T_{min}), mean temperature (T_{mean}) factors, are important factors and used for mid-term load forecasting of the country. They are recorded from Thai Meteorological Department (TMD). Other factors in economic group such as consumer price index (CPI) and industrial index (IDI) are also determined for the correlation.

3.2.2 Substation Control Central Area (SCCA) Relative Factors

In SCCA, the relative factors are separated and followed in each region based on substation control center area of EGAT. They are derived from TMD recorded in that area. The relative factors in each area will identify the behavior and trend components of ECLD of the each region.

Table 3.2: Substation control central area relative factors.

Order	Factors (samples)
1	Electricity consumption load demand (kWh) in each SCCA (ref.1)
2	Maximum temperature (T_{max}) in each SCCA (ref.2)
3	Minimum temperature (T_{min}) in each SCCA (ref.2)
4	Mean temperature (T_{mean}) in each SCCA (ref.2)
5	Consumer price index (CPI) (ref.3)
6	Industrial index (IDI) in WA (ref.4)

Since SCCA does not have IDI in each subarea, it is necessary to employ or claim the industrial index (IDI) as the same as the whole area.

3.3 Relationship between Electricity Demand (kWh) and Factors

3.3.1 Whole Area (WA) Relationship

Figure 3.5 shows scatter plots of electric load demand with relative factors such as maximum temperature, minimum temperature, mean temperature, and industrial index. In Figure 3.5 (a), maximum temperature (T_{max}), is a peak of the temperature in each month from year 1997 to 2006. They have the value interval 35 Celsius degree to 44 Celsius degree when the kWh varies between 0.7×10^{10} kWh to 1.3×10^{10} kWh. This maximum of these kWh occurred by load demand, especially, air conditioner, heater, and load motor from industrial sectors. In Figure 3.5 (b), minimum temperature (T_{min}) is the lowest of temperature in all month in the similar year. They have the value interval 7 Celsius degree to 22 Celsius degree.

In addition, Figure 3.5 (c) shows the mean temperature (T_{mean}) in each month from year 1997 to 2006, with the values interval 23 Celsius degree to 31 Celsius degree. These values came from maximum and minimum temperatures. Finally, industrial index (IDI) is shown in Figure 3.5 (d). The values of IDI increasing in every month and year from 75 percent to 190 percent are presented.

The system dynamic of the correlation [8] in Figure 3.6 shows the relationship between essential factors and load demand with segregations of weather and economic factor boundaries. In each branch shows the correlation value, such as, ELCD

(E) with maximum temperature is 0.097 (9.7 percent), and ELCD (E) with industrial index (IDI) is 0.939 (93.9 percent).This figure will be explained in the detail again in Chapter 7.

Figure 3.7 (a-d) illustrates the relationship between humidity, rainfall, consumer price index, wind speed and electricity load demand (kWh) of the country.

3.3.2 Substation Control Central Area (SCCA) Relationship

In Figure 3.8(a) illustrates scatter plots of the consumer price index (CPI) and electricity load demand (kilo-watt hour, kWh) in each area, while Figure 3.9 shows scatter plot of electricity demand with maximum temperature in NSCCA, NESCCA, SSCCA, and CSCCA of the country.

3.4 Summary

In this chapter, electric load demand and the relationship with essential factors for load forecasting are discussed. The electric load demand of Thailand can be analyzed into two ways; the whole area and the substation control central areas. The whole area load demand is used data in the whole country from 1997 to 2006. While the substation control central area, the electric load demand data is clearly carried out by 4 areas (Northern, Northeastern, Southeastern, and Central). The energy load demand data in both ways can be found the correlations with load forecasting factors involved with weather and economics. Weather factors such as Tmax, Tmin, Tmean are closely related to cyclical components of electric load demand, while economic factors like GDP, IDI, CPI are totally followed trend components. The appropriate factors will be used as input features in each mid-term load demand forecasting models described in the next chapters.

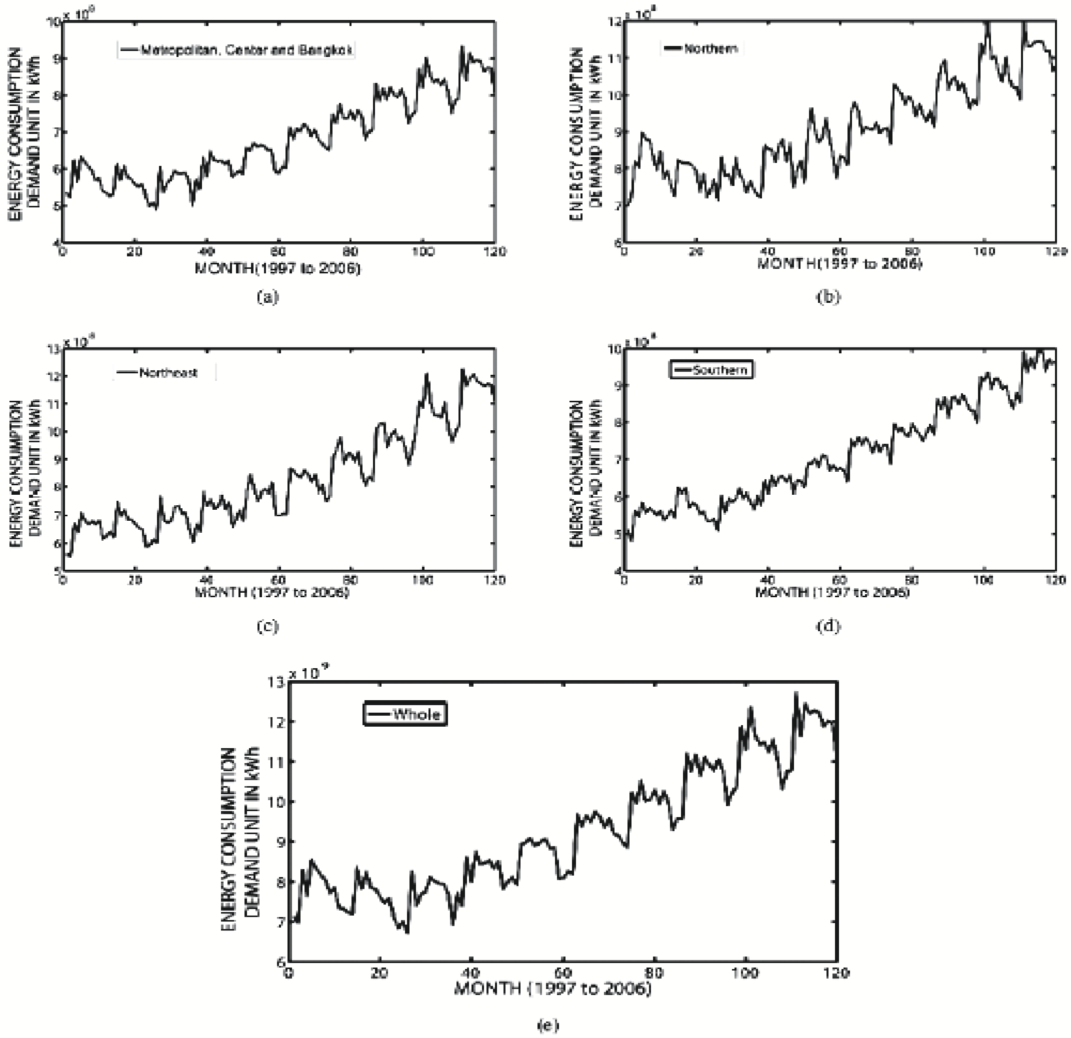

Figure 3.3: Electricity consumption load demand of a) Center b) Northern c) Northeastern d) Southern e) whole area of the country.

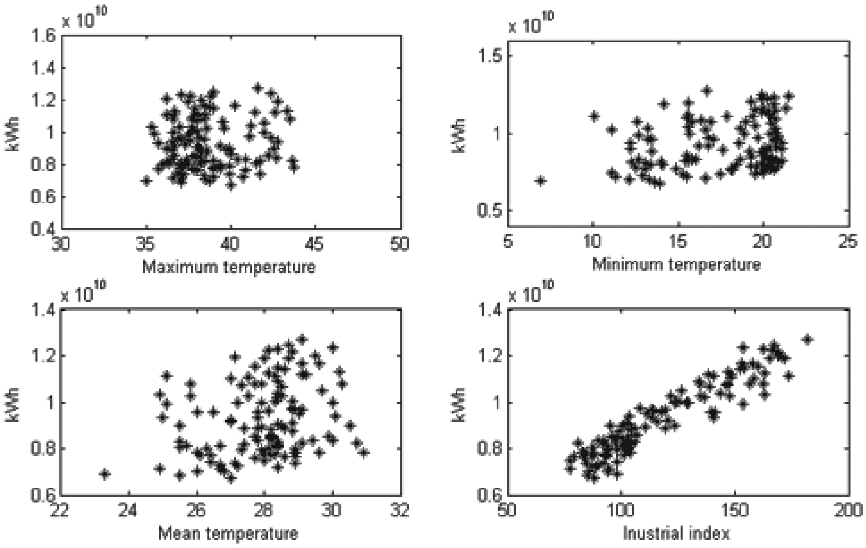

Figure 3.4: Scatter plot of electricity demand with (a) maximum temperature (b) minimum temperature (c) mean temperature (d) industrial index in whole area.

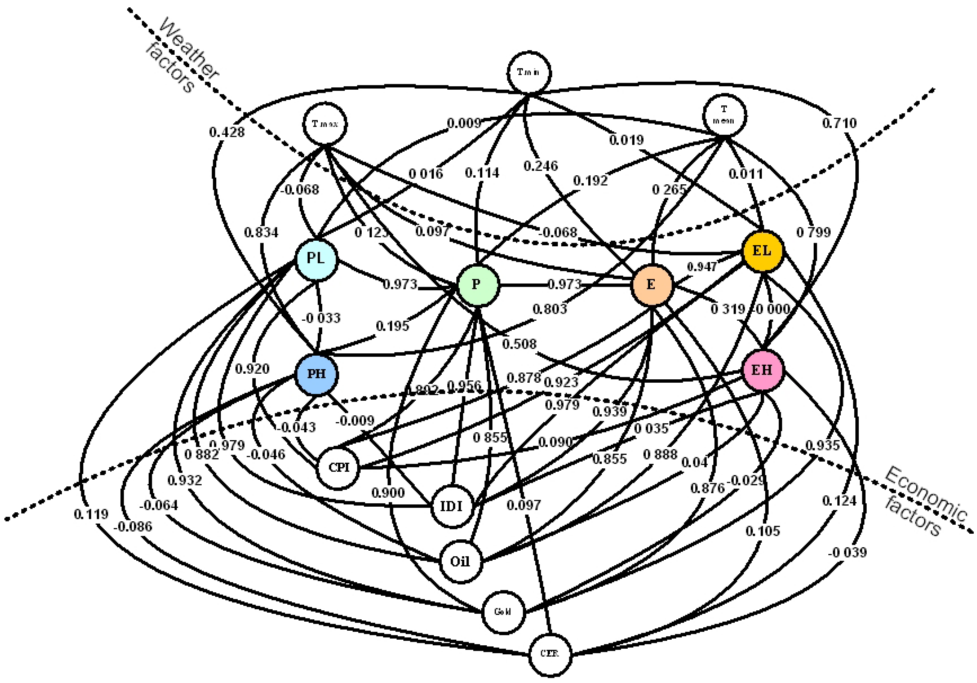

Figure 3.5: System dynamic of correlation.

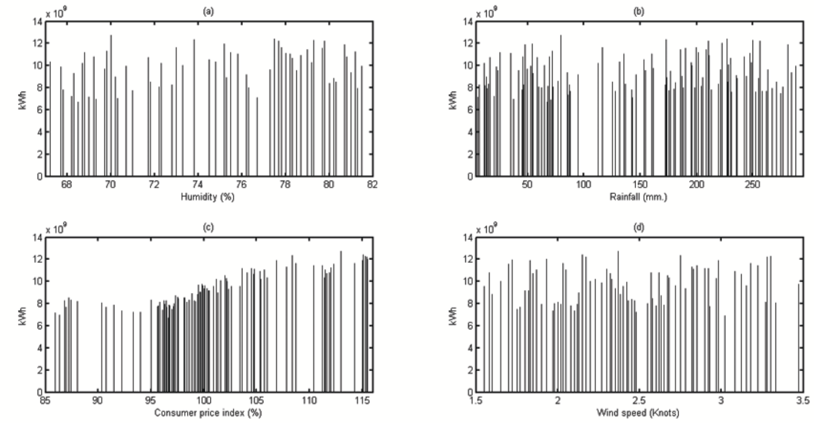

Figure 3.6: Correlations: humidity, rainfall, consumer price index, wind speed with electricity consumption load demand of the country.

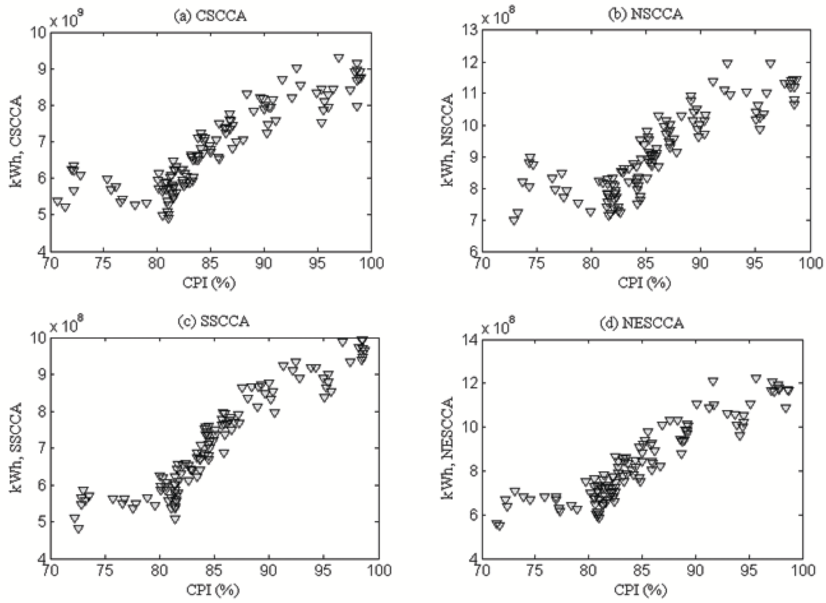

Figure 3.7: Scatter plots of electricity load demand with consumer price index in each subarea.

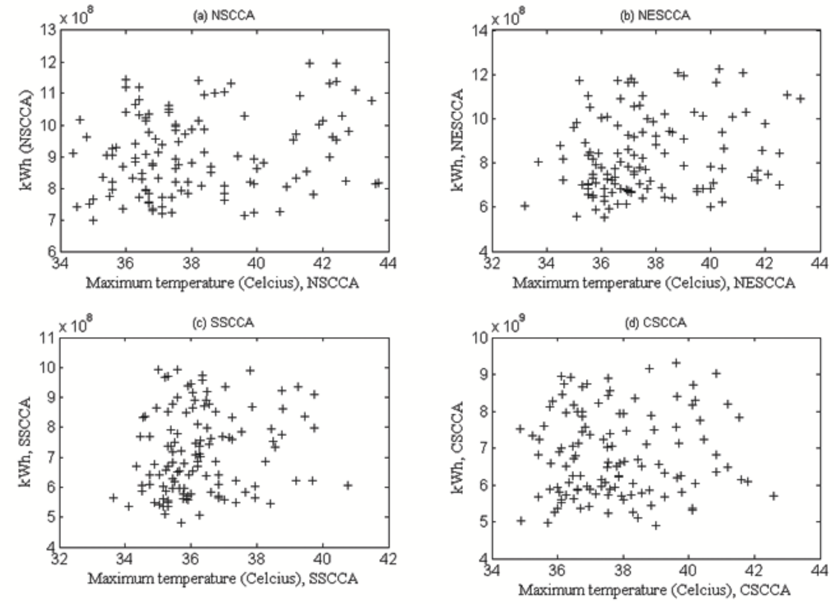

Figure 3.8: Scatter plots of electricity load demand with maximum temperature (a) NSCCA (b) NESCCA (c) SSCCA (d) CSCCA of the country, Thailand.

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th.

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," Energy procedia 14, pp.438-444, 2012.

[7] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Wavelet and neural network approach to demand forecasting based on whole and electric sub-control center area," IJ Soft Computing and Engineering, vol.1, issue. 6, pp.81-86, 2012.

[8] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Improving the model for energy consumption load demand forecasting," IEEJ Transactions on Power and Energy, vol.132," issue. 3, pp.235-243, 2012.

Chapter 4

MLR and SARIMA Approach for Electricity Load Demand Forecasting

4.1 Time Series Suggestion

In statistics a time series is a sequence of data points, measured typically at successive time instants spaced at uniform time intervals such as in signal processing, econometrics and mathematical finance. Examples of time series are the daily closing value of the stock market prediction or the annual flow volume of the any river. Time series analysis comprises of methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values [14].

Time series data have a natural temporal ordering. This makes time series analysis distinct from other common data analysis problems, in which there is no natural ordering of the observations (e.g. explaining people's wages by reference to their education level, where the individuals' data could be entered in any order). Time series analysis is also distinct from spatial data analysis where the observations typically relate to geographical locations (e.g. accounting for house prices by the locations as well as the intrinsic characteristics of the houses). A time series model will generally reflect the fact that observations close together in time will be more closely related than observations further apart. In addition, time series models will often make use of the natural one-way ordering of time so that values for a given period will be expressed as deriving in some way form past values, rather than from future values. Methods for time series analysis may be divided into two classes: frequency domain methods and time domain methods. The former include spectral analysis and recently wavelet analysis; the latter include auto-correlation and cross correlation analysis [14]. In economic, industrial, and electric load demand situations, are necessary to predict, or forecast, the value of a particular quantity over some future time period. Often, data are available describing the past behavior of the process whose future behavior we wish to predict. In order to be able to make accurate forecasts, it is desirable to be able to relate the future behavior of the process to its past behavior. In other words, we need a model of the process.

Statistical methods have many methodologies for example exponential smoothing (ES), regression (R) [11], multi-linear regression (MLR) [10], autoregressive integrated moving average (ARIMA) [7] and seasonal autoregressive integrated moving average (SARIMA). These methods are used in various models for forecasting; for example economic forecasting, energy forecasting, or load forecasting in the power system. In decade, statistical methodologies had been used for load prediction. Many researchers used them to predict electricity load demand such as [7], [8], [10], and [11].

Chapter 4 demonstrates a load forecasting based on multiple linear regression (MLR) and seasonal autoregressive integrated moving average (SARIMA) which approach to the demand forecasting based on data information in Thailand.

4.2 Multiple Linear Regression (MLR) Forecasting Models

Linear regression is a form of regression analysis in which the relationship between one or more independent variables and another variable, called the dependent variable, and it is modeled by a least squares functions, called a linear regression equation. This function is a linear combination of one or more model parameters, called regression coefficients. A linear regression equation with one independent variable represents a straight line when the predicted value is plotted against the independent variables: this is called a simple linear regression. However, note that "linear" does not refer to this straight line, but rather to the way in which the regression coefficients occur in the regression equation.

If the information is created by numerous variables, the analytical methods become very complicated. If a function is a linear function containing two or more variable, the multiple linear regression method is an excellent approach to fit engineering data. Multiple linear regressions is a statistical method to model the relationship between variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable. Multiple linear regression method can be used for forecasting under the assumption of continuing the correlation between the variables in the future. Multiple linear regressions model is as follows:

Definitions:

1. Multiple linear regression is a quantitative method of forecasting in-

volving the use of more than one variable to predict some criterion.

2. Multiple linear regression is a method of determining the relationship between a continuous process output (Y) and several factors (Xs).

3. Multiple regression differs from simple regression in that it studies the relationship between a single dependent variable and two or more independent variables.

Thus the model takes the form $y_i = \beta_1 x_{i1} + ... + \beta_p x_{ip}, \ (i = 1, ..., n)$ that the relationship between the dependent variable (y_i) and indepent (x_i), β is regression coefficients, and ε_i is called error term, disturbance, or noise[15].

4.2.1 Empirical Results of WAF

4.2.1.1 Data Sets and Models

Multiple linear regression (MLR) approach is performed by using the whole area (WA) data of EGAT, from year 1997 to 2006 [1]. The load demand given in Figure 4.1 is the time series which the detail demonstrated in the Chapter 3. In order

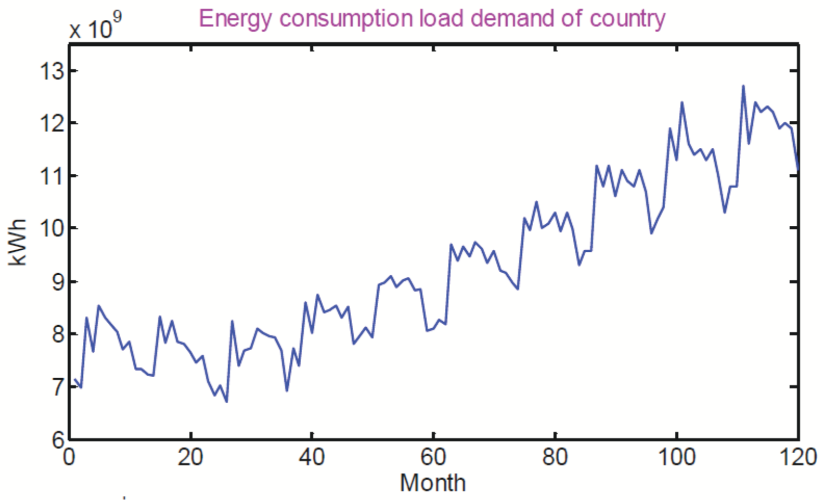

Figure 4.1: Time series of the electricity consumption load demand (ECLD) of Thailand.

to obtain a suitable statistical platform for mid-term load forecasting, MLR approach is performed in 8 different models depended on factor selection. Table 4.1 shows different factors selection in model 1 to model 8. For example model 5 is performed with all factors in both temperature and economic features, whereas model 2 uses only two factors of mean temperature and consumer price index. Table 4.2 shows the appropriate coefficients of each model.

Table 4.1: Statistical MLR forecasting models.

	Selected factor						
Model no.	T_{max}	T_{min}	T_{mean}	$Humidity$	CPI	IDI	GDP
Model 1	×	×	×		×		
Model 2			×		×		
Model 3	×	×	×			×	
Model 4			×			×	
Model 5	×	×	×	×	×	×	×
Model 6			×		×	×	×
Model 7	×	×	×		×	×	
Model 8	×	×	×	×	×	×	

$$
\begin{aligned}
Equation_W\ 1 \quad (\ell og E_t) &= 4.705 + 0.040\ell og T_{\max_t} + 0.032\ell og T_{\min_t} + 0.705\ell og T_{mean_t} \\
&\quad +2.064\ell og CPI_t \\
Equation_W\ 2 \quad (\ell og E_t) &= 4.612 + 0.841\ell og T_{\mathrm{mean}_t} + 2.064\ell og CPI_t \\
Equation_W\ 3 \quad (\ell og E_t) &= 7.612 - 0.107\ell og T_{\max_t} + 0.102\ell og T_{\min_t} + 0.694\ell og T_{mean_t} \\
&\quad +0.675\ell og IDI_t \\
Equation_W\ 4 \quad (\ell og E_t) &= 7.194 + 0.954\ell og T_{mean_t} + 0.674\ell og IDI_t \\
Equation_W\ 5 \quad (\ell og E_t) &= 5.796 + 0.016\ell og T_{\max_t} + 0.062\ell og T_{\min_t} + 0.757\ell og T_{mean_t} \\
&\quad +0.143\ell og H_t + 0.256\ell og CPI_t + 0.506\ell og IDI_t + 0.194\ell og GDP_t \\
Equation_W\ 6 \quad (\ell og E_t) &= 5.915 + 0.986\ell og T_{mean_t} + 0.354\ell og CPI_t + 0.504\ell og IDI_t \\
&\quad +0.148\ell og GDP_t \\
Equation_W\ 7 \quad (\ell og E_t) &= 7.244 - 0.069\ell og T_{\max_t} + 0.099\ell og T_{\min_t} + 0.663\ell og T_{mean_t} \\
&\quad +0.243\ell og CPI_t + 0.612\ell og IDI_t \\
Equation_W\ 8 \quad (\ell og E_t) &= 7.006 - 0.025\ell og T_{\max_t} + 0.072\ell og T_{\min_t} + 0.702\ell og T_{mean_t} \\
&\quad +0.090\ell og H_t + 0.226\ell og CPI_t + 0.616\ell og IDI_t
\end{aligned}
\tag{4.1}
$$

The model explanatory power as reflected by the adjusted R^2 and model estimation expectations of the coefficient signs are the model selection criteria. The

Table 4.2: Results of OLS regression analysis between ECLD and factors.

Independent	Coefficient values							
Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
T_{max}	0.013		-0.034		0.005		-0.022	-0.008
	(0.284)		(0.118)		(0.127)		(0.115)	(0.128)
T_{min}	0.040		0.128***		0.077		0.125***	0.090
	(0.084)		(0.035)		(0.048)		(0.034)	(0.048)
T_{mean}	0.215	0.257***	0.212***	0.219***	0.231***	0.301***	0.203***	0.214***
	(0.476)	(0.133)	(0.198)	(0.062)	(0.197)	(0.067)	(0.193)	(0.199)
H					0.052			0.033
					(0.113)			(0.112)
CPI	0.856***	0.856***			0.106***	0.147***	0.101***	0.094**
	(0.100)	(0.098)			(0.087)	(0.093)	(0.086)	(0.088)
IDI			0.939***	0.939***	0.705***	0.701***	0.852***	0.858***
			(0.012)	(0.014)	(0.056)	(0.061)	(0.025)	(0.026)
GDP					0.150**	0.114		
					(0.087)	(0.094)		
$AdjustR^2$	0.801	0.804	0.965	0.957	0.968	0.961	0.967	0.967

results of OLS regression analysis show the relationship between ECLD and variables-temperature, humidity, consumer price index, industrial index, and gross domestic product in term of beta coefficients and standard errors. (*) defines a significant value at 0.1 or 90 percent. (**) and (***) define the significant values at 0.05 (95 percent) and 0.01 (99 percent), respectively. Adjusted R^2 of each model shows in the table.

The Ordinary Least Squares (OLS) regression analysis are employed to test the equations. The data both dependent and independent were transformed by taking the logarithm (log_{10}) into each variable. Those coefficients will be placed in equation 4.1 for forecasting result determination.

4.2.1.2 Results of MLR forecasting for WA

The monthly data from January 1997 through December 2006 are used to estimate the electricity energy consumption load demand from January 1997 through December 2007. The forecasting results of ECLD from January 2007 to December 2007 are illustrated in Table 4.3 .

To evaluate the accuracy of an electricity load demand forecasting, different criterions are used. The accuracy can be obtained as a percent of error (PE) in function of the actual electricity load demand as seen in Equation (4.2). The averaged percent error can be determined by adding up all of the absolute percent errors at each

Table 4.3: Results of electricity demand forecast by using MLR for WAF.

				PE				
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	2.06	2.04	4.34	3.87	3.34	2.13	3.19	4.12
February	-4.16	-4.89	-1.04	-4.50	-3.25	-6.41	-2.27	-1.87
March	7.19	6.67	3.45	1.02	2.77	0.44	2.91	3.42
April	-4.84	-4.73	2.75	1.33	-0.33	-1.54	0.96	1.48
May	0.16	0.60	1.74	2.82	-0.14	1.28	0.52	0.71
June	-0.59	-0.39	1.04	2.29	-0.46	0.50	-0.08	0.40
July	0.44	0.49	2.88	3.77	0.76	1.71	1.72	1.98
August	2.28	2.66	-0.92	1.59	-1.31	0.52	-1.56	-0.85
September	0.24	0.64	-2.97	-0.41	-3.47	-1.35	-3.66	-3.15
October	1.02	1.64	-0.97	1.66	-2.98	-0.35	-1.89	-1.39
November	-1.56	-1.48	-0.92	-0.39	-2.37	-2.32	-2.09	-1.39
December	-2.02	-2.04	-2.46	-1.87	-3.42	-3.72	-3.50	-2.57
MAPE	2.21	3.36	2.12	2.13	2.05	1.86	2.03	1.94

time period point and divided them by the number of time points as seen in Equation (4.3). It is called the Mean Absolute Percent Error (MAPE).

$$PE = \frac{(Actual\,elctricity\,demand - Forecated\,value)}{Actual\,demand} \times 100 \tag{4.2}$$

$$MAPE = \frac{1}{N}\sum_{1}^{N} \frac{|(Actual\,elctricity\,demand_N - Forecated\,value_N)|}{Actual\,elctric\,demand_N} \times 100 \tag{4.3}$$

Where N is a number of months for forecasting, $MAPE$ is the mean absolute percent error and PE is the percent error.

The percent error (PE) and mean absolute percent error ($MAPE$) in some months, the forecasted value of the electricity demand are illustrated in negative values whereas some months are shown in positive.

From Table 4.3, the maximum $MAPE$ of 3.36 % is given by model 2 while the minimum $MAPE$ of 1.86 % occurred in model 6 of which 4 selected factors of mean temperature, consumer price index, industrial index, and gross domestic product are used. Figure 4.2 shows the electricity load demand forecasting results given by model 6 in comparison with the actual load demand.

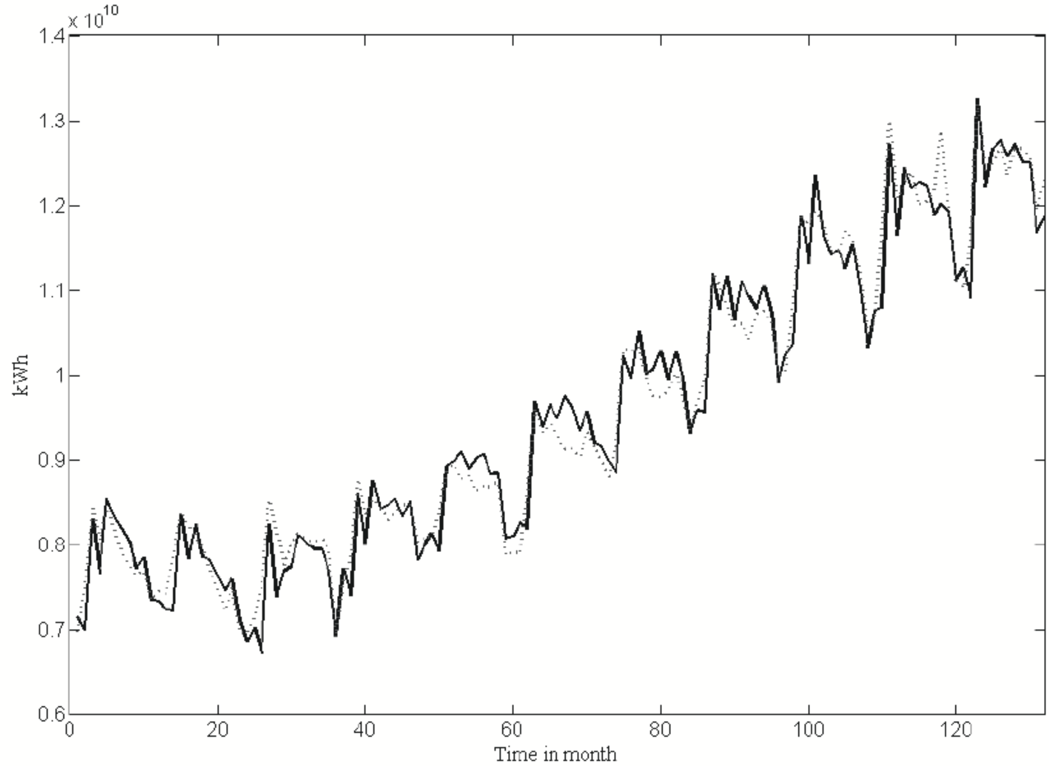

Figure 4.2: ECLD forecasting result based on model (dot)and actual demand (dash).

4.2.2 Empirical Results of SCCAF

4.2.2.1 Data Sets and Models

The SCCAF proposes the models of MLR using data of substation control central areas of EGAT recorded from year 1997 to 2007 [1]. The electricity consumption load demand (ECLD) of each area (NSCCA, NESCCA, SSCCA and CSCCA) is given in Figure 3.3(a-d) in Chapter 3. Each model of SCCAF is obtained the weather and economic factors from different regions of Thailand which demonstrates in previous section [2-4]. SCCAF using MLR will be performed with the 8 models of factors selection similarly obtained in WAF.

By using the same procedure, the results of OLS regression analysis of NSCCA are summarized in Table 4.4. The model explanatory power as reflected by the adjusted R^2 and model estimation expectations of the coefficient signs are the model selection criteria. The table shows eight models of multiple linear regressions and selected factors of each model such as maximum temperature, minimum temperature, mean tem-

Table 4.4: Results of OLS regression analysis of NSCCA.

Independent	Coefficient values							
Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
T_{max}	0.165**		0.060		0.170**		0.066	0.153**
	(0.176)		(0.106)		(0.155)		(0.107)	(0.161)
T_{min}	0.129		0.097		0.015		0.097	0.054
	(0.030)		(0.018)		(0.019)		(0.018)	(0.020)
T_{mean}	0.069	0.278***	0.233***	0.339***	0.257***	0.382***	0.215***	0.189***
	(0.213)	(0.077)	(0.128)	(0.046)	(0.130)	(0.051)	(0.128)	(0.130)
H					0.132**			0.099
					(0.080)			(0.082)
CPI	0.817***	0.816***			0.065	0.095	0.062	0.034
	(0.089)	(0.090)			(0.116)	(0.113)	(0.115)	(0.119)
IDI			0.906***	0.910***	0.503***	0.505***	0.851***	0.874***
			(0.016)	(0.016)	(0.076)	(0.076)	(0.035)	(0.036)
GDP					0.361***	0.337***		
					(0.118)	(0.118)		
$AdjustR^2$	0.753	0.748	0.912	0.911	0.920	0.917	0.912	0.913

perature, humidity, consumer price index, industrial index, and gross domestic product.

$$
\begin{aligned}
&Equation_N\ 1 \quad (\ell ogE_t) = 5.073 + 0.376\ell ogT_{\max_t} + 0.039\ell ogT_{\min_t} + 0.115\ell ogT_{mean_t} \\
&\qquad\qquad\qquad +1.594\ell ogCPI_t \\
&Equation_N\ 2 \quad (\ell ogE_t) = 5.213 + 0.466\ell ogT_{\mathrm{mean}_t} + 1.592\ell ogCPI_t \\
&Equation_N\ 3 \quad (\ell ogE_t) = 7.061 + 0.136\ell ogT_{\max_t} + 0.029\ell ogT_{\min_t} + 0.373\ell ogT_{mean_t} \\
&\qquad\qquad\qquad +0.541\ell ogIDI_t \\
&Equation_N\ 4 \quad (\ell ogE_t) = 7.026 + 0.567\ell ogT_{mean_t} + 0.544\ell ogIDI_t \\
&Equation_N\ 5 \quad (\ell ogE_t) = 4.224 + 0.387\ell ogT_{\max_t} + 0.005\ell ogT_{\min_t} + 0.431\ell ogT_{mean_t} \\
&\qquad\qquad +0.178\ell ogH_t +\ 0.127\ell ogCPI_t + 0.301\ell ogIDI_t + 0.389\ell ogGDP_t \\
&Equation_N\ 6 \quad (\ell ogE_t) = 4.913 + 0.639\ell ogT_{mean_t} + 0.186\ell ogCPI_t + 0.302\ell ogIDI_t \\
&\qquad\qquad\qquad +0.363\ell ogGDP_t \\
&Equation_N\ 7 \quad (\ell ogE_t) = 6.891 + 0.151\ell ogT_{\max_t} + 0.029\ell ogT_{\min_t} + 0.360\ell ogT_{mean_t} \\
&\qquad\qquad\qquad +0.121\ell ogCPI_t + 0.509\ell ogIDI_t \\
&Equation_N\ 8 \quad (\ell ogE_t) = 6.484 + 0.348\ell ogT_{\max_t} + 0.016\ell ogT_{\min_t} + 0.316\ell ogT_{mean_t} \\
&\qquad\qquad\qquad +0.133\ell ogH_t\ + 0.066\ell ogCPI_t + 0.522\ell ogIDI_t
\end{aligned} \tag{4.4}
$$

Table 4.4 presents the results of OLS regression analysis of the relationship between ECLD of NSSCA and variables: temperature, humidity, consumer price index, industrial index, and gross domestic product. In the table shows a beta coefficient

Table 4.5: Results of OLS regression analysis of NESCCA.

Independent	Coefficient values							
Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
T_{max}	0.049		-0.020		0.028		0.001	0.013
	(0.218)		(0.139)		(0.132)		(0.101)	(0.136)
T_{min}	-0.038		0.007		-0.043		-0.009	-0.018
	(0.044)		(0.028)		(0.022)		(0.020	(0.022)
T_{mean}	0.265**	0.264***	0.333***	0.327***	0.337***	0.329***	0.313***	0.308***
	(0.286)	(0.081)	(0.182)	(0.051)	(0.132)	(0.041)	(0.131)	(0.135)
H					0.045			0.016
					(0.075)			(0.075)
CPI	0.893***	0.892***			0.329***	0.333***	0.323***	0.320***
	(0.088)	(0.088)			(0.086)	(0.083)	(0.087)	(0.088)
IDI			0.932***	0.931***	0.463***	0.482***	0.646***	0.649***
			(0.018)	(0.018)	(0.059)	(0.057)	(0.027)	(0.028)
GDP					0.187***	0.163***		
					(0.095)	(0.090)		
$AdjustR^2$	0.882	0.880	0.952	0.952	0.977	0.977	0.975	0.975

and standard error in all models. (*) is meant a significant value at 0.1 or 90 %. Also (**), (***) are meant the significant values at 0.05 (95 %) and 0.01 (99 %), respectively. Adjusted R^2 of each model are also shown in the Table.

$Equation_N1$ to $Equation_N8$, the coefficient values are used by OLS regression method in SPSS program. Then all models are used to forecast ECLD of the northern substation control central area of the country based on equation (4.4). The forecasting results are illustrated in Table 4.8.

$$
\begin{aligned}
&Equation_{NE}\ 1 \quad (\ell og E_t) = 2.877 + 0.178 \ell og T_{\max_t} - 0.019 \ell og T_{\min_t} + 0.679 \ell og T_{mean_t} \\
&\qquad\qquad + 2.492 \ell og CPI_t \\
&Equation_{NE}\ 2 \quad (\ell og E_t) = 3.147 + 0.675 \ell og T_{mean_t} + 2.489 \ell og CPI_t \\
&Equation_{NE}\ 3 \quad (\ell og E_t) = 6.111 - 0.071 \ell og T_{\max_t} + 0.004 \ell og T_{\min_t} + 0.852 \ell og T_{mean_t} \\
&\qquad\qquad + 0.818 \ell og IDI_t \\
&Equation_{NE}\ 4 \quad (\ell og E_t) = 6.027 + 0.837 \ell og T_{mean_t} + 0.817 \ell og IDI_t \\
&Equation_{NE}\ 5 \quad (\ell og E_t) = 2.993 + 0.100 \ell og T_{\max_t} + 0.021 \ell og T_{\min_t} + 0.863 \ell og T_{mean_t} \\
&\qquad + 0.101 \ell og H_t + 0.918 \ell og CPI_t + 0.406 \ell og IDI_t + 0.297 \ell og GDP_t \\
&Equation_{NE}\ 6 \quad (\ell og E_t) = 3.514 + 0.841 \ell og T_{mean_t} + 0.929 \ell og CPI_t + 0.423 \ell og IDI_t \\
&\qquad\qquad + 0.259 \ell og GDP_t \\
&Equation_{NE}\ 7 \quad (\ell og E_t) = 4.858 + 0.004 \ell og T_{\max_t} - 0.005 \ell og T_{\min_t} + 0.802 \ell og T_{mean_t} \\
&\qquad\qquad + 0.900 \ell og CPI_t + 0.567 \ell og IDI_t \\
&Equation_{NE}\ 8 \quad (\ell og E_t) = 4.759 + 0.047 \ell og T_{\max_t} - 0.009 \ell og T_{\min_t} + 0.788 \ell og T_{mean_t} \\
&\qquad\qquad + 0.035 \ell og H_t + 0.893 \ell og CPI_t + 0.570 \ell og IDI_t
\end{aligned} \tag{4.5}
$$

Table 4.5 presents the results of OLS regression analysis of the relationship between ECLD of NESSCA and variables: temperature, humidity, consumer price index, industrial index, and gross domestic product. In the table shows a beta coefficient and standard error in all models. (*) is meant a significant value at 0.1 or 90 %. Also (**), (***) are meant the significant values at 0.05 (95 %) and 0.01 (99 %), respectively. Adjusted R^2 of each model are also shown in the Table.

$Equation_N 1$ to $Equation_N 8$, the coefficient values are used by OLS regression method in SPSS program. Then all models are used to forecast ECLD of the northeastern substation control central area of the country based on equation (4.5). The forecasting results are illustrated in Table 4.9.

Table 4.6: Results of OLS regression analysis of SSCCA.

Independent	Coefficient values							
Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
T_{max}	0.030		-0.121**		-0.070		-0.080*	-0.068
	(0.366)		(0.273)		(0.235)		(0.213)	(0.233)
T_{min}	0.009		0.066**		0.028		0.041	0.030
	(0.074)		(0.054)		(0.052)		(0.042)	(0.051)
T_{mean}	0.130	0.148***	0.229***	0.162***	0.215***	0.160***	0.207***	0.210***
	(0.515)	(0.193)	(0.383)	(0.160)	(0.320)	(0.137)	(0.296)	(0.298)
H					0.021			0.020
					(0.136)			(0.134)
CPI	0.921***	0.942***			0.360***	0.392***	0.362***	0.358***
	(0.082)	(0.071)			(0.107)	(0.108)	(0.103)	(0.105)
IDI			0.960***	0.964***	0.614***	0.606***	0.635***	0.636***
			(0.020)	(0.021)	(0.073)	(0.075)	(0.034)	(0.034)
GDP					0.021	-0.008		
					(0.113)	90.114)		
$AdjustR^2$	0.875	0.897	0.932	0.922	0.959	0.955	0.959	0.959

$$
\begin{aligned}
&Equations\ 1 \quad (\ell og E_t) = 2.907 + 0.156\ell og T_{\max_t} + 0.014\ell og T_{\min_t} + 0.816\ell og T_{mean_t} \\
&\qquad\qquad +2.332\ell og CPI_t \\
&Equations\ 2 \quad (\ell og E_t) = 2.865 + 0.998\ell og T_{\mathrm{mean}_t} + 2.352\ell og CPI_t \\
&Equations\ 3 \quad (\ell og E_t) = 5.986 - 0.618\ell og T_{\max_t} + 0.109\ell og T_{\min_t} + 1.432\ell og T_{mean_t} \\
&\qquad\qquad +0.783\ell og IDI_t \\
&Equations\ 4 \quad (\ell og E_t) = 5.794 + 1.013\ell og T_{mean_t} + 0.771\ell og IDI_t \\
&Equations\ 5 \quad (\ell og E_t) = 4.245 - 0.357\ell og T_{\max_t} + 0.046\ell og T_{\min_t} + 1.348\ell og T_{mean_t} \\
&\qquad\qquad +0.092\ell og H_t + 0.911\ell og CPI_t + 0.501\ell og IDI_t + 0.031\ell og GDP_t \\
&Equations\ 6 \quad (\ell og E_t) = 4.535 + 1.001\ell og T_{mean_t} + 0.993\ell og CPI_t + 0.494\ell og IDI_t \\
&\qquad\qquad -0.012\ell og GDP_t \\
&Equations\ 7 \quad (\ell og E_t) = 4.692 - 0.411\ell og T_{\max_t} + 0.067\ell og T_{\min_t} + 1.297\ell og T_{mean_t} \\
&\qquad\qquad +0.915\ell og CPI_t + 0.518\ell og IDI_t \\
&Equations\ 8 \quad (\ell og E_t) = 4.446 - 0.350\ell og T_{\max_t} + 0.049\ell og T_{\min_t} + 1.316\ell og T_{mean_t} \\
&\qquad\qquad +0.087\ell og H_t + 0.906\ell og CPI_t + 0.519\ell og IDI_t
\end{aligned} \tag{4.6}
$$

Table 4.6 presents the results of OLS regression analysis of the relationship between ECLD of SSCCA and variables: temperature, humidity, consumer price index, industrial index, and gross domestic product. In the table shows a beta coefficient and standard error in all models. (*) is meant a significant value at 0.1 or 90 %. Also

Table 4.7: Results of OLS regression analysis of CSCCA.

Independent	Coefficient values							
Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
T_{max}	-0.077		-0.123**		-0.87**		-0.115**	-0.100**
	(0.322)		(0.141)		(0.151)		(0.140)	(0.146)
T_{min}	0.044		0.079*		0.037		0.078*	0.038
	(0.066)		(0.029)		(0.038)		(0.029)	(0.038)
T_{mean}	0.283**	0.258***	0.319***	0.286***	0.313***	0.288***	0.312***	0.317***
	(0.444)	(0.134)	(0.196)	(0.068)	(0.194)	(0.073)	(0.194)	(0.194)
H					0.055			0.043
					(0.094)			(0.091)
CPI	0.847***	0.859***			0.071*	0.124***	0.077**	0.063
	(0.087)	(0.085)			(0.086)	(0.095)	(0.081)	(0.085)
IDI			0.928***	0.941***	0.783***	0.787***	0.861***	0.877
			(0.013)	(0.015)	(0.061)	(0.069)	(0.028)	(0.029)
GDP					0.093	0.046		
					(0.096)	(0.104)		
$Adjust R^2$	0.807	0.799	0.961	0.948	0.962	0.950	0.962	0.962

(**), (***) are meant the significant values at 0.05 (95 %) and 0.01 (99 %), respectively. Adjusted R^2 of each model are also shown in the Table.

$Equation_N 1$ to $Equation_N 8$, the coefficient values are used by OLS regression method in SPSS program. Then all models are used to forecast ECLD of the southern substation control central area of the country based on equation (4.6). The forecasting results are illustrated in Table 4.10.

$$
\begin{aligned}
Equation_C\ 1 \quad (\ell ogE_t) &= 5.482 - 0.271\ell ogT_{\max_t} + 0.032\ell ogT_{\min_t} + 0.921\ell ogT_{mean_t} \\
&\quad +1.759\ell ogCPI_t \\
Equation_C\ 2 \quad (\ell ogE_t) &= 5.162 + 0.838\ell ogT_{\mathrm{mean}_t} + 1.785\ell ogCPI_t \\
Equation_C\ 3 \quad (\ell ogE_t) &= 7.571 - 0.433\ell ogT_{\max_t} + 0.057\ell ogT_{\min_t} + 1.037\ell ogT_{mean_t} \\
&\quad +0.660\ell ogIDI_t \\
Equation_C\ 4 \quad (\ell ogE_t) &= 7.0970.929\ell ogT_{mean_t} + 0.669\ell ogIDI_t \\
Equation_C\ 5 \quad (\ell ogE_t) &= 6.401 - 0.305\ell ogT_{\max_t} + 0.027\ell ogT_{\min_t} + 1.018\ell ogT_{mean_t} \\
&\quad +0.139\ell ogH_t + \ 0.147\ell ogCPI_t + 0.557\ell ogIDI_t + 0.119\ell ogGDP_t \\
Equation_C\ 6 \quad (\ell ogE_t) &= 6.468 + 0.936\ell ogT_{mean_t} + 0.258\ell ogCPI_t + 0.560\ell ogIDI_t \\
&\quad +0.058\ell ogGDP_t \\
Equation_C\ 7 \quad (\ell ogE_t) &= 7.351 - 0.403\ell ogT_{\max_t} + 0.056\ell ogT_{\min_t} + 1.013\ell ogT_{mean_t} \\
&\quad +0.159\ell ogCPI_t \ + 0.612\ell ogIDI_t \\
Equation_C\ 8 \quad (\ell ogE_t) &= 7.116 - 0.353\ell ogT_{\max_t} + 0.027\ell ogT_{\min_t} + 1.031\ell ogT_{mean_t} \\
&\quad +0.107\ell ogH_t + 0.130\ell ogCPI_t + 0.624\ell ogIDI_t
\end{aligned}
\tag{4.7}
$$

Table 4.7 presents the results of OLS regression analysis of the relationship between ECLD of CSSCA and variables: temperature, humidity, consumer price index, industrial index, and gross domestic product. In the table shows a beta coefficient and standard error in all models. (*) is meant a significant value at 0.1 or 90 %. Also (**), (***) are meant the significant values at 0.05 (95 %) and 0.01 (99 %), respectively. Adjusted R^2 of each model are also shown in the Table.

$Equation_N 1$ to $Equation_N 8$, the coefficient values are used by OLS regression method in SPSS program. Then all models are used to forecast ECLD of the central substation control central area of the country based on equation (4.7). The forecasting results are illustrated in Table 4.11.

4.2.2.2 Results

In this study, SCCAF using MLR are analyzed and results of each model (NSCCAF, NESCCAF, SSCCAF and CSCCAF) are explained in Table 4.8, 4.9, 4.10, and 4.11 in terms of the percent error (PE) and mean absolute percent error (MAPE) of ECLD of each substation control central area. Each table especially show only the forecasted results from January to December 2007.

Table 4.8: Northern substation control central area forecasting.

	PE							
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	0.384	1.828	2.551	3.547	2.181	3.185	1.294	2.498
February	-4.032	-2.680	-3.088	-2.893	-4.873	-3.765	-4.354	-3.482
March	7.900	8.265	2.710	2.537	4.197	3.082	1.873	4.035
April	1.730	3.369	5.927	6.356	3.446	4.265	4.538	5.757
May	0.164	0.568	0.637	1.040	-1.827	1.224	-0.598	-1.212
June	3.295	2.078	3.568	3.504	2.561	3.102	2.375	3.297
July	4.046	3.480	5.405	5.746	3.443	4.700	4.172	4.806
August	7.727	6.372	5.696	5.653	5.074	5.865	4.676	5.332
September	4.543	2.013	1.134	0.712	1.622	1.375	0.140	1.345
October	1.842	-0.124	-0.533	-0.611	-2.388	-2.316	-1.629	-0.314
November	0.512	-1.978	-0.184	-0.872	-1.273	-1.821	-1.412	-0.041
December	-0.448	0.293	0.070	0.983	-0.481	0.529	-1.215	-0.036
MAPE	3.051	2.753	2.625	2.871	2.780	2.935	2.356	2.679

Table 4.9: Northeastern substation control central area forecasting.

	PE							
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	3.310	3.647	10.895	10.865	6.142	5.691	6.650	6.378
February	-5.447	-3.048	0.626	0.065	-5.133	-4.797	-3.296	-3.646
March	5.795	7.279	3.853	3.437	1.858	1.642	2.478	2.116
April	-5.571	-3.429	8.236	7.644	0.215	0.113	2.066	1.705
May	-2.515	-1.897	3.712	3.553	-1.525	-1.047	-0.349	-1.062
June	-1.884	-1.833	5.040	5.121	-0.589	-0.287	0.732	0.165
July	0.424	-0.296	6.833	7.184	1.343	1.175	2.714	2.362
August	-0.304	-1.237	2.000	2.401	-1.890	-1.675	-0.947	-1.553
September	-1.530	-2.577	0.958	1.414	-2.871	-2.569	-2.180	-2.825
October	-3.143	-3.618	1.440	1.641	-5.032	-4.468	-2.432	-3.152
November	-5.766	-7.124	-0.234	0.320	-4.722	-5.850	-4.575	-4.688
December	-0.427	-0.699	3.471	3.655	-1.055	-1.471	-0.201	-0.543
MAPE	3.009	3.056	3.941	3.941	2.698	2.565	2.384	2.516

Table 4.8 are the results of northern substation control central area (NSCCAF). The maximum accuracy of NSCCAF with minimun MAPE of 2.356% is occurred model 7. This model is built from five selected factors such as maximum temperature, minimum temperature, mean temperature, consumer price index, and industrial index. Similarly, the NESCCAF has the maximum accuracy at model 7. The accuracy of this model is 2.384%. In SCCAF, model 8 has a maximum accuracy of 2.36%. Finally, the CSCCAF has a maximum accuracy (MAPE) at 2.082% from model 6. It can be conclude that the accuracy of the SCCAF models may be occurred in the different models for forecasting following the influencing factor in each area.

Table 4.10: Southern substation control central area forecasting.

	PE							
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	2.445	2.627	8.049	4.376	5.256	3.675	4.741	4.420
February	-4.285	-3.638	4.073	-3.852	0.498	-3.092	-0.039	-0.366
March	5.454	6.196	7.638	1.609	5.720	2.697	5.615	4.799
April	-1.097	-0.708	11.293	11.204	5.571	3.982	5.798	4.833
May	0.553	0.911	6.223	5.415	2.747	2.413	2.786	1.786
June	-0.379	-0.259	4.428	6.792	1.215	1.360	1.194	0.276
July	2.560	2.984	6.709	7.401	4.204	4.550	3.923	3.285
August	4.962	5.177	3.043	6.085	2.843	3.437	2.270	1.832
September	1.715	1.911	0.364	5.522	-0.257	0.508	-0.770	-1.346
October	1.509	1.886	1.843	3.891	0.201	1.030	0.117	-0.712
November	-3.943	-3.444	1.205	-1.650	-2.117	-3.417	-2.309	-3.065
December	-2.150	-1.751	2.160	96.414	-0.674	-2.201	-1.113	-1.613
MAPE	2.587	2.624	4.752	12.851	2.608	2.696	2.556	2.361

Table 4.11: Center substation control central area forecasting.

	PE							
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	1.860	0.907	2.669	0.447	2.351	0.190	2.331	2.596
February	-2.682	-5.372	-1.547	-6.795	-3.976	-7.015	-1.937	-3.267
March	7.563	5.974	2.149	-1.519	1.273	-0.858	2.266	1.440
April	-4.382	-5.606	1.147	-1.319	-0.635	-2.324	0.380	0.215
May	1.463	2.302	1.561	2.379	0.304	2.172	1.248	0.547
June	-0.559	0.544	-0.451	0.696	-1.285	0.405	-0.714	-1.136
July	2.791	1.391	5.411	2.194	1.695	1.760	5.047	2.087
August	1.842	3.379	-2.006	-0.064	-2.406	0.167	-1.963	-2.377
September	-0.016	1.401	-3.837	-2.095	-4.616	-1.818	-3.830	-4.624
October	0.955	2.788	-2.710	-0.254	-3.785	-0.437	-2.701	-3.152
November	-0.765	-0.178	-1.817	-1.515	-2.331	-1.817	-2.045	-2.189
December	-3.285	-3.341	-4.983	-5.787	-5.559	-6.020	-5.215	-5.310
MAPE	2.346	2.765	2.523	2.088	2.518	2.082	2.473	2.411

Table 4.12: Total substation control central area forecasting.

	PE							
Month	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
January	1.916	1.407	3.923	2.086	2.955	1.303	2.864	3.113
February	-3.222	-4.740	-0.996	-5.492	-3.796	-6.155	-2.138	-3.078
March	7.245	6.335	2.816	-0.391	1.962	0.046	2.522	2.020
April	-3.613	-4.072	3.205	1.420	0.384	-0.883	1.426	1.312
May	0.880	1.618	2.069	2.619	0.130	1.790	1.048	0.330
June	-0.311	0.387	0.874	1.897	-0.648	0.669	-0.122	-0.470
July	2.652	1.552	5.663	3.476	2.035	2.215	4.631	2.477
August	2.466	3.370	-0.446	1.242	-1.191	0.815	-0.865	-1.196
September	0.412	1.108	-2.544	-0.856	-3.491	-1.397	-3.038	-3.608
October	0.688	1.827	-1.739	0.231	-3.453	-0.881	-2.348	-2.691
November	-1.377	-1.268	-1.263	-1.295	-2.437	-2.327	-2.244	-2.293
December	-2.637	-2.602	-3.058	4.462	-4.218	-4.627	-3.991	-4.025
MAPE	2.284	2.523	2.383	2.122	2.225	1.925	2.269	2.217

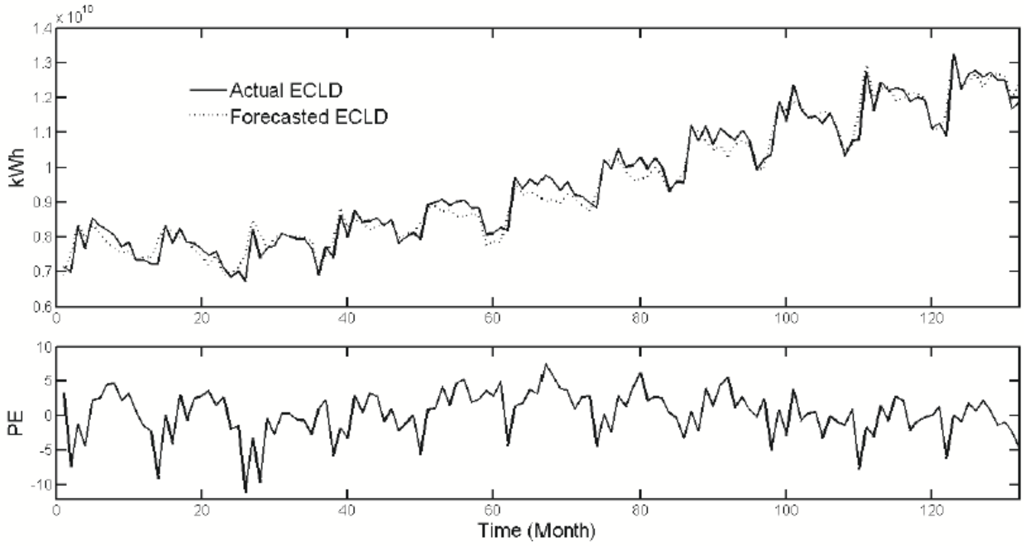

Figure 4.3: ECLD forecasting based of SCCAF based on model 6 and actual demand.

Table 4.12 presents the total SCCAF to whole region of the country which came from the summarization of the forecasted values in each region. In summarization, similar models in each SCCAF are computed together, for instance model 1 of NSCCAF, NESCCAF, SSCCAF, and CSCCAF to whole area based on each equation in model 1.

From Table 4.12 ,the best accuracy shown in $MAPE$ is 1.925% in Model

6, in which there are four factors used in this model as illustrated in Figure 4.3. These are mean temperature, consumer price index, industrial index and gross domestic product of the country, Thailand. The worst accuracy in term of $MAPE$ is 2.523% in Model 2. The selected factors of this model is explained in Table 4.1.

4.3 Seasonal Autoregressive Integrated Moving Average Approach

The autoregressive integrated moving average (ARIMA) approach was first developed in the late 60s but was systemized by Box and Jenkins in 1976. In the ARIMA method, the model consists of three main components, namely, Autoregressive model, Differences, and Moving Average. Each component has two sub-components i.e. simple sub-component and seasonal sub-component. An ARIMA [p, d, q] model can account for temporal dependence in several ways. First, the time series is differenced to render it stationary, by taking d differences. Second, the time dependence of the stationary process is modeled by including p auto-regressive and q moving-average terms, in addition to any time-varying covariates. For a cyclical time series, these steps can be repeated according to the period of the cycle, whether quarterly or monthly or another time interval. ARIMA models are extremely flexible for continuous data. ARIMA can be more complex to use than other statistical forecasting techniques, although when implemented properly can be quite powerful and flexible [7].

This general linear modeling approach follows the following procedure:

1. Determination of seasonality of time series and application of seasonal differencing.

2. Application of further differencing transformations to make the time series stationary.

3. Investigation of significant inputs to use as causal variables with the model.

4. Determination of order of seasonal and non-seasonal regressions.

5. Identification of model parameters.

In the ECLD, the analysis performed by ARIMA procedure is divided into four stages, corresponding to stages described by Box and Jenkins, which are summarized as follows:

1. Identification, the identify statement is used to specify the response series and identify candidate ARIMA models for it. The identity statement reads time series that are to be used in later statement, possibly differencing them, and computes autocorrelations, inverse autocorrelations, partial autocorrelations, and cross-correlations. Stationary tests can be performed to determine if differencing is necessary. The analysis of the identity statement output usually suggests one or more ARIMA models that could be fit. Options enable you to test for stationary and tentative ARMA order identification.

2. Estimation, in this stage, the estimate statement is used to specify the ARIMA model to fit to the variable specified in the previous identify statement and to estimate the parameters of that model. The estimate statement also produces diagnostic statistics to help you judge the adequacy of the model. Significant tests for parameter estimates indicate whether some terms in the model might be unnecessary. Goodness-of-fit statistics aid in comparing this model to others. Tests for white noise residuals indicate whether the residual series contains additional information that might be used by a more complex model.

3. Diagnostic Checking, the outlier statement, provides another useful tool to check whether the currently estimated model accounts for all the variation in the series. If the diagnostic tests indicate problems with the model, you try another model and then repeat the estimation and diagnostic checking stage.

4. Forecasting, in this stage, the forecast statement is used to forecast future values of the time series and to generate confidence intervals for these forecasts from the ARIMA model produced by the preceding estimate statement.

In this research, SARIMA, the seasonal autoregressive integrated moving average (SARIMA) with non-stationary of data is proposed.

4.3.1 Empirical Results of WAF

4.3.1.1 Data Sets and Models

From load demand data of Thailand in Chapter 3 or Figure 4.1, whole electricity load demand of the country are shown a time series in monthly types, 120 points [1], [5]. It is noted that two years or twenty-four months of the data from 1997 to 1998 are different from other months because of the problems with the economy in the world. The accuracy of the forecast may necessarily need to cut the some problems out. The ECLD curve of the whole area is given in Figure 4.1.

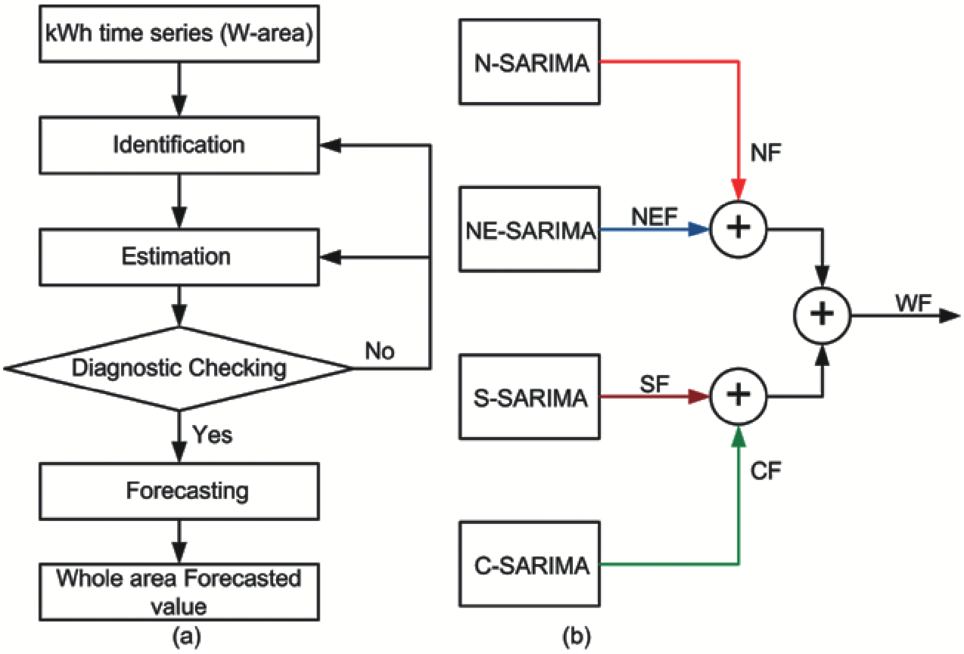

Figure 4.4: Box-Jenkins flow chart (ARIMA) a)whole area b)subarea.

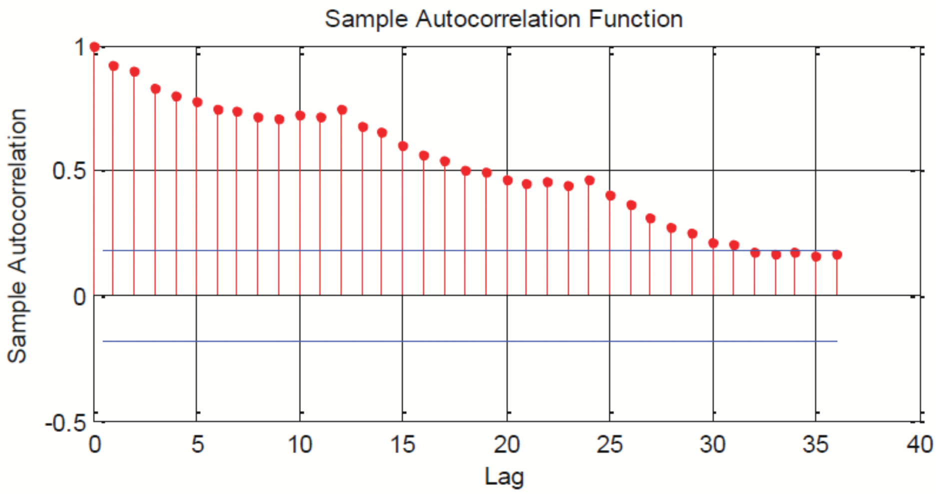

Figure 4.5: Correlogram of $r_k(y_t)$.

In this study, ACF and PACF considerations in Figure 4.4 and 4.5 were used to find the appropriate model based on SARIMA approach. From these considerations, the correlogram of ECLD obtaining in Figure 4.4 and 4.5 is analyzed. Figure 4.4 is autocorrelation (ACF) and Figure 4.5 is partial autocorrelation of the demand. From Figure 4.4, $|r_k(y_t)|$ has slowly decreased. Since the time series is non-stationary from a trend and seasonal factor, it is necessary to build a new time series that is $\{z_t)$ which is $Z_t = \Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$. It is moving around zero. The $|r_k(z_t)|$

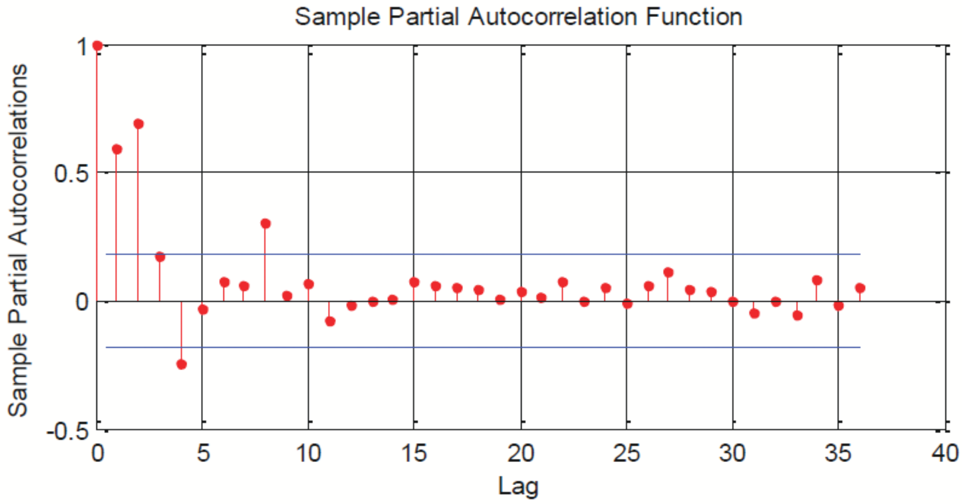

Figure 4.6: Correlogram of $r_{kk}(y_t)$.

and $|r_{kk}(z_t)|$ correlogram are rapidly decreased to zero if time series is stationary. In this study, the appropriate model is nearly $MA(1) \times SMA(1)_{12}$. The model prescribes MA(1) form because the $|r_1(z_t)|$ is large value while the $|r_{kk}(z_t)|$ decreases to zero. The model prescribes $SMA(1)_{12}$ form because the $|r_{12}(z_t)|$ has a large value whereas the $|r_{kk}(z_t)|$ rapidly decreases to zero. The approximate coefficient value of $\hat{\theta}_1 = 0.85$ and $\hat{\theta}_{12} = 0.492$. To analyze the model $MA(1) \times SMA(1)_{12}$, or $SARIMA(0,1,1)(0,1,1)_{12}$, the test results are $H_0 : \theta_1 = 0;\ H_1 : \theta_1 \neq 0$ and $H_0 : \theta_{12} = 0;\ H_1 : \theta_{12} \neq 0$, and their results have $t = 13.581\,(P = 0.000)$ and $t = 4.236\,(P = 0.000)$, respectively. All testing results show the model declining H_0 at 0.05 (significant value). In testing outcome, the model found that $|r_k(e_t)|$ is nearly zero for other k. Testing a hypothesis H_0 and H_1 is $H_0 : \rho_1(e_t) = \rho_2(e_t) = .. = \rho_m(e_t) = 0$ and $H_1 : \rho_i(e_t)\ for\ i = 1, 2, .., m$ at least one non-zero value. These are $Q_m = 4.769$ $(P = 0.574)$ and $Q_m = 9.004$ $(P = 0.703)$ for $m = 6, 12$ respectively, and it accepted H_0 at 0.05 (sig.). Therefore, the testing results that show the $MA(1) \times SMA(1)_{12}$ model is appropriate model for time series Y_t , and can predict the load demand by using $\hat{Z}_t = (1 - 0.850B)(1 - 0.492B^{12})e_t$, the result of whole area based on SARIMA illustrate in Figure 4.7, actual and forecsted value of the country.

4.3.1.2 Results

In the whole area forecasting, the approach for analysis and building a time series model in whole area is the method of finding for a model (SARIMA)

Table 4.13: WAF by SARIMA .

Model	Description	MAPE(%)
1	$SARIMA(1,1,0)(1,1,0)_{12}$	2.836
2	$SARIMA(0,1,1)(1,1,0)_{12}$	2.214
3***	$SARIMA(0,1,1)(0,1,1)_{12}$	2.282
4	$SARIMA(1,1,0)(0,1,1)_{12}$	2.637
5	$SARIMA(1,1,1)(1,1,0)_{12}$	2.233
6	$SARIMA(1,1,1)(0,1,1)_{12}$	2.351
7	$SARIMA(1,1,1)(1,1,1)_{12}$	2.339

that adequately represent a data generating process for a given set of data information. The test time series data was processed by taking the first order regular difference and the first seasonal difference in order to remove the varied trend and seasonal characteristics. Stationary of time series of each seasons, can be identified the model $SARIMA(p,d,q)(P,D,Q)_{12}$ by considering the characteristic of the ACF (Autocorrelation factor) and PACF (Partial autocorrelation factor) based on a correlogram testing step. This approach estimated parameters by Ordinary Least-Squares Method (OLS), and checked hypothesis of the model are validate was used to determine the appropriate model and then the best model were taken to final forecast in year 2007, whole region.

Table 4.13 presents theWAF by using SARIMA in different model description and MAPE results of ECLD forecasting performed by the proposed SARIAM approach. In this chart, each model will come from ACF and PACF graph consideration in SPSS program. Table 4.14 shows in detail of MAPE of Model 3

4.3.2 Empirical Results of SCCAF

4.3.2.1 Data Sets and Models

SCCA electricity consumption load demand of the country shows a time series in monthly types of 120 points [1], [5]. It is noted that two years or twenty-four months of the data from 1997 to 1998 are different from other months because the problems with the economy in the world. The accuracy of the forecast may necessarily

Table 4.14: SARIMA for WAF, $SARIMA(0,1,1)(0,1,1)_{12}$.

Month	Model	
	$SARIMA(0,1,1)(0,1,1,)_{12}$	
	Forecasted (kWh)	APE
Jan	11456088324	1.78
Feb	11489461685	5.37
March	13211965152	0.33
April	12439466430	1.77
May	13209954265	4.27
Jane	12802505843	0.27
July	12860596884	2.25
Aug	12842522452	0.94
Sep	12561858467	0.46
Oct	12764428765	2.11
Nov	12506211702	7.09
Dec	11778947563	0.74
MAPE		2.28

need to cut some problems out. The ECLD curve of NSCCA is given in Figure 4.8. In this study, ACF and PACF considerations in Figure 4.9 and 4.10 are used to find the appropriate model based on SARIMA approach.

Figure 4.9 is autocorrelation (ACF) and Figure 4.10 is partial autocorrelation of the demand. From Figure 4.9, $|r_k(y_t)|$ has slowly decreased. Due to the time series is non-stationary from a trend and seasonal factor. Then it is necessary to build a new time series that is $\{z_t)$ which is $Z_t = \Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$. It is moving around zero. The $|r_k(z_t)|$ and $|r_{kk}(z_t)|$ correlogram are rapidly decreased to zero that time series is stationary. In this study, the appropriate model is nearly $MA(1) \times SMA(1)_{12}$, $AR(1) \times SMA(1)_{12}$, , and $AR(1) \times SAR(1)_{12}$. The model prescribes MA(1) form because

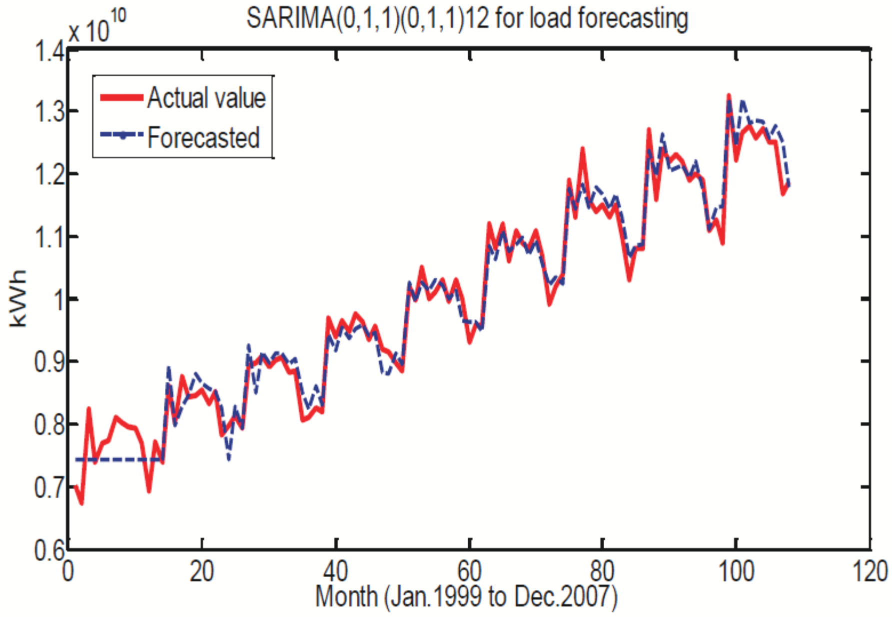

Figure 4.7: Actual and forecasted value of the country.

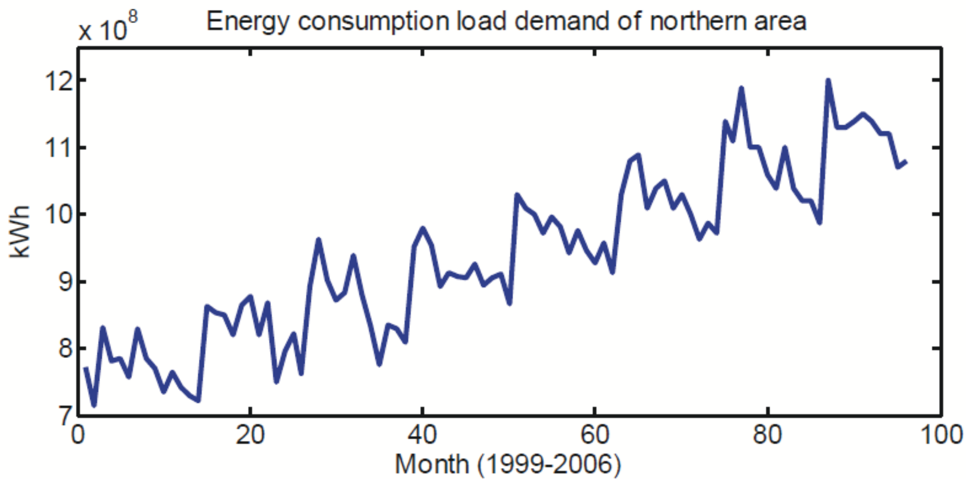

Figure 4.8: Time series of the NSCCA electricity consumption load demand of Thailand.

the $|r_1(z_t)|$ is large value while the $|r_{kk}(z_t)|$ decreases to zero, and the model prescribes $SMA(1)_{12}$ form because the $|r_{12}(z_t)|$ has a large value whereas the $|r_{kk}(z_t)|$ rapidly decreases to zero. For example, the approximate coefficient value of $\hat{\theta}_1 = 0.964$ and $\hat{\theta}_{12} = 0.780$. To analyze the model,$MA(1) \times SMA(1)_{12}$ or $SARIMA(0,1,1)(0,1,1)_{12}$, the test results are $H_0 : \theta_1 = 0$; $H_1 : \theta_1 \neq 0$ and $H_0 : \theta_{12} = 0$; $H_1 : \theta_{12} \neq 0$, and their results have $t = 10.610(P = 0.000)$ and $t = 5.163\,(P = 0.000)$, respectively. All testing results show the model declining H_0 at 0.05 (significant value). In testing outcome, the model found that $|r_k(e_t)|$ is nearly zero for other k. Testing a hypothesis H_0 and H_1

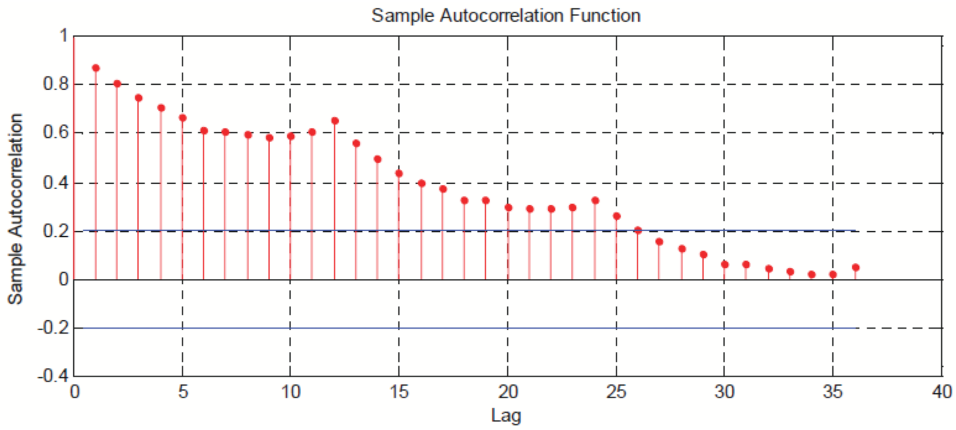

Figure 4.9: Correlogram of $r_k(y_t)$ of NSCCA.

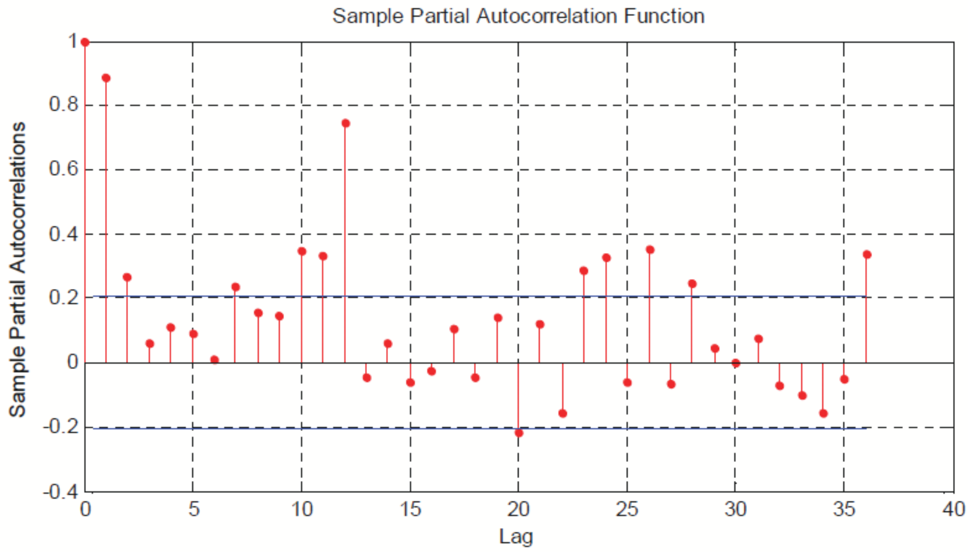

Figure 4.10: Correlogram of $r_{kk}(y_t)$ of NSCCA.

is $H_0 : \rho_1(e_t) = \rho_2(e_t) = .. = \rho_m(e_t) = 0$ and $H_1 : \rho_i(e_t)\ for\ i = 1, 2, .., m$ at least one non-zero value. These are $Q_m = 5.993$ $(P = 0.424)$ and $Q_m = 10.834$ $(P = 0.543)$ for $m = 6,\ 12$ respectively, and it accepted H_0 at 0.05 (sig.). Therefore, the testing results that show the $MA(1) \times SMA(1)_{12}$ model or $\hat{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} + \varepsilon_t - \theta_{12}\varepsilon_{t-12}$ is appropriate model for time series Y_t , and can predict the load demand by using $\hat{Z}_t = (1 - 0.964B)(1 - 0.780B^{12})e_t$ or $\hat{Y}_t = \hat{Z}_t + \hat{Y}_{t-1} + \hat{Y}_{t-12} - \hat{Y}_{t-13}$.

In Northeastern region with time series of ECLD shown in Figure 4.11, the analytical of the correlogram of ECLD are obtained in Figure 4.12 and 4.13. Figure 4.12 is autocorrelation (ACF) and Figure 4.13 is partial autocorrelation of the demand. From Figure 4.12, $|r_k(y_t)|$ has slowly decreased. Due to the non-stationary

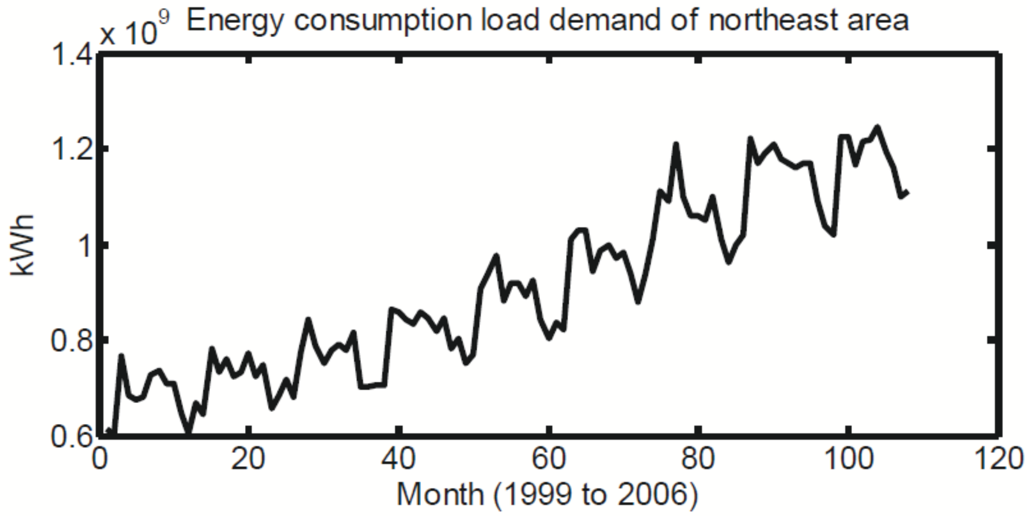

Figure 4.11: Time series of the NESCCA electricity consumption load demand of Thailand.

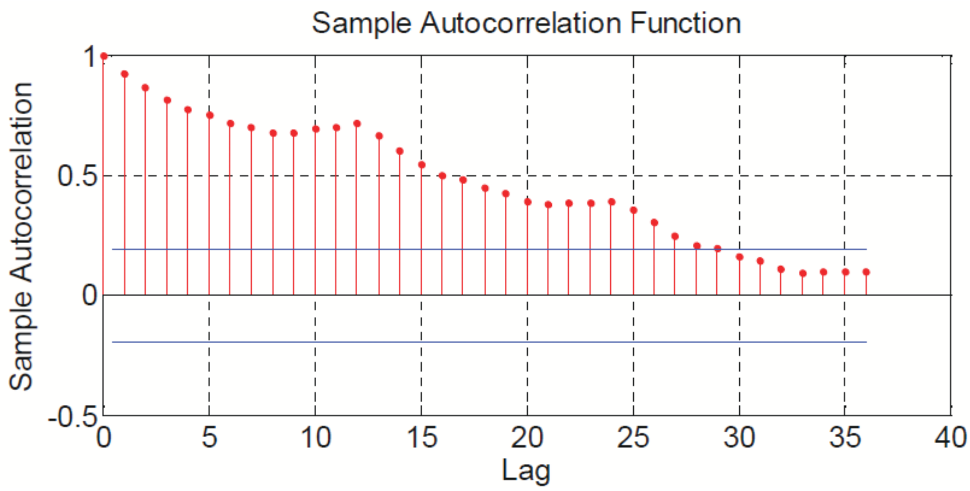

Figure 4.12: Correlogram of $r_k(y_t)$ of NESCCA.

time series from a trend and seasonal factor, it is necessary to build a new time series $\{z_t)$ in which $Z_t = \Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$. It is moving around zero. The $|r_k(z_t)|$ and $|r_{kk}(z_t)|$ correlogram are rapidly decreased to zero if time series is stationary. In this study, the appropriate model is nearly $MA(2) \times SMA(1)_{12}$. The model prescribes MA(1) form because the $|r_1(z_t)|$ is large value while the $|r_{kk}(z_t)|$ decreases to zero, and the model prescribes $SMA(1)_{12}$ form because the $|r_{12}(z_t)|$ has a large value whereas the $|r_{kk}(z_t)|$ rapidly decreases to zero. The approximate coefficient value of $\hat{\theta}_1 = 0.593$, $\hat{\theta}_2 = 0.374$ and $\hat{\theta}_{12} = 0.600$. To analyze the model, $MA(2) \times SMA(1)_{12}$ or $SARIMA(0,1,2)(0,1,1)_{12}$, the test results are $H_0 : \theta_1 = 0$; H_1 :

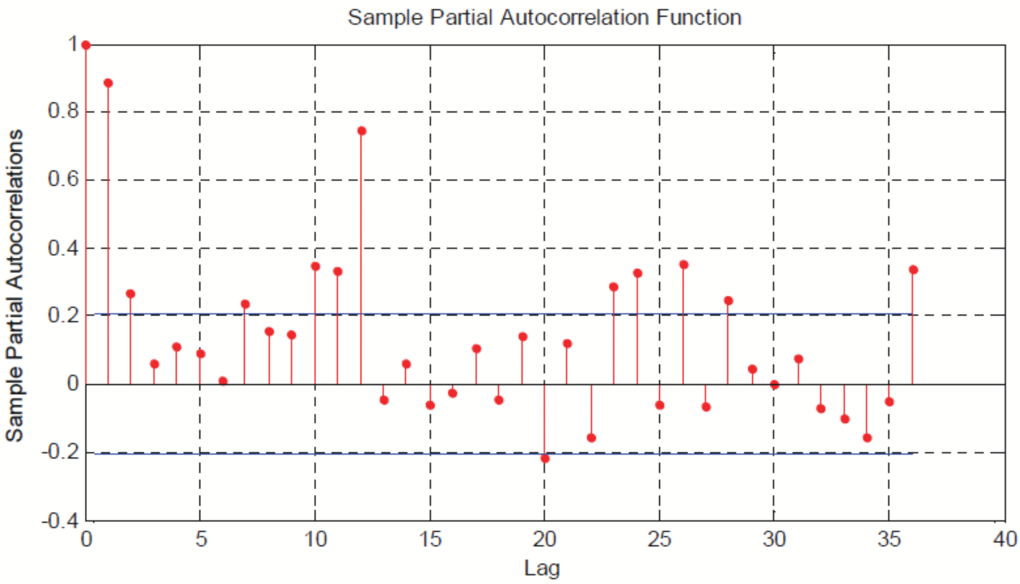

Figure 4.13: Correlogram of $r_{kk}(y_t)$ of NESCCA.

$\theta_1 \neq 0$ and $H_0 : \theta_{12} = 0$; $H_1 : \theta_{12} \neq 0$, and their results have $t = 3.427\,(P = 0.000)$, $t = 2.890\,(P = 0.004)$ and $t = 5.132(P = 0.000)$, respectively. All testing results show the model declining H_0 at 0.05 (significant value). In testing outcome, the model found that $|r_k(e_t)|$ is nearly zero for other k. Testing a hypothesis H_0 and H_1 is $H_0 : \rho_1(e_t) = \rho_2(e_t) = .. = \rho_m(e_t) = 0$ and $H_1 : \rho_i(e_t)\ for\ i = 1, 2, .., m$ at least one non-zero value. These are $Q_m = 5.700$ $(P = 0.453)$ and $Q_m = 15.193$ $(P = 0.231)$ for $m = 6, 12$ respectively, and it accepted at 0.05 (sig.). Therefore, the testing results of the $MA(2) \times SMA(1)_{12}$ or $\hat{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} + \varepsilon_t - \theta_{12}\varepsilon_{t-12}$ is appropriate model for time series Y_t .

In Southern region with time series of ECLD shown in Figure 4.14, the analytical of the correlogram of ECLD are obtained in Figure 4.15 and 4.16. Figure 4.15 is autocorrelation (ACF) and Figure 4.16 is partial autocorrelation of the demand. From Figure 4.15, $|r_k(y_t)|$ has slowly decreased. Due to the non-stationary time series is from a trend and seasonal factor, it is necessary to build a new time series $\{z_t)$ in which $Z_t = \Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$. It is moving around zero. The $|r_k(z_t)|$ and $|r_{kk}(z_t)|$ correlogram are rapidly decreased to zero if time series is stationary. In this study, the appropriate model is nearly $MA(1) \times SMA(1)_{12}$, $AR(1) \times SMA(1)_{12}$, and $ARMA(1, 1) \times SMA(1)_{12}$. The model prescribes MA(1) form because the $|r_1(z_t)|$ is large value while the $|r_{kk}(z_t)|$ decreases to zero, and the model prescribes $SMA(1)_{12}$

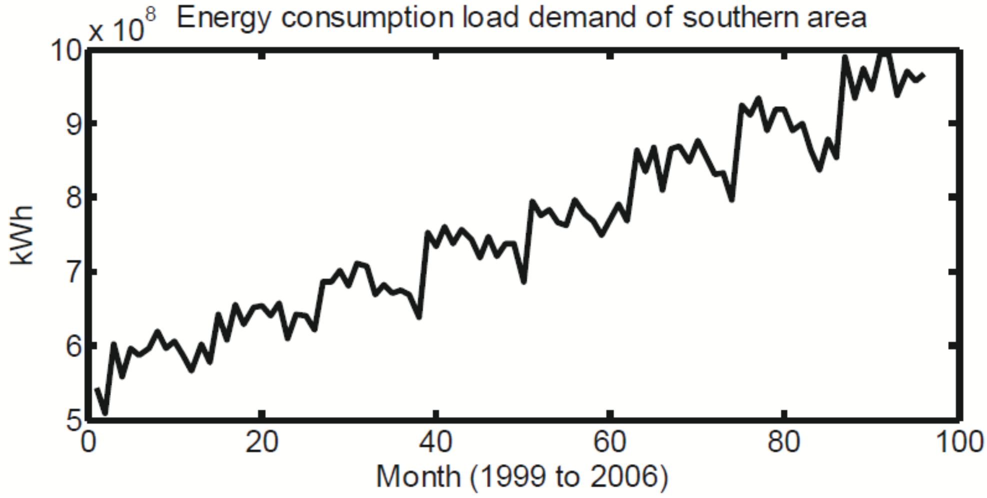

Figure 4.14: Time series of the SSCCA electricity consumption load demand of Thailand.

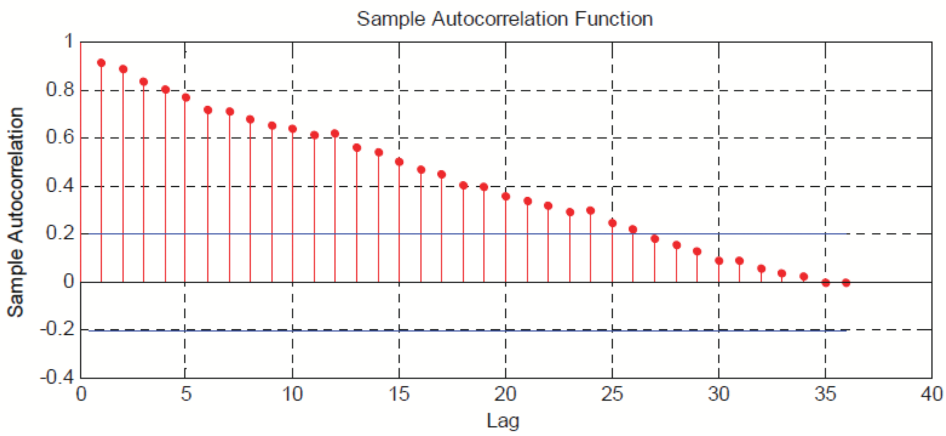

Figure 4.15: Correlogram of $r_k(y_t)$ of SSCCA.

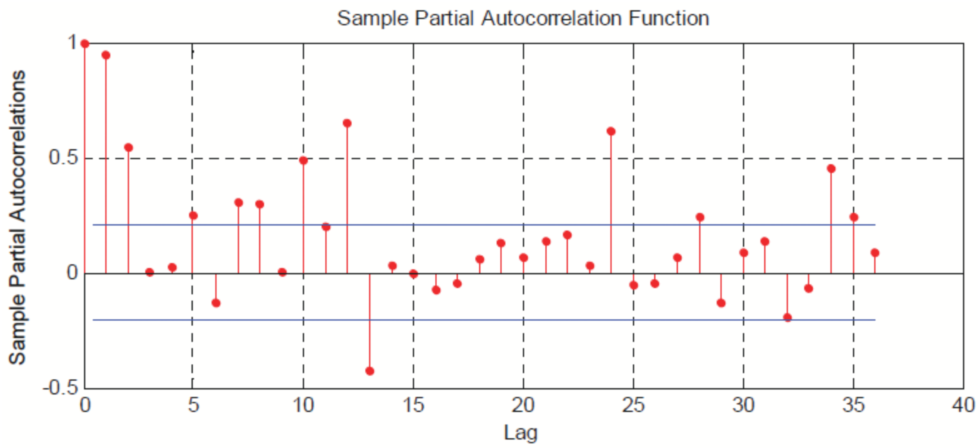

Figure 4.16: Correlogram of $r_{kk}(y_t)$ of SSCCA.

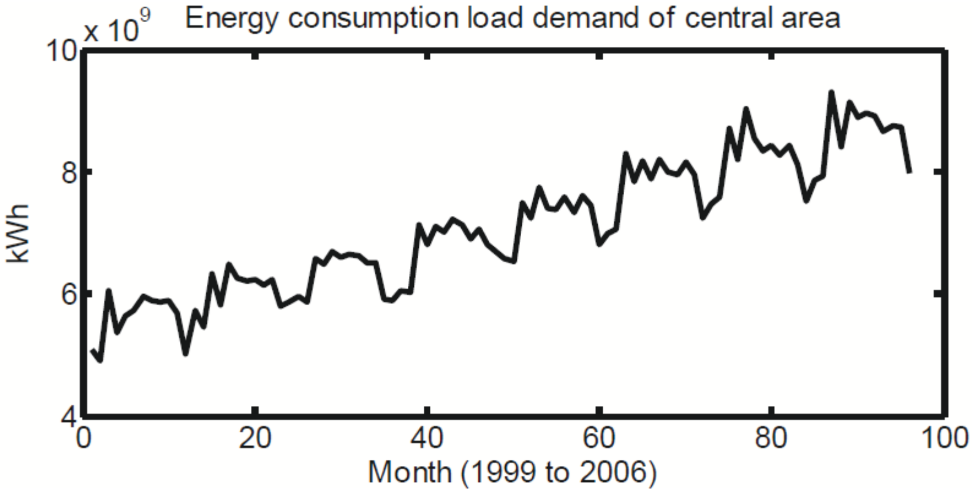

Figure 4.17: Time series of the CSCCA electricity consumption load demand of Thailand.

form because the $|r_{12}(z_t)|$ has a large value whereas the $|r_{kk}(z_t)|$ rapidly decreases to zero. For example, the approximate coefficient value of $\hat{\theta}_1 = 0.636$ and $\hat{\theta}_{12} = 0.748$. To analyze the model, $MA(1) \times SMA(1)_{12}$ or SARIMA(0,1,1)(0,1,1)12, the test results are $H_0 : \theta_1 = 0$; $H_1 : \theta_1 \neq 0$ and $H_0 : \theta_{12} = 0$; $H_1 : \theta_{12} \neq 0$, and their results have $t = 7.583\,(P = 0.000)$ and $t = 5.042\,(P = 0.000)$, respectively. All testing results show the model declining H_0 at 0.05 (significant value). In testing outcome, the model found that $|r_k(e_t)|$ is nearly zero for other k. Testing a hypothesis H_0 and H_1 is $H_0 : \rho_1(e_t) = \rho_2(e_t) = .. = \rho_m(e_t) = 0$ and $H_1 : \rho_i(e_t)\ for\ i = 1, 2, .., m$ at least one non-zero value. These are $Q_m = 5.344$ $(P = 0.501)$ and $Q_m = 7.038$ $(P = 0.855)$ for respectively, and it accepted H_0 at 0.05 (sig.). Therefore, the testing results of $MA(1) \times SMA(1)_{12}$ model or $\hat{Z}_t = \varepsilon_t - \theta_1\varepsilon_{t-1} + \varepsilon_t - \theta_{12}\varepsilon_{t-12}$ is appropriate model for time series Y_t, and can predict the load demand by using $\hat{Z}_t = (1-0.636B)(1-0.748B^{12})e_t$ or $\hat{Y}_t = \hat{Z}_t + \hat{Y}_{t-1} + \hat{Y}_{t-12} - \hat{Y}_{t-13}$.

In Central region with time series of ECLD shown in Figure. 4.17. The analytical of the correlogram of ECLD are obtained in Figure4.18 and 4.19. Figure 4.18 is autocorrelation (ACF) and Figure 4.19 is partial autocorrelation of the demand. From fig.4.18, $|r_k(y_t)|$ has slowly decreased. Due to the non-stationary time series is from a trend and seasonal factor, it is necessary to build a new time series $\{z_t)$ in which $Z_t = \Delta\Delta_{12}Y_t = Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$. It is moving around zero. The $|r_k(z_t)|$ and $|r_{kk}(z_t)|$ correlogram are rapidly decreased to zero if time series is stationary. In this study, the appropriate model is nearly $MA(1) \times SMA(1)_{12}$. The model prescribes

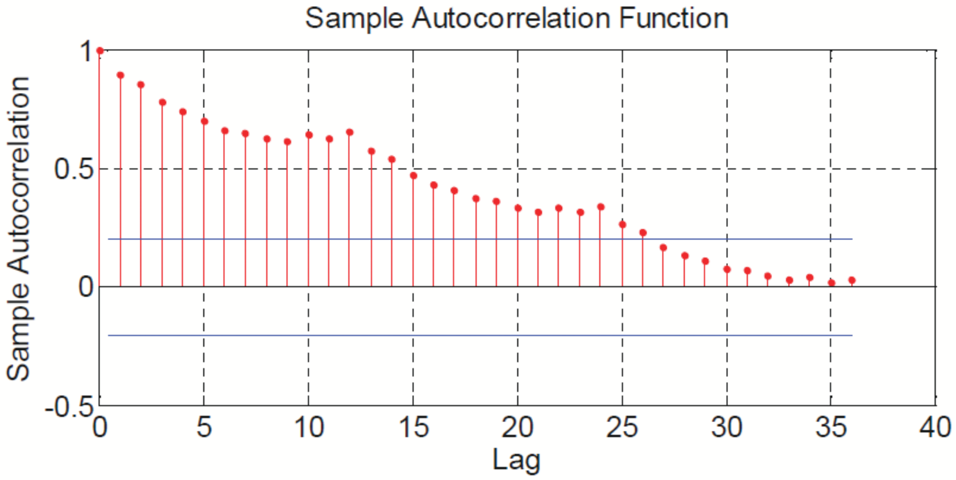

Figure 4.18: Correlogram of $r_k(y_t)$ of CSCCA.

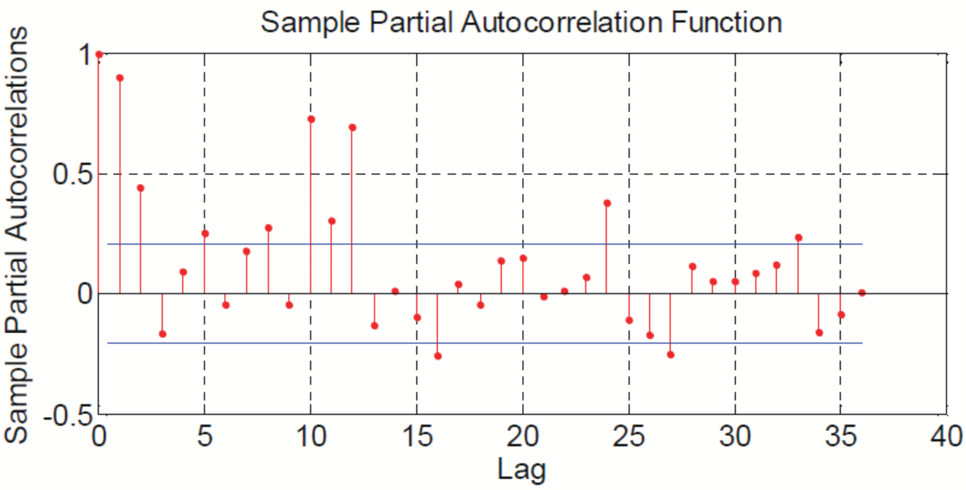

Figure 4.19: Correlogram of $r_{kk}(y_t)$ of CSCCA.

MA(1) form because the $|r_1(z_t)|$ is large value while the $|r_{kk}(z_t)|$ decreases to zero, and the model prescribes $SMA(1)_{12}$ form because the $|r_{12}(z_t)|$ has a large value whereas the $|r_{kk}(z_t)|$ rapidly decreases to zero. The approximate coefficient value of $\hat{\theta}_1 = 0.836$ and $\hat{\theta}_{12} = 0.532$. To analyze the model, $MA(1) \times SMA(1)_{12}$ or $SARIMA(0,1,1)(0,1,1)_{12}$, the test results are $H_0 : \theta_1 = 0$; $H_1 : \theta_1 \neq 0$ and $H_0 : \theta_{12} = 0$; $H_1 : \theta_{12} \neq 0$, and their results have $t = 12.863\,(P = 0.000)$ and $t = 4.511\,(P = 0.000)$, respectively. All testing results show the model declining H_0 at 0.05 (significant value). In testing outcome, the model found that $|r_k(e_t)|$ is nearly zero for other k. Testing a hypothesis H_0 and H_1 is $H_0 : \rho_1(e_t) = \rho_2(e_t) = .. = \rho_m(e_t) = 0$ and $H_1 : \rho_i(e_t)\ for\ i = 1,2,..,m$ at least one non-zero value. These are $Q_m = 4.864$ $(P = 0.561)$ and $Q_m = 8.800$ $(P = 0.720)$ for

Table 4.15: SARIMA for the northern region .

Model	Description	MAPE(%)
1***	$SARIMA(0,1,1)(0,1,1)_{12}$	2.862
2***	$SARIMA(1,1,0)(0,1,1)_{12}$	2.887
3***	$SARIMA(0,1,1)(1,1,0)_{12}$	2.851
4***	$SARIMA(1,1,0)(1,1,0)_{12}$	2.858
5	$SARIMA(1,1,1)(1,1,0)_{12}$	2.880
6	$SARIMA(1,1,1)(0,1,1)_{12}$	2.872
7	$SARIMA(1,1,1)(1,1,1)_{12}$	2.796

$m = 6, 12$ respectively, and it accepted H_0 at 0.05 (sig.). Therefore, the testing results of $MA(1) \times SMA(1)_{12}$ model or $\hat{Z}_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} + \varepsilon_t - \theta_{12} \varepsilon_{t-12}$ is appropriate model for time series Y_t, and can predict the load demand by using $\hat{Z}_t = (1 - 0.836B)(1 - 0.532B^{12})e_t$ or $\hat{Y}_t = \hat{Z}_t + \hat{Y}_{t-1} + \hat{Y}_{t-12} - \hat{Y}_{t-13}$.

4.3.2.2 Results

Table 4.15 to Table 4.18 show the results of SARIMA forecasting for SCCA in northern, northeastern, southern, and central regions. In northern region, Model 1 to 4 approaches are appropriate for prediction in the area. The $MAPE$ is about 2.851 to 2.887 % while other Models of 5, and 7 are not appropriate. In northeastern area, model 1 approach is good model for prediction which the $MAPE$ of the accuracy is 4.412 %. The southern region, the appropriate Models are 1 to 3 and 6, and Model 6 is the better model with $MAPE$ of 1.320 %. Finally, the central area, the appropriate model is model 2 which the $MAPE$ is 2.546 %. It can be concluded that, for all the results of SCCAF using SARIMA, the accuracy of NSCCAF is lower than that of the other regions. Because NSCCA is large region and has different weather when compared with other regions. The $MAPE$ of whole are with summarization of SCCA forecasting, is 2.36 %.

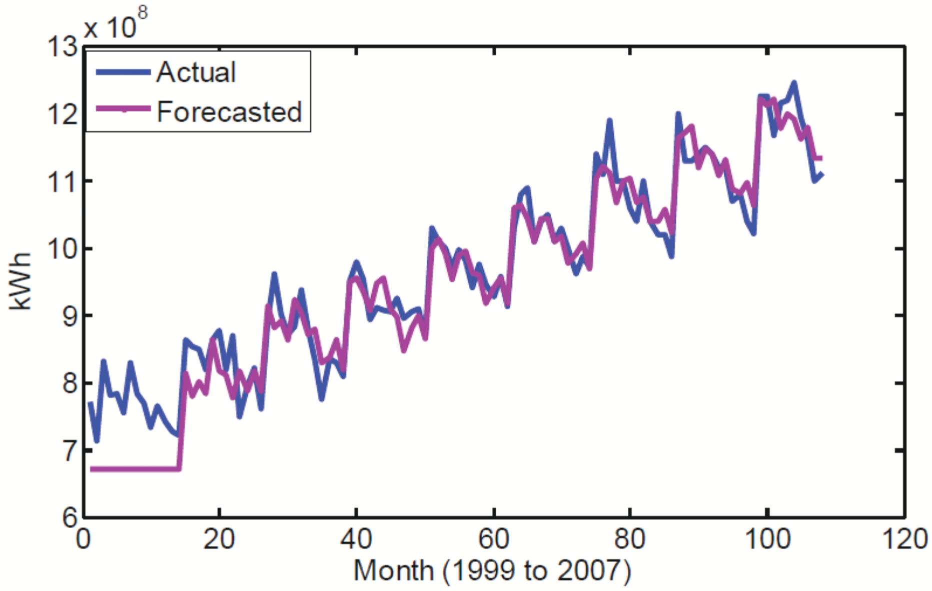

Figure 4.20: Actual and NSCCA forecasted value of the country.

Table 4.16: SARIMA for the northeastern region .

Model	Description	MAPE(%)
1***	$SARIMA(0,1,2)(0,1,1)_{12}$	4.412
2	$SARIMA(1,1,0)(0,1,1)_{12}$	5.500
3	$SARIMA(0,1,1)(1,1,0)_{12}$	4.861
4	$SARIMA(1,1,0)(1,1,0)_{12}$	6.469
5	$SARIMA(1,1,1)(1,1,0)_{12}$	4.879
6	$SARIMA(1,1,1)(0,1,1)_{12}$	4.460
7	$SARIMA(1,1,1)(1,1,1)_{12}$	4.762

4.4 Discussion

In these demonstrations, multiple linear regression (MLR) and seasonal autoregressive integrated moving average (SARIMA), they are basic approaches for studying the mid-term load demand forecasting in the research. MLR approach demonstrates electricity consumption demand prediction from many influencing factors such as maximum temperature, minimum temperature, maximum temperature, humidity, consumer price index, industrial index, and gross domestic product. It studies about the

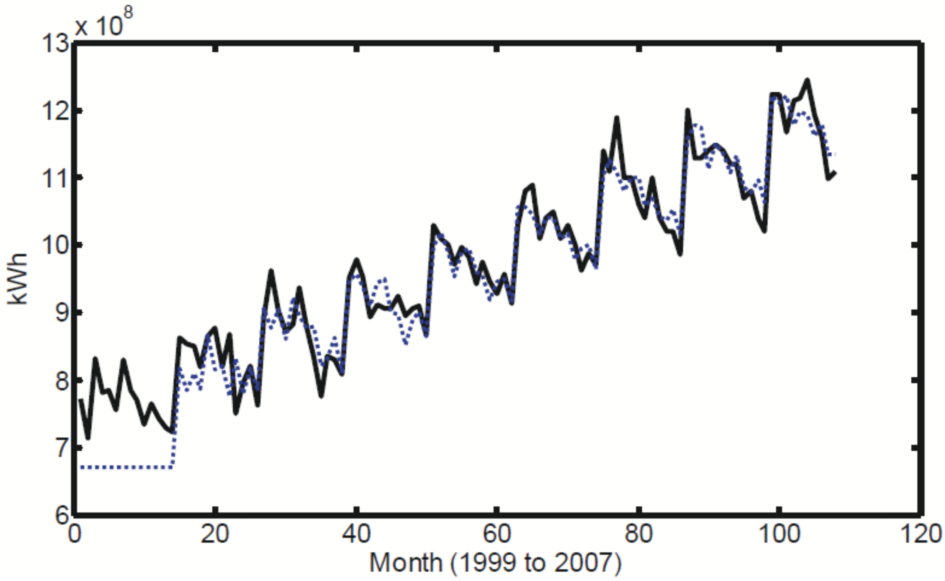

Figure 4.21: Actual and NESCCA forecasted value of the country.

Table 4.17: SARIMA for the southern region .

Model	Description	MAPE(%)
1***	$SARIMA(1,1,0)(0,1,1)_{12}$	2.564
2***	$SARIMA(0,1,1)(0,1,1)_{12}$	1.703
3***	$SARIMA(0,1,1)(1,1,0)_{12}$	1.372
4	$SARIMA(1,1,0)(1,1,0)_{12}$	3.040
5	$SARIMA(1,1,1)(1,1,0)_{12}$	1.296
6***	$SARIMA(1,1,1)(0,1,1)_{12}$	1.320
7	$SARIMA(1,1,1)(1,1,1)_{12}$	1.296

relationship between load demand and influencing variables. In this approach, there are a variety of models to test the demand predictions. Each model will have to adjust the influencing variables to find the appropriate model. On other hand, SARIMA approach shows the ability to predict the load demand without influencing factors. This approach uses the principle of the regression method of ECLD time series and seasonality of it, and the period of seasonality is 12.

Forecasting approach by using the MLR method is applied to the whole

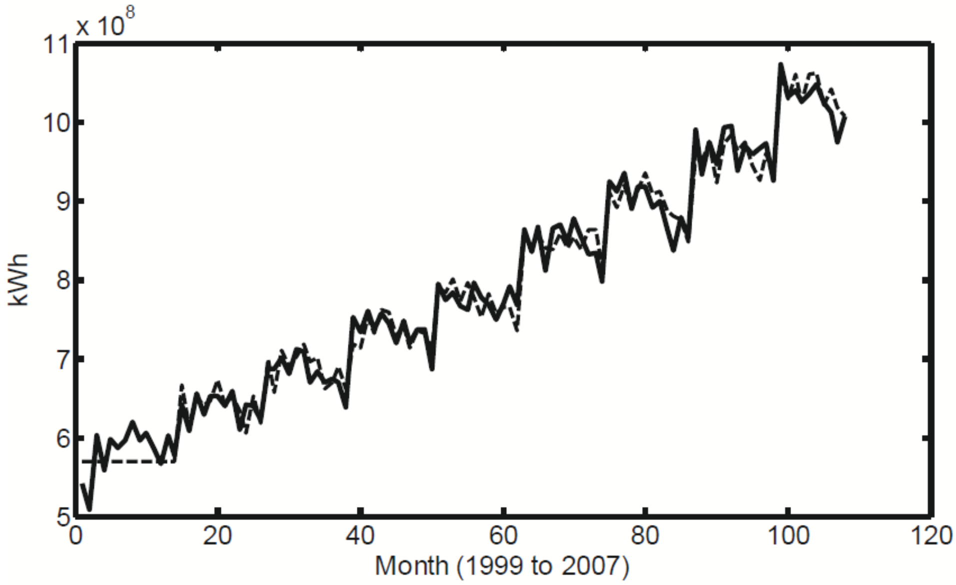

Figure 4.22: Actual and SSCCA forecasted value of the country.

Table 4.18: SARIMA for the central region .

Model	Description	MAPE(%)
1	$SARIMA(0,1,2)(0,1,1)_{12}$	2.587
2***	$SARIMA(0,1,1)(0,1,1)_{12}$	2.546
3	$SARIMA(0,1,1)(1,1,0)_{12}$	2.416
4	$SARIMA(1,1,0)(1,1,0)_{12}$	2.647
5	$SARIMA(1,1,1)(1,1,0)_{12}$	2.450
6	$SARIMA(1,1,1)(0,1,1)_{12}$	2.592
7	$SARIMA(1,1,1)(1,1,1)_{12}$	2.585

area and subarea to the total area or whole area. The forecasting result, the model 6 is that it is better accurate than the other models. This model is suitable for predicting both the whole area forecasting and subarea forecasting. The whole area forecasting the MAPE is 1.86%, while the subarea forecasting is 1.926%. The variables in this model are the mean temperature, the consumer price index, industrial index, and gross domestic product used in model 6 both WA and SCCA predictions. Also, the gross domestic product (GDP) of whole area represents the GDP in SCCA forecasting but this factor

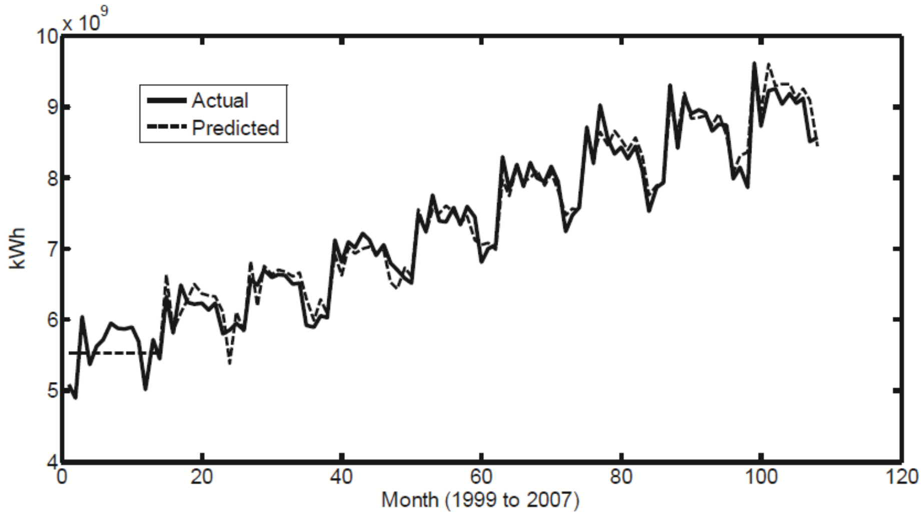

Figure 4.23: Actual and CSCCA forecasted value of the country.

used in SCCA forecasting may not reflect the GDP of each area for subarea forecasting. The accuracy of SCCA forecasting may decrease from this problem.

The following approach is the forecasting by using SARIMA method to WA and subarea similarly, but this approach does not lead variables to build the model, it forecast the load demand by learning the demand patterns. The results of the prediction of whole area are 2.282% for the appropriate model of $(0,1,1)(0,1,1)_{12}$, and the subarea is 2.36% in the summation of the different appropriate models in each area. The forecasting result of the total area is slightly better accurate than SCCA forecasting.

The predictions of two approaches that are both the MLR and SARIMA covered both whole area and subarea. However, MLR is more accurate than SARIMA, but it is the complexity of the variables to build the model. Also, there are many parts of the SARIMA approach. It is a difficult to understand in each part of the SARIMA algorithm. But it does not need to interfere in the matter variables. At the same time, if we look at the predictions of the two methods, it can be summarized that the whole area of the country forecasting provides the best accuracy in this work.

4.5 Summary

Multiple linear regression (MLR) and seasonal autoregressive integrated moving average (SARIMA) approaches that are basically and widely used for demand forecasting. These approaches can give good forecasting results especially in linear time series. They also used in several fields such as economics, accounting, and etc. These methods deploy different implementations: exponential smoothing, multiple linear regression, seasonal autoregressive integrated moving average, and other approaches. MLR approach analyzed the equation through the relationship between demand and influencing factors whereas SARIMA analyzed the model by time series regression method. These approaches may not be the appropriate methods for non-linear mid-term load time series. Thus, next chapter will propose the approaches for supporting both non-linear and linear time series.

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] K. Metaxiotis, A. Kagiannas, D. Askounis, and J.Psarras, "Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the research," Energy Conversion and Management 44, pp.1525-1534, 2003.

[7] Jian-Chang Lu, Dong-xiao niu, and zheng-yuan jia, "A study of short term load forecasting based on arima-ann," Machine Learning and Cybernetics, Proceeding of 2004 International Conference on Volume 5, pp.3183-3187, 2004.

[8] Kyung-bin song, young-silk, dug hun hong, and gilsoo jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," IEEE, pp.589-594, 2004.

[9] H.A. Al-Hamadi, and S.A. Soliman, "Long-term/Mid-term electric load forecasting based on short-term correlation and annual growth," Electric Power Systems Research, science direct, pp.353-361, 2005.

[10] Zone-Ching Lin, and Wen-Jang Wu, "Multiple Linear Regression Analysis of the Overlay Accuracy Model," IEEE Trans. on semiconductor manufacturing, Vol.12, No. 2, pp.229-237, 1999.

[11] Kyung-Bin Song, Young-Sik Baek, Dug Hun Hong, and Gilsoo Jang, "Short-Term Load Forecasting for the Holidays Using Fuzzy Linear Regression Method," IEEE Trans. on power systems, Vol. 20, No. 1 pp.96-101, 2005.

[12] J.V. Ringwood, D. Bofelli, and F.T. Murray, "Forecasting Electricity Demand on Short, Medium and Long Time Scales Using Neural Networks," Journal of Intelligent and Robotic System 31, kluwer academic publishers, Netherlands, pp.129-147, 2001.

[13] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid term load forecasting of the country using statistical methodology: case study in Thailand," 2009 International conference on signal processing systems, pp.924-928, 2009.

[14] http://www.en.wikipedia.org/wiki/Time series.

[15] http://www.en.wikipedia.org/wiki/Linear regression.

Chapter 5

WA and SCCA Forecasting using Neural Network(NN)

5.1 Introduction

The power and usefulness of artificial neural networks (ANN) have been widely demonstrated in several applications including speech synthesis, diagnostic problems, medicine, business and finance, robotic control, signal processing, computer vision and many other problems that fall under the category of pattern recognition. For some application areas, neural network models also show promise in achieving human-like performance over more traditional artificial intelligence techniques.

Several energy load demands forecasts are major prediction in whole area by using many methods in the artificial intelligent approach such as neural network, and fuzzy logic approaches. It may reflect the demand that is inconsistent with the factors that affect an energy consumption load demand. References [6] to [16] are used to predict the total load based on an area of the country. For the present, this thesis will focus on the factors that predict the overall impact on the overall demand. This contrasts with some forecasting about the proposed the sub-control area to the total area (whole area) [18].

Many researches used artificial neural network approach to load demand forecasting. Kittipong Methaprayoon [8] proposed Multistage artificial neural network short-term load forecasting engine with front-end weather forecast. Demien Fay [19] offered the influence of weather forecast errors in short-term load forecasting models and proposed short-term [6-16], mid-term [18] and long-term [20] load forecasting using neural network (NN). Furthermore, many researchers used these methodologies for forecasting in the power system can demonstrate in Chapter 2. Therefore, this chapter proposes an ANN approach to whole and multi-area forecasting described as follow.

5.2 Artificial Neural Network (ANN)

An artificial neural network (ANN), usually called a neural network (NN), is a mathematical model or computational model that is inspired by the structure and/or functional aspects of biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are usually used to model complex relationships between inputs and outputs or to find patterns in data [17]. The tasks artificial neural networks are applied to tend to fall within the following broad categories:

- Function approximation, or regression analysis, including time series prediction, fitness approximation and modeling.
- Classification, including pattern and sequence recognition, novelty detection and sequential decision making.
- Data processing, including filtering, clustering, blind source separation and compression.
- Robotics, including directing manipulators, computer numerical control.

Artificial neural network structure, ANN are models of biological neural structures. The starting point for most neural networks is a model neuron, as showed in Figure 5.1. This neuron consists of multiple inputs and a single output. Each input is modified by a weight, which multiplies with the input value. The neuron will combine these weighted inputs and, with reference to a threshold value and activation function, use these to determine its output. This behavior follows closely our understanding of how real neurons work [17]. Generally, the input layer is considered a distributor of the signals from the external world. Hidden layers are considered to be categorizers or feature detectors of such signals. The output layer is considered a collector of the features detected and producer of the response. While this view of the neural network may be helpful in conceptualizing the functions of the layers, you should not take this model too literally as the functions described may neither be specific nor localized.

The neural network proposed in this thesis has three-layer based on feed-forward back propagation algorithm (FFBP). The fundamental structure of this algorithm can be seen in Figure 5.1. In this figure, the three layers of ANN consist of input layer, hidden layer and output layer. The hidden neuron can be varied to find the optimal weight and bias before simulation or forecasting.

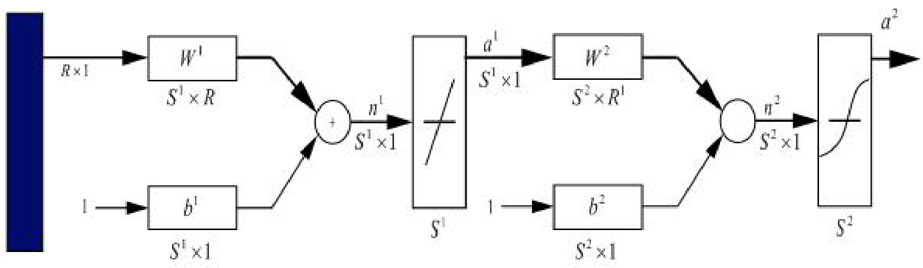

Figure 5.1: Structure of a neural network .

5.3 Proposed Method

5.3.1 Empirical Results of WAF

5.3.1.1 Data Sets and Models

The forecasts were made according to the following procedures:

- The electric load data (kWh) were recorded from EGAT [1], [5]. Table 5.1 shows the sample of electricity consumption load demand in whole area in year 2000, 2005, and 2007.
- Factors were supported from several organization, for instance TMD, and the economic factors of Thailand [2-4].
- All data must be normalized before using in the forecast process (equation (5.1)).
- An artificial neural network was used in prediction (figure 5.1).
- The appropriate factors were chosen to forecast load demand.
- To Summarize, and discuss.

Figure 5.2 shows features for neural network, for instance economic factors and weather factors. All selected factors in a model have to compute a correlated values before taking it into neural network algorithm to prediction. These are demonstrated in Table 5.2 to 5.3; however, the factors which should have correlated values between 45% to100% were chosen to be feature inputs for neural network.
The forecasts were made according to the following procedures:

Step 1, whole area forecasting (WAF), we employed the electricity consumption load demand in kWh in whole area of the country that are recored from EGAT (Monthly type).

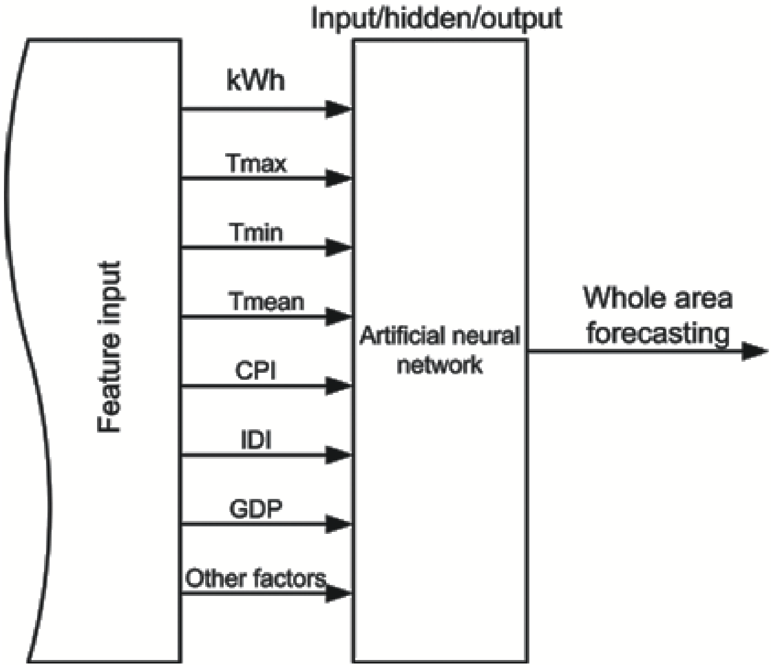

Figure 5.2: Features and neural network .

Table 5.1: Sample of electricity consumption load demand in whole are.

	kWh											
Y/M	Jan.	Feb.	Mar.	Apr.	May	Jan.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
2000, ($\times 10^9$)	7.72	7.39	8.61	8.01	8.75	8.42	8.46	8.54	8.32	8.51	7.82	7.98
2005, ($\times 10^9$)	1.02	1.04	1.19	1.13	1.24	1.15	1.13	1.15	1.10	1.03	1.08	1.08
2007, ($\times 10^9$)	1.13	1.09	1.33	1.22	1.27	1.28	1.26	1.27	1.25	1.25	1.17	1.19

Table 5.2: Samples of economic factors in whole area in 2000.

	%											
Y/M	Jan.	Feb.	Mar.	Apr.	May	Jan.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec
CPI^{2000}	97.10	97.50	97.50	97.2	97.3	97.5	97.6	98.20	98.60	98.10	98.2	98.2
IDI^{2000}	93.53	97.99	106.90	93.77	99.11	100.61	96.86	98.10	103.2	102.59	104.25	103.07

Step 2, influencing factors, economic and weather factors, are employed to identify the load behavior and the trend component of electricity load demand. Economic factors such as the industrial index (IDI), the consumer price index (CPI), and gross domestic product are used to identify the cycle or behavior of load demand. They are came form recording in monthly type. Furthermore, weather factors for example, maximum, minimum, and mean temperature are recorded from TMD of the country. Table 5.3 shows weather sample in whole area year 2000.

Step 3, all data above must be demonstrated by curve fitting and normalized by using an equation(5.1):

Table 5.3: Samples of weather factors in whole area in 2000.

						^oC						
Y/M	Jan.	Feb.	Mar.	Apr.	May	Jan.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec
$Tmax^{2000}$	37.1	39.5	40.0	41.5	38.6	37.9	38.2	37.9	37.4	36.7	37.0	37.5
$Tmin^{2000}$	13.1	11.1	15.5	19.9	19.9	20.4	20.4	20.1	18.7	19.3	12.5	13.4

Table 5.4: Samples of the correlated value between factors and electricity load demand.

	Correlation (%)	kWh						
Energy	kWh	100	CPI					
	CPI	87.89	100	IDI				
Economic	IDI	93.90	90.10	100	GDP			
	GDP	86.5	83.8	96.10	100	T_{max}		
	T_{max}	9.760	-9.50	-5.65	-14.20	100	T_{min}	
Weather	T_{min}	24.60	5.53	-3.23	-15.80	24.25	100	T_{mean}
	T_{mean}	26.50	2.22	-1.36	-14.50	72.33	78.24	100

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5.1}$$

where y is the normalized value, x is variable.

Step 4, the appropriate factors are considered from the high correlated values, such as 70%, up to 90% correlation. These factors will be used to be feature inputs for the neural network.

Step 5, the forecasting model shown in Figure 5.2 is used for learning data and forecasting the energy load demand. The historical kWh data from 1997 to 2006 in monthly type will be taken to the neural network for learning and finding a good performance of network to build the forecasting model.

In forecasting by using neural network algorithm, it composes of three layers of the neural network; an input layer, a hidden layer, and an output layer. Types of input factors are economic and weather factors. In addition, we varied the hidden nodes in the hidden layers from five, eight, twenty, and twenty-five, for finding the suitable number of node in hidden layer. According to the results, eight-nodes of hidden layer is chosen for the model prediction.

Step 6, analyses and summary are carried out.

5.3.1.2 Simulation Results

Table 5.5 shows the results of whole area forecasting by using neural network algorithm, in January to December 2007, unit in kWh. The table shows the

Table 5.5: Results of whole area forecasting, year 2007.

Whole-area	E-kWh	E-kWh	Accuracy	
Month	Actual	Forecasted	Error	APE
January	11255610455	11164176393	0.008	0.812
February	10903740111	11393099648	-0.045	4.488
March	13256202895	13109249168	0.011	1.109
April	12222850690	12160212602	0.005	0.512
May	12668470924	12704436327	-0.003	0.284
June	12768025563	12571058166	0.015	1.543
July	12577211640	12463415497	0.009	0.905
August	12722562991	12585177620	0.011	1.080
September	12504929463	12413918715	0.007	0.728
October	12500303814	12411062678	0.007	0.714
November	11678682183	12105865076	-0.037	3.658
December	11866899916	11814428981	0.004	0.442
		MAPE (%)		1.356

actual values, the forecasted values, the percentage errors, and the absolute percent errors in each month of 2007. An actual value came from EGAT, and the forecasted values in the third column (3^{rd}) were obtained from ANN forecasting model. The fourth column explains an error in each month, such as 0.008 in January, 0.003 in May, and 0.004 in December. The fifth column shown low accuracy appeared in February and November, are 4.488% and 3.654%, respectively.

5.3.2 Empirical Results of SCCAF

5.3.2.1 Data Sets and Models

The forecasts were made according to the following procedures.

Step 1, multi-substation control center area forecasting (MSCCAF) are employed for the ECLD in kWh in multi-substation control center area of the country.

Step 2, influencing factors, economic and weather factors, are employed to identify the load behavior and the trend component of electricity load demand. Economic factors such as the industrial index (IDI), the consumer price index (CPI), and gross domestic product are used to identify the cycle or behavior of load demand. They are came form recording in monthly type. Furthermore, weather factors for example,

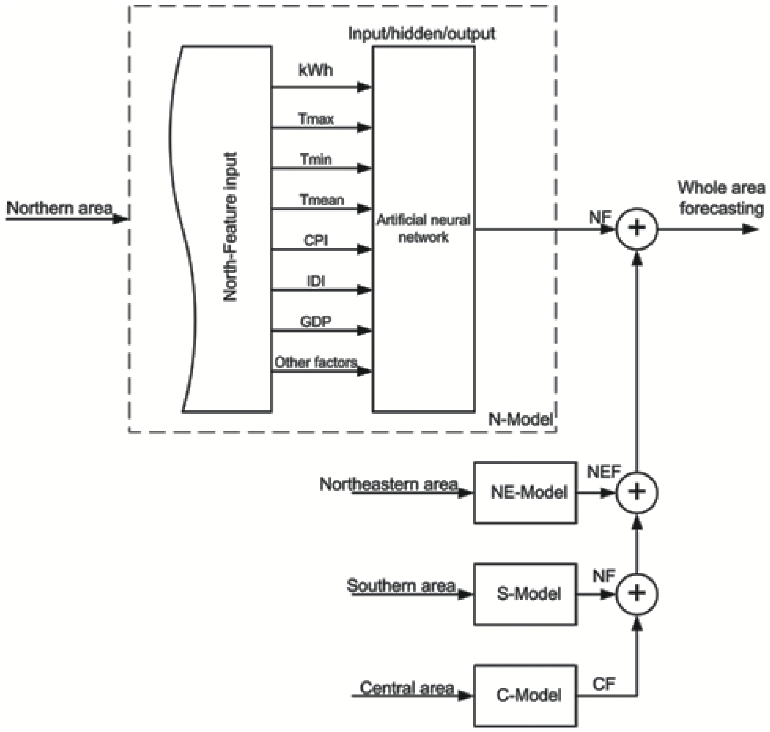

Figure 5.3: Multi-substation control center area forecasting (MSCCAF) model.

maximum, minimum, and mean temperature are recorded from TMD of the country. Table 5.3 shows weather sample in whole area year 2000.

Step 3, all data above must be processed by curve fitting and normalized by using equation 5.1.

Step 4, all appropriate data analyzed in **Chapter 3** selected according to a high correlated values, such as 70%, 90 up% correlation. The chosen factor will be used as feature inputs for the neural network.

Step 5, this step is similar **step 4**. It is just changed the whole to multi-area forecasting illustrated in Chapter 3. Then there are four-model of neural networks for the forecasting, and showed in Figure 5.3.

Neural network in this model is also used feed forward back propagation algorithm. Furthermore, there are three layers that are input, hidden, and output layers. The period of learning learning algorithm, the data from 1997 to 2006 in monthly type will be taken to the neural network for learning and finding a good performance of network to build the forecasting model. In forecasting by using neural network algorithm, there are three layers of the neural network that are an input layer, a hidden layer, and an output layer. Input layer uses the factors such as economic and weather. In addition,

Table 5.6: Results of Northern sub-control area forecasting (NSCCA), year 2007.

NSCCA	E-kWh	E-kWh	Accuracy	
Month	Actual	Forecasted	Error	APE
January	1038481065	1051124015	-0.012	1.217
February	1020435007	1052296876	-0.031	3.122
March	1225352263	1192311201	0.027	2.696
April	1224067238	1156038142	0.056	5.558
May	1167594892	1159890736	0.007	0.660
June	1215331075	1166267961	0.040	4.037
July	1218343368	1162320010	0.046	4.598
August	1245747126	1168295162	0.062	6.217
September	1193942663	1160998484	0.028	2.759
October	1160576854	1158695793	0.002	0.162
November	1099262401	1114631318	-0.014	1.398
December	1110833578	1126460151	-0.014	1.407
		MAPE (%)		2.819

the hidden nodes of hidden layer are varied from five, eight, twenty, and twenty-five. Finally, the eight-node of hidden layer was chosen for the model prediction.

Step 6, analyses and summary are carried out.

5.3.2.2 Simulation Results

Table 5.6 to Table 5.9 show the results of multi-substation. They forecasts by using the neural network algorithm, and the feature data such as the consumer price index (CPI), the industrial index (IDI) are employed from EGAT in each area of substation and from the government of Thailand.

After forecasting, northern substation control central area (NSCCA), the results are contained in the Table 5.6 which can be explained as follow. The forecasted results are from January 2007 to December 2007. The factors, for instance $t_{max,min,mean}$, CPI, IDI covered the northern area of the country are used in the model. The investigated results show the actual values, the forecasted values, the percentage errors, and the absolute percent errors in each month of 2007. The actual values came from EGAT and, forecasted value in the third column came from forecasting. In particular, The fourth column (4^{th}) explains an error in each month, such as -0.012 in January, 0.007 in

Table 5.7: Results of Northeastern sub-control area forecasting (NESCCA), year 2007.

NESCCA	E-kWh	E-kWh	Accuracy	
Month	Actual	Forecasted	Error	APE
January	1094940168	1065330308	0.027	2.704
February	1092987541	1130241773	-0.034	3.408
March	1337842093	1289253729	0.036	3.632
April	1244967156	1225636988	0.016	1.553
May	1237520099	1250701389	-0.011	1.065
June	1280766830	1258396356	0.017	1.747
July	1285211547	1239574305	0.036	3.551
August	1239810725	1244616591	-0.004	0.388
September	1235876599	1247772521	-0.010	0.963
October	1204380426	1246388351	-0.035	3.488
November	1089843587	1198025040	-0.099	9.926
December	1187655478	1201438269	-0.012	1.161
		MAPE (%)		2.798

May, and -0.014 in December. In this model, the fifth column shows that low accuracy appearing in April and in August are 5.558% and 6.217%, respectively. Consequently, in Northern substation control center area forecasting, the MAPE value in 2007, is 2.819%.

Table 5.7 demonstrates the northeastern substation control central area (NESCCA) forecasting in year 2007. Te factors which were recorded from the northeastern area are taken to be the feature inputs for neural network in northern substation control center area forecasting model.

The results of northeastern area forecasting (NESCCA) can be demonstrated in Table 5.7 from January to December 2007. The actual values, the forecasted values, the errors, the APE values, and forecasted values are illustrated in each column. It also shows a performance of a forecast model in MAPE. Maximum percent error is about 9.926% in November; and minimum percent error is about 0.388% in August. As a result, the mean absolute percentage error (MAPE) is 2.798%. In northeastern substation control center model used especially the factors in northeastern area, for instance $T_{northeastern}$ max, min, mean, and $CPI_{northeast}$; however, $IDI_{W-claim}$ that its data come from the government of the country is instead of an IDI of an area.

The results of southern substation control central area (SSCCA) fore-

Table 5.8: Results of Southern sub-control area forecasting (SSCCA), year 2007.

SSCCA	E-kWh	E-kWh	Accuracy	
Month	Actual	Forecasted	Error	APE
January	973180849	933257296.7	0.041	4.102
February	926033159	917002876.5	0.010	0.975
March	1072148004	1019986267.0	0.049	4.865
April	1032229554	968731440.7	0.062	6.152
May	1040381230	1003760331.0	0.035	3.520
June	1026516877	991822750.5	0.034	3.380
July	1035113272	1003861452.0	0.030	3.019
August	1047354952	1018190302.0	0.028	2.785
September	1023527786	993914300.5	0.029	2.893
October	1012443297	1006173363.0	0.006	0.619
November	974956145	1001669862.0	-0.027	2.740
December	1006716301	1013114612.0	-0.006	0.636
		MAPE (%)		2.974

casting show in Table 5.8. The table shows the actual values, the forecasted values, the percentage errors, and the absolute percent errors in each month of 2007. The actual values have come from EGAT and forecasted values in the third column came from forecasting. The fourth column explains an error in each month, such as maximum 0.049 in March, 0.041 in January, and 0.062 in April. In this forecasting model, high accuracy appearing in February, October, and December are 0.975%, 0.619%, and 0.636% respectively.

Finally in the results of central substation control central area (CSCCA) forecasting show in Table 5.9. In the table has shown the actual values, the forecasted values, the percentage errors, and the absolute percent errors in each month of 2007. The actual values have come from EGAT of the country and forecasted values in the third column have come from forecasting. The fourth column explains an error in each month, such as maximum -0.069 in October, -0.078 in February. In this forecasting model, high accuracy giving in March and in July that are 0.051% , 0.286% respectively.

In Table 5.10 the forecasted value of all regions are combined. Final forecasted values are integrated from Area-1, Area-2, Area-3, and Area-4 that are northern, northeastern, southern, and central area, respectively. Finally, the mean absolute per-

Table 5.9: Results of Central sub-control area forecasting (CSCCA), year 2007.

CSCCA	E-kWh	E-kWh	Accuracy	
Month	Actual	Forecasted	Error	APE
January	8149008373	8401599236	-0.031	3.100
February	7864284404	8476123963	-0.078	7.780
March	9620860535	9586065349	0.004	0.362
April	8721586742	8825022932	-0.012	1.186
May	9222974703	9395597458	-0.019	1.872
June	9245410781	9240689596	0.001	0.051
July	9038543453	9064359696	-0.003	0.286
August	9189650188	9333840414	-0.016	1.569
September	9051582415	9211167112	-0.018	1.763
October	9122903237	9252767547	-0.014	1.423
November	8514620050	9105548277	-0.069	6.940
December	8561694559	8740254986	-0.021	2.086
		MAPE (%)		2.368

Table 5.10: Results of Whole area from sub-control area forecasting (SCCA), year 2007.

region	Area-1	Area-2	Area-3	Area-4	Whole-F	Actual-W	PE	APE
Name	Northern	Northeast	Southern	Central	kWh	kWh	%	%
Month	1051124015	1065330308	933257297	8401599236	11451310855	11255610455	-1.739	1.739
January	1052296876	1130241773	917002876	8476123963	11575665488	10903740111	-6.162	6.162
February	1192311201	1289253729	1019986267	9586065349	13087616547	13256202895	1.272	1.272
March	1156038142	1225636988	968731441	8825022932	12175429503	12222850690	0.388	0.388
April	1159890736	1250701389	1003760331	9395597458	12809949913	12668470924	-1.117	1.117
May	1166267961	1258396356	991822751	9240689596	12657176664	12768025563	0.868	0.868
June	1162320010	1239574305	1003861452	9064359696	12470115462	12577211640	0.852	0.852
July	1168295162	1244616591	1018190302	9333840414	12764942469	12722562991	-0.333	0.333
August	1160998484	1247772521	993914300	9211167112	12613852418	12504929463	-0.871	0.871
September	1158695793	1246388351	1006173363	9252767547	12664025055	12500303814	-1.310	1.310
October	1114631318	1198025040	1001669862	9105548277	12419874498	11678682183	-6.347	6.347
November	1126460151	1201438269	1013114612	8740254986	12081268018	11866899916	-1.806	1.806
December	1051124015	1065330308	933257297	8401599236	11451310855	11255610455	-1.739	1.739
						MAPE (%)		1.922

centage error (MAPE) of multi-station control central area is 1.922%. Maximum error occurs in October about 6.347%, and in February about 6.162%.

Table 5.11: Comparison between WAF and MSCCAF models in 2007.

Month	Actual	WAF	MAPE	MSCCAF	MAPE(%)
	kWh	kWh	%	kWh	%
January	11255610455	11164176393		11451310855	
February	10903740111	11393099648		11575665488	
March	13256202895	13109249168		13087616547	
April	12222850690	12160212602		12175429503	
May	12668470924	12704436327		12809949913	
June	12768025563	12571058166	1.35	12657176664	1.92
July	12577211640	12463415497		12470115462	
August	12722562991	12585177620		12764942469	
September	12504929463	12413918715		12613852418	
October	12500303814	12411062678		12664025055	
November	11678682183	12105865076		12419874498	
December	11866899916	11814428981		12081268018	

5.4 Discussion

Table 5.11 demonstrates the data came from the results in the previous sections of whole area forecasting (WAF) and multi-substation control central area forecasting (MSCCAF). In each column of the table shows month, actual, WAF, MAPE, MSCCAF, and MAPE.

In comparison between WAF and MSCCAF models in 2007, the WAF model confirmed better result than of that MSCCAF model. There are two reasons for supporting the given results.

Firstly, there are no data of industrial index (IDI) specified in each substation control area. Therefore, MSCCAF must be used IDI that is specified for the whole country. As the results, some low accuracy will be observed in substation control area that are not crowned with industrial.

The second reason is consumption of regions. The region of central substation is much larger than other regions. This area is covered by three sub-region of central area, Metropolitan, and Bangkok. It has many industrial sectors and the number of population. The electric demand in this area is very high compared with other areas. Therefore, the final forecasted value of SCCA will be depended on the large area of the country mainly. In conclusion, factors in each area, the size of area, and consumption demand of each region are influential to predict the results of the energy demand in the power system.

5.5 Summary

This chapter proposes artificial neural network approach to midterm load forecasting based on whole area and sub-control area predictions. Whole area forecasting used information data of the country; also, sub-control center area forecasting used data from sub-control center area of the country based on EGAT. The neural network is used to forecast for all models by using three layers-input, hidden, and output layers. As a result, the whole area forecasts shows better result than the multi-subarea model.

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] George gross, franscisco D.galiana, "Short-term load forecasting," Preceeding of the IEEE, vol.75, no.12, pp.1558-1573, December 1987.

[7] Tomonoobu senjyu, hitoshi takara,katsumi ueezato,toshihisa funabashi, "One-hour ahead load forecasting using neural network," IEEE Trans. on power systems, vol.17, no.1, pp.113-118, February 2002.

[8] Kittipong Methaprayoon, Wei-Jen Lee, Sothaya Rasmiddatt, James R.Liao, Richard J.Ross, "Multistage artificial neural short-term load forecasting engine with front-end weather forecast," IEEE Trans.Ind.app., vol.43, no.6, Nov.2007.

[9] G.A.Adepoju, S.O.A.ogaunjuyigbe, K.O.Alawode, "Application of neural network to load forecasting in nigerian electrical power system," The pacific journal of science and technology ,vol.8, no.1, pp.68-72, May 2007.

[10] Mohsen Hayati, Yazdan Shirvany, "Artificial neural network approach for short term load forecasting for Illam region," International journal of electrical,computer,and system engineering ,vol.1, no.2, pp.121-125.

[11] Bhavesh kumar chauha,amit sharma, m.hanmandlu, "Neuro-fuzzy approach based short term electric load forecasting," IEEE-PES Transmission and Distribution Conference-Exhibition:Asia and Pacific Dalian, Chaina, 2005.

[12] Jian-Chang Lu,Dong-xiao niu,zheng-yuan jia, "A study of short term load forecasting based on arima-ann," Machine Learning and Cybernetics, Volume 5, pp.3183 - 3187 vol.5, August 2004.

[13] Ummuhan basaran filik,Mehmet kurban, "A new approach for the short term load forecasting with autoregressive and artificial neural network models," International journal of computational intelligence research , vol.3, no.1, pp.66-71.

[14] P.K.Dash, S.Mishra, S.Dash, A.C.Liew, "Genetic optimization of a self organizing fuzzy-neural network for load forecasting," IEEE 2000.

[15] Gwo-ching liao, Ta-peng tsao, "Novel GA-Based Approach and Neural fuzzy networks application in short-term load forecasting," Power Engineering Society General Meeting, vol.1, pp.589-594, June 2004.

[16] Kyung-bin song,young-sik baek,dug hun hong and gilsoo jang., "Short-term load forecasting for the holidays using fuzzy linear regression method," IEEE Trans. on power systems, vol.20, no.1, pp.96-101, February 2005.

[17] Ben krose, and Patrick van der Smagt, "An Introduction Neural Network," Eight edition, November 1996.

[18] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "The comparision of mid term load forecasting between multi-regional and whole country area using artificial neural network," International journal of computer and electrical engineering, vol. 2, no. 2, pp.338-343, 2010 (Singapore).

[19] Damien Fay and John V.Ringwood, "On the influence of weather forecast errors in short-term load forecsting models, " IEEE Trans.power syst., vol.25, no.3, Aug.2010.

[20] Mohammad Moradi Dalvand, Seyed Bahram Azami, Hadi Tarimoradi, "Long-term load forecasting of Iranian power grid using fuzzy and artificial neural network," Universities power engineering conference, pp.1-4, 2008.

Chapter 6

WA and SCCA Forecasting using Wavelet and NN

6.1 Introduction

Most signals are represented in the time domain. More information about the time signals can be obtained by applying signal analysis. The time signals can be transformed using analysis functions. The Fourier transform is the most commonly known method to analyze a time signal for its frequency content. A relatively new analysis method is the wavelet analysis. The wavelet analysis differs from the Fourier analysis by using short wavelets instead of long waves for the analysis function. The wavelet analysis has some major advantages over Fourier transform which makes it an interesting alternative for many applications. The use and fields of applications of wavelet analysis have grown rapidly in the last years.

The section proposes a wavelet transform technique in preprocessing stage and integrated the wavelet transform to neural network algorithm for forecasting. Many algorithms, used for preprocessing, were used in forecasting such as self organizing map (SOM), K-mean Clustering, etc. This section focuses on preprocessing stage because it related with the objective of research (feature selection). Furthermore the preprocessing method will be investigates this method with a whole area and sub-control center area forecasting described in Chapter 3. Several papers used these approaches before predicting such as Qi Wu [6] presented the hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts, furthermore a short-term load forecasting by using similar day-based wavelet neural network [7] was proposed. Ajay Shekhar [8] presented the intelligent hybrid wavelet models for prediction. In 2009, the combination of wavelet transform and neuro-evolutionary algorithm approach to demand forecasting [9]. In 2008, the research proposed an adaptive wavelet neural network-based energy price forecasting in electricity markets [10] and in the year 2006 the wavelet based nonlinear multi-scale decomposition model [11]. In ref. [12], in the same year, Tai Nengling proposed techniques of applying wavelet transform into combined model for demand forecasting in electricity. Bai-Ling Zhang [13] presented an adaptive neural wavelet model, and Tongxin Zheng [14] proposed a hybrid wavelet-Kalman filter method for load forecasting. Lastly, the wavelets transform and neural networks for short-term electric load forecasting were proposed [15].

This chapter the whole area and sub-control area forecasting are performed by using wavelet transform in the preprocessing stage of all areas and neural network for forecasting in the last process. In the preprocessing stage, wavelet transform is used to separate the original signal of demand into one to four levels. Then, each sub-signal will be taken to find the relationship with specific factors before choosing the suitable factors to be feature inputs for neural network in prediction process. The results of the whole and sub-control center area forecasting are presented in the comparison based on mean absolute percentage error (MAPE). The analysis of a non-stationary signal using the FFT or the STFT does not give satisfactory results. Better results can be obtained by using wavelet analysis. One advantage of wavelet analysis is the ability to perform local analysis. Wavelet analysis is able to reveal signal aspects that other analysis techniques miss, such as trends, breakdown points, discontinuities, etc. In comparison to the STFT, wavelet analysis makes it possible to perform a multi-resolution analysis.

6.2 Theoretical Background

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Wavelets were developed independently in the fields of mathematics, quantum physics, electrical engineering, and seismic geology. Interchanges between these fields during the last ten years have led to many new wavelet applications such as image compression, turbulence, human vision, radar, and earthquake prediction. This paper introduces wavelets to the interested technical person outside of the digital signal processing field. The theory describes the history of wavelets beginning with Fourier, compare wavelet transforms with Fourier transforms, state properties and other special aspects of wavelets, and finish with some interesting applications such as image compression, musical tones, and de-noising noisy data.

The time-frequency resolution problem is caused by the Heisenberg uncertainty principle and exists regardless of the used analysis technique. For the STFT, a fixed time-frequency resolution is used. By using an approach called multi-resolution analysis (MRA), it is possible to analyze a signal at different frequencies with different resolutions. The change in resolution is schematically displayed in Figure 6.1.

For the resolution of Figure 6.1, it is assumed that low frequencies last for the entire duration of the signal, whereas high frequencies appear from time to time as short burst. This is often the case in practical applications.

The wavelet analysis calculates the correlation between the signal under consideration and a wavelet function $\psi(t)$. The similarity between the signal and the analyzing wavelet function is computed separately for different time intervals, resulting in a two dimensional representation. The analyzing wavelet function $\psi(t)$ is also referred to as the mother wavelet.

6.2.1 Wavelet Transform (WT)

In comparison to the Fourier transform, the analyzing function of the wavelet transform $\psi(t)$ can be chosen with more freedom, without the need of using sine-forms. A wavelet function is a small wave, which must be oscillatory in some way to discriminate between different frequencies. The wavelet contains both the analyzing shape and the window. For the CWT several kind of wavelet functions are developed which all have specific properties [19].

An analyzing function is classified as a wavelet if the following mathematical criteria are satisfied [20]:

1. A wavelet must have finite energy, in equation (6.1)

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty \tag{6.1}$$

The energy E equals the integrated squared magnitude of the analyzing function and must be less than infinity.

2. If $\psi(t)$ is the Fourier transform of the wavelet, the following condition must hold, in equation (6.2)

$$C_\psi = \int_0^{\infty} \frac{\left|\hat{\psi}(f)\right|^2}{f} df < \infty \tag{6.2}$$

This condition implies that the wavelet has no zero frequency component ($\psi(t) = 0$), i.e. the mean of the wavelet $\psi(t)$ must equal zero. This condition is known

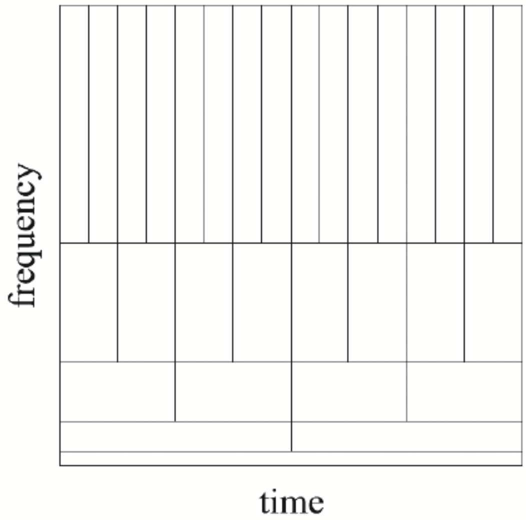

Figure 6.1: Multi-resolution time-frequency plane.

as the admissibility constant. The value of C_ψ depends on the chosen wavelet.

3. For complex wavelets the Fourier transform must be both real and vanish for negative frequencies.

- Continuous wavelet transforms (CWT)

  It is used to divide a continuous-time function into wavelets. Unlike Fourier transform, the continuous wavelet transform possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. The continuous wavelet transform is defined as follow:

$$X_{WT(\tau,s)} = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\psi * \left(\frac{t-\tau}{s}\right) dt. \tag{6.3}$$

  This transformed signal $X_{WT(\tau,s)}$ is a function of the translation parameter τ and the scale parameter s. The mother wavelet is denoted by ψ , the * indicates that the complex conjugate is used in case of a complex wavelet. The signal energy is normalized at every scale by dividing the wavelet coefficients by $1 \big/ \sqrt{|s|}$. This ensures that the wavelets have the same energy at every scale.

  The mother wavelet is contracted and dilated by changing the scale parameter s . The variation in scale s changes not only the central frequency f_c of the wavelet, but also the window length. Therefore the scale s is used instead of the frequency for representing the results of the wavelet analysis. The translation parameter τ specifies the location of the wavelet in time, by changing τ the wavelet can be shifted over the signal. For constant scale s and varying translation τ the rows of the time-scale plane are filled, varying the scale s and keeping the translation τ constant fills the columns of the time-scale plane. The elements in $X_{WT(\tau,s)}$ are called wavelet coefficients, each wavelet coefficient is associated to a scale (frequency) and a point in the time domain.

The WT also has an inverse transformation, as was the case for the FT and the STFT. The inverse continuous wavelet transformation (ICWT) is defined by

$$x(t) = \frac{1}{C_{\psi}^{2}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X_{WT(\tau,s)} \left(\frac{t-\tau}{s} \right) d\tau ds. \tag{6.4}$$

Note that the admissibility constant C_{ψ} must satisfy the second wavelet condition.

A wavelet function has its own central frequency f_c at each scale; the scale s is inversely proportional to that frequency. A large scale corresponds to a low frequency, giving global information of the signal. Small scales correspond to high frequencies, providing detail signal information.

For the WT, the Heisenberg inequality still holds, the bandwidth-time product $\Delta t \Delta f$ is constant and lower bounded. Decreasing the scale s, i.e. a shorter window, will increase the time resolution Δt, resulting in a decreasing frequency resolution Δf. This implies that the frequency resolution Δf is proportional to the frequency f, i.e. wavelet analysis has a constant relative frequency resolution [19]. The Morlet wavelet, is obtained using a Gaussian window, where f_c is the center frequency and f_b is the bandwidth parameter

$$\psi(t) = g(t)e^{-j2\pi f_c t}, g(t) = \sqrt{\pi f_b e^{t^2/f_b}} \tag{6.5}$$

The center frequency f_c and the bandwidth parameter f_b of the wavelet are the tuning parameters. For the Morlet wavelet, scale and frequency are coupled as

$$f = \frac{f_c}{s} \tag{6.6}$$

The calculation of the continuous wavelet transform is usually performed by taking discrete values for the scaling parameter s and translation parameter τ. The resulting wavelet coefficients are called wavelet series. For analysis purposes only, the discretization can be done arbitrarily, however if reconstruction is required, the wavelet restrictions mentioned become important.

The constant relative frequency resolution of the wavelet analysis is also known as the constant Q property. is the quality factor of the filter, defined as the center-frequency f_c divided by the bandwidth f_b [19]. For a constant Q analysis (constant relative frequency resolution), a dyadic sample-grid for the scaling seems suitable. A dyadic grid is also found in the human hearing and music. A dyadic grid discretizes the scale parameter on a logarithmic scale. The time parameter is discretized with respect to the scale parameter. The dyadic grid is one of the most simple and efficient discretization methods for practical purposes and leads to the construction of an orthonormal wavelet basis [20]. Wavelet series can be calculated as

$$X_{WT_{m,n}} = \int_{-\infty}^{\infty} x(t)\psi_{m,n}(t)dt, \quad with \quad \psi_{m,n} = s_0^{-m/2}\psi(s_0^{-m}t - n\tau_0) \qquad (6.7)$$

The integers m and n control the wavelet dilation and translation. For a dyadic grid, $s_0 = 0$ and $\tau_0 = 1$. Discrete dyadic grid wavelets are chosen to be orthonormal, i.e. they are orthogonal to each other and normalized to have unit energy. This choice allows the reconstruction of the original signal by

$$x(t) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} X_{WT_{m,n}}\psi_{m,n}(t) \qquad (6.8)$$

- Discrete wavelet transforms (DWT)

The wavelet series is just a sampled version of CWT and its computation may consume significant amount of time and resources, depending on the resolution required. The Discrete wavelet transform (DWT), which is based on sub-band coding is found to yield a fast computation of wavelet transform. It is easy to implement and reduces the computation time and resources required. In CWT above, the signals are analyzed using a set of basis functions which relate to each other by simple scaling and translation. In the case of DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales. The discrete wavelet transform (DWT) uses filter banks for the construction of the multi-resolution time-frequency plane. Filter banks will be introduced in next section. The DWT uses multi-resolution filter banks and special wavelet filters for the analysis and reconstruction of signals. The DWT will be discussed in next section.

- Filter Banks

  A filter bank consists of filters which separate a signal into frequency bands [21]. An example of a two channel filter bank is shown in Figure 6.2. A discrete time signal $x(k)$ enters the analysis bank and is filtered by the filters $L(z)$ and $H(z)$ which separate the frequency content of the input signal in frequency bands of equal width. The filters $L(z)$ and $H(z)$ are therefore respectively a low-pass and a high-pass filter. The output of the filters each contains half the frequency content, but an equal amount of samples as the input signal. The two outputs together contain the same frequency content as the input signal; however the amount of data is doubled. Therefore down sampling by a factor two, denoted by ($\downarrow 2$), is applied to the outputs of the filters in the analysis bank.

  Reconstruction of the original signal is possible using the synthesis filter bank [19], [21]. In the synthesis bank the signals are up sampled ($\uparrow 2$) and passed through the filters $L'(z)$ and $H'(z)$. The filters in the synthesis bank are based on the filters in the analysis bank. The outputs of the filters in the synthesis bank are summed, leading to the reconstructed signal $y(k)$.

  The different output signals of the analysis filter bank are called subbands, the filter-bank technique is also called subband coding [19].

- Down and up sampling

  The low and high-pass filters $L(z)$ and $H(z)$ split the frequency content of the signal in half. It therefore seems logical to perform a downsampling with a factor two to avoid redundancy. If half of the samples of the filtered signals $c_l(k)$ and $c_h(k)$ are reduced, it is still possible to re-

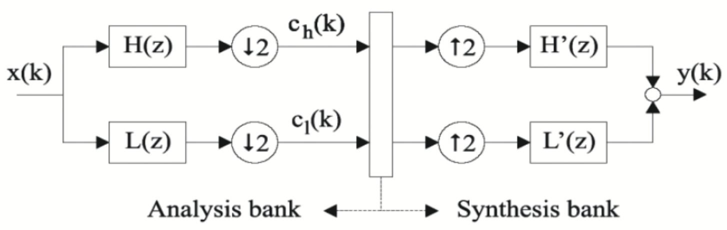

Figure 6.2: Two-channel filter bank.

construct the signal $x(t)$. The downsampling operation ($\downarrow 2$) saves only the even-numbered components of the filter output, hence it is not invertible. In the frequency domain, the effect of discarding information is called aliasing. If the Shannon sampling theorem is met, no less of information occurs. The sampling theorem of Shannon states that downsampling a sampled signal by a factor M produces a signal whose spectrum can be calculated by partitioning the original spectrum into M equal bands and summing these bands [19].

In the synthesis bank the signals are first upsampled before filtering. The upsampling by a factor two ($\uparrow 2$) is performed by adding zeros in between the samples of the original signal. Note that first downsampling a signal and than upsampling it again will not return the original signal.

$$[x] = \begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ x(3) \\ x(4) \\ . \end{bmatrix} \qquad (\downarrow 2)x = \begin{bmatrix} x(0) \\ x(2) \\ x(4) \\ . \end{bmatrix} \qquad (\uparrow 2)(\downarrow 2)x = \begin{bmatrix} x(0) \\ 0 \\ x(2) \\ 0 \\ x(4) \\ . \end{bmatrix} \tag{6.9}$$

The transpose of ($\downarrow 2$) is ($\uparrow 2$). Since transposes come in reverse order, synthesis can be preformed as the transpose of the analysis. Furthermore $(\downarrow 2)(\uparrow 2) = I$, since ($\uparrow 2$) is the right-inverse of ($\downarrow 2$) [21]. This indicates that it is possible to obtain the original signal again with up and downsampling. By first inserting zeros and then removing them, the original signal is obtained again.

- Perfect reconstruction

For perfect reconstruction to be possible, the filter bank should be biorthogonal. Furthermore some design criteria for both the analysis and synthesis filters should be met to prevent aliasing and distortion and to guarantee a perfect reconstruction [21].

In the two channels filter bank of Figure 6.3, the filters $L(z)$ and $H(z)$ split the signal into two frequency bands, i.e. the filters are respectively a low-pass and a high-pass filter. If the filters were perfect brick-wall filters, the downsampling would not lead to loss of information. However ideal filters cannot be realized in practice, so a transition band exists. Besides aliasing, this leads to an amplitude and phase distortion in each of the channels of the filter band [19].

For the two channel filter bank of Figure 6.3, aliasing can be prevented by designing the filters of the synthesis filter bank as [21]

$$L'(z) = H(-z) \tag{6.10}$$

$$H'(z) = -L(-z) \tag{6.11}$$

To eliminate distortion, a product filter $P_0(z) = L'(z)L(z)$ is defined. Distortion can be avoided if

$$P_0(z) - P_0(-z) = 2z^{-N}, \tag{6.12}$$

Where N is the overall delay in the filter bank. Generally an N^{th} order filter produces a delay of N sample [19]. The perfect reconstruction filter bank can be designed in two steps [21]:

1. Design a low-pass filter P_0 satisfying in equation (6.11).

2. Factor P_0 into $P_0(z) = L'(z)L(z)$ and use an equation (6.9) and (6.10) to calculate $H(z)$ and $H'(z)$.

The design of the product filter P_0 of the first step and the factorization of the second step can be done in several ways. More information about the wavelet filter design can be found in [21].

- Multiresolution filter banks

  The CWT of the previous section performs a Multiresolution analysis which makes it possible to analyze a signal at different frequencies with different resolutions. For low frequencies (high scales), a good frequency resolution is more important. The CWT has a time-frequency resolution as shown in Figure 6.1. This Multiresolution can also be obtained using filter banks, resulting in the discrete wavelet transform (DWT). Note that the discretized version of the CWT is not equal to the DWT, the DWT uses filter banks, whereas the discretized CWT uses discretized versions of the scale and dilatation axes.

  The low-pass and high-pass filtering branches of the filter bank retrieve respectively the approximations and details of the signal $x(k)$. In Figure 6.2(a), a three level filter bank is shown. The filter bank can be expanded to an arbitrary level, depending on the desired resolution. The coefficients $C_l(k)$ (see Figure 6.3(a)) represent the lowest half of the frequencies in $x(k)$, downsampling doubles the frequency resolution. The time resolution is halved, i.e. only half the number of samples are present in $C_l(k)$.

  In the second level, the outputs of $L(z)$ and $H(z)$ double the time resolution and decrease the frequency content, i.e. the width of the window is increased. After each level, the output of the high-pass filter represents the highest half of the frequency content of the low-pass filter of the previous level, this leads to a pass-band. The time-frequency resolution of the analysis bank of Figure 6.3 (a) is similar to the resolution shown in Figure 6.1. For a special set of filters $L(z)$ and $H(z)$ this structure is called the DWT, the filters are called wavelet filters.

- Wavelet filters

  The relationship between the CWT and the DWT is not very obvious. The wavelets in the CWT have a center frequency and act as a band-pass filter in the convolution of the wavelet function with the signal $x(t)$. The sequence of low-pass filter, downsampling and high-pass filter also acts as a band-pass filter. Finaly, in this research, discrete wavelet transform is used to separate an original signal that is a load demand.

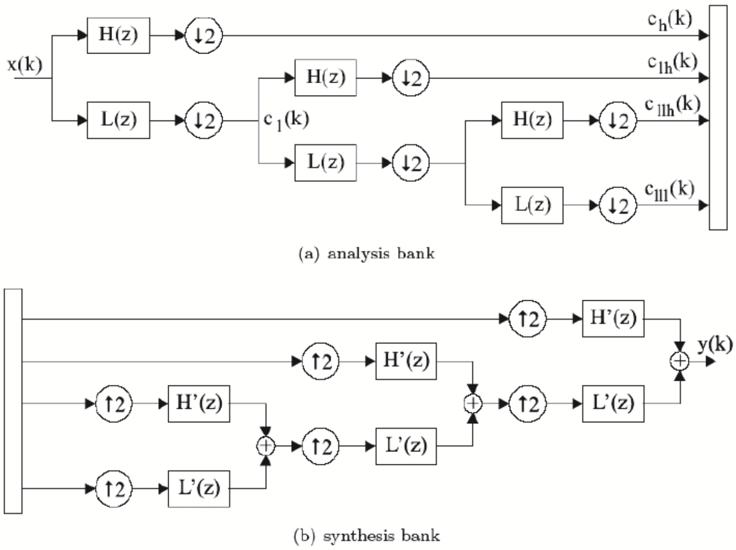

Figure 6.3: Three-level filter bank.

6.2.2 Neural Network (NN)

Neural Network Artificial neural network is one of the good approaches to apply to the load demand forecasting problem because this technique does not require explicit models to represent the complex relationship between the load demand and its factors. The neural network algorithm presented in this research composes of three layers: input layer, hidden layer, and output layer based on a feed-forward back propagation algorithm (FFBP). The input variables come from historical and existing data of factors affecting the load demand. The number of inputs, the number of hidden nodes (neurons), and the activate transfer function in each layer affect the forecasted performance. Hence, they need to be chosen carefully.

6.3 Proposed Method

6.3.1 Empirical Results of WAF

6.3.1.1 Data Sets and Models

By implementing the wavelet transform in preprocessing stage, the design of the wavelet transform in preprocessing was carried out in three stages, as follows:

1. Simulates the wavelet transform in MATLAB to understand a basic of wavelet transformation.

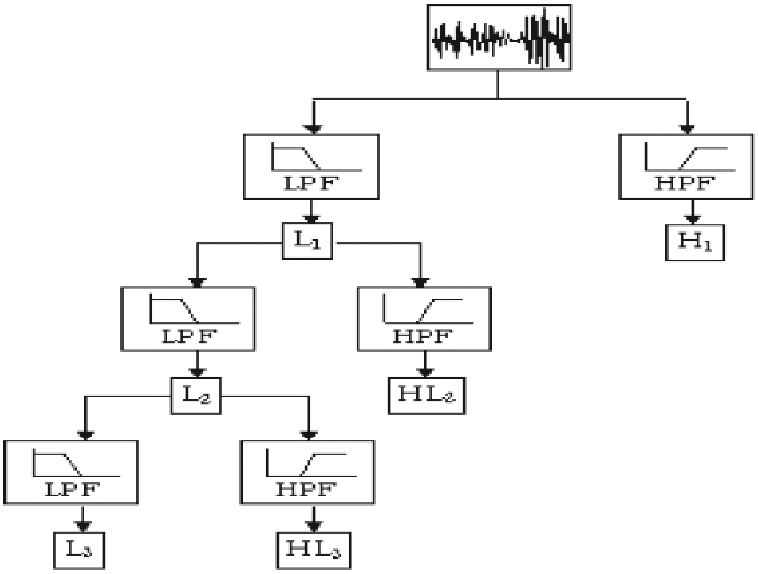

Figure 6.4: Discrete wavelets transform structure.

2. Designs the wavelet filter to decompose and reconstruct the original signal (kWh).

3. In this stage, we used daubechies2 mother wavelet (db2) to decompose kWh-signal into multi-resolutions (multi-components).

4. Each component is reconstructed to an original signal of kWh.

as illustrate in Figure 6.5, the whole area of the country forecasting had been performed seven stages as follow:

First stage, an original signal of load demand is decomposed to high and low frequency by using dB2 mother wavelet (dB2) for calculating the coefficient of detail (D) and approximate (A) components. The level of wavelet varies from 1 level, 2 levels, 3 levels, and 4 levels based on wavelet discrete transform (DWT). In each level of decomposition by wavelet method will obtain the details and the approximate. For example, there are detail (D1) and approximate (A1) components for 1 level. There are the two details (D1, D2) and a approximate (A2) components for the two levels. In 4 levels, there are the three details (D1, D2, and D3) and a approximate (A3) components. Lastly, there are the four details (D1, D2, D3, and D4) and a approximate (A4) components for the 4 levels of wavelet transform. After decomposing all data above are recorded and taken to stage 2.

Second stage, coefficient components from stage 1 are reconstructed to the

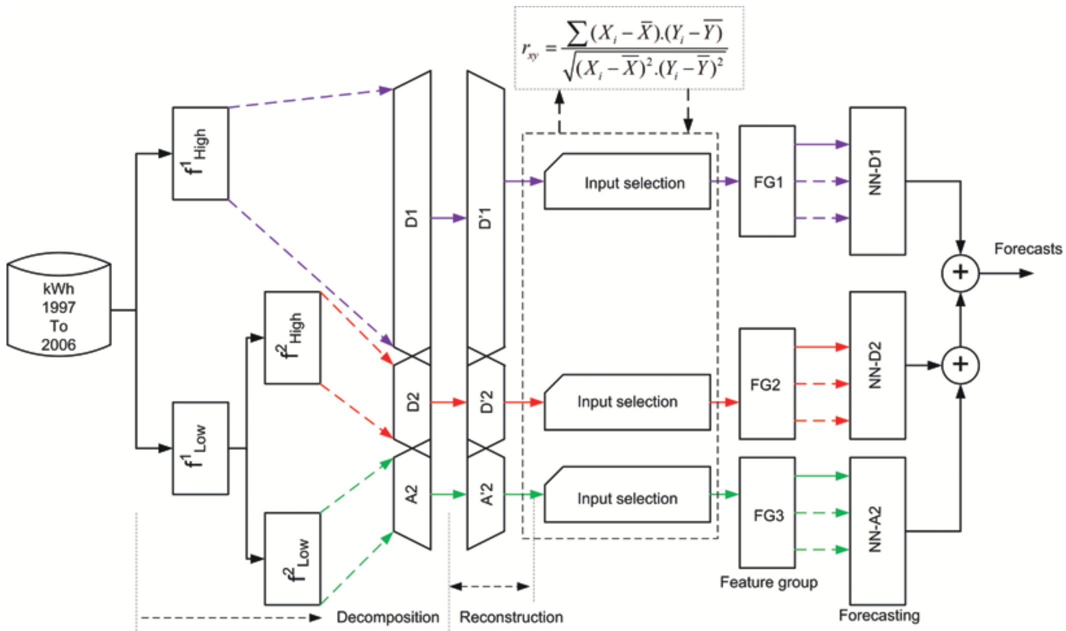

Figure 6.5: Whole area forecasting based on two levels.

actual components using similar mother wavelet, these are actual detail and approximate components of each level.

Third stage, actual detail and approximate components are taken to find the relationship between each component in each level and factors such as temperature, humidity, rainfall, consumer price index, and industrial index.

Fourth stage, input selection factors which closely related to the components more than 40% (up) will be chosen to be quality inputs for a neural network model. Note that, the factors chosen in each component of wavelet differing may be different depended on values.

Fifth stage, feature grouping inputs of a model NND forecast consists of the detail component and factors are selected from stage 4 while the feature inputs of a model NNA forecast consists of an approximate component.

Sixth stage, forecasting after partial forecasts in each sub-model, for example, the 4-level sub-model consists of A4, D4, D3, D2, and D1 models, the outputs for all sub-models are integrated to the final forecasting value.

Finally, result analyses and summary of all levels of wavelet and ANN are carried out.

Forecasting stage using neural network will be deployed in 4 sub-model. Each model of NN will predict based on group of feature inputs. For instance, the feature

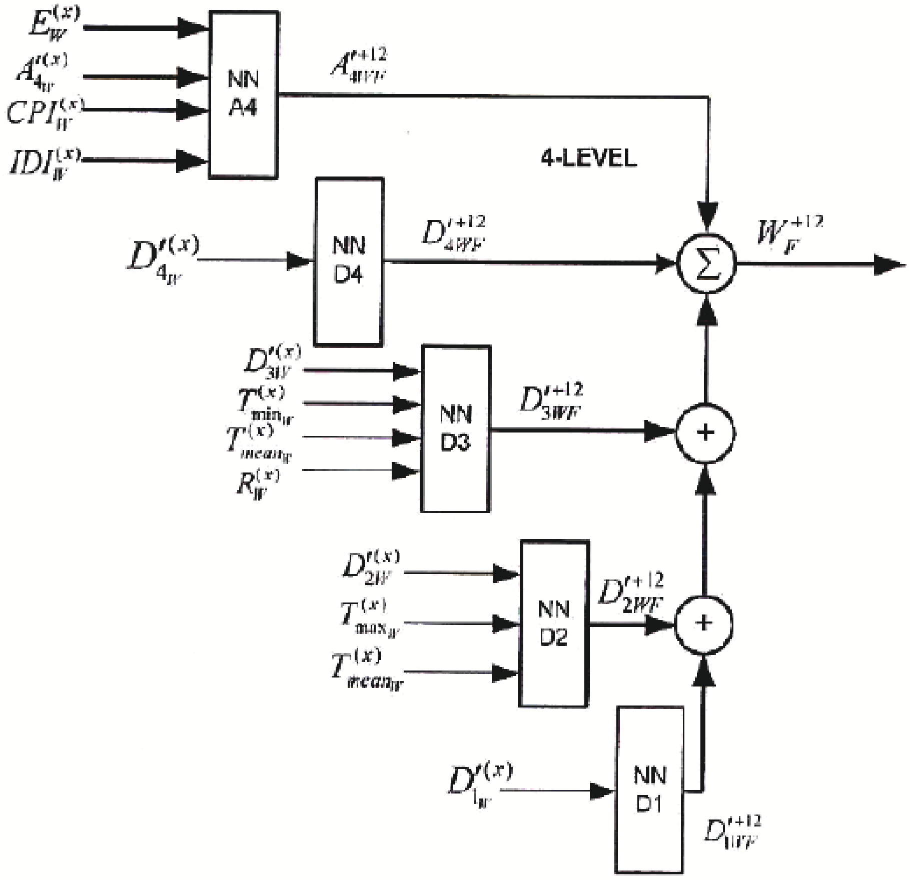

Figure 6.6: Whole area forecasting based on four levels.

inputs of NNA4 compose of energy demand (kWh), the approximate of energy demand in level four (A4), a consumer price index (CPI), and industrial index (IDI). While, neural network model for NND3 forecasting.The inputs of this model are the detail or cyclical component in level 4, but shift in the year, for example 1990 to 2004 in channel one, 2000 to 2005 in channel two. Figure 6.7 shows whole area forecasting in time series based on two-level model of wavelet-NN.

6.3.1.2 Simulation Results

The results of electricity load demand forecasting using wavelet transform and neural network algorithm are presentd in Table 6.1. The Mean Absolute Percent Error (MAPE) in one level, two levels, three levels, and four levels are 2.437%, 2.254%, 2.833%, and 3.465%, respectively. It also shows the percent error in each month from January to December 2007. It can be concluded that the two-level model of wavelet-NN

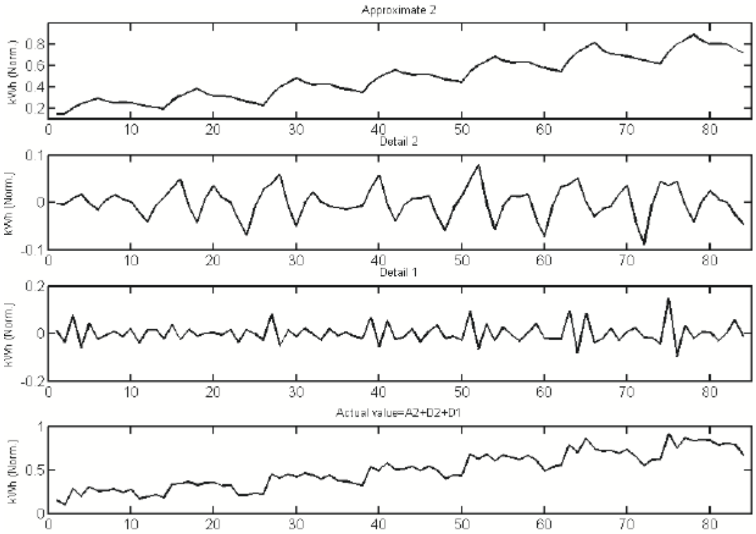

Figure 6.7: Whole area forecasting based on two levels.

Table 6.1: Results of whole area forecasting based on wavelet and neural network algorithm.

W-Level	1	2**	3	4
Month/error	error	error	error	error
January	-0.201	0.653	-0.100	-1.555
February	-4.216	-3.413	-4.035	-5.674
March	1.352	1.336	0.197	0.661
April	3.122	1.901	4.475	3.329
May	-2.639	-3.410	-2.077	-1.972
June	3.613	2.454	5.968	5.765
July	0.925	0.539	3.376	3.887
August	1.032	1.077	1.710	2.820
September	3.079	3.131	3.991	5.465
October	1.039	1.192	1.918	4.112
November	-5.239	-4.645	-3.932	-0.910
December	2.798	3.307	2.219	5.326
MAPE	2.437	2.254	2.833	3.456

give better accuracy than other models.

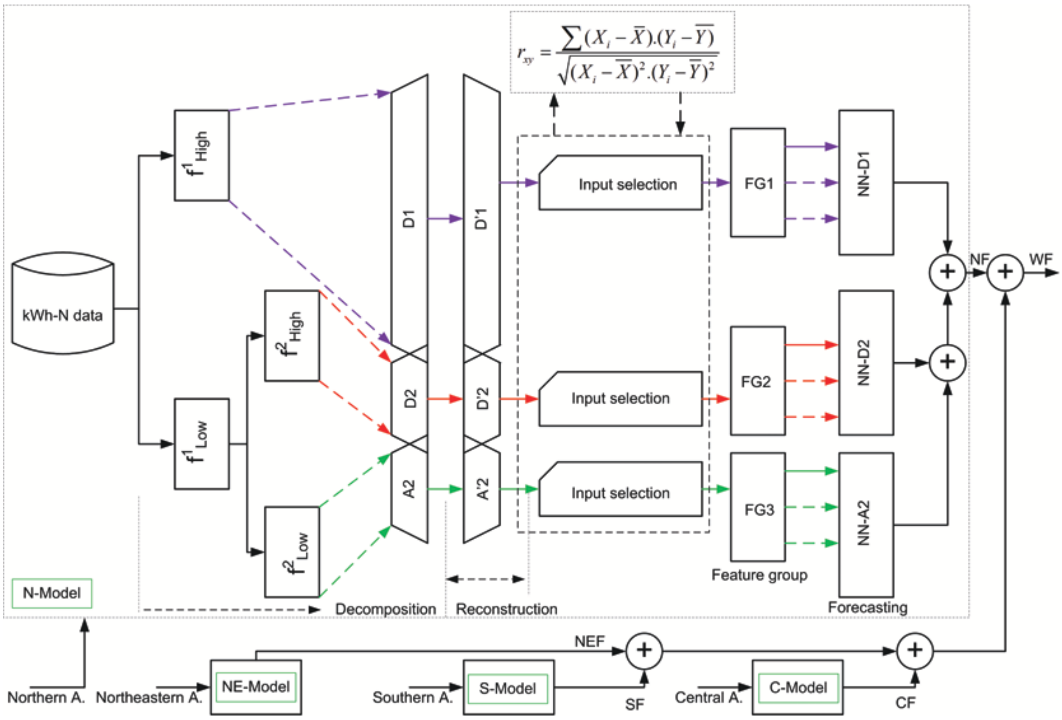

Figure 6.8: SCCA to WA forecasting model based on two levels.

6.3.2 Empirical Results of SCCAF

6.3.2.1 Data Sets and Models

Figure 6.8 illustrates an electric sub-control center area forecasting of the country; for overall structure of the electricity load demand forecasting. The wavelet transform and neural network algorithms are also used. The main steps proposed for the load demand forecasting model are as follows:

1. The Northern, Northeast, Southern, and Central sub-control center area forecasted using six previous stages.

2. The outputs of sub-model areas are integrated to the final forecasted value.

3. Analyses and summary of all forecast results are carried out.

6.3.2.2 Simulation Results

Figure 6.8 illustrates sub-area to whole area forecasting based on two-level of wavelet transform. Figure 6.9 presents the electricity substation control central load area based on three level. For two-level of wavelet transform, the signal after decomposition are shown in Figure 6.10 to 6.13. The results of SCCA in each area are

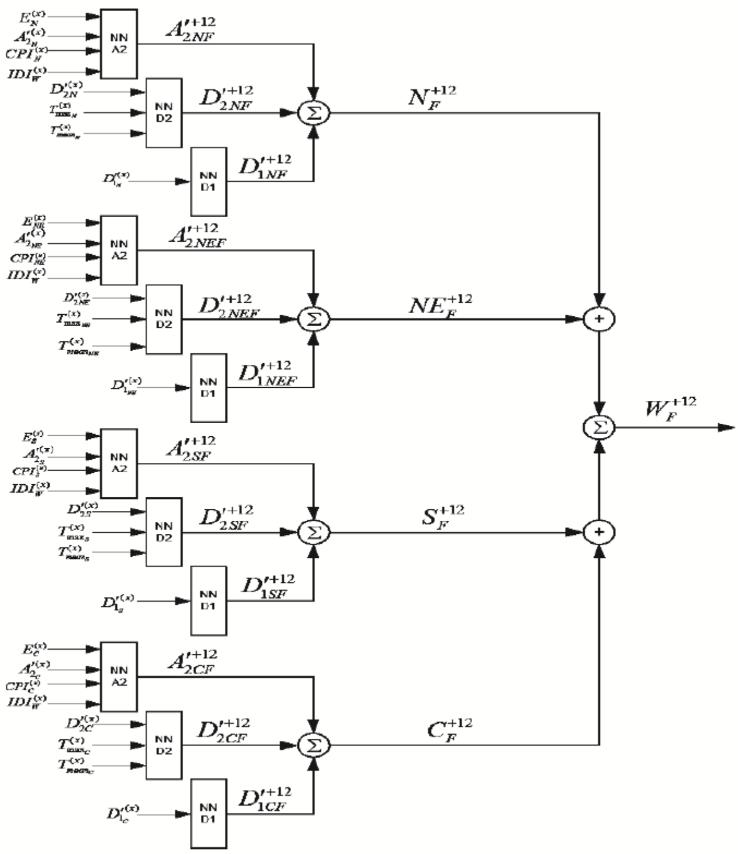

Figure 6.9: Electricity sub-control central area based on two levels.

also illustrated in Table 6.2 to Table 6.5.

The minimum MAPE of NSCCA, SSCCA and CSCCA are 3.258%, 1.735% and 2.280%, respectively. They are occured from two-level model. Yet, the minimum MAPE of NESCCA is 3.635% occured in one-level model. In conclusion, the two-level of wavelet transform using in preprocessing gives the best result in each area.

6.4 Discussion

Table 6.1 demonstrates the whole area load demand forecasting using one level to four levels of wavelet transform. The MAPE results of all levels are 2.437%, 2.254%, 2.833%, and 3.456%, respectively. The appropriate level is level 2. The signal componenent after preprocessing in level 2 has been shown in Figure 6.7.

Table 6.2 to 6.5 show the results of SCCA forecasting using NSCCA, NESCCA, SSCCA, and CSCCA of the country. One level to four levels are used for preprocessing. The appropriate level of all areas is level 2. The MAPE of each area

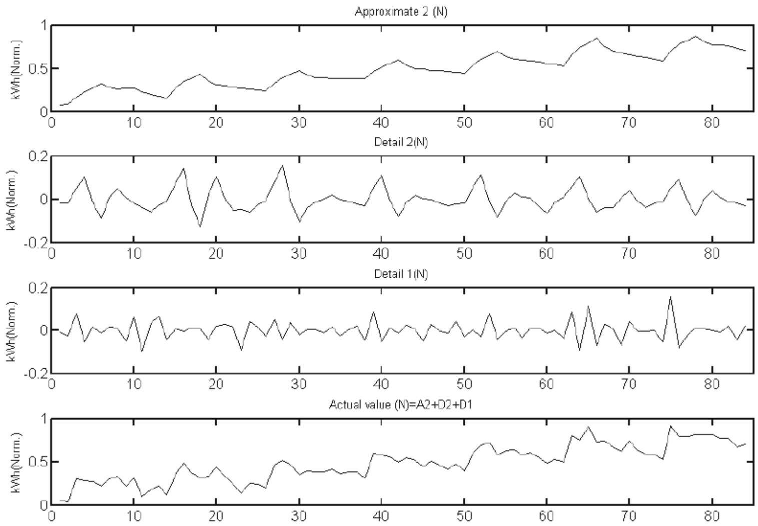

Figure 6.10: Northern substation control central area based on two levels.

Table 6.2: Results of Northern sub-control central area forecasting based on wavelet and neural network algorithm.

W-Level	1	2	3	4
Month/error	error	error	error	error
January	-4.186	-3.526	-2.937	-5.389
February	-2.059	-2.480	-2.510	-5.546
March	2.980	3.421	2.462	0.968
April	6.327	4.364	6.751	4.928
May	-2.271	-3.546	-0.131	-2.075
June	6.073	4.816	8.772	6.886
July	4.423	3.783	6.137	5.472
August	6.178	5.845	7.226	7.372
September	5.015	4.841	6.345	6.890
October	0.593	0.798	2.382	3.581
November	-1.720	-1.488	1.149	2.221
December	-0.374	0.193	1.202	2.541
MAPE	3.516	3.258	4.000	4.489

in level 2 is 3.288%, 3.662%, 1.735%, and 2.280 respectively. Figure 10 to 13 show substation control central area in detail and approximate components based on level 2. Table 6.6 to 6.9 demonstrate the results of SCCA forecasting based on one level to four levels of wavelet transform. The result shows the combination of forecasted value in all

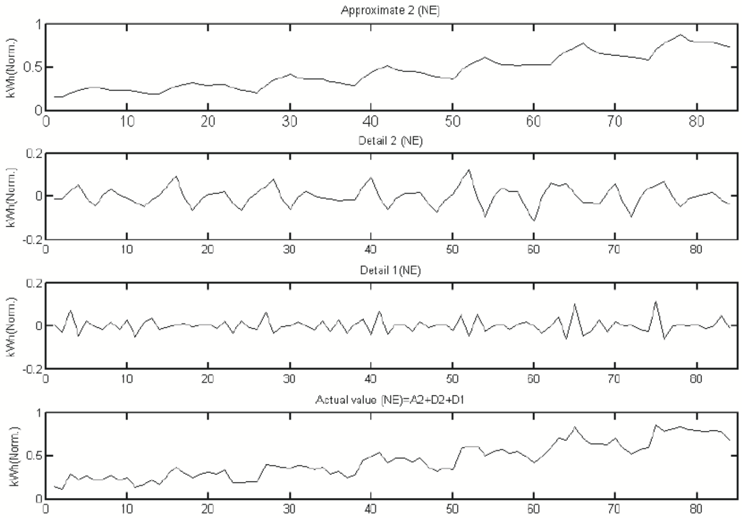

Figure 6.11: Northeastern substation control central area based on two levels.

Table 6.3: Results of Northeastern sub-control central area forecasting based on wavelet and neural network algorithm.

W-Level	1	2	3	4
Month/error	error	error	error	error
January	-0.140	1.623	1.950	-1.438
February	-3.663	-3.271	-3.690	-8.074
March	4.683	5.130	2.761	2.437
April	1.219	1.568	4.160	1.248
May	-6.828	-5.087	-4.115	-5.668
June	2.604	5.303	7.309	5.522
July	2.861	5.129	6.652	6.002
August	-1.813	0.345	0.289	1.297
September	0.198	2.555	2.036	4.235
October	-5.266	-2.801	-3.170	0.309
November	-13.554	-9.470	-9.514	-4.625
December	-0.792	1.667	-0.206	4.479
MAPE	3.635	3.662	3.821	3.777

areas. Finally, the appropriate level is level 2 which MAPE is 2.072%. However, the forecasted result performed by SCCAF give more accuracy than the result from WAF.

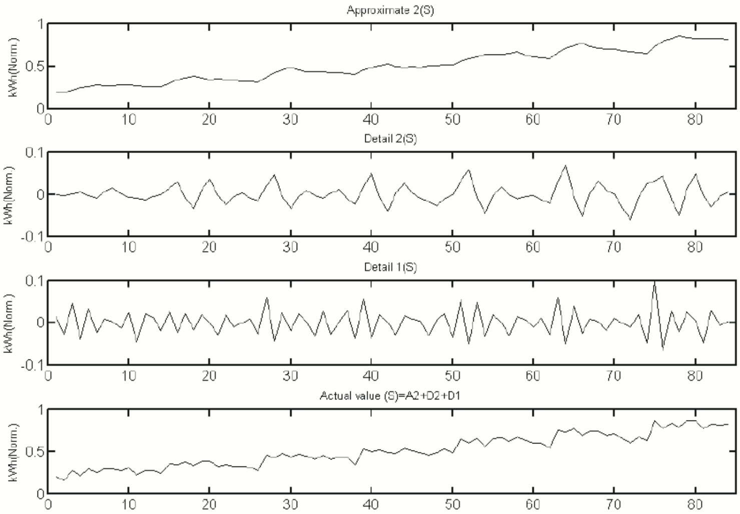

Figure 6.12: Southern substation control central area based on two levels.

Table 6.4: Results of Southern sub-control central area forecasting based on wavelet and neural network algorithm.

W-Level	1	2	3	4
Month/error	error	error	error	error
January	1.999	1.644	2.540	0.332
February	1.498	0.375	0.697	-1.914
March	3.316	1.537	2.067	1.561
April	7.293	4.484	7.162	6.656
May	1.810	-0.431	1.399	1.979
June	4.654	2.242	5.581	6.314
July	2.490	0.255	2.943	4.109
August	3.176	0.771	2.501	3.966
September	4.399	2.137	3.715	5.879
October	0.521	-1.243	0.792	3.957
November	-3.127	-4.040	-0.984	2.234
December	0.570	-1.671	0.867	3.904
MAPE	2.904	1.735	2.604	3.567

6.5 Summary

The energy load demand forecasting based on the suitable level of wavelet transform, and a neural network method. Approaching to the whole and sub-control center area forecasting are proposed. The load demand data which employs the historical

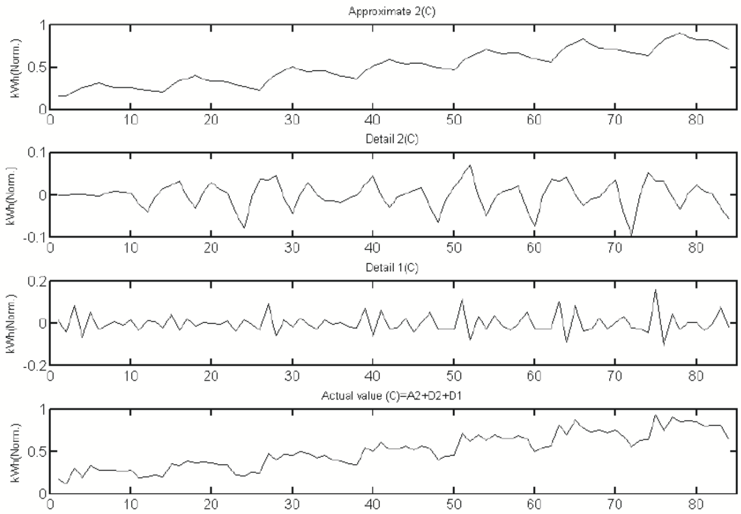

Figure 6.13: Central substation control central area based on two levels.

Table 6.5: Results of Center sub-control central area forecasting based on wavelet and neural network algorithm.

W-Level	1	2	3	4
Month/error	error	error	error	error
January	-1.718	-0.645	0.308	-1.951
February	-6.641	-5.684	-3.706	-5.996
March	-1.001	-1.216	-0.461	-0.510
April	2.156	1.230	3.390	1.755
May	-3.860	-3.888	-2.711	-2.930
June	2.684	2.340	5.058	4.714
July	-0.555	-0.490	7.573	7.806
August	-0.346	-0.410	0.904	1.793
September	2.048	2.012	3.231	4.522
October	0.820	1.170	2.657	4.522
November	-7.209	-5.901	-3.477	-1.204
December	1.966	2.383	2.337	4.535
MAPE	2.583	2.280	2.984	3.519

data from Electricity Generating Authority of Thailand (EGAT) is decomposed into one, two, three, and four levels of different frequencies. The correlation technique is used to choose the suitably affecting factors for each frequency component of load demand. The factors which affected in each level are chosen for neural network inputs. As a result,

Table 6.6: Results of sub-control central area forecasting based on one-level wavelet and neural network algorithm.

Area1	Area2	Area3	Area4	Whole area	Actual	PE	APE
North(kWh)	South(kWh)	Northeast(kWh)	M+C(kWh)	Total(kwh)	kWh	%	%
1081947371	953724227.7	1096470696	8289005034	11421147328.3428	11255610455.00	-1.47071	1.470705
1041447068	912156804.5	1133028812	8386525339	11473158023.8008	10903740111.00	-5.22223	5.222226
1188837751	1036596486	1275185636	9717179294	13217799166.4511	13256202895.00	0.289704	0.289704
1146615873	956946877.1	1229790003	8533531941	11866884693.7006	12222850690.00	2.912299	2.912299
1194108430	1021545799	1322013064	9578941817	13116609109.7468	12668470924.00	-3.53743	3.537429
1141523147	978741416	1247420898	8997309721	12364995182.2453	12768025563.00	3.15656	3.15656
1164458973	1009336861	1248435679	9088664232	12510895745.3813	12577211640.00	0.52727	0.52727
1168781013	1014095715	1262285708	9221448830	12666611265.3775	12722562991.00	0.439783	0.439783
1134065384	978505540.4	1233426810	8866171744	12212169478.0785	12504929463.00	2.341157	2.341157
1153691229	1007166245	1267797312	9048086308	12476741093.2552	12500303814.00	0.188497	0.188497
1118165834	1005446492	1237564202	9128442614	12489619141.7124	11678682183.00	-6.94374	6.943737
1114983789	1000978629	1197058971	8393361793	11706383181.8997	11866899916.00	1.352643	1.352643
					MAPE		2.365168

Table 6.7: Results of sub-control central area forecasting based on two-level wavelet and neural network algorithm.

Area1	Area2	Area3	Area4	Whole area	Actual	PE	APE
North(kWh)	South(kWh)	Northeast(kWh)	M+C(kWh)	Total(kwh)	kWh	%	%
1075093371	957181049.6	1077169089	8201597500	11311041010	11255610455.00	-0.49247	0.49247
1045742068	922562069.5	1128740397	8311321439	11408365973	10903740111.00	-4.62801	4.628007
1183437582	1055669068	1269217307	9737832943	13246156900	13256202895.00	0.07578	0.075783
1170653124	985943376.8	1225451106	8614344213	11996391820	12222850690.00	1.85275	1.85275
1208994772	1044868947	1300477499	9581544372	13135885590	12668470924.00	-3.68959	3.68959
1156806151	1003498236	1212845423	9029086557	12402236367	12768025563.00	2.864885	2.864885
1172248674	1032470582	1219291795	9082861653	12506872704	12577211640.00	0.559257	0.559257
1172928726	1039283893	1235539411	9227282909	12675034938	12722562991.00	0.373573	0.373573
1136143558	1001659113	1204296549	8869439549	12211538770	12504929463.00	2.3462	2.3462
1151319875	1025028408	1238110625	9016178064	12430636972	12500303814.00	0.557321	0.557321
1115622431	1014345595	1193052028	9017041202	12340061255	11678682183.00	-5.66313	5.663131
1108689919	1023543274	1167855877	8357653449	11657742519	11866899916.00	1.762528	1.762528
					MAPE		2.072125

"two levels" is given the Mean Absolute Percentage Error (MAPE) better than that other level and "sub-control center area" can show better result than "whole area forecasting". The objective of load forecasting has significantly forecast for fuel reserve planning in the power system.

Table 6.8: Results of sub-control central area forecasting based on three-level wavelet and neural network algorithm.

Area1	Area2	Area3	Area4	Whole area	Actual	PE	APE
North(kWh)	South(kWh)	Northeast(kWh)	M+C(kWh)	Total(kwh)	kWh	%	%
1068982378	948458844.8	1073588609	8123934440	11214964272	11255610455.00	0.361119	0.361119
1046051938	919579777.1	1133320981	8155707244	11254659939	10903740111.00	-3.21834	3.218344
1195187285	1049990990	1300899619	9665202582	13211280476	13256202895.00	0.338878	0.338878
1141426314	958303917.8	1193175629	8425946390	11718852250	12222850690.00	4.123412	4.123412
1169127900	1025827711	1288449141	9473052504	12956457256	12668470924.00	-2.27325	2.273253
1108721012	969222579.5	1187155323	8777782821	12042881736	12768025563.00	5.679373	5.679373
1143574300	1004647888	1199723215	8354014776	11701960180	12577211640.00	6.959026	6.959026
1155727449	1021159143	1236230956	9106602857	12519720405	12722562991.00	1.594353	1.594353
1118190388	985498973.3	1210720321	8759131938	12073541620	12504929463.00	3.449742	3.449742
1132930772	1004426345	1242563068	8880482677	12260402863	12500303814.00	1.919161	1.919161
1086635029	984554411.2	1193534037	8810682697	12075406173	11678682183.00	-3.39699	3.396993
1097482245	997983291	1190102474	8361599661	11647167671	11866899916.00	1.85164	1.85164
					MAPE		2.930441

Table 6.9: Results of sub-control central area forecasting based on four-level wavelet and neural network algorithm.

Area1	Area2	Area3	Area4	Whole area	Actual	PE	APE
North(kWh)	South(kWh)	Northeast(kWh)	M+C(kWh)	Total(kwh)	kWh	%	%
1094443586	969953693.1	1110688235	8307978904	11483064418	11255610455.00	-2.02081	2.020805
1077024200	943757059.3	1181237576	8335799085	11537817920	10903740111.00	-5.81523	5.815232
1213495057	1055411335	1305239924	9669973347	13244119663	13256202895.00	0.091152	0.091152
1163741642	963519744.7	1229429949	8568501615	11925192951	12222850690.00	2.435256	2.435256
1191827898	1019795660	1307657548	9493222393	13012503498	12668470924.00	-2.71566	2.71566
1131642054	961702817.1	1210044372	8809572646	12112961889	12768025563.00	5.130501	5.130501
1151670652	992575767.9	1208068220	8332950287	11685264927	12577211640.00	7.091768	7.091768
1153916760	1005820057	1223733347	9024916406	12408386571	12722562991.00	2.469443	2.469443
1111681263	963351427.3	1183543141	8642268097	11900843928	12504929463.00	4.830779	4.830779
1119015075	972377437.4	1200664745	8710392539	12002449796	12500303814.00	3.982735	3.982735
1074850265	953173474.3	1140250383	8617139940	11785414063	11678682183.00	-0.9139	0.913903
1082611504	967418412.8	1134463000	8173382008	11357874925	11866899916.00	4.289452	4.289452
					MAPE		3.482224

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] Qi Wu, "Hybrid model based on wavelet support vector machine and modified genetic algorithm penalizing Gaussian noises for power load forecasts," International journal of expert systems with applications pp.379-385, 2011.

[7] Ying Chen, Peter B. Luh, Che Guan, Yige Zhao, Laurent D. Michel et.al., "Short-term load forecasting: Similar day-based wavelet neural networks," IEEE Trans.on power syst., vol.25, pp.322-330, 2010.

[8] Ajay Shekhar Pandey, Devender Singh, and Sunil Kumar Sinha, "Intelligent hybrid wavelet models for short-term load forecasting," IEEE Trans.on power syst., vol. 25, pp.1266-1273, 2010.

[9] N. Amjady, and F. Keynia, "Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm," International journal of energy, vol.34, pp.46-57, 2009.

[10] N. M. Pindoriya, S. N. Singh, and S. K. Singh, "An adaptive wavelet neural network based energy price forecasting in electricity markets," IEEE Trans.on power syst., vol.23, pp.1423-1432, 2008.

[11] D. Benaouda, F. Murtagh, J.-L. Starck, and O. Renaud, "Wavelet-based nonlinear multi-scale decomposition model for electricity load forecasting," International journal of neuro computing, vol.70, pp.139-154, 2006.

[12] Tai Nengling, Jurgen Stenzel, and Wu Hongxiao, "Techniques of applying wavelet transform into combined model for short-term load forecasting," International journal of electric power systems research, vol.76, pp.525-533, 2006.

[13] Bai-Ling Zhang, and Zhao-Yang Dong, "An adaptive neural-wavelet model for short term load forecasting," International journal of electric power systems research, vol.59, pp.121-129, 2001.

[14] Tongxin Zheng, Adly A. Girgis, and Elham B. Makram, "A hybrid wavelet-Kalman filter method for load forecasting," International journal of electric power systems research, vol.54, pp.11-17, 2000.

[15] S. J. Yao, Y. H. Song, L. Z. Zhang, and X. Y. Cheng, " Wavelet transform and neural network for short-term electrical load forecasting," International journal of energy conversion and management, vol.41, pp.1975-1988, 2000.

[16] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, Mid-tem load forecasting: Level suitably of wavelet and neural network based on factor selection, Energy procedia 14, pp.438-444, 2012.

[17] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, Wavelet and neural network approach to demand forecasting based on whole and electric sub-control center area, IJ Soft Computing and Engineering, vol.1, issue. 6, pp.81-86, 2012.

[18] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," proceedings of SPIE second International conference on digital image processing, vol. 7546, pp.75460B-1-75460B-9, 2010.

[19] M.G.E. Schneiders, "Wavelets in control engineering," Master's thesis Eindhoven University of Technology, August 2001.

[20] P.S. Addison, "The illustrated wavelet transform handbook," IOP Publishing Ltd.,2002.

[21] G.Strang and T.Nguyen, "Wavelet and Filter banks, wellesley combridge press," second edition, 1997.

Chapter 7

WA and SCCA Forecasting using HP-Filter and NN

7.1 Introduction

This chapter proposes Multi-area and Whole-area using HP-NN for mid-term load Forecasting based on the electricity load consumption (kWh) of Thailand in whole-area and substation control center area of Electricity Generating Authority of Thailand (EGAT). It consists of introduction, theory, method, results and conclusion as follows. The Hodrick-Prescott Filter [25-27] is widely used in applications and has been embodied in various statistical packages. The time aggregation properties of the Hodrick-Prescott (HP) filter to decompose a time series into trend and cycle are analyzed for the case of annual, quarterly, and monthly data. It is seen that aggregation of the disaggregated component estimators cannot be obtained as the exact result from applying an HP filter to the aggregate series and vice versa. Nevertheless, using several criteria, one can find HP decompositions for different levels of aggregation that provide similar results. The approximation works better for the case of temporal aggregation than for systematic sampling.

Several approaches were used for electricity load demand forecasting such as statistical approach [7] and artificial intelligence [8-22], [24, 28-32], [34, 36-37]. Some approaches were used for short-term load forecasting and some were used fore mid-term load forecasting. However, this chapter, proposes a new approach which it used for mid-term load forecasting based on two-type of area that are whole area and sub-area of the country. The new approach is the combination of Hodrick-Prescott filter and neural network algorithm. The HP filter is used for preprocessing stage and also neural network is used for forecasting stage.

7.2 Theoretical Background

7.2.1 Hodrick-Prescott (HP) Filter

Trend-cycle decompositions are routine in modern macroeconomics but this research is just started to take it into load decomposition in power sytem. The basic idea is to decompose the load demand series of interest (kWh) into the sum of a slowly evolving secular trend and a transitory deviation, which is classified as 'cycle'.

The Hodrick-Prescott filter separates a time series into a trend component identified as economic growth factors which include the industrial index and the consumer price index; and a cyclical component identified as behaviors of weather factors inclusive of temperature, humidity, rainfall, and wind speed, such that [26],

$$y_t = \tau_t + c_t \tag{7.1}$$

Because these constituent parts, trend and cycle, are not readily observed, any decomposition must necessarily be built on a conceptual artifact. Thus, any de-trending method must start out by arbitrarily defining what shall be counted as trend and as cycle, before these elements can be estimated from the data.

The equation is equivalent to a cubic spline smoother, with the smoothed portion on the inside. The loss function in equation (7.2) could be minimized and can be written in the form of:

$$\min_{\{\tau_t\}} \left\{ \sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=2}^{T-1} \left((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}) \right)^2 \right\}, \tag{7.2}$$

where t is a number of samples (t = 1, 2, 3,.., T), τ_t is a trend component, c_t is a cyclical component, and is a smoothing parameter.

There are two terms in the loss function. The first term in loss function is a well-known measure of the goodness of-fit, the error sum of squares. The second term punishes variations in the long-term trend component.

Definition: The loss function (LF), to obtain a HP type smoother for the observations in model (1) we define the loss function that yields the smoother by using equation (7.3) to (7.6) as following [26-27]:

$$\hat{y} = \min_{\{\tau\}} LF(\tau) \quad \text{with } LF(\tau) = ESS(\tau) + \lambda * smooth(\lambda), \tag{7.3}$$

where the ESS is defined as error sum of squares:

$$ESS(\tau) = \sum_t (y_t - \tau_t)^2 = \sum_t c_t^2 \tag{7.4}$$

Substituting (7.4) into (7.3), we obtain.

$$\begin{aligned} \hat{y} = \min\ LF(\tau) \quad \text{with} \quad LF(\tau) \\ = \sum_t^T (y_t - \tau_t)^2 + \lambda \sum_{t=3}^T (\Delta^2 \tau_t)^2 \end{aligned} \tag{7.5}$$

Thus,

$$\min_{\{\tau\}} LF(\tau) = \min_{\{\tau\}} \left\{ \underbrace{\sum_{t=1}^{T} c_t^2}_{Goodness-of-Fit} + \lambda \underbrace{\sum_{t=2}^{T-1} ((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2}_{Penelty-of-Roughness} \right\} \tag{7.6}$$

The original HP filter problem can be defined as a minimized of the loss function (LF) which has two components, the goodness of fit and the smooth.

The smoothing parameter controls the smoothness of the adjusted trend series,τ_t , for example, as λ= 0. The trend approximates the actual series, or the observed series (y_t), whereas when $\lambda = infinity$ the trend becomes linear. It is customary to set λ to 1,600 for quarterly data. For monthly and annual data it has been recommended to take the derivatives of the loss function in equation (7.2) and rearrange them [26-27, 35]. Then, the trend component and the cyclical component can be identified by using equation (7.7) and (7.8). These components are shown in Figure 7.1.

$$\tau_t = (\lambda F + I_T)^{-1} y_t \tag{7.7}$$

$$c_t = y_t - \tau_t \tag{7.8}$$

$$\mathbf{K}_{ij} = \begin{cases} 1, & \text{if} \quad i = j \quad or \quad i = j+2; \\ -2, & \text{if} \quad i = j+1; \\ 0, & \text{if} \quad \text{otherwise}; \end{cases} \tag{7.9}$$

Where $\mathbf{I_T}$ is a $T \times T$ identity matrix, and $\mathbf{F}$ is a matrix of $\mathbf{K}'\mathbf{K}$ which $\mathbf{K} = \{\mathbf{K}_{ij}\}$ is a $(T-2) \times T$ or number of membership. for example :

$$\mathbf{K} = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & . & . & . & 1 & -2 & 1 \end{bmatrix}$$

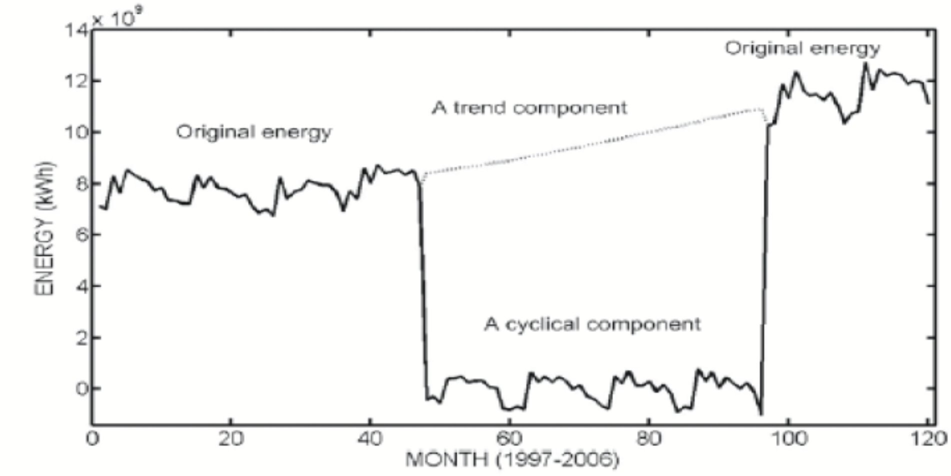

Figure 7.1: Electricity consumption demand before and after decomposition using Hodrick-Prescott (HP) filter. Data record during January 1997 to December 2006 from EGAT (The Electricity Generating Authority of Thailand).

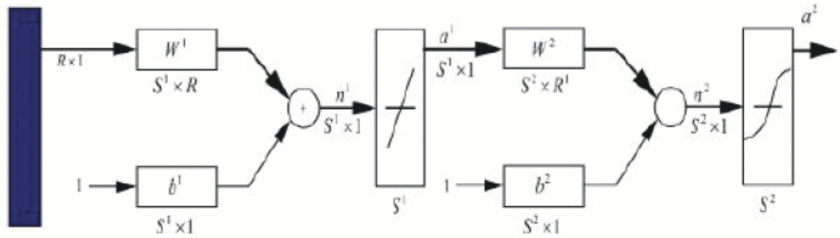

Figure 7.2: Structure of FFBP neural network.

7.2.2 Neural Network (NN)

Neural Network Artificial neural network is one of the good choices to apply to the load demand forecasting problem because this technique does not require explicit models to represent the complex relationship between the load demand and its factors. The neural network algorithm presented in this thesis composes of three layers: input layer, hidden layer, and output layer based on a feed-forward back propagation algorithm (FFBP). The input variables come from historical and existing data of factors affecting the load demand. The number of inputs, the number of hidden nodes (neurons), and the activate transfer function in each layer affect the forecasted performance. Hence, they need to be chosen carefully (Chapter 5 and 6) [33].

Table 7.1: Percent of correlations between electricity consumption load demand and variables, before and after decomposition by using HP-filter.

	Weather/a cyclical						Economic/ a trend			
variable	Max.Temp.		Min.Temp.		Mean.Temp.		Consumer price.		Industrial index	
	before	after	before	after	before	after	before	after	before	after
ECLD	9.7	50.8	24.6	71.0	26.5	79.7	87.8	92.3	93.9	97.9

7.3 Proposed Method

7.3.1 Experimental Results of WAF

7.3.1.1 Data Sets and Models

The Hodrick-Prescott (HP-filter) separates a time series of energy consumption load demand (y_t) [1, 5] into the trend component which identified as a growth (economic factor such as the industrial index, the consumer price index) [3, 4] and the cyclical component which identified as behavior of weather factors (temperature, humidity, rainfall, and wind speed) [2].

Model in Figure 7.3, for the whole country with preprocessing, the main steps of the DNNs proposed for the load demand forecasting model, as flow charted in the figure are as follow:

Forecasting steps:

1) Historical data variables are collected.

2) Historical data of the load demand is decomposed to trend and cyclical components by using HP-filter in equation (7.1) to equation (7.9).

3) Data are normalized for the trend and the cyclical components.

4) Variables that are related to the trend component of the load demand are selected feature inputs of neural network model (NN1).

5) Variables that are related to the cyclical component of the load demand are selected feature inputs of neural network model (NN2).

6) After partial forecasts in the model NN1 and the model NN2, the outputs from NN1 and NN2 are integrated, using equation (7.1), to the final forecasting value.

7) Analyses and summary are carried out.

Table 7.1 shows percent of correlations between the ECLD and selected

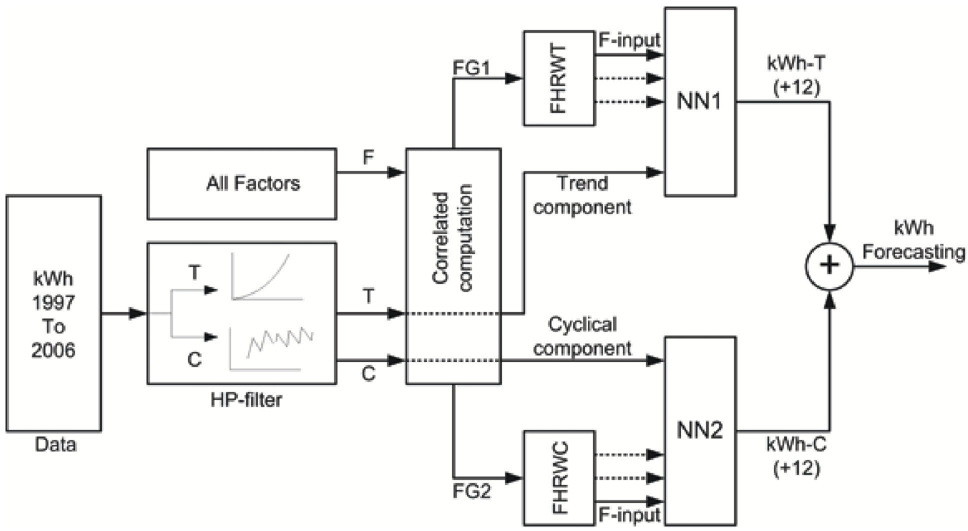

Figure 7.3: Double neural networks model for the whole country area forecasting with preprocessing (HP-filter).

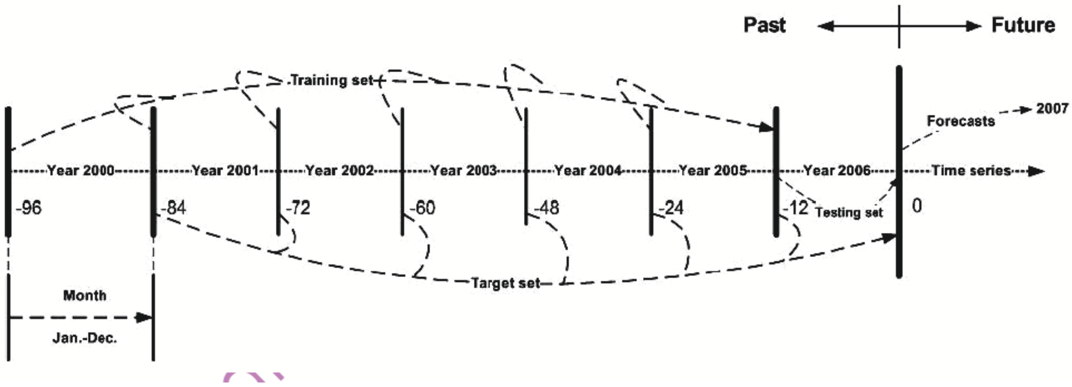

Figure 7.4: Data period for training, testing, and forecasting.

variables, before and after decomposition by using HP-filter approach. The table presents the correlations of actual value (kWh) and cyclical components with weather factors (maximum temperature, minimum temperature, and maximum temperature). For instance, from table 7.1, the correlated values of mean temperature before and after decomposition are 26.5% and 79.7% respectively. Additionaly, the correlation of the trend component and economic factors, such as, consumer price index before and after decomposition are 87.8% and 92.3% respectively.

Figure 7.4 illustrates data period for training, testing, and forecasting time. For training, the data of energy load demand are used from year 2000 to 2005 and factors from year 2001 to 2006. Also, the target of the model has the data of ECLD

Table 7.2: Result of whole area using HP-filter and neural network.

Factors	Trend E, GDP,CPI,Cyclical E, Tmax,Tmin,Tmean	Trend E, CPI, IDI, Cyclical E, Tmax,Tmin,Tmean	Trend E, IDI, Cyclical E, Tmax,Tmin,Tmean
January	-1.50922	-0.61532	-1.12010
February	-4.19519	-2.98928	-3.61653
March	-2.26348	-1.28652	-1.97945
April	-2.13400	-0.87751	-1.37130
May	-0.67125	0.54627	-0.08624
June	0.63294	2.03077	1.36693
July	-1.06039	0.53227	-0.19923
August	0.56334	2.36679	1.48163
September	-1.34836	0.57104	-0.38451
October	-0.20772	1.70699	0.71176
November	-2.04030	0.10825	-0.99298
December	-0.43284	1.86458	0.70194
MAPE	1.421587	1.291298	1.167716

from year 2001 to 2006 in monthly type. In this research, we forecast the ECLD in year 2007 from January to December by using ECLD in 2006 and all factors in 2007.

7.3.1.2 Simulation Results

Table 7.2 explains the results of whole area forecasting by using HP-filter in preprocessing and neural network in forecasting stage. This table shows the results of three models based on the selected feature inputs:

- Model 1: Trend ECLD, GDP, CPI, cyclical ECLD, Tmax, Tmin, Tmean
- Model 2: Trend ECLD, CPI, IDI, cyclical ECLD, Tmax, Tmin, Tmean
- Model 3: Trend ECLD, IDI, cyclical ECLD, Tmax, Tmin, Tmean

The mean absolute percent error (MAPE) of these models is 1.42%, 1.29%, and 1.16% respectively. Model 2 can be the appropriate approach because the error in each month is decreased to be under +3% accept in February.

Table 7.3: Results of whole area forecasting (correction).

Whole-area	E-kWh	PE			
Month/Fac.	Actual	All Factor	IDI^{only}	IDI^{wo}	$IDI^{correction}$
January	11255610455	-0.615	-1.120	-0.458	-0.912
February	10903740111	-2.989	-3.617	-2.771	-3.386
March	13256202895	-1.287	-1.979	-1.037	-1.680
April	12222850690	-0.878	-1.371	-0.713	-1.172
May	12668470924	0.546	-0.086	0.766	0.202
June	12768025563	2.031	1.367	2.268	1.656
July	12577211640	0.532	-0.199	0.800	0.106
August	12722562991	2.367	1.482	2.701	1.725
September	12504929463	0.571	-0.385	0.930	-0.113
October	12500303814	1.707	0.712	2.078	1.006
November	11678682183	0.108	-0.993	0.517	-0.666
December	11866899916	1.865	0.702	2.299	0.935
MAPE (%)		1.291	1.168	1.445	1.130

7.3.2 Experimental Results of SCCAF

7.3.2.1 Data Sets and Models

Model in Figure 7.5 is the sub-control center area to whole country forecasting with preprocessing. The main steps of the DNNs proposed for the load demand forecasting model, are as follow:

Forecasting steps:

1. Historical data variable are collected.

2. Historical data of the load demand is decomposed to trend and cyclical components by using HP-filter in equation (7.1) to equation (7.9).

3. Data are normalized for the trend component (norm.) and for the cyclical component (norm.).

4. Variables that are related to the trend component of the load demand are selected feature inputs for neural network model NN1.

5. Variables that are related to the cyclical component of the load demand for a feature inputs for neural network model (NN2).

6. After partial forecasts in the model NN I (TNN) and the model NN

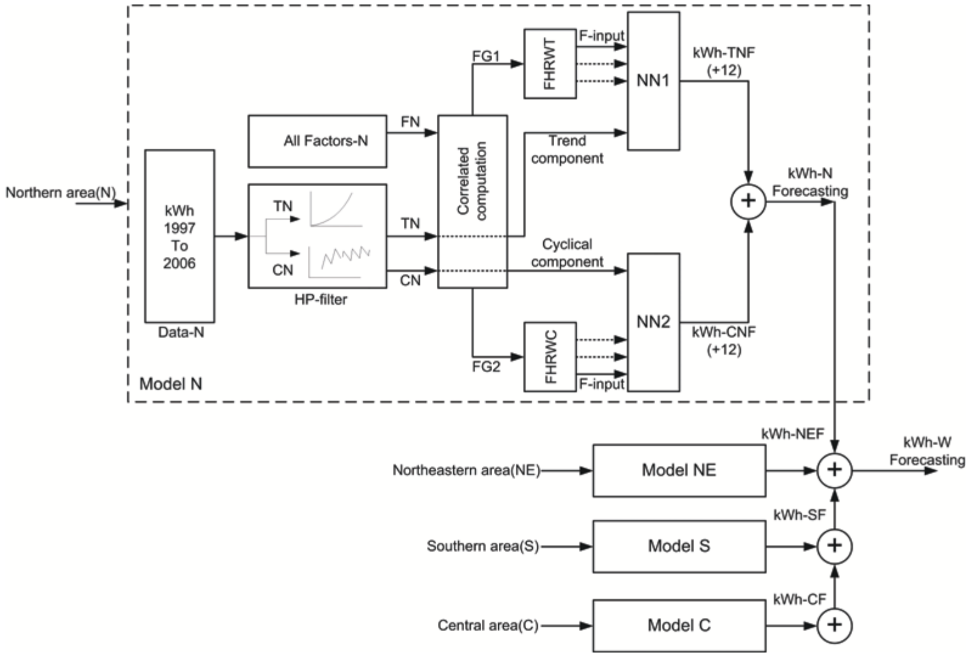

Figure 7.5: Double neural networks model for the sub-control center area to whole country area forecasting with preprocessing (HP-filter).

Table 7.4: Substation control central area (SCCA) of the country.

Order	Region	Code
R1	Northern substation control central area	$NSCCA$
R2	Northeastern substation control central area	$NESCCA$
R3	Southern substation control central area	$SSCCA$
R4	Center substation control central area	$CSCCA$

Table 7.5: Factors used for electricity consumption load demand in each SCCA.

Order	Variable	Code
E1	Electricity consumption load demand (kWh) in each SCCA	$ECLD$
F1	Maximum temperature	T^{max}
F2	Minimum temperature	T^{min}
F3	Mean temperature	T^{mean}
F4	Consumer price index	CPI
F5	Industrial index	IDI

II (CNN), the outputs from NN I and NN II are integrated, using equation (7.1), to the final forecasting value.

7. Analyses and summary are carried out.

Table 7.6: The correlation between electricity consumption load demand (kWh) and factors.

Factor	Percentage of factor correlation			
Region	N-sub 1	NE-sub 2	S-sub 3	C-sub 4
Maximum temperature	20.49	21.24	4.10	7.30
Minimum temperature	26.54	21.75	11.70	25.40
Mean temperature	29.18	28.24	15.92	26.10
Humidity	6.5	6.15	14.43	18.45
Consumer price index	83.89	91.34	93.40	91.70

Table 7.7: Thailand sub-control electricity consumption load demand (Unit in kWh).

Month/Year	Area 1	Area 2	Area 3	Area 4	Whole
Jan. 97	699573782	560051620	511621811	5368322371	7139569584
Feb. 98	726641682	623223859	545055665	5322395321	7217316527
March 99	831006444	766596724	601893552	6043231343	8242728063
April 00	853367909	734071004	608244154	5813658714	8009341781
May 01	902302609	788174291	701132134	6698443503	9090052537
June 02	892995540	833051145	737059241	7021193734	9484299660
July 03	996804076	919892118	762215248	7376018389	10054929831
Aug. 04	1049728022	1003447919	870271190	7999517628	10922964759
Sep. 05	1035962657	1051040169	890892925	8273771810	11251667561
Oct. 06	1120332659	1172400906	969859701	8750546104	12013139370
Nov. 07	1099262401	1089843587	974956145	8514620050	11678682183
Dec. 08	1083803180	1088841635	988992532	7377175128	10538812475

Table 7.3 presents the results of whole area forecasting based on correction approaches: all factors, only IDI, without IDI, and using IDI correction. As seen from the table, the forecasting model with IDI correction using HP-filter can give the best MAPE of 1.130% compared to other approaches. Yet, the MAPE in i.e February reaches the limit of +3%. While in the forecasting without IDI approach, all the MAPE in each month are under the limitation of +3% but the total MAPE is 1.445% which is larger than that in IDI correction approach.

Table 7.4 explains the substation control central area (SCCA) of the country. R1 is the northern substation control central area represented NSCCA. This area is located at Pitsanulok province. It is used for monitoring ECLD in this area

Table 7.8: The correlation coefficients between SCCA load demand samples and factors (before preprocessing by using HP-filter).

1	Correlation coefficients with load samples							
Region	NSCCA		NESCCA		SSCCA		CSCCA	
Factor	-1 to 1	sig.	-1 to 1	sig.	-1 to 1	sig.	-1 to 1	sig.
$ECLD$	1	-	1	-	1	-	1	-
T^{max}	0.20*	0.02	0.21*	0.020	0.17	0.06	0.04	0.62
T^{min}	0.26**	0.00	0.21*	0.017	0.11	0.20	0.29**	0.00
T^{mean}	0.29**	0.00	0.28**	0.002	0.15	0.08	0.24**	0.00
CPI	0.83**	0.00	0.91**	0.000	0.93**	0.00	0.87**	0.00
IDI	0.90**	0.00	0.92**	0.000	0.95**	0.00	0.93**	0.00

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 7.9: The correlation coefficients between SCCA load demand samples and factors (after preprocessing by using HP-filter).

Region	NSCCA				NESCCA				SSCCA				CSCCA			
kWh	Tr.	sig.	Cyc.	sig.	Tr.	sig.	Cyc.	sig.	Tr.	sig.	Cyc.	sig.	Tr.	sig.	Cyc.	sig.
E^{trend}	1	-	-0.00	0.95	1	-	-0.00	0.91	1	-	-0.01	0.84	1	-	0.00	0.96
$E^{cyc.}$	-0.00	0.95	1	-	-0.00	0.91	1	-	-0.01	0.84	1	-	0.00	0.96	1	-
T^{max}	-0.01	0.83	0.55*	0.00	0.01	0.83	0.58*	0.00	0.07	0.42	0.45*	0.00	-0.11	0.22	0.45*	0.00
T^{min}	0.03	0.71	0.58*	0.00	-0.00	0.95	0.67*	0.00	0.01	0.84	0.45*	0.00	0.05	0.54	0.71*	0.00
T^{mean}	-0.01	0.83	0.77*	0.00	-0.00	0.96	0.86*	0.00	0.01	0.85	0.65*	0.00	-0.01	0.90	0.75*	0.00
CPI	0.91*	0.00	0.00	0.95	0.95*	0.00	0.01	0.84	0.95*	0.00	-0.01	0.91	0.92*	0.00	0.02	0.82
IDI	0.90*	0.00	0.01	0.89	0.97*	0.00	0.01	0.90	0.97*	0.00	0.01	0.88	0.93*	0.00	0.04	0.65

* Correlation is significant at the 0.01 level (2-tailed)

covered 19 provinces in the northern area of the country. Next, R2 is the northeastern substation control central area represented NESCCA located at Khonkan province. It is used for monitoring ECLD in this area covered 20 provinces in the northeastern area of the country. In addition, R3 is the southern substation control central area represented SSCCA. It is used for monitoring the ECLD in this area covered 16 provinces and the location of SSCCA is krabi province. Finally, R4 is center substation control central area represented CSCCA. It locates at nontaburi province (Bangkreay). It is used for monitoring the ECLD in this area covered 29 provinces. These SCCA can illustrate in Chapter 3.

Table 7.5 shows the factors used for electricity consumption load demand

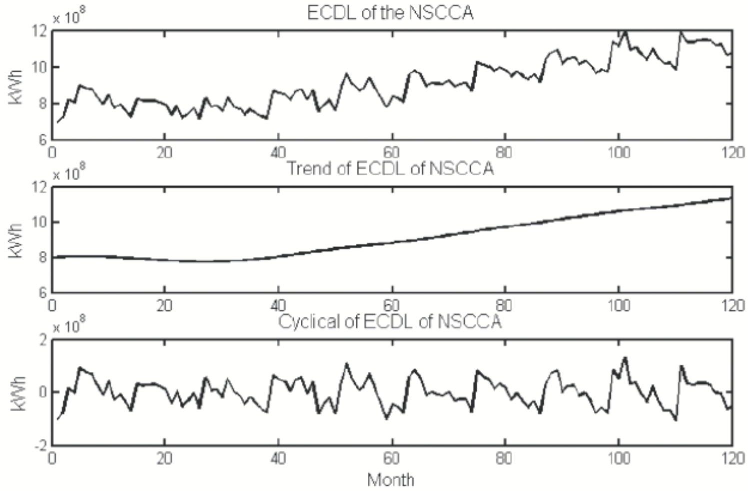

Figure 7.6: ECLD of the NSCCA of the country.

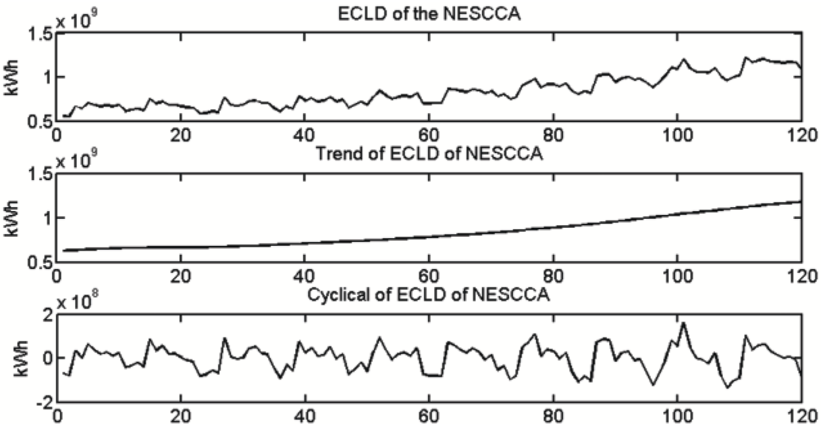

Figure 7.7: ECLD of the NESCCA of the country.

forecasting in each SCCA. These factors are the ECLD unit in kWh, maximum temperature, minimum temperature, mean temperature, consumer price index, and industrial index in each area. The code of each factor shows in this table.

Table 7.6 shows the correlation between the ECLD and factors in each region: the north, northeastern, south, and the center. In maximum temperature, the northeastern area shown high correlation with ECLD is 21.24%, and also the southern area shown minimum value is 4.10%. Next, in minimum temperature with ECLD, the northern area shows high correlated value and southern area shows low correlated value. Also, mean temperature with ECLD, the northern, northeastern, and central are high correlated value but the southern area is low correlated value. In addition, humidity with ECLD, the southern area shows high correlated value rather than other areas. Finally,

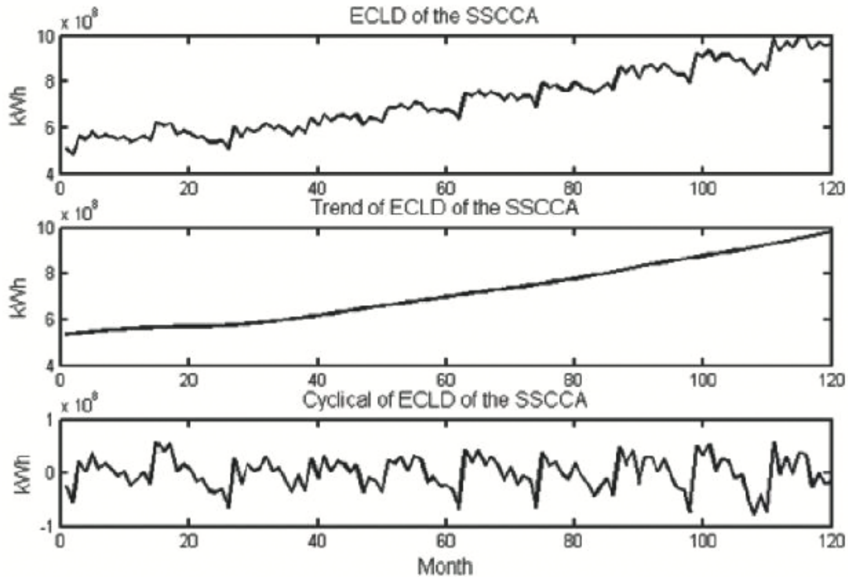

Figure 7.8: ECLD of the SSCCA of the country.

the consumer price index in all areas are high correlated value when compared with ELCD in that area: 83.89% in the north, 91.34% in the northeast, 93.40% in the south, and 91.70% in the center of the country.

Table 7.7 shows the samples of ECLD substation control central area (SCCA) (unit in kWh) and ECLD in whole area of the country. Table 7.8 and 7.9 explain the coefficient correlation between SCCA load demand samples and factors before and after preprocessing by using HP-filter approach.

Figure 7.6 to Figure 7.9 illustrate the ECLD and a trend and a cyclical decomposition using HP-filter in four regions. Lamda is 14,400 for monthly type. Each figure shows different characteristics of the trend and cyclical component, especially the southern area.

7.3.2.2 Simulation Results

Double neural networks (DNNs) are used in the prediction stage are used. There are three layers: input layer, a hidden layer, and output layer. The input layer in the part of the trend component forecast has three feature inputs such as historical trend component, consumer price index, and industrial index; input layer in the part of the cyclical component forecast has three feature inputs such as maximum, minimum and mean temperature in each area. The hidden layer can be varied neuron nodes shown

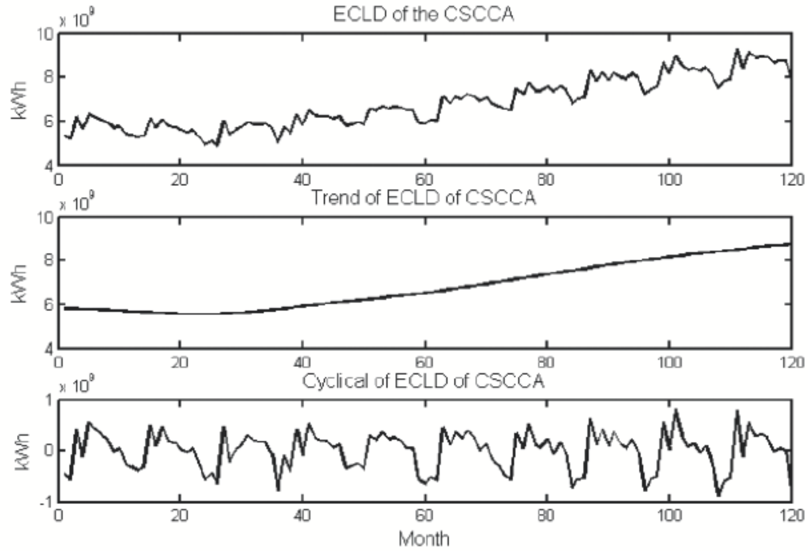

Figure 7.9: ECLD of the CSCCA of the country.

Table 7.10: Results of forecasting.

Hidden	N			NE			S			C			whole
H_T, H_C	T_f	C_f	N_f.	T_f	C_f	NE_f	T_f	C_f	S_f	T_f	C_f	C_f	$MAPE$
3, 3	0.45	84.4	3.26	0.96	193.64	2.40	0.45	175.76	1.98	0.57	73.72	1.66	1.42
8, 8	1.60	151.43	5.42	1.19	177.42	3.48	0.44	290.14	4.86	0.53	97.97	3.39	3.20
15, 15	1.74	125.31	4.72	1.24	174.52	2.10	0.48	180.70	2.86	0.29	150.12	3.70	3.18
20, 20	1.47	164.87	5.60	1.22	245.29	2.71	1.32	285.78	5.86	0.41	104.22	2.85	2.95

T_f is the trend forecasted error(MAPE)

C_f is the cyclical forecasted error(MAPE)

N_f, NE_f, S_f, MC_f are meanly Mean Absolute Percent Error(MAPE) in each area

in Table 7.10. Finally, the output is the last layer of the neural network. It is going to show the results of this forecasting model. MAPE in the northern area is about to 3.36% at the number of neuron (3; 3), the northeast area is about to 2.40% at the number of neuron (3; 3), the southern area is about to 1.98% at the number of neuron (3; 3), and the central area is 1.66% at the number of neuron (3; 3). Factors in each area will be used the different economic factors and weather factors. To conclude, this research focuses on two stages as following: the first stage shows signal load demand decomposition by using HP-filter and separates the signal to the trend component that it correlates with economic factors, and the cyclical component that it correlates with weather factors. Second stage, offers DNNs model for the load forecast such as the trend forecasting, and the cyclical forecasting in multi-area.

7.4 Discussion and Summary

The HP-filter in preprocessing stage combined with neural network in forecasting stage is presented. It is used for electricity load demand forecasting based on whole area (WA) and substation control central area (SCCA). The results of whole area forecasting show in Table 7.2 and Table 7.3, and of sub-area shows in Table 7.10. Then the comparison between these forecasting show the accuracy of whole area forecasting with the approach better than sub-area forecasts, can see in these table.

The electricity consumption load demand (kWh) using filter method (HP) and neural network in the power system is presented. The accuracy of forecasts is very vital for fuel reserve planning and unit commitment. A high-quality relationship between load demand and factors will be guaranteed the accuracy of forecasts. The research employs the historical data of demand from EGAT, whole and four substation areas, in Thailand. This investigation, HP-filter used for preprocessing stage will decompose the load demand signal into two components. One is the trend component and the other is the cyclical component. The trend identifies an economic growth and the cyclical component identifies the behavior of load demand. ECLD in each substation control central area covering whole area of the country is used and decomposed it into trend and cyclical components and investigated for prediction in forecasting stage by using "DDNs". Feature input will be chosen by Pearson correlation technique. The appropriate factors are used for feature input for neural network. Experimental results, HP-filter with "DNNs" approach to "whole area" and "SCCA" forecasting, is in good accuracy and also is a new method for SCCA load forecasting. In the research, the results of each area must be integrated to whole area by linear summation. Then the comparison between whole area and sub-area forecasting show the accuracy of whole area forecasting with this approach better than sub-area forecasts.

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th.

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "A computing model of artificial intelligent approaches to mid-term load forecasting: a state-of-the-art-survey for the researcher," IACSIT International Journal of Engineering and Technology, 2, pp.95-100, 2010.

[7] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid Term Load Forecasting of the Country Using Statistical Methodology: Case study in Thailand," International Conference on Signal Processing System, pp.924-928, 2009.

[8] Tomonobu Senjyu, Paras Mandal, Katsumi Uezato, and Toshihisa Funabashi, "Next day load curve forecasting using hybrid correction method," IEEE Trans. Power Syst., 20, pp.102-109, 2005.

[9] Kittipong Methaprayoon, Wei-Jen Lee, Sothaya Rasmiddatta, James R.Liao, and Richard J.Ross, "Multistage artifcial neural network short-term load forecasting engine with front-end weather forecasting," IEEE Trans. Ind. Appl., vol.43, pp.1410-1416, 2007.

[10] Z.A. Bashir and M.E. El-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural net works," IEEE Trans. Power Syst., vol.24, pp.20-27, 2009.

[11] Shu Fan, Kittipong Methaprayoon, and Wei-Jen Lee, "Multiregion load forecasting for system with large geographical area," IEEE Trans. Ind. Appl., vol.45, pp.1452-1459, 2009.

[12] Che-Chiang Hsu, Chia-Yon Chen, "Regional load forecasting in Taiwan-applications of artificial neural networks," International Journal of Energy Conversion and Management 44, pp.1941-1949, 2003.

[13] Kyung-Bin Song, Seong-Kwan Ha, Jung-Wook Park, Dong-Jin Kweon, and Kyu-Ho Kim, "Hybrid load forecasting method with analysis of temperature sensitivities". IEEE Trans. Power Syst., vol.21, pp.869-876, 2006.

[14] Bo-Juen Chen, Ming-Wei Chang, and Chih-Jen Lin, "Load forecasting using support vector machines: A Study on eunite competition 2001," IEEE Trans. Power Syst., vol.19, pp.1821-1830, 2004.

[15] Wei-Chiang Hong, "Chaotic particle swarm optimization algorithm in a support vector regression electric load forecast ing model," International Journal of Energy Conversion and Management 50, pp.105-117, 2009.

[16] Redwan E. Abdel-Adel, "Short-term hourly load forecasting using abductive networks," IEEE Trans. Power Syst., vol19, pp.164-173, 2004.

[17] Kyung-Bin Song, Young-Silk Baek, Dug Hun Hong, and Gilsoo Jang, "Short-term load forecasting for the Holidays using fuzzy linear regression method," IEEE Trans. Power Syst., vol.20, pp.96-101, 2005.

[18] Agnaldo J. Rocha Reis, and Alexandre P. Alves da Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," IEEE Trans. Power Syst., vol.20, pp.189-198, 2005.

[19] Gwo-Ching Liao, and Ta-Peng Tsao, "Application of a fuzzy neural network combined with a chaos genetic algorithm and simulated annealing to short-term load forecasting," IEEE Trans. Evol. Comput., vol.10, pp.330-340, 2006.

[20] Zhang Yun, Zhou Quan, Sun Caixin, Lei Shaolan, Liu Yum ing, and Song Yang, "RBF neural network and anns-based short-term load forecasting approach in real-time price environment", IEEE Trans. Power Syst., vol.23, pp.853-858, 2008.

[21] Shu Fan, Luonan Chen, and Wei-Jen Lee, "Short-term load forecasting using comprehensive combination based on multimeteorological Information," IEEE Trans. Ind. Appl., vol.45, no.4, pp.1460-1466, July-Aug 2009

[22] Huina Mao, Xio-Jun Zeng, Gang Leng, Yong-Jie Zhai, and John A. Keane, "Short-term and midterm load forecasting using a bilevel optimization model," IEEE Trans. Power Syst., vol.24, pp.1080-1090, 2009.

[23] Rui Bo, and Fangxing Li, "Probabilistic LMP forecasting considering load uncertainty," IEEE Trans. Power Syst., vol.24, pp.1279-1289, 2009.

[24] Ying Chen, Peter B. Luh, Che Guan, Yige Zhao, Laurent D. Michel, Matthew A. Coolbeth, Peter B. Friedland, and Stephen J. Rourke, "Short-term load forecasting: Similar day based wavelet neural neurals," IEEE Trans. Power Syst., vol.25, no.1, pp.323-330, 2010.

[25] Emi Mise, Tae-Hwan Kim, Paul Newbold, "On suboptimality of the Hodrick-Precott filter at time series endpoint," Journal of Macroeconomics 27, pp.53-67, 2005.

[26] Robert J. Hodrick, and Edward C. Prescott, "Postwar U.S. business cycles: An empirical investigation, Journal of Money," Credit and Banking 29, pp.1-16, 1997.

[27] Agustin Maravall, and Ana del Rio, "Time aggregation and the Hodrick-Prescott filter," Banco de Espana, 2001.

[28] Jose Ramon Cancelo, Antoni Espasa, Rosmarie Grafe, "Forecasting the electricity load from one day to one week ahead for the Spanish system operator," International Journal of Forecasting 24, pp.588-602, 2008.

[29] Nedaa Agami, Amir Atiya, Mohamed Saleh, and Hisham El- Shishiny, "A neural network based dynamic forecasting model for Trend Impact Analysis," International of Technological Forecasting and Social Change 76, pp.952-962, 2009.

[30] James W. Taylor, "An evaluation of methods for every short-term load forecasting using minute-by-minute British data," International Journal of Forecasting 24, pp.645-648, 2008.

[31] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," The 2nd International Conference on Digital Image Processing, pp.75460B-1-75460B-9, 2010.

[32] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "The comparision of mid term load forecasting between multi-regional and whole country area using artificial neural network," International Journal of Computer and Electrical Engineering 2, pp.338-343, 2010.

[33] Ben krose, and Patrick van der Smagt, "An Introduction Neural Network," Eight edition, 1996.

[34] Phatchakorn Areekul, Tomonobu Senjyu, Naomitsu Urasaki, and Atsushi Yona, "Next day price forecasting in deregu lated market be combination of artificial neural network and ARIMA times series," IEEJ Transactions on Power and Energy, vol.129, pp.1267-1274, 2009.

[35] Wolfgang Polasek, "MCMC Estimation of extended Hodrick-Prescott (HP) filtering models," RCEA Journal, wp25-11 , 2011.

[36] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Improving the model for energy consumption load demand forecasting," IEEJ Transactions on Power and Energy, vol.132, issue. 3, pp.235-243, 2012.

[37] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "multi substations control central load forecasting by using HP filter and double neural networks (HP-filter) model," International Journal of Electrical Power and Energy Engineering, vol.44, issue. 1, pp.561-570, 2013.

Chapter 8

Research Syntheses and Discussion

8.1 Background

The analysis of the research in the previous chapter presented the mid-term load demand forecasting (MTLDF) especially. The case studies employ the data information recorded from the Electricity Generating Authority of Thailand (EGAT), from year 1997 to 2007 [1]-[2]. The previous chapters, many forecasting for example MLR, SARIMA, ANN, WL-ANN, and HP-ANN methodologies are performed in order to find the optimize platform for mid-term load forecasting. Moreover, other parameters such as feature inputs and area divisions are also varied to obtain the better goal of MAPE which is less than +/-3%. Therefore, in this chapter, the experimental results from each chapter will be synthesized and analyzed in various points of view. The discuss will point to the advantages/disavantage of any forecasting approach and also lead to the feature research directions.

There are several methods and focus points in energy forecasting research for the present. Majority of forecasts are short-term load forecasting methods that they have interval in thirty minutes to forty-eight hours. The data for forecasting study came from the competitive information or from several countries in the world. Each method presents the strength of itself: different algorithms, different implementations, and different feature inputs. The complexity, difficulty, and simplicity are different. An index of accuracy of forecasting is an error at that time.

For this research, it is presented in different time period, methods, modeling, regions, and some variables for comparing the results. A few and wide range data may have a problem in forecasting and affect on accuracy. So, it is necessary to analyze several dimensions of the inputs before used. Several influencing factors, which are presented in Chapter 3, will be summarized in this chapter. The synthesis of the variables will be obtained in section 8.2.1. The affecting feature inputs to the energy load forecasting will be concluded. In section 8.2.2, the proposed models mentioned previously are comparatively analyzed. Additionally, the last synthesis, in section 8.2.3, is about the synthesis of area divisions between whole and substation control central areas. Finally, the discussion in the last section of 8.3 are performed to obtain the conclusion in this thesis.

There are several forecasts such as load forecasting, economic forecasting or weather forecasting. In this research, the forecast of energy consumption load demand (ECLD) forecasting interval mid-term period and 12-month ahead, a case study in Thailand are performed. The energy consumption load demand of the country shows in Chapter 3, which is composed of ECLD in whole area (WA) and substation control central area (SCCA). All data of SCCA can be integrated to whole area of the country. The ECLD in each area are different demand following the influencing factors in that area which are economic and weather factors [3-5]. In the research, we divide factors into two major groups. The first group is economic group that presents the components of a trend, such as industrial index, consumer price index, and gross domestic product. The second group is the weather group that proposes the behavior of load demand in detail. It characterizes as cyclical component. Both groups are the main essential points for mid-term consumption load demand forecasting of the country. These ideas are the conceptual model and assumption of researcher. These components (trend and cyclical components) are the essential components in the energy consumption load demand in whole and each area. The quantity of trend and detail components in each area may not be equal followed area. For instance, the northern area components may differ from the northeastern area component although the weather and economic factors are quite alike. For the southern area, the weather factors are different from other regions in the country. It is rainy throughout the year. The temperature is medium level. Therefore, the behavior of a detail component in this area is different from other regions. The trend component has been increased in all areas following the economic factors of each area such as consumer price index or industrial index. Therefore, the trend and detail components of energy consumption load demand are very important for analyzing the factors in mid-term load demand of the country. In this research, the variables used in this approach are confirmed by using the Pearson's correlation method, which is used for finding the correlated value between factors and the trend or detail component of ECLD. This method is in statistical theory and used in several researches. It is used to evaluate the related factors between all variables and demand. The results show in percent correlation (%) or interval -1 to 1. In the case that the Pearson correlation is +1, it will signify a perfect linearly positive increasing correlation trend. If the Pearson correlation is -1, it will indicate a perfect linearly negative declining correlation trend. If the Pearson correlation is between 1 and -1, it will indicate a degree of linear dependence between the two variables. Lastly, if the Pearson correlation is zero, it shows that there is no correlation between the variables [9], [10]. Before a load prediction, all variables must be correlated with the energy consumption load demand including the trend ECLD, and the detail ECLD components in order to choose the appropriate variables for the best feature inputs to the research model. For example, the correlation between temperature variable and ECLD is 0.1 that is meant 10 percent correlation. Also, the correlation between the consumer price index and ECLD is 0.96 that is meant 96 percent correlation. Then the number of percent correlation will present the relationship between variables in statistical approach. It can be seen in Figure 3.6 in Chapter 3.

Before going to other section, we will concisely conclude the relationship between factors and ECLD in each region of Thailand, referred to Chapter 3. The important factors referred from the chapter that are weather and economic factors such as max-min-mean temperature, humidity, rainfall, consumer price index, industrial index, and gross domestic product. The weather factors of the country are come from an averaged degree based on all substations (regions) of TMD. Also, the economic factor like the consumer price index is come from each area (north, northeast, south, and center). Other economic factors for example industrial index is come from whole area, and it is claimed the data to the subarea of the country because the rate of growth can obviously indicate the trend of energy consumption load demand both in whole and substation control central areas. All data used for the forecasting research is in monthly type recorded in kilowatt-hour from 1997 to 2007 (120 data times series). The weather data recorded from TMD in each region were computed to average data to whole area of the country. The economic data were recorded from Ministry of Commerce of Thailand and Ministry of industry of Thailand. Also, the correlation analysis between factor and demand in SCCA is quite clearly in weather factor and consumer price index (CPI).

The temperature affect to the behavior of electricity customer as well. It has affected both low and high temperature. Low temperature, the customer need to use the types of the electric heating appliance such as water heater, kettle, and others. High temperature, air condition is important for several customers for instance offices, companies, and home. It shows that the highest demand happens when the temperature is high. Some countries, a heater is more essential than an air condition because of a long winter time, for example, in U.S., Canada, and Russia. In conclusion, variables and electricity demand of Thailand are necessary for analyzing in demand forecasting by using statistical methods for finding the correlated value before used. If it changes the region to other country for forecasting, it will be found the new correlated between variables and demand again. However, we can use this method, although it has a different area.

8.2 Contributed Syntheses

This section shows the synthesis of the variables affecting the forecasts, the proposed models and the difference of WA and SCCA. Moreover, the suggestion of researcher are. concluded in the last

8.2.1 Variables Affecting the Forecasts

Electricity demand in each country in the world cannot supply the electric power to customer without a thinking of the influencing factors and customer demand. It must plan the future electric demand on each area of the country. The customer demand depended on two major important factors which are economic variables such as consumer price index and industrial index, and weather factors such as max-min-mean temperature, humidity, rainfall, could cover, and ultra violet. Some forecasting researches, forecast may take other factors, for instance, week day, week end, holiday, and seasons for supporting the data in each region.

For the research, we synthesize all factors in this forecasting research by finding the correlate values between factors and load demand. After that, the best factors indicated by the number or percent correlation were used to forecast the electric demand.

Figure 8.1 and 8.2 show examples of relative diagram between electricity load demand(kWh) and factors. It is taken from Table 8.1 which presents the factors and load demand relationship. Figure 8.1 illustrates the correlated values in whole area (WA) and Figure 8.2 shows in subarea (SCCA): Northern area, Northeastern area, Southern area, and Central area. In Whole area, the appropriate factors are an industrial index (IDI), consumer price index (CPI), mean temperature, minimum temperature, maximum temperature. Table 8.1 also shows the appropriate factors both whole area and substation control central area (SCCA). Note that, the minus(-) in the Table 8.1 is the negative

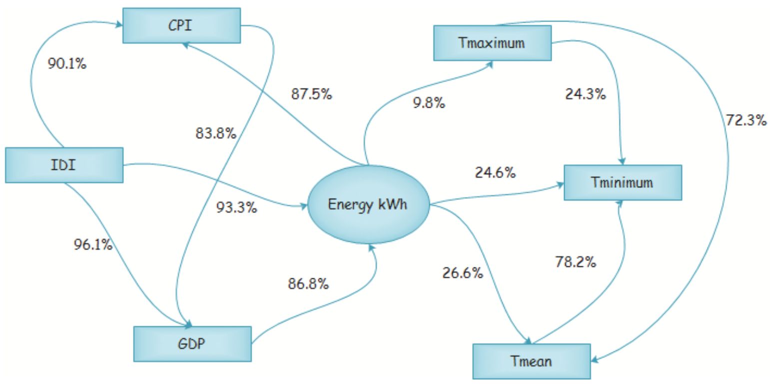

Figure 8.1: Relationship between electricity load demand in mid-term and factors such as economic and weather factors of Thailand (Whole area).

Table 8.1: Appropriate factors in whole and substation control central area.

Relationship between factor and demand in WA and SCCA of Thailand															
	Before					After									
Area	WA	SCCA				HP-WA		HP-SCCA							
Factor	WA	N	NE	S	C	T	C	NT	NC	NET	NEC	ST	SC	CT	CC
T_{max}	9.8	20.5	21.2	17.1	4.6	6.9^{-}	50.5	1.9^{-}	55.5	1.9	58.9	7.4	45.1	11.3^{-}	45.6
T_{min}	24.6	26.5	21.8	11.7	29.0	2.0	71.1	3.3	58.6	0.5^{-}	67.2	1.8	45.3	5.5	71.3
T_{max}	26.6	29.2	28.2	15.9	24.2	1.1	79.8	1.9^{-}	77.1	0.4^{-}	86.6	1.7	65	1.1^{-}	75.9
H	12.3	6.5	6.2	4.4	15.3	2.3	31.7	6.6	1.2	1.8^{-}	23.8	5.4	3.8^{-}	0.7^{-}	47.9
R	14.8	20.8	15.8	10.2	15.1	0.0	46.5	6.3	37.4	1.0	44.8	1.3	40.9	0.4^{-}	46.5
W	16.1	24.6	31.3	30.2	25.6	9.0	23.6	2.2^{-}	66.5	37.4	12.5^{-}	28.3	11.3	10.9	46.3
CPI	87.9	83.9	91.3	93.4	87.9	92.4	1.2	91.1	0.7	95.8	1.8	95.6	1.0^{-}	92.6	2.1
IDI	93.9	90.1	92.9	95.4	93.6	97.9	3.5	97.6	1.4	97.7	1.1	97.1	1.3	97.9	4.2
GDP	86.8	83.2	85.2	89.1	86.5	95.5	11.6^{-}	95.7	11.5^{-}	94.6	13.4^{-}	93.9	12.9^{-}	95.6	10.7^{-}

relationship, and also other values are shown in positive relationship between electricity load demand and factors.

The orders of important factors are shows in Figure 8.3 which shows the relationship between electricity load demand and factors. Whole area and substation control central area, the first to the third important factors are industrial index, consumer price index, and gross domestic product. These are factors in economic group, and also importantly identifies the growth at rate of the country. The consumer price index(CPI) composes of food and drink, apparel, medical care and personal services, transportation and communication, recreation and education, tobacco and alcohol, living expense, inflation, and fresh food and energy. All of these seem to cover the infrastructure which quite relates with an economic growth of Thailand. Thus, this factor is vary significant for whole area and SCCA because it related with electricity load demand. Next, an

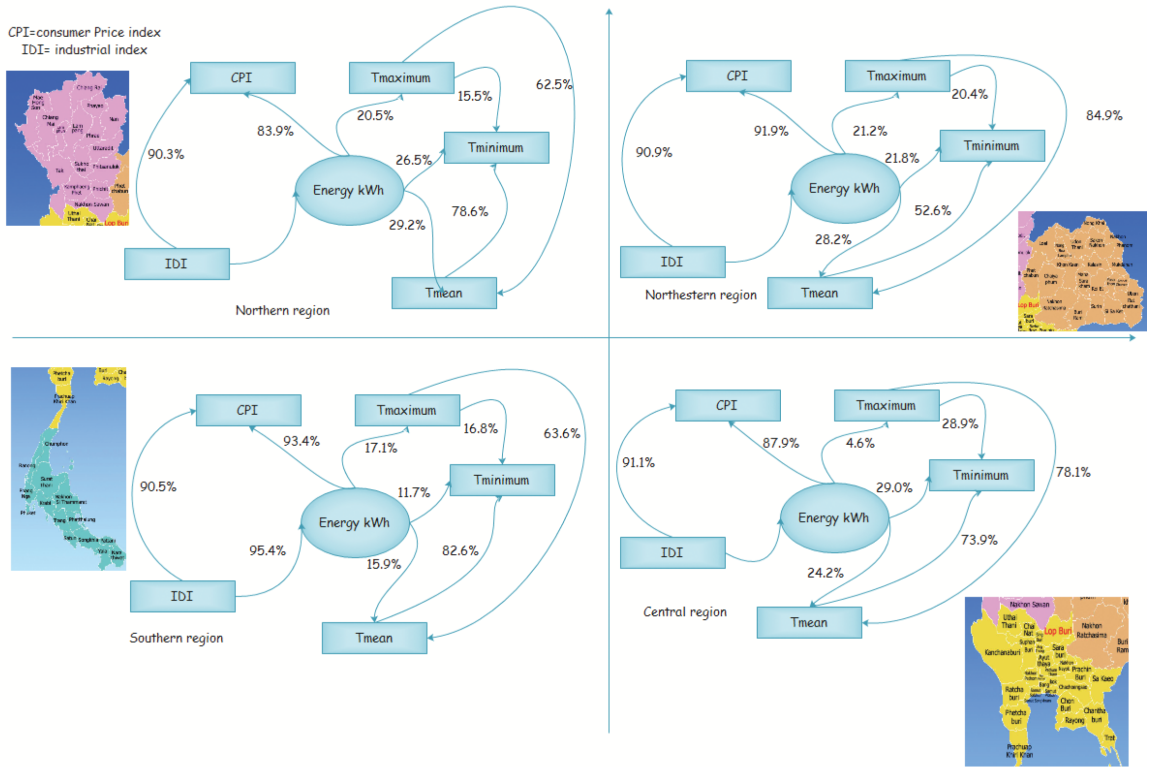

Figure 8.2: Relationship between electricity load demand in mid-term and factors such as economic and weather factors in each area (Substation control central area).

industrial index(IDI), used for whole area forecasting and claimed for SCCA forecasting. It is one of factor that is important for midterm load forecasting. It composes of many components and identifies the growth rate of industry in some areas where have industrial location, for example in central area. In Thailand, the industrial index(IDI) especially shows the total index of the country. Other factors shows the order in Figure 8.3 that shows the factor order in whole area and substation control central area.

Table 8.1 also shows relationship between electricity load demand and factors before and after using Hodrick Prescott filter (HP) in preprocessing stage, in whole area and substation control central area. After preprocessing, the trend and the cyclical load demand components show better correlations between factors. The order of important factors is illustrated in Figure 8.4 and Figure 8.5. Figure 8.4, the first to the third important factors with the trend component are an industrial index(IDI), consumer price index(CPI), and domestic product (GDP). On other hand, the first to the third important factors with the cyclical component are mean temperature, minimum

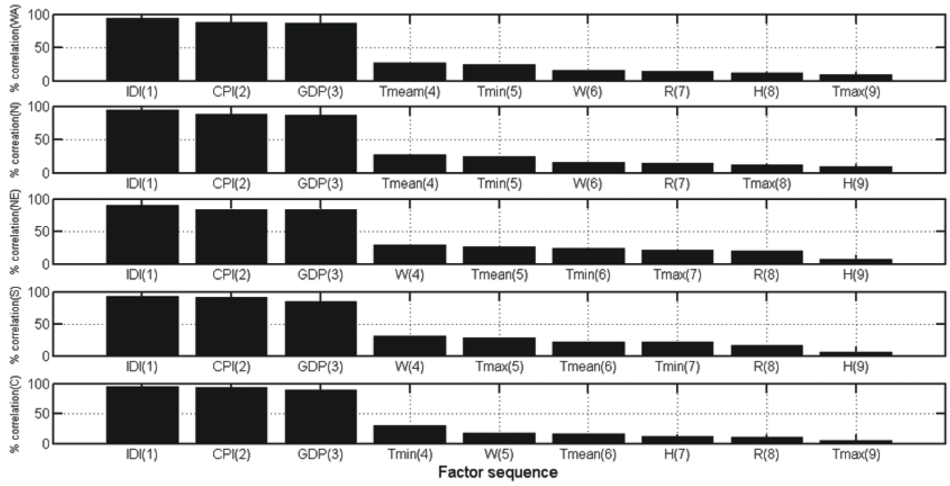

Figure 8.3: Relationship between electricity load demand and factors.

temperature, and maximum temperature exempted the central area factors that are mean temperature, min temperature, and humidity. Central area shows humidity in the third important factors because the load of air condition may be normally used in almost all time be parallel the electricity load demand in this area.

8.2.2 Proposed Models

The investigation of forecasting research in peak load and energy demand prediction, are obtained in several methods such as statistical methods: exponential smoothing (ES), multi-linear regression (MLR), autoregressive integrated moving average (ARIMA), and seasonal autoregressive integrated moving average (SARIMA), and artificial intelligent methods: expert system, neural network, and fuzzy logic. Statistical methods are widely used in accounting, economics, and some in engineering. While, artificial intelligent (AI) methods are widely used in engineering field and others. In the research, we mainly used the neural network for load demand forecasting. This method shows the ability for learning and predicting the non-linear data of load demand well. For this section, we conclude all methods as following.

This research studies electricity load demand by using five approaches: multiple linear regression(MLR)(Chapter 4), seasonal autoregressive moving average (SARIMA)(Chapter 4), neural network (NN)(Chapter 5 [12]), wavelet transform and

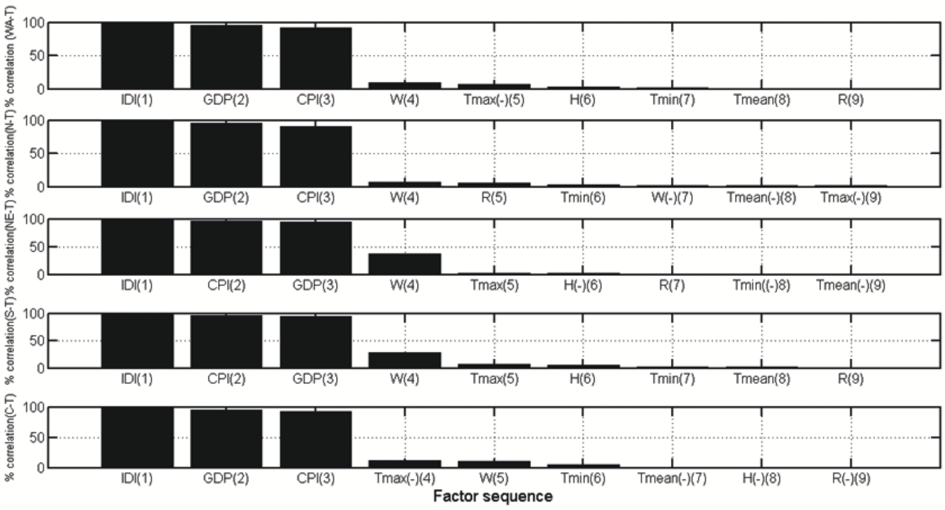

Figure 8.4: Relationship between the trend of load demand and factors.

neural network (W-NN)(Chapter 6 [6],[8][14]), and Hodrick-Prescott filter and neural network (HP-NN)(Chapter 7 [9],[10]). The proposed methods are presented in Figure 8.6 to Figure 8.10. In the first, from Figure 8.6 illustrates the best results from each approach (Chapter 4 to Chapter 7). It shows the Mean Absolute Percent Error (MAPE) in each approach of whole area(WA). The details of figure show MAPE of each approach: MLR, SARIMA, NN, W-NN, and HP-NN, respectively. MLR approach, the appropriate MAPE of this approach is 1.86% from model no.6 in wholw area. The factors used in this model are mean temperature, consumer price index, industrial index, and gross domestic product (Chapter 4). There are eight models in statistical model. The results are summarized in Table in Chapter 4. The model explanatory power as reflected by adjusted R^2 and model estimation expectations of the coefficient signals are the model selection criteria. This study shows eight models of multiple linear regressions and factors such as maximum temperature, minimum temperature, mean temperature, humidity, consumer price index, industrial index, and gross domestic product[3], [4]. Therefore, MAPE of whole area forecasting is 3.36% that is a maximum error of Model No.2, and also the minimum error is 1.86%. For this approach, MAPE of substation control central areas in northern, northeastern, southern, and central area are 2.356% (Model No.7), 2.384% (Model No.7), 2.361% (Model No.8), respectively, and also MAPE of total SCCA is 1.925% based on Model No.6 which can be seen in Chapter 4 and Figure 8.6. Next approach, forecasting by using seasonal autoregressive moving

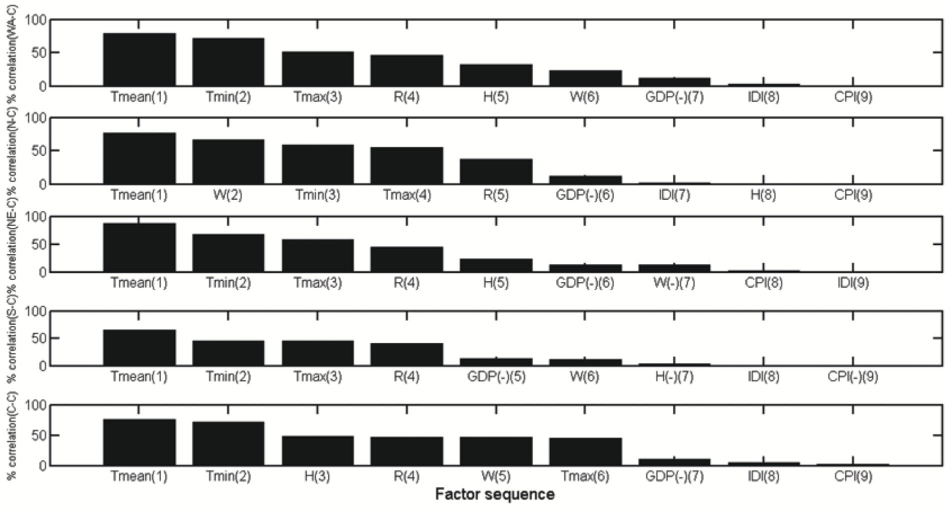

Figure 8.5: Relationship between the cyclical of load demand and factors.

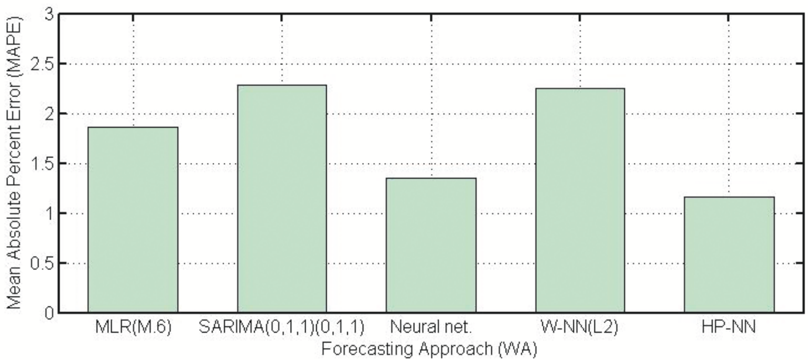

Figure 8.6: MAPE in each approach of whole area (WA).

average (SARIMA), is also known as the Box-Jenkins model and is used in time series analysis. It goes a step further than the Autoregressive Moving Average (ARMA) model which is the best suited to stationary time series, by modeling and forecasting non stationary time series. The ARIMA[p,d,q] model consists of these elements: p is autoregressive, d is differencing involving the first derivative of the following procedure: determination of seasonality of time series, differencing transformations to make the time series stationary, investigation of significant inputs, determination of order of seasonal and non-seasonal regression, and identification of model parameter. For this approach, whole area, the MAPE is 2.28% by using SARIMA$(0,1,1)(0,1,1)_{12}$ model, and also substation control central area: NSCCA, NESCCA, SSCCA, and CSCCA, which

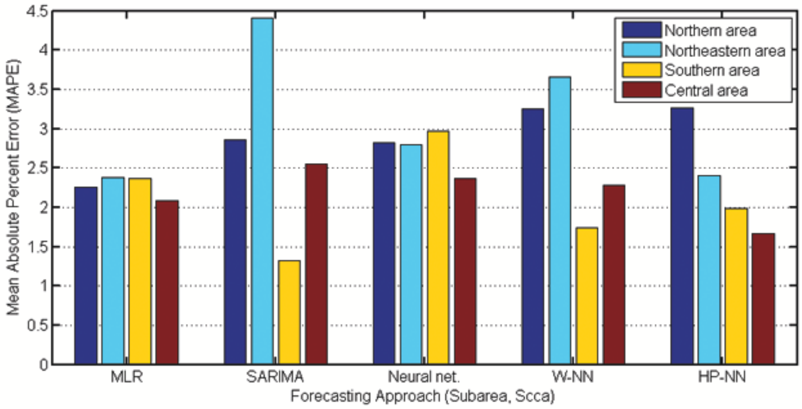

Figure 8.7: MAPE in each approach of each subarea (SCCA).

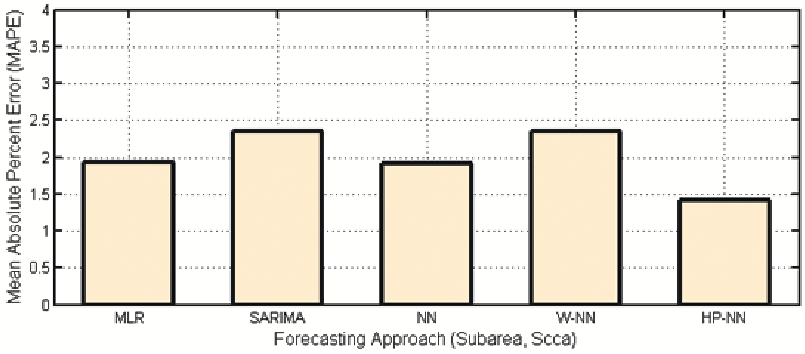

Figure 8.8: MAPE in each approach from SCCA in each area.

the MAPE are 2.862% (SARIMA(0,1,1)((0,1,1)$_{12}$), 4.412% (SARIMA(0,1,2)(0,1,1)$_{12}$), 1.320% (SARIMA(1,1,1)(0,1,1)$_{12}$), and 2.546% (SARIMA(0,1,1)(0,1,1)$_{12}$), respectively. Next approach, artificial neural network approach, usually called neural network(NN), is a mathematical model or computational model that is inspired by the structure and/or function aspects of biological neural networks. A neural consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases, an ANN is an adaptive system that changes its structure based on external or internal information that flow through the network during the learning phase. Modern neural networks are non-linear statistical data modeling tools. They are usually used to model complex relationship between inputs and outputs or to find patterns in data (Chapter 5). For this approach, Figure 8.9 and 8.10 (the top right), show neural network model which has three-layer based on feed-forward backpropagation algorithm. The neuron to find the optimal weight and bias before simulation or forecasting.

An accuracy, the MAPE of whole area is 1.35%, and also SCCA: northern,northeastern, southern, and central area are 2.819%, 2.798%, 2.974%, and 2.368%, respectively. Thus, the result of whole area from SCCA is 1.922%. Next approach, whole and substation control central area forecasting using wavelet transform and neural network based on model, is proposed. The wavelet transform technique is used in preprocessing stage, and also neural network is used for forecasting stage. The wavelet analysis calculates the correlation between the signal under consideration and a wavelet function. The similarity between signal and analyzing wavelet function is computed separately for different time interval, resulting in a dimensional representation. The analyzing wavelet function is also referred to as the mother wavelet. For this approach, investigates four levels that are a level, two-level, three-level, and four-level. For the result, in whole area, the appropriate MAPE is 2.254% based on two-level. Also the result of northern, northeastern, southern, and central area, are 3.258% (2-level), 3.662% (2-level), 1.735% (2-level), and 2.280% (2-level), respectively. So, the best MAPE result of whole area from SCCA forecasting based on 2-level is 2.072%. To sum up, as a result, "two-level" is given the Mean Absolute Percent Error (MAPE) better than other levels and "sub-control central area" shows MAPE better than "whole area forecasting". Final approach, whole and substation control central area using Hodrick-Prescott filter (HP-filter) and neural network based on model, is proposed. The trend and cyclical decomposition are routine in modern macroeconomics but this research is just started to take it into electricity load demand decomposition in power system. The basic idea is to decompose the load demand time series of interest (kWh) into the sum of a slowly evolving secular trend and a transitory deviation from it which is classified as "cycle". The Hodrick-Prescott filter separates a time series into a trend component which identified as economic growth factors included the industrial index and consumer price index, and a cyclical component indentified as behaviors of weather factors included temperature, humidity, rainfall, and wind speed. For the result of whole area, the MAPE is 1.16% that the factors used in model are the trend and cyclical load demand component, maximum temperature, minimum temperature, mean temperature, and industrial index. The results of SCCA: NSCCA, NESCCA, SSCCA, and CSCCA, are 3.26%, 2.40%, 1.98%, and 1.66%, respectively. Thus, whole area form SCCA, the MAPE is 1.42%.

As a result, Table 8.2, the forecasting approach comparison shows the eight basic comparison of these approaches: multiple linear regression (MLR), seasonal autoregressive integrated moving average (SARIMA), neural network (NN), wavelet and neural network (W-NN), and Hodrick-Prescott filter and neural network (HP-NN). The first, model complexity model, the MLR, SARIMA, NN, and HP-NN are medium complexity, especially, HP-NN which it additionally study in the trend and cyclical components. The wavelet transform is very high complexity, and also composes of one or more components for comparison. Secondly, the number of factors, are essential for learning data. If the model has several factors, it may be taken a lot of time for work. In this study, there are many factors in MLR and W-NN approaches, and few factors in HP-NN approach. SARIMA approach is no factor required because it learn from the pattern of energy load demand in the past. SARIMA and W-NN are quite complited while MLR, NN, HP-NN are less complexity. Fourth, computation time in each approach, it shows long time study and investigation in MLR, SARIMA, and W-NN. Medium time study and investigation are shown in NN and HP-NN. The fifth consideration is easy understanding of algorithm and model. The MLR, SARIMA, and W-NN are difficult to understand but NN and HP-NN approaches are quite easy to understand. The sixth, all models can be similar applied to other works. The seventh is accuracy which is necessary for the prediction. The MLR and SARIMA may be less accurate than other approaches because they are more suitable with a more linear time series than non-linear. The W-NN approach, in this study, shows medium precision which may be due to the complexity of the model. The NN approach, shows quite high accuracy but less than HP-NN approach. Finally, the reliability of each approach shown that MLR, SARIMA, NN, and W-NN are medium reliability, and also HP-NN is high reliability of the model. Thus, HP-NN is less complexity, a few number of factors, relative easy algorithm, easy to understand, applied to other works, high accuracy, and high reliability.

8.2.3 Differences between WA and SCCA

For this research proposes the differential areas that are whole area (WA) and substation control central area (SCCA) (4 SCCAs) for mid-term load demand forecasting. Forecasting employs the data recorded from the Electricity Generating Authority of Thailand. The weather information recorded from Thailand Meteorological department in each region such as the northern area, the northeastern area, the southern area, and the central area.

The economic and weather factors in each region are important to en-

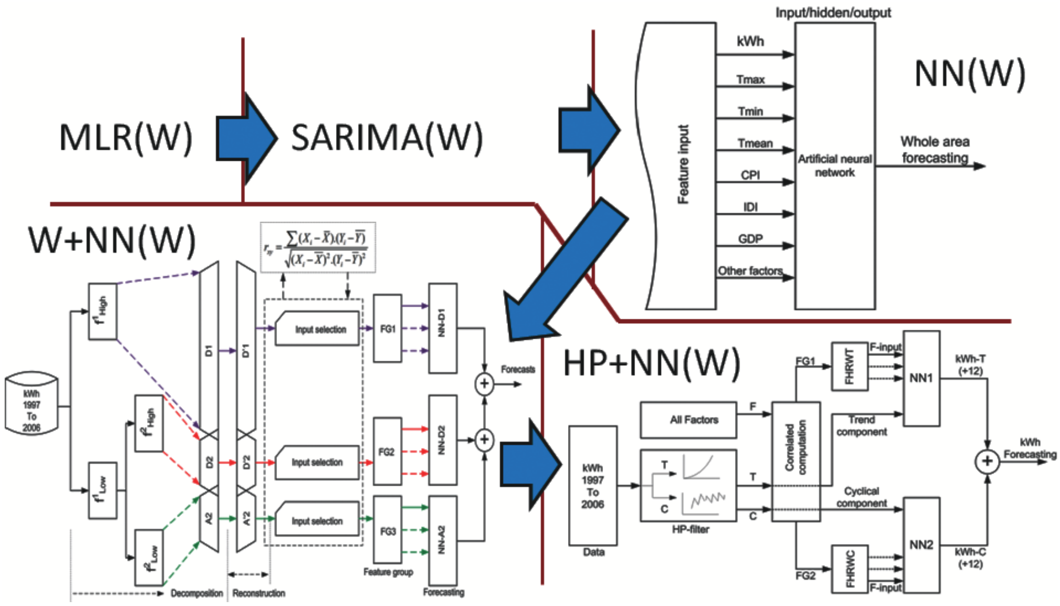

Figure 8.9: All models for electricity load demand forecasting in the research, whole area.

ergy consumption load demand. Large areas may require less or more electricity than a small area which are depended on these influencing factors. In this research, we present both whole area and substation control central area forecasting which show as following. Figure 8.9 shows five models for energy load demand forecasting for the country. Figure 8.10 also shows five models for the substation control central area forecasting. There are many different variables for the region factors such as weather and economic factors. The weather factors in whole area are the average of the weather in total subarea of the country. The SCCA weather factors show different values in the time. The economic factors in whole are total index values of the country such as the industrial index, consumer price index, gross domestic product, and other factors. The industrial index does not cover all region of the country. It has especially some areas used for the industrial zone, for instance chonburi, rayong, and ayuttaya provinces. The consumer price index which comes from many factor for identifying the growth at rate of economy (Section 8.2.1). Thus, the whole area factors show the behavior and trend of electricity load demand of the country, and also the substation control central area deep down in the behavior and trend of the electricity load demand of the subareas: northern, northeastern, southern, and central area. Figure 8.11 illustrates the Mean Absolute Percent Error (MAPE) in each model for whole area(WA), subarea to whole area(WA-SCCAs), northern area(N),

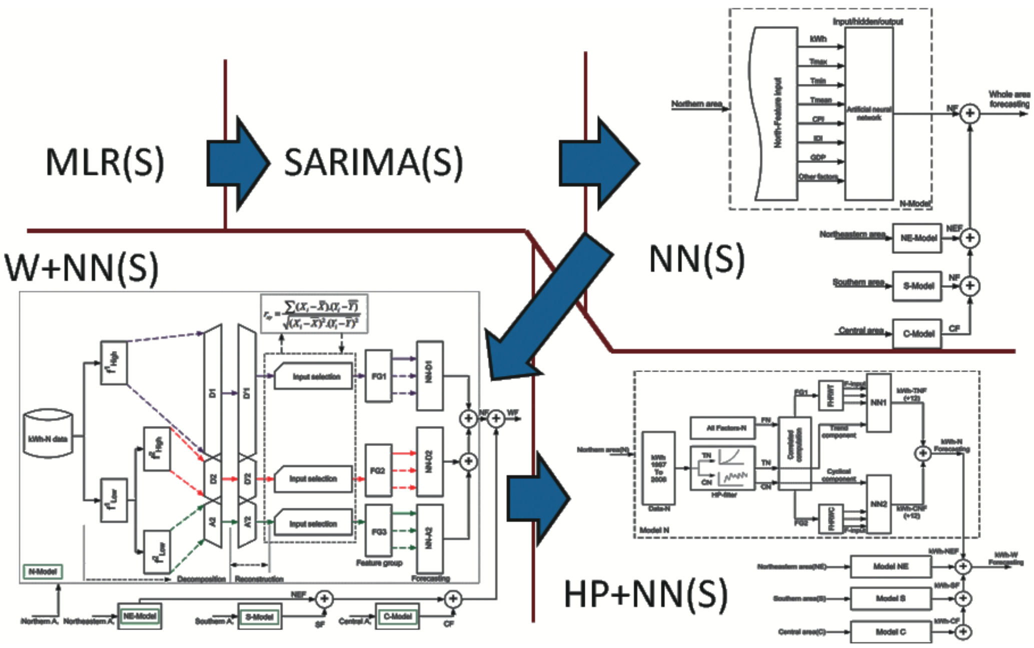

Figure 8.10: All models for electricity load demand forecasting in the research, subarea.

northeastern area(NE), southern area(S), and central area(C). All approaches almost give the best result in whole and SCCA forecasts. In subarea forecastings, some areas may be low accuracy such as northeastern area in SARIMA and W-NN. Summarily, the comparison of whole area and substation control central area forecastings, the best results happens in whole area forecasting.

8.2.4 Suggestions

Our suggestions for the research is that the accuracy of forecast is depended on choices of influencing variables directly affected on the energy load demand. The variables that are not related to the load demand should not be used for this forecasting. This research presents the appropriate variables as the parameters are used for this model. All factors are the main variables for driving the demand in the future. In the next research, we would like to study in multi-region, in the basic structure of each area, the development of the region, the growth of the population, and the number of households. This may take a lot of times, and researchers to collect these data. This research studying substation control central area (SCCA), we only employ four substations that are the NSCCA, NESCCA, SSCCA, and CSCCA. If the researcher wants to study in multi-substation, the research can use the data of substation of EGAT which has the substations to 200 substations. Also, it will show the behavior and the trend

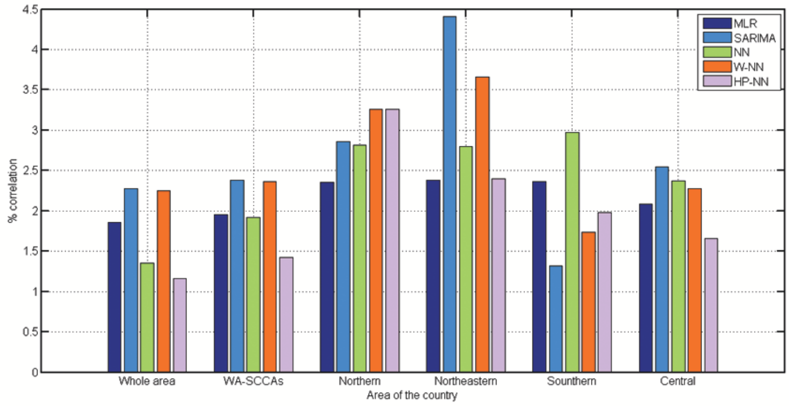

Figure 8.11: The MAPE of area comparison in Thailand.

components of the energy load demand clearly. If forecasting has more substations, it may decrease the accuracy as the result of more influencing factors. The advantage of SCCA forecasting is that it has found the authentic influencing factor in each region. Finally, this method is easier approach and non complexity, and it can precisely forecast the mid-term load demand.

8.3 Discussion

In the previously research synthesis, we found that the influencing factors to demand forecast have difference from other period forecast: very short-term, short-term, and long-term load demand forecasting. Firstly, in very short-term and short-term forecast, the important variables are min-max-mean temperature in hourly and daily, and other factors: humidity, wind speed weekend, weekday, and holiday. Secondly, long-term load forecast, the period of this approach can predict the demand up to 25 year ahead. The main influencing factors are economic variables that are industrial index, consumer price index, gross domestic product, population, and household, yearly. The weather factors may not used for feature inputs for forecasting due to the time series with long term (up to 25 years). Thus, mid-term load forecasting, there are different influencing factors related with the electric demand. There are two groups: economic group and weather group. It cannot nelect any group because the demand depends on both groups: economic and weather groups. It may reduce the accuracy and reliability of demand forecasting. Previously, although this research has five to seven factors, but these variables are the main essential factors which are an infrastructure of the country like consumer price index (CPI) and industrial index (IDI). The economic factors and weather factors are obviously identified the trend component and cyclical (detail or behavior) component, respectively. Next, we study several methods to be used for mid-term load demand forecasting, twelve months ahead. There are four approaches studied in the research, for instance multiple linear regressions (MLR) , seasonal autoregressive integrated moving average (SARIMA), neural network, wavelet transform (WT) with neural network, and Hodrick-Prescott (HP-filter) with neural network. The first two methods are statistical approaches that are widely used in economic, social, and engineering fields. The accuracy of each method is low. Neural network and wavelet transform are widely used in engineering field. However, the proposed method, HP-filter with neural network is the first approach for integrating for demand forecasting. HP-filter approach is easier to use for filtering the demand signal. Each signal of trend and detail components shows higher relationship between each demand component and related factors. This approach can greatly reduce the complexity of the wavelet transform method as well. Finally, this research also presents a forecasting model both whole (WA) and substation control central area (SCCA). Whole area forecasting employs the demand data of the country, averaged weather data, and economic factors of the country. SCCA employs the demand and factors from each area except the industrial index that is claimed from whole region. The multi-region forecasting will show the trend and behavior of electricity which are related with influencing factors in each area and will lead to the multi-area prediction of the electricity demand in the future. Although an accuracy of the demand prediction in region is less than the total area forecasting. But it ostensibly shows the load demand in each region corresponding to the influencing factors in that area. Researcher who interested this approach may bring it to build the mid-term load demand model more efficiently.

8.4 Conclusion

Mid-term load demand forecasting of Thailand, studying and analyzing the data employs demand information from the Electricity Generating Authority of Thailand: EGAT, and the factors recorded from many ministry of the country, in similar time series. There are three concepts used in the research. The first is influencing factor analysis with load demand by use Pearson's correlation. Secondly, preprocessing approach

Table 8.2: Forecasting approach comparison.

Discription	MLR	SARIMA	NN	W-NN	HP-NN
Model Complexity	medium	medium	medium	high	medium
Number of Factor	many factors	non	a few of factors	many factors	a very few of factors
Algorithm	difficult	difficult	easy	difficult	easy
Time	long times	long times	short times	long times	medium times
Easy Understanding	difficult	difficult	easy	difficult	medium
Applied to other works	easy	easy	easy	easy	easy
Accuracy	medium	medium	high	medium	quite high
Reliability	medium	medium	high	medium	quite high

by HP-filter method which is presented for reducing the complexity of the method such as wavelet transforms. Finally, the electricity load demand shows the forecasting implementation in whole area and substation control central area of the country. Influencing factor analysis, HP-filter in preprocessing approach, and forecasting area implementation are particularly useful for a new mid-term load forecasting model. It shows high accuracy, and easy to learning about the implementation.

References

[1] Electricity Generating Authority of Thailand (EGAT), http://www.egat.go.th

[2] Thai Meteorological Department, Ministry of Transport and Communications, http://www.tmd.go.th.

[3] The Office of the National Economic and Social Development Board, Ministry of Commerce of Thailand, http://www.moc.go.th.

[4] Ministry of Industry Thailand, http://www.industry.go.th.

[5] Ministry of Energy, Thailand, http://www.energy.go.th.

[6] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Mid-tem load forecasting: Level suitably of wavelet and neural network based on factor selection," Energy procedia 14, pp.438-444, 2012.

[7] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Wavelet and neural network approach to demand forecasting based on whole and electric sub-control center area," International Journal of Soft Computing and Engineering, vol.1, issue. 6, pp.81-86, 2012.

[8] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," proceedings of SPIE second International conference on digital image processing, vol. 7546, pp.75460B-1-75460B-9, 2010.

[9] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Improving the model for energy consumption load demand forecasting," IEEJ Transactions on Power and Energy, vol.132, issue. 3, pp.235-243, 2012.

[10] Pituk Bunnoon, Kusumal Chalermyanont, Chusak Limsakul, "Multi-substations control central load demand forecasting by using HP filter and double neural networks (HP-filter)," International Journal of Electrical Power and Energy Engineering, vol.44, issue.1, pp.561-570, 2013.

[11] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid term load forecasting of the country using statistical methodology: case study in Thailand," 2009 International conference on signal processing systems, pp.924-928, 2009.

[12] P. Bunnoon, K. Chalermyanont, and C. Limsakul, "The comparision of mid term load forecasting between multi-regional and whole country area using artificial neural network," International journal of computer and electrical engineering, vol. 2, no. 2, pp.338-343, 2010.

[13] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "A computing model of artificial intelligent approaches to mid-term load forecasting: a state-of-the-art-survey for the research," IACSIT International journal of engineering and technology, vol.2, no.1, pp.94-100, Feb.2010.

[14] Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," International conference on advance energy engineering, pp.438-444, 2011.

Chapter 9

Conclusion and Future Works

9.1 Conclusion

The major contributions of the thesis was summarized in this final chapter. This investigation presented the approaches to mid-term load forecasting based on hybrid correction methods which are the combination between the artificial neural and appropriate approaches. In particular, chapter 2, the state-of-the-art of this study are proposed. The demand forecasting approaches are started by using statistical method such as multiple linear regression (MLR) and seasonal autoregressive integrated moving average (SARIMA). Next approaches are combined the preprocessing methods with neural network, for instance wavelet transform and Hodrick-Presscott filter for demand forecasting. These five methods appled to mid-term electricity load demand forecasting are covered two cases studies that were whole area and substation control central areas. The highlight of research shows the new approach to the mid-term load demand forecasting by using HP-filter combined with neural network. This dissertation obtained the contributions as following:

- Firstly, the research proposes the application of neural network (NN) for forecasting with the non-stationary time series of mid-term electricity load demand of EGAT, and applied several approaches to solve the problem: MLR, SARIMA, NN, W-NN, and HP-NN. Also, the predictions show the results of the investigation both whole area and substation control central area.
- The second is to apply the new approach of HP-filter to feature correction or improvement before using as the feature input for neural network. The model is the combination of HP-filter algorithm and ANN algorithm. The HP-filter works in preprocessing stage, and ANN also works in forecasting stage. These approaches can be applied to other works related the trend and cyclical components in both stationary and non-stationary time series. This is a new model for applying and investigating the electriclity load forecast in the future.
- Next, several factors used in the research were computed the correlation value between factors and electricity load demand-before and after preprocessing using

HP-filter. The trend and cyclical components after decomposition quite showed very high relationship between economic factors and the trend component, and weather factors and the cyclical component of electricity load demand.

- Finally, all results of these approaches found that the HP-filter with Neural network forecasting was the appropriate approach when compared with other models. It was better both whole area and substation control central area forecasting.

9.2 Future Works

This thesis mainly investigates and applies the neural network (NN) to electricity mid-term load demand forecasting which combined the forecast approaches more than one model: W-NN and HP-NN. They lead to the improvement of the forecasting performance for non-stationary time series. The appropriate approaches, Hodrick-Prescott filter and neural network show the best method for factor separation and prediction. Then in the future these approaches may be led to development to short-term, mid-term or long-term electricity load demand forecasting or other works.
The future researches:

- Improve the best model (Chapter 7) to short-term, mid-term, and long-term electricity load demand forecasting, and use more relative factors such as day-type, weekday, holiday, population, oil price, gold price, and other significant factors, to increase the forecasting accuracy.

- Apply the feature inputs for neural network based on seasonal factors to forecasting.

- Vary the electricity load demand forecasting for mid-term demand from three months to three years for comparison.

- Apply the HP-filter optimization for correcting patterns and improving the appropriate factors for feature inputs of the model.

Conferences/Journals

.1 International conferences

There are three international conferences as follows.

1. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid term load forecasting of the country using statistical methodology: case study in Thailand," 2009 International conference on signal processing systems, pp.924-928, 2009 (Singapore).

2. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Peak load demand forecasting using 2-level discrete wavelet decomposition and neural network algorithm," proceedings of SPIE second International conference on digital image processing, vol. 7546, pp.75460B-1-75460B-9, 2010 (Singapore).

3. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," International conference on advance energy engineering, pp.438-444, 2011 (Thailand, Hongkong).

.2 International journals

There are six international journals as follows.

1. P. Bunnoon, K. Chalermyanont, and C. Limsakul, "The comparision of mid term load forecasting between multi-regional and whole country area using artificial neural network," International journal of computer and electrical engineering, vol. 2, no. 2, pp.338-343, 2010.

2. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "A computing model of artificial intelligent approaches to mid-term load forecasting: a state-of-the-art-survey for the research," IACSIT International journal of engineering and technology, vol.2, no.1, pp.94-100, Feb.2010.

3. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Mid-term load forecasting: Level suitably of wavelet and neural network based on factor selection," International journal on energy procedia, vol.14, pp.438-444, 2011 (scopus, sciencedirect, SJR=0.088).

4. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Improving the model for energy consumption load demand forecasting," IEEJ Transactions on power and energy, vol.132, no.3, pp.235-243, 2012 (scopus, SJR=0.028).

5. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Wavelet and Neural Network Approach to Demand Forecasting based on Whole and Electric Sub-Control Center Area," International Journal of Soft Computing and Engineering(IJSCE), vol.1, issue 6, pp.81-86, 2012 (DOAJ, J-stage).

6. Pituk Bunnoon, Kusumal Chalermyanont, and Chusak Limsakul, "Multi substation control central load forecasting by HP Filter and double neural networks (HP-DNNs) model," International Journal of Electrical Power and Energy System, vol.44, issue 1, pp.561-570, 2013 (ISI impact factor 2.247).

VITAE

Name Mr. Pituk Bunnoon

Student ID 5110130008

Educational Attainment

Degree	Name of Institution	Year of Graduation
B-Eng (Electrical Engineering)	Kingmongkut's Institute of Technology Ladkrabang	1998
M-Eng (Electrical Engineering)	Prince of Songkla University	2002

Work-Position and Address

Lecturer at Rajamangala University of Technology Srivijaya

address 1 Raddumniennok Road, Boiyang Vilage, Songkhla District, Songkhla Province, 90000, Thailand

Scholarship during Enrolment

This work was funded by the Office of the Higher Education Commission, Thailand. Pituk Bunnoon was supported by CHE510382 Ph.D. Scholarship.

Lists
International Conferences/Journals