

raf

TK1005 .C77 2010.



0000073274

Short term load forecasting with time series analysis / Che
Roslee Che Remle.

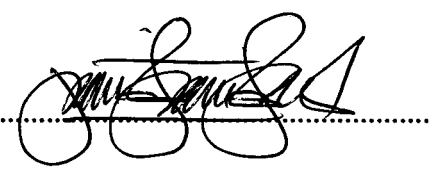
**SHORT TERM LOAD FORECASTING
WITH TIME SERIES ANALYSIS**

Che Roslee Bin Che Remle

Bachelor of Electrical Engineering (Industrial Power)

April 2010

“ I hereby declare that I have read through this report entitle “Short Term Load Forecasting with Time Series Analysis” and found that it has comply the partial fulfillment for awarding the degree of Bachelor of Electrical Engineering (Industrial Power)”

Signature : 

Supervisor's Name : AIMIE NAZMIN BINA AZMI

Date : 22 APRIL 2010

SHORT TERM LOAD FORECASTING WITH TIME SERIES ANALYSIS

CHE ROSLEE BIN CHE REMLE


**A report submitted in partial fulfillment of the requirements for the degree of
Bachelor of Electrical Engineering (Industrial Power)**

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2009/10

I declare that this report entitle “Short Term Load Forecasting with Time Series Analysis” is the result of my own research except as cited in the references. The report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : 

Name : CHE ROSLEE B. CHE KEMLE

Date : 21 APRIL 2010

To my beloved family and friends

ACKNOWLEDGEMENT

Sometimes words fall short to show gratitude, the same happened with me during this Final Year Project process. The immense help and support received from my supervisor and friends overwhelmed me during the process of the project.

It was a great opportunity for me to be supervised by my supervisor Mrs Intan Azmira Binti Abdul Razak and Mr. Aimie Nazmin bin Azmi. I am extremely grateful during the entire project, they had shared their expertise and knowledge with me and without whom the completion of this first part of the project would have been virtually impossible.

I am highly indebted to panels, Mr Alias Bin Khamis and Mrs Elia Erwani Binti Hassan who have provided me the valuable suggestion and a good support to improve the project during the first seminar. I am thankful to everyone who always inspires me to take right step in the journey of success in my life.

ABSTRACT

Load forecasting is vitally important for the electric industry in the deregulated economy. It has many applications including energy purchasing and generation, load switching, contract evaluation, and infrastructure development. A large variety of mathematical methods have been developed for load forecasting. Short-term load forecasting plays an important role in electric power system operation and planning. An accurate load forecasting not only reduces the generation cost in a power system, but also provides a good principle of effective operation. In this project, the Autoregressive Integrated Moving Average (ARIMA) of Time-Series model will be applied to the short-term load forecasting for the Peninsular Malaysia load data. ARIMA is a practical forecasting method in the electric short-term load forecasting fields for linear prediction. The choice of the forecasting model becomes the important factor to improve load forecasting accuracy. The aim of this project is to achieve forecasting error that is equal or less than 1.5% using Minitab and XLSTAT statistical software. The data collected is 7 weeks of half an hourly load data for Peninsular Malaysia.

ABSTRAK

Peramalan beban adalah sangat penting untuk industri elektrik dalam keadaan ekonomi yang tidak menentu. Ia mempunyai banyak penggunaan termasuk dalam pembelian tenaga dan generasi, pensuisan beban, penilaian kontrak dan pembangunan infrastruktur. Pelbagai kaedah matematik telah dibangunkan untuk tujuan peramalan. Peramalan beban jangka pendek memainkan peranan penting dalam sistem operasi kuasa elektrik dan perancangan. Satu peramalan beban yang tepat bukan sahaja mengurangkan kos penjanaan dalam satu sistem kuasa, tetapi ia juga dapat beroperasi dengan efektif. Dalam kertas ini, *Autoregressive Integrated Moving Average* (ARIMA) dari model Siri-Masa akan diaplikasikan dalam proses ramalan beban jangka pendek untuk data beban Semenanjung Malaysia. ARIMA adalah satu kaedah peramalan yang praktikal dalam peramalan linear beban elektrik jangka pendek. Pilihan model peramalan menjadi faktor mustahak untuk meningkatkan ketepatan ramalan. Tujuan projek ini adalah bagi mencapai ralat iaitu sama atau kurang daripada 1.5% menggunakan perisian statistik Minitab dan XLSTAT. Data yang digunakan adalah data 7 minggu yang diambil dengan sela setiap setengah jam untuk beban di Semenanjung Malaysia.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE
	ACKNOWLEDGEMENT	v
	ABSTRACT	vi
	ABSTRAK	vii
	TABLE OF CONTENTS	viii
	LIST OF TABLES	xii
	LIST OF FIGURES	xiii
	LIST OF ABBREVIATIONS	xv
	LIST OF APPENDICES	xvii
1	INTRODUCTION	1
	1.1 Problem Statement	1
	1.2 Objectives	2
	1.3 Scope	3
2	LITERATURE REVIEW	4
	2.1 Journal review 1	5
	2.2 Journal review 2	6
	2.3 Journal review 3	7
	2.4 Journal review 4	8
	2.5 Conclusion of paper review	9
	2.6 Theory	9

2.6.1	Autoregressive Integrated Moving Average (ARIMA)	10
3	METHODOLOGY	14
3.1	The steps of the Methodology	14
3.1.1	Research on Time Series Forecasting method	16
3.1.2	Choose the specific method to be applied.	17
3.1.3	Data Collection	18
3.1.4	Data Analysis	18
3.1.4.1	The Analysis of the graph	24
3.1.5	Grouping the Data	27
3.1.6	Data Simulation	29
3.1.7	Trend Analysis	29
3.1.8	Spectral Plot	30
3.1.9	Box Cox Plot	30
3.1.10	Data Transformation	31
3.2	Forecasting Procedures	32
3.2.1	ARIMA Models	34
3.2.2	Forecasting Process	35
3.2.3	Summary of rules for identifying ARIMA models	35
3.2.4	Diagnostic Checking	37
3.2.5	Model Identification	38
3.2.6	The summary of result	38
3.3	Differencing	39

4	RESULTS & ANALYSIS	40
4.1	Trend Analysis	40
4.2	Data transformation (Box Cox Plot)	41
4.3	Result of forecasting	43
	4.3.1 Group1 (MON)	44
	4.3.2 Group2 (TUE, WED)	45
	4.3.3 Group3 (THU)	46
	4.3.4 Group4 (FRI)	47
	4.3.5 Group5 (SAT)	48
	4.3.6 Group6 (SUN)	49
4.4	Forecasting Error	50
	4.4.1 Group1 (MON)	51
	4.4.2 Group2 (TUE, WED)	52
	4.4.3 Group3 (THU)	53
	4.4.4 Group4 (FRI)	54
	4.4.5 Group5 (SAT)	55
	4.4.6 Group6 (SUN)	56
4.5	Differencing	57
	4.5.1 Group1 (MON)	58
	4.5.2 Group2 (TUE, WED)	59
	4.5.3 Group3 (THU)	60
	4.5.4 Group4 (FRI)	60
	4.5.5 Group5 (SAT)	61
	4.5.6 Group6 (SUN)	61

5	DISCUSSION OF RESULT	62
5.1	Discussion	62
5.1.1	Data analysis	62
5.1.2	Trend analysis	63
5.1.3	Box Cox plot	63
5.1.4	Spectral plot	63
5.1.5	Forecast	63
5.1.6	Percent error	64
5.1.7	Differencing	64
6	CONCLUSSION	65
6.1	Summary	65
6.2	Conclusion	66
6.2.1	Recommendation	67
	REFERENCES	68
	APPENDICES	70

LIST OF TABLES

TABLE	TITLE	PAGE
1.1	Data analysis	19
2.1	Grouping Result	27
3.1	The summary of result	38
4.1	Box Cox Transformation	41
5.1	Chosen Models	43
6.1	Forecasting Error	50

LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Seasonal ARIMA	6
1.2	Flow chart of Methodology	15
2	Time series plot of data.	19
2.1	Time series plot of Monday load data.	20
2.2	Time series plot of Tuesday load data.	20
2.3	Time series plot of Wednesday load data.	21
2.4	Time series plot of Thursday load data.	21
2.5	Time series plot of Friday load data.	22
2.6	Time series plot of Saturday load data.	22
2.7	Time series plot of Sunday load data.	23
3.0	ARIMA methodology	34
4.1	Results for Group1 (Time Series Plot)	44
4.2	Results for Group2 (Time Series Plot)	45
4.3	Results for Group3 (Time Series Plot)	46
4.4	Results for Group4 (Time Series Plot)	47

4.5	Results for Group5 (Time Series Plot)	48
4.6	Results for Group6 (Time Series Plot)	49
5.1	Forecasting Error of Group1	51
5.2	Forecasting Error of Group2	52
5.3	Forecasting Error of Group3	53
5.4	Forecasting Error of Group4	54
5.5	Forecasting Error of Group5	55
5.6	Forecasting Error of Group6	56

LIST OF ABBREVIATIONS

ACF	-	Autocorrelation function
ARIMA (p,d,q)	-	Autoregressive Integrated moving average of order p, d and q, where p order of autoregressive terms and q order of moving average terms and d number of differenced required.
AR	-	Autoregressive components
MA	-	Moving averages components.
B	-	Backward operator
Δ	-	Difference operator
ARIMA (p, d, q)	-	Seasonal autoregressive integrated moving model of normal components
(P, D, Q)S		p,q and seasonal components P and Q and differenced d for normal components and D for seasonal component respectively.
CHI	-	Chi-squared statistic
MW	-	Mega Watt
PACF	-	Partial autocorrelation function
T (t)	-	Trend
S (t)	-	Seasonal
R (t)	-	Irregular / random

p	-	the order of the non seasonal AR process
d	-	the order of the non seasonal MA process
q	-	the number of non seasonal differences
P	-	the number of multiplicative autoregressive coefficient
D	-	the number of seasonal differences
Q	-	the number of multiplicative moving average coefficient
s	-	the seasonal period

EQUATION	TITLE	PAGE
1.1	Analysis of time series	10
2.1	The forecast error	50
3.1	Backshift operator	57
3.2	Backshift operator	57
3.3	Non-seasonal backshift operator	57
3.4	Seasonal backshift operator	57
3.5	General stationarity transformation	57

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Project Planning Gantt chart	70
B	Data (provided by TNB)	71
C	Transformed Data	86
D	Chi-Square Table	99
E	Differencing	100
F	Grouping the data	148
G	Time series plot of data in group	150
I	Trend analysis	153
J	Box cox plot	156
K	Forecasting process using Minitab	162
L	Result of forecasting	165
M	Diagnostic checking	183
N	Forecasting error	189

CHAPTER 1

INTRODUCTION

This project is generally to achieve the forecasting result a week ahead for the Peninsular Malaysia electricity load based on 6 weeks previous data. This chapter also discusses the objective and the scope of the project.

1.1 Problem Statement

The amount of electricity used by consumer need to be arranged wisely to avoid the short or waste of load supplied. The problems that occurred due to lack of supervising of the load supply are the waste of time and the cost for equipment and worker to fix the problem.

The goal of this project is to predict daily load demand of a week ahead based on previous load data. This is to ensure the electricity is always available although there is event held such as ceremony or public holiday at the site. Thus, the variations of usage by the users would not affect the load supplied.

Load Forecasting leads to better quality of supply from utilities with convenient amount of electricity supplied to the users and the daily electricity operation run smoothly.

Time series method is used due to the accurate result since the implementation of time as alignment. The accuracy of forecast values is necessary for economically efficient operation and also for effective control and maintenance.

The method ARIMA is chosen because the model able to provide appropriate result of the project with less error. The model is an advance model in its class because it does not base on exponential smoothing which is inefficient and sometime inappropriate for time series forecasting.

1.2 Objectives

The project is based on a clear view of objectives to be achieved based on the problem statement and scope that are instructed aligning with the project title. The aims of this project are:

a) To analyze the electricity load data of Peninsular Malaysia.

The interest of modeling the time series for real world data can be exposed by examine the behavior of that series. In order to achieve the task of modeling the first step is to evaluate it visually, which is accomplished through various graphical methods. This analysis will be helpful to verify conclusion that we make. One can reach to conclusion by careful visual inspection of the original series. (Cooray, 2008)

b) To provide forecasted value of a week ahead with minimum error.

The specific aimed value of the error for this project is less or equal 1.5 percent. Unfortunately, all forecasting situation involve some degree of uncertainty. (Connell, 1993) We recognize this fact by including an irregular component, which is represent unexplained or unpredictable fluctuation in the data, means that some error in forecasting must be expected.

c) To prove time series (ARIMA) is a good method of forecasting.

The other methods of short term load forecasting are Neural Network and Fuzzy Logic Algorithm. However the data provided is suitable for time series analysis method because the data is not assist with other information such as weather data or temperature data that mostly require by the other methods. The time based data is proven suitable for ARIMA, one of the time series methods instead of other time series method such as ARMA. There is fairly extensive coverage of ARIMA models because they are widely used. (Janacek, 2001)

1.3 Scope

The project is generally focused on Load Forecasting to provide a forecasted future value based on half hourly 7 weeks of history load data in Peninsular Malaysia provided by utility of Malaysia.

The analysis is on short term period that is for a week ahead. This means that the desired result is the forecasted data of week 7.

Time series analysis is chosen due to the time oriented data in the investigation of the project. The specified analysis of time series that is called Auto Regression Integrated Moving Average (ARIMA) model is implemented in the project.

The target of the forecasting error is less than or equal 1.5% to ensure the accuracy of the analysis. The Minitab software is the most important tool in this project to obtain the desired result of analysis.

CHAPTER 2

LITERATURE REVIEW

A time series is a sequence of value $\{x_1, x_2 \dots x_t \dots\}$ observed through time. For most of this project it can be assumed that the observations are made at integer $\{1, 2, \dots, t, \dots\}$. Time series are ubiquitous; they arise in almost all situations where we keep records or make measurements. The interest in them differs across applications, thus in some situations the main aim may be to forecast while in others the structure of the series or its interrelation with other time series may be the main interest. To look at series in any useful way we need to think in terms of a model. This may be a model which reflects some knowledge of the generating mechanism, for example fish population over time, or more usually a model which is capable of generating the behavior observed. (Janacek, 2001)

A time series is an ordered sequence of observations. Although the ordering is usually through time, particularly in terms of some equally spaced time interval, the ordering may also be taken through other dimensions, such as space. The utility of time series analysis can be well documented by producing a partial listing of the diverse fields in which important time series problems may arise. For example many familiar time series arise in the field of economics, where we are continually exposed to daily stock market quotations or to monthly unemployment figures. Social scientists and demographers are interested in following population series such as birth rates or school enrollment. Time series occur in a variety of fields. In engineering, we observe sound, electric signals and voltage. A time series, such as electric signals and voltage, which can be recorded continuously in time, is said to be continuous. (Cooray, 2008)

2.1 Journal review 1

“Customer Short Term Load Forecasting by using ARIMA transfer function model”

(M.Y Cho, T. H. 2009)

Summary:

In this paper the ARIMA model and transfer function model are applied to the short term load forecasting by considering weather-load relationship.

To demonstrate the effectiveness of the proposed method this paper compares result of the transfer function model and the ARIMA model with conventional regression. To improve the accuracy of ARIMA and other method, the temperature effect is considered in the transfer function.

Result (Comparison):

There are 4 types of data by different customers have been used in the process of forecasting. In the end, both ARIMA and transfer function can provide satisfactory load forecasting accuracy. The transfer function model considered the effect of temperature can achieve better accuracy of load forecast than ARIMA model and can provide guidance of short term load forecasting.

2.1 Journal review 2

“Adaptive Load Forecasting of the Hellenic Electric Grid” (S. Sp. PAPPAS, L. E. 2008)

Summary:

In this paper, load forecasting is required to plan for future expansion and also to estimate the grid’s future utilization. The proposed method uses past data to learn and model the normal periodic behavior of the partitioning electric grid. The applied adaptive multi-model partitioning or filtering algorithm is successfully both normal periodic behavior and unusual activity of the electric grid. The performance of the proposed method is also compared to the result that produced by ARIMA model.

Result (Comparison):

Multi Model Partitioning Algorithm (MMPA) is obviously predict the load demand and identifies the anomalies in a successful manner. ARIMA successfully identifies and forecasts the normal periodic load behavior but it is not as capable of forecasting sudden anomalies

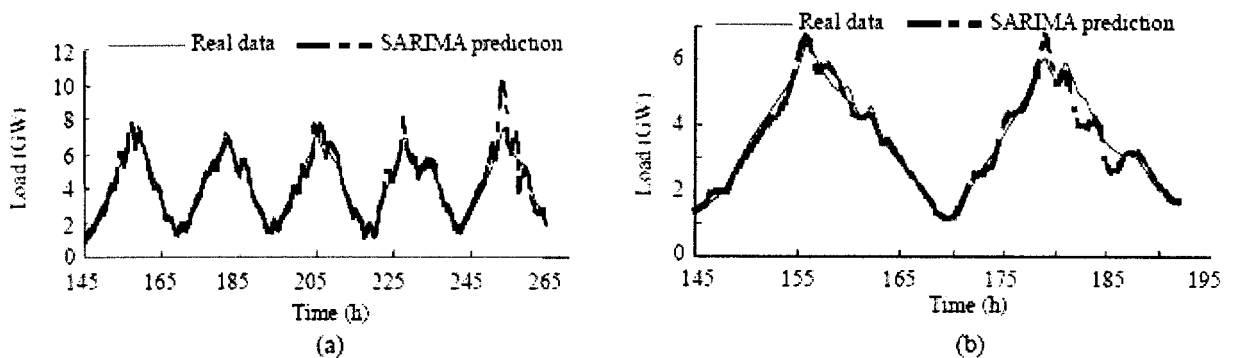


Figure 1.1: Seasonal ARIMA = SARIMA

- a) Weekdays (Monday- Friday) b) Weekend (Saturday and Sunday)