

## Status of Thesis

### A NOVEL FORWARD BACKWARD LINEAR PREDICTION ALGORITHM FOR SHORT TERM POWER LOAD FORECAST

I ZUHAIRI B HJ BAHARUDIN hereby allow my thesis to be placed at the Information Resource Center (IRC) of Universiti Teknologi PETRONAS (UTP) with the following conditions:

1. The thesis becomes the property of UTP
2. The IRC of UTP may make copies of the thesis for academic purposes only
3. This thesis is classified as

Confidential

Non-confidential

If this thesis is confidential, please state the reason(s):

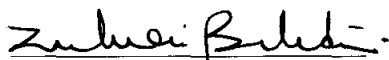
---

The contents of the thesis will remain confidential for \_\_\_\_\_ years.

Remarks on disclosure:

---

Endorsed by



Zuhairi B Hj Baharudin, Author

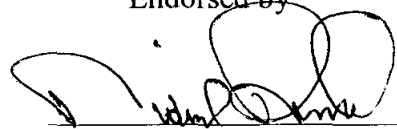
8 Lorong Changkat Mewah 3

Taman Changkat Mewah

31000 Batu Gajah, Perak

MALAYSIA

Date: 20-05-2010



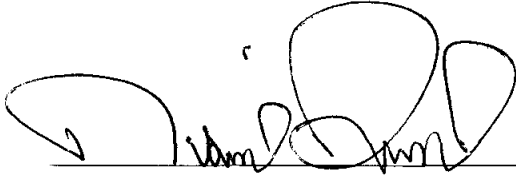
Nidal S. Kamel, Supervisor

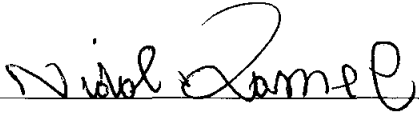
Date: 20.05.10

**Approval by Supervisors**

The undersigned certify that they have read, and recommend to the Postgraduate Studies Programme for acceptance, a thesis entitled “A Novel Forward Backward Linear Prediction Algorithm for Short Term Power Load Forecast” submitted by Zuhairi B Hj Baharudin in fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.).

\_\_\_\_\_  
Date : 20.05.10

Signature :  \_\_\_\_\_

Main Supervisor :  \_\_\_\_\_

Date : 20 MAY 2010 \_\_\_\_\_

Co-Supervisor 1 : \_\_\_\_\_

Co-Supervisor 2 : \_\_\_\_\_

UNIVERSITI TEKNOLOGI PETRONAS

**A NOVEL FORWARD BACKWARD LINEAR PREDICTION ALGORITHM  
FOR SHORT TERM POWER LOAD FORECAST**

By

ZUHAIRI B HJ BAHARUDIN

Bachelor of Engineering Hons. (Electrical Engineering)

UiTM (ITM), Malaysia, 1998

Master of Engineering (Electrical Power)

UniSA, Australia, 2001

A DISSERTATION

Submitted to the Postgraduate Studies Programme

as a Requirement for the

Degree of Doctor of Philosophy (Ph.D.)

(Electrical & Electronics Engineering)

BANDAR SERI ISKANDAR, 31750 TRONOH

PERAK, MALAYSIA

May, 2010

## Declaration

I hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degrees at UTP or other institutions.

Signature: Zuhairi Baharudin.

Name: Zuhairi B. Hj. Baharudin

Date: 20-05-2010.

## **DEDICATION**

This thesis is dedicated to my beloved wife Suriati Binti Sufian and to my lovely children 5-year-old Nur Afiqah and 3-year-old Muhammad Syahmi. It is also dedicated to my mother and all families. Last but not least, I dedicate this thesis to my late father, Allahyarham Haji Baharudin Bin Haji Sidik.

## ABSTRACT

Electrical load forecast is an important part of the power system energy management system. Reliable load forecast technique will help the electric utility to make unit commitment decisions, reduce spinning reserve capacity, and schedule device maintenance plan properly. Thus, besides being a key element in reducing the generation cost, power load forecast is an essential procedure in enhancing the reliability of the power systems. Generally speaking, power systems worldwide are using load forecast as an essential part of off-line network analysis. This is in order to determine the status of the system, and the necessity to implement corrective actions, such as load shedding, power purchases or using peaking units.

Short term load forecast (STLF), in terms of one-hour ahead, 24-hours ahead, and 168-hours ahead is a necessary daily task for power dispatch. Its accuracy will significantly affect the cost of generation and the reliability of the system. The majority of the single variable based techniques are using autoregressive-moving average (ARMA) model to solve the STLF problem.

In this thesis, a new AR algorithm especially designed for long data records as a solution to STLF problem is proposed. The proposed AR-based algorithm divides long data record into short segments and searches for the AR coefficients that simultaneously model the data with the least means squared errors. In order to verify the proposed algorithm as a solution to STLF problem, its performance is compared with other AR-based algorithms, like Burg and the seasonal Box-Jenkins ARIMA (SARIMA). In addition to the parametric algorithms, the comparison is extended towards artificial neural networks (ANN). Three years data power demand record collected by NEMMCO in four Australian states, NSW, QLD, SA, and VIC, between the beginning of 2005 and the end of 2007 are used for the comparison. The results show the potential of the proposed algorithm as a reliable solution to STLF.

## ABSTRAK

Ramalan beban elektrik adalah sebahagian daripada pengurusan tenaga sistem kuasa yang sangat penting. Teknik ramalan beban dipercayai boleh membantu utiliti elektrik untuk membuat keputusan komitmen unit, mengurangkan keupayaan simpanan, dan menjadualkan pelan penyelenggaraan peranti dengan betul. Maka, selain menjadi satu elemen penting dalam mengurangkan kos penjanaan, mendayai ramalan beban adalah satu prosedur penting dalam mempertingkatkan kebolehpercayaan sistem-sistem tenaga. Umumnya, sistem-sistem tenaga dunia menggunakan ramalan beban ini sebahagian daripada analisis rangkaian luar talian. Ini merupakan suatu tahap dimana status system dapat ditentukan, dan keperluan untuk melaksanakan tindakan pembetulan, seperti beban gugur, kuasa membeli atau menggunakan unit-unit puncak.

Ramalan beban tempoh singkat (STLF), dalam soal satu jam di hadapan, 24-jam di hadapan, dan 168-jam di hadapan adalah satu tugas seharian yang perlu untuk penghantaran kuasa. Ketepatannya akan nyata sekali menjejaskan kos generasi dan kebolehpercayaan sistem. Majoriti pembolehubah tunggal berpengkalan oleh teknik-teknik dengan menggunakan model purata bergerak autoregresif (ARMA) untuk menyelesaikan masalah STLF.

Dalam tesis ini algoritma AR yang baharu dicadangkan sebagai satu penyelesaian bagi STLF terutama untuk jangka masa rekod data yang panjang. Algoritma berdasarkan AR mencadangkan membahagikan rekod data yang panjang kepada segmen-segmen yang pendek, seterusnya mencari pekali-pekali AR yang serentak serta mengaplikasikan data dengan cara yang paling kurang kesilapan-kesilapan. Untuk mengesahkan algoritma yang dicadangkan sebagai satu penyelesaian bagi masalah STLF, prestasinya dibandingkan dengan algoritma yang berdasarkan AR yang lain, seperti Burg dan Box- Jenkins ARIMA yang bermusim (SARIMA). Tambahan kepada algoritma berparameter, perbandingan diperluaskan lagi kepada jaringan saraf tiruan (ANN). Tiga tahun rekod permintaan kuasa data dikumpul oleh

NEMMCO dalam empat negeri-negeri di Australia, iaitu NSW, QLD, SA, dan VIC, antara permulaan tahun 2005 dan penghujung tahun 2007 digunakan untuk tujuan perbandingan. Keputusan menunjukkan potensi algoritma yang dicadangkan bersesuaian sebagai satu penyelesaian yang boleh dipercayai untuk menyelesaikan masalah STLF.



## AFFIRMATION

### Conference Proceedings

1. Z. Baharudin, N. S. Kamel (August 2007). One Week Ahead Short Term Load Forecasting. Proceeding of the 7<sup>th</sup> IASTED International Conference on Power and Energy Systems, 29-31 Aug 2007, Palma, Spain.
2. Z. Baharudin, N. S. Kamel (November 2007). Short Term Load Forecast Using Burg Autoregressive Technique. Proceeding of the 2<sup>nd</sup> ICIAS International Conference, 25-28 Nov 2007, Kuala Lumpur, Malaysia.
3. Z. Baharudin, N. S. Kamel (March 2008). AR Burg in Short Term Load Forecast. Proceeding of the 1<sup>st</sup> National Postgraduate Symposium, 31 Mar 2008 Universiti Teknologi PETRONAS, Malaysia.
4. Z. Baharudin, N. S. Kamel (April 2008). AR Modified Covariance in One Week Ahead Short Term Load Forecast. Proceeding of the 9<sup>th</sup> IASTED International Conference on Power and Energy Systems, 2-4 Apr 2008, Langkawi, Malaysia.
5. Z. Baharudin, N. S. Kamel. ARMA versus AR in Short Term Load Forecast. Proceeding of the 2<sup>nd</sup> PEOCO International Power Engineering & Optimization Conference, 4-5 June 2008 Shah Alam, Malaysia.
6. Z. Baharudin, N. S. Kamel. Autoregressive Models in Short Term Load Forecast: A Comparison of AR and ARMA. Proceeding of the 28<sup>th</sup> International Symposium on Forecasting, 22-25 June 2008, Nice, France.
7. Z. Baharudin, N. S. Kamel. Auto-regressive in Short Term Load Forecast. Proceeding of the PECON International Conference of Power Energy, 1-3 Dec 2008, Johor Bharu, Malaysia.
8. Z. Baharudin, N. S. Kamel, MFBLP Method Forecast for Regional Load Demand System, Proceeding of the 3rd PEOCO International Power Engineering & Optimization Conference, June 2009 Shah Alam, Malaysia.

### **Published Journal Paper**

1. Z. Baharudin, N. S. Kamel, ARMA versus AR in Short Term Load Forecast. International Journal of Power Energy and Artificial Intelligence IJPEAI, Aug 2008.
2. Z. Baharudin, N. S. Kamel, MFBLP Method Forecast for Regional Load Demand System, International Journal of Engineering, Computer Science Journal, Vol. 3, Issue 3, pp 280-292, Oct 2009.

### **Submitted Journal Paper**

1. Z. Baharudin, N. S. Kamel. Auto-regressive Modeling of New South Wales Load Demand System. International Journal of Power System, Acta-Press. Submitted on March 2009. Under revision.
2. Z. Baharudin, N. S. Kamel. MFBLP Method Forecast for Regional Load Demand System. International Journal of Forecasting (IJF). Submitted on August 2009. Under revision.
3. Z. Baharudin, N. S. Kamel. A New Method of MFBLP Forecast for Regional Load Demand System. International Journal of Power System IEEE. Submitted on September 2009. Under revision.

## **ACKNOWLEDGEMENTS**

In the name of Allah, the Beneficent and the most Merciful, it is with the deepest sense of gratitude to Allah, and Peace and Blessing to Prophet Muhammad (PBUH) who has given the strength and ability to complete this research work and the thesis as it is today.

First and foremost, I would like to express my deepest thankful to my supervisor, Associate Professor Dr. Nidal S. Kamel for his excellent way of supervising the work, abundance of guidance and assistance throughout this research. Many thanks to the Head of Department Electrical and Electronics, Dr. Nor Hisham Bin Hamid and Director of Postgraduate Studies, Associate Professor Dr. Mohd Noh Bin Karsiti for their consistent moral support and positive ideas for me to complete this research.

My special gratitude goes to my sponsor, Universiti Teknologi PETRONAS (UTP), Malaysia for giving me the opportunity and the scholarship for my studies. I would also want to thank the Human Resources Management of UTP in funding the local and abroad conference participation.

Warm appreciation to all UTP staffs who have either directly or indirectly assist in completing the research. Not to forget all my colleagues in Block 22 and 23, especially the office mates in level 3, Block 22.

## TABLE OF CONTENTS

ABSTRACT .....	vii
ABSTRAK .....	viii
AFFIRMATION.....	x
ACKNOWLEDGEMENTS .....	xii
LIST OF FIGURES.....	xvi
LIST OF TABLES .....	xix
LIST OF ABBREVIATIONS .....	xxii
CHAPTER 1 INTRODUCTION.....	1
1.1 The Problem.....	1
1.2 The STLF Techniques .....	2
1.3 Research Objectives.....	3
1.4 Research Contributions.....	4
1.5 Organization of Thesis.....	5
CHAPTER 2 LITERATURE REVIEW: PRINCIPLES OF SHORT TERM LOAD FORECASTING.....	6
2.1 Background.....	6
2.2 Recent Methods of STLF.....	8
2.2.1 Time Series Model .....	8
2.2.2 Regression Model.....	9
2.2.3 Expert System .....	9
2.2.4 Neural Network.....	9
2.3 Summary.....	10
CHAPTER 3 STOCHASTIC MODELS FOR POWER LOAD FORECAST .....	11
3.1 Autoregressive Moving Average (ARMA) Models .....	11
3.2 The Autoregressive (AR) Models.....	15
3.3 Lattice Methods for AR Modeling .....	16
3.3.1 Burg's Method .....	18
3.3.2 The Modified Covariance Method .....	19
3.4 Box-Jenkins ARIMA Method.....	21

3.5 Conclusion .....	22
CHAPTER 4 A MODIFIED FORWARD BACKWARD LINEAR PREDICTION (MFBLP) ALGORITHM .....	23
4.1 Introduction.....	23
4.2 Modified Forward-Backward Linear Prediction (MFBLP) Algorithm.....	24
4.3 The Least-Squares Solution.....	30
4.4 Summary of the MFBLP Algorithm.....	32
4.5 The Optimum Length of the Data Segments and Predictor Order .	32
4.6 Conclusion .....	37
CHAPTER 5 POWER LOAD TIME SERIES: CHARACTERISTICS AND PROCESSING.....	39
5.1 Introduction.....	39
5.2 Analyzing the Seasonal Behavior of the Power Load Time Series	39
5.2.1 New South Wales (NSW) .....	40
5.2.2 Queensland (QLD) .....	42
5.2.3 South Australia (SA).....	44
5.2.4 Victoria (VIC) .....	46
5.3 Demand Curves Affected by the Anomalous Day .....	48
5.4 Data Processing for Redundancy Reduction .....	51
5.4.1 Box-Jenkins ARIMA Algorithm.....	51
5.4.2 Box-Jenkins SARIMA processes.....	52
5.5 Summary.....	59
CHAPTER 6 RESULTS AND DISCUSSION .....	60
6.1 Metrics of Performance .....	61
6.2 Forecast Horizons of Forecast .....	63
6.3 The Performance of the MFBLP Algorithm with Raw and Differenced Data.....	64
6.4 Conclusion.....	99
CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS.....	101
7.1 Conclusions.....	101
7.2 Summary of the Main Contribution.....	102
7.3 Recommendations for Future Work .....	103

REFERENCES .....	105
Appendix A DATA FILES .....	117
Appendix B AUTOCORRELATION FUNCTION (acf) PLOTS.....	118
Appendix C ARTIFICIAL NEURAL NETWORK (ann) .....	124
Appendix D MATLAB FILES .....	125
Appendix E .....	130

## LIST OF FIGURES

Figure 3.1 Modeling a random process $x(n)$ as the response of a linear shift-invariant filter to unit variance white noise.....	12
Figure 4.1 MAPE values of the MFBLP algorithm as a function of predictor order 36	
Figure 5.1 Two years of hourly power load demand of NSW .....	40
Figure 5.2 Three weeks of hourly power load demand of NSW.....	41
Figure 5.3 Seven days of hourly power load demand of NSW .....	41
Figure 5.4 Two years of hourly power load demand of QLD .....	42
Figure 5.5 Three weeks of hourly power load demand of QLD .....	43
Figure 5.6 Seven days of hourly power load demand of QLD.....	43
Figure 5.7 Two years of hourly power load demand of SA .....	44
Figure 5.8 Three weeks of hourly power load demand of SA.....	45
Figure 5.9 Seven days of hourly power load demand of SA.....	45
Figure 5.10 Two years of hourly power load demand of VIC .....	46
Figure 5.11 Three weeks of hourly power load demand of VIC.....	47
Figure 5.12 Seven days of hourly power load demand of VIC.....	47
Figure 5.13 The load demand over the days of the week of year 2006 in NSW, initiated by the New Year day.....	48
Figure 5.14 The initiated load demand by Anzac day in QLD over one week in April 2005 .....	49
Figure 5.15 The initiated load demand by Australia day in SA over one week in January 2005 .....	50
Figure 5.16 The VIC load demand curves for one week during the Queen’s Birthday in June 2005 .....	50

Figure 5.17 The ACF of NSW time series: a) ACF of raw data, b) ACF of processed data.....	54
Figure 5.18 The ACF of QLD time series: a) ACF of raw data, b) ACF of processed data.....	55
Figure 5.19 The ACF of SA time series: a) ACF of raw data, b) ACF of processed data.....	56
Figure 5.20 The ACF of VIC time series: a) ACF of raw data, b) ACF of processed data.....	57
Figure 6.1 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for NSW. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	66
Figure 6.2 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for QLD. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	67
Figure 6.3 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for SA. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	69
Figure 6.4 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for VIC. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	70
Figure 6.5 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for NSW. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ ....	73
Figure 6.6 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for QLD. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	74
Figure 6.7 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for SA. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	75
Figure 6.8 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for VIC. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	76
Figure 6.9 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for NSW. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d =$ $168$ .....	78



Figure 6.10 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for QLD. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	79
Figure 6.11 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for SA. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	80
Figure 6.12 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for VIC. a) Raw data, b) Differenced data with $d = 1$ , $d = 24$ , and $d = 168$ .....	81

## LIST OF TABLES

Table 4.1 Q and N for the NSW two-year hourly load demand data .....	33
Table 4.2 The MAPE values of the MFBLP as a function of L in 1-hour ahead forecast .....	34
Table 4.3 The MAPE values of the MFBLP as a function of L in 24 hours-ahead forecast .....	34
Table 4.4 The MAPE values of the MFBLP as a function of L in 168 hours-ahead forecast .....	35
Table 4.5 The computational time of the MFBLP algorithm with $L = 0.25N$ and $N = 17520$ samples as a function of the different segmentations .....	37
Table 6.1 Data specification and the MFBLP parameters.....	63
Table 6.2 The MAPE values of the MFBLP in 1-hour forecast scheme.....	71
Table 6.3 The MAPE values of the MFBLP in 24-hours forecast scheme .....	77
Table 6.4 The MAPE values of the MFBLP in 168-hours forecast scheme .....	82
Table 6.5 Estimated parameters coefficients for Box-Jenkins SARIMA model.....	83
Table 6.6 The performance of the different techniques with the raw data of NSW..	85
Table 6.7 The performance of the different techniques with the raw data of QLD ..	85
Table 6.8 The performance of the different techniques with the raw data of SA .....	86
Table 6.9 The performance of the different techniques with the raw data of VIC....	86
Table 6.10 The performance of the different techniques with the differenced (differencing) data of NSW.....	87
Table 6.11 The performance of the different techniques with the differenced (differencing) data of QLD .....	88
Table 6.12 The performance of the different techniques with the differenced (differencing) data of SA .....	88

Table 6.13 The performance of the different techniques with the differenced (differencing) data of VIC.....	89
Table 6.14 The data for the seasons in Australia.....	90
Table 6.15 The performance of the different techniques with the summer data of NSW .....	91
Table 6.16 The performance of the different techniques with the fall data of NSW.....	91
Table 6.17 The performance of the different techniques with the winter data of NSW.....	92
Table 6.18 The performance of the different techniques with the spring data of NSW.....	92
Table 6.19 The performance of the different techniques with the summer data of QLD.....	93
Table 6.20 The performance of the different techniques with the fall data of QLD .....	93
Table 6.21 The performance of the different techniques with the winter data of QLD .....	94
Table 6.22 The performance of the different techniques with the spring data of QLD.....	94
Table 6.23 The performance of the different techniques with the summer data of SA.....	95
Table 6.24 The performance of the different techniques with the fall data of SA .....	95
Table 6.25 The performance of the different techniques with the winter data of SA .....	96
Table 6.26 The performance of the different techniques with the spring data of SA.....	96

Table 6.27	The performance of the different techniques with the summer data of VIC .....	97
Table 6.28	The performance of the different techniques with the fall data of VIC.....	97
Table 6.29	The performance of the different techniques with the winter data of VIC.....	98
Table 6.30	The performance of the different techniques with the spring data of VIC.....	98

## LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ACS	Autocorrelation Sequence
AI	Artificial Intelligent
ANN	Artificial Neural Network
AR	Auto-Regressive
ARIMA	Auto-Regressive Integrated Moving Average
ARMA	Auto-Regressive Moving Average
ES	Exponential Smoothing
FL	Fuzzy Logic
HW	Holt-Winters
LTLF	Long Term Load Forecast
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MFBLP	Modified Forward Backward Linear Predictor
MM	Multivariate Models
MTLF	Medium Term Load Forecast
NEM	National Electricity Market
NEMMCO	National Electricity Market Management Company Ltd
NSW	New South Wales
PAR	Pure Auto-Regressive
PCA	Principal Component Analysis
QLD	Queensland
SA	South Australia

SARIMA	Seasonal Auto-Regressive Integrated Moving Average
STLF	Short Term Load Forecast
TS	Time-Series
UM	Univariate Models
VIC	Victoria

# **CHAPTER 1**

## **INTRODUCTION**

Electrical load forecast is an important part of the power system energy management system. Reliable load forecast technique will help the electric utility to make unit commitment decisions, reduce spinning reserve capacity, and schedule device maintenance plan properly. Thus, besides being a key element in reducing the generation cost, power load forecast is an essential procedure in enhancing the reliability of the power systems. Generally speaking, power systems worldwide are using the load forecast as an essential part of off-line network analysis. This is in order to determine the status of the system, and the necessity to implement corrective actions, such as load shedding, power purchases or using peaking units.

### **1.1 The Problem**

Short term load forecast (STLF), in terms of one-hour ahead, 24-hours ahead, and 168-hours ahead is a necessary daily task for power dispatch. Its accuracy will significantly affect the cost of generation and the reliability of the system. Under forecast of STLF leads to insufficient reserve capacity preparation, and consequently increases the operating cost by using expensive peaking units. On the other hand, over forecast of STLF leads to unnecessarily large reserve capacity, which also means high operating cost. It is estimated in the British power system that every 1% increase in the forecasting error will lead to an increase in the operating costs by about 10 million pounds yearly [1, 2].

In addition to the STLF, which is the major issue being addressed in, this thesis, the medium term load forecast (MTLF) is necessary for scheduling of fuel supply and maintenance operation [3, 4]; and the long-term load forecast (LTLF) is

important for system planning in order to meet the expected long-term growth in demand [5-8].

## 1.2 The STLF Techniques

Many techniques have been proposed for STLF solution over the last three decades, with different degrees of accuracy. The majority of these techniques are based on the time series modeling [9-15], exponential smoothing [16-21], and artificial neural networks [ANN] [22-36]. Generally speaking, the implemented STLF techniques are either based on a single variable, which is the historical record of the power load, or a combination of the weather variable with the power load, in what is called the multivariate approach. The majority of the single variable based techniques are using autoregressive moving average (ARMA) model [37-41] and Kalman filter [41-44] to solve the STLF problem. On the other hand, the multivariate STLF techniques are mainly based on ANN, genetic algorithms, and Fuzzy logic [45-51] or combinations between them, like ANN-Fuzzy [31, 52-61] or ANN-GA [58, 58, 62, 63, 63, 64]. Combinations between ANN, GA, Fuzzy logic and ARMA are also used to solve the STLF problem [65-70].

Owing to the importance of ARMA for the STLF, a large number of estimation methods for ARMA model parameters have been proposed over the last 40 years. ARMA model has more degrees of freedom than the autoregressive, so greater latitude in its ability to generate diverse time-series shapes is therefore expected of its estimators. Unfortunately, this is not always the case, because of the nonlinear nature requirement of the algorithms that must simultaneously estimate the moving average and autoregressive parameters of the ARMA model.

Indeed, all existing solutions to this problem appear to suffer from one or more drawbacks, as explained briefly in the followings:

- Several methods may end up in a hard failure mode. This means that the identification algorithm may return an invalid model or the algorithm cannot be carried out to completion because as a result of a step during its execution,



a parameter set outside the class of permissible ones arises for which no provisions have been adopted. This may happen for moment fitting procedures, as well as to methods that first estimate the AR parameters and then the MA.

- Maximum-likelihood methods that depend on search over the parameter space involve significant computations and are not guaranteed to converge, or they may converge to the wrong solution.
- Finally, some methods may be inaccurate (e. g., significantly biased) in finite samples. This is the case with Durbin's two stage least-squares method [71-73], and with the approximate subspace methods that enforce positivity of the estimated MA spectrum based on [74].

This type of drawback also affects the methods based on higher order statistics [75, 76], which usually need large data samples to achieve satisfactory STLF accuracy.

### **1.3 Research Objectives**

There are three main objectives in this research. The objectives are as follows:

- To provide adequate analysis to describe the data using graphical methods. Before trying to implement and forecast a given time-series, it is desirable to have a preliminary observation at the data. The aim is to determine the optimum parameters of the designed model.
- To design and develop the methodology and algorithm for the model that can represent the data generating process. The research work concentrates on univariate model (UM), which the historical data are the only input to the model algorithm.
- To forecast the future values of the given data-series. The forecasting process implements the obtained filter coefficients generated by the estimation process. The forecast experiment is done by offline computation and the validation with the actual out-of-sample data is measured by mean absolute percentage error (MAPE).

## 1.4 Research Contributions

Though the all-pole models have less degree of freedom than ARMA, they exhibit major advantages. First of all, in some applications the physical process by which a signal is generated will result in an all-pole (AR) signal. However, even in those applications for which it may not be possible to justify all-pole models, one often finds an all-pole model being used. One reason for this is that all-pole models have been found to provide a sufficiently accurate representation for many different types of signals in many different applications. Another reason for the popularity is the special structure, which leads to fast and efficient algorithms for finding the all-pole parameters especially in case of short data records. However, in long data records, as is the case with most power load time series, the required number of all-pole coefficients dramatically increases making most of all-pole algorithm less efficient in estimating their values and accordingly less accurate in modeling the power load data for proper solution of STLF.

In this thesis, a new AR algorithm particularly designed for long data records as a solution to STLF problem is proposed. This algorithm is inspired by a multiple array snapshots for direction of arrival estimation technique [73]. The proposed AR-based algorithm divides long data record into short segments then searches for the AR coefficients that simultaneously model the data with the least means squared errors.

In order to verify the proposed algorithm as a solution to STLF problem, its performance is compared other with other AR-based algorithms, like Burg [77, 78] and the Modified Covariance (MCOV) [73, 79], as well as with Durbin's as ARMA-based algorithm [73, 74]. The proposed algorithm is also compared with recent model based methods suggested by Box-Jenkins [80-82], and artificial neural networks (ANN) [81, 83, 84].

Three years data load demand record collected by NEMMCO in four Australian

states, NSW, QLD, SA, and VIC, between the beginning of 2005 and the end of 2007, are used for the comparison. The algorithms are run with raw data and with differenced data. The results show the potential of the proposed algorithm as a reliable solution to STLF.

## **1.5 Organization of Thesis**

To make the representation clear, this thesis is organized as follows:

Chapter two reviews the stochastic models of time-series and outlines the backgrounds of AR-based algorithms such as Burg and the modified covariance, as well as ARMA-based algorithms such as Box-Jenkins and Durbin.

Chapter three explains thoroughly the proposed Modified Forward Backward Linear Prediction (MFBLP) algorithm, and finds the optimum number of segments ( $Q$ ) as well as the predictor order ( $L$ ).

In Chapter four, the major characteristics of the four Australian power load time series (NSW, QLD, SA, and VIC) are addressed. The daily, weekly, and seasonal patterns are investigated, and the differencing scheme as a way to mitigate the patterns is described.

In Chapter five, the performance of MFBLP is compared with AR Burg, Box-Jenkins SARIMA and Artificial Neural Network (ANN) in terms of hourly, daily and one week-ahead forecasts. The comparisons are carried out with raw data as well as with differenced data. Sorted data in seasonal form are also considered for the performance of the different algorithms.

Chapter six concludes the thesis and highlights the major contributions, in addition to suggestions for future work.

## **CHAPTER 2**

# **LITERATURE REVIEW: PRINCIPLES OF SHORT TERM LOAD FORECASTING**

In this chapter a summary of the recently proposed techniques for power load forecast is presented. The main difference between the used techniques for power load forecast is the type of the used variables. The variables for power load forecast are historical load, metrological conditions, seasonal effects (daily and weekly cycles), special events (holidays, weekends, etc) and other random variables. Though the difference in nature and the type of the used variables, the ultimate goal of the proposed algorithms is to produce better estimation of the power load.

### **2.1 Background**

Power load forecast for the STLF is an essential process in electrical power system operation and planning. Many economic implications of power utility, such as economic scheduling of generating capacity, scheduling of fuel purchases, security analysis, planning of energy transactions, short term maintenance scheduling, and dispatching of generating units are mainly operated based on accurate load forecasting. Hence, many approaches and methods have been suggested and applied to improve the forecast accuracy. Higher forecast accuracy keeps the utilities operation production cost at minimum level and maintains the energy system supply reliability [37, 85, 86].

One main aspect that plays an important role in determining the STLF models is the type of variables used. The historical variables for any STLF models can be

classified into several types namely; historical load, meteorological condition, seasonal effects (daily and weekly cycles), special events and other random variables [42, 87, 88].

The STLF procedures and relationships of the variables used have been discussed [1] and can be divided into three broad groups:

- Load models using no weather data
- Weather load models
- Composite load forecasting models.

The methods which use no weather information, essentially extrapolate the past load behavior [89]. Previously a wide variety of models have been developed which emphasize the probabilistic features of the STLF. The stochastic process with a Kalman filter algorithm [42, 43], and autoregressive moving average (ARMA) models [37, 38, 90] had shown that the STLF model is developed by using only historical load data. These stochastic models are based on determining the linear and nonlinear filter which could have generated the results of load demand. The methods show a relatively good in forecast results, thus it become one of the practical approach since then. Others had discussed the methods with regression techniques [49, 91-94] and an enhanced of time series approaches [15, 95, 96].

Load models which use weather variables are largely reported in the literature. The methods can be classified into several techniques such as expert system based algorithm [54, 97-100], rule-based algorithm [7, 60, 101], regression-based approach [92, 93], adaptive STLF by using Kalman filter and exponentially weighted recursive-least-squares [42, 102, 103], and priority vector based technique [104, 105]. These previously discussed techniques utilize weather variables in developing the STLF models with encouraging results. The advent of fast computational process in the 90s had guided many researchers to broaden up the investigation of STLF. The most accepted approach in the study is Artificial Intelligence (AI). Large number of literatures has been reported of using this approach to STLF problem. By

utilizing the historical and weather data, many of the researchers used Artificial Neural Network (ANN) to develop their STLF models [26-28, 32, 33, 83, 84, 106]. Other than ANN approach, Fuzzy Logic (FL) [45, 50, 107-109], Genetic Algorithm (GA) [62-64], combined approach of ANN-Fuzzy [58-60, 110, 111], ANN-ARIMA [112], ANN-GA [58, 62, 63] and Regression-Fuzzy [47, 49, 91] are also well accepted by the previous researchers. The mentioned approaches had shown an improvised results and the results been compared to others previous approaches. However some aspects of future works can still be considered to implement the develop models in STLF.

So far, there is no single model or algorithm that is superior for all utilities [1, 79, 81, 82, 113-115]. The reason is that utility service areas vary in deferent mixtures of industrial, commercial, and residential customers. They also vary in geographic, climatologic, economic, and social characteristics. Selecting the most suitable algorithm by a utility can be done by testing the algorithms on real data.

## **2.2 Recent Methods of STLF**

Due to the importance of load forecast to the utilities, recently numerous methods for power load forecast have been suggested. These methods can generally be classified into four categories, i.e. time series, linear and non-linear regression, expert system and neural network. The earlier suggested methods mainly try to develop the algorithm or model based on the past historical load demand data. Such models can then be used to estimate the future load prediction.

### ***2.2.1 Time Series Model***

One of the most practical and commonly used models for the load forecast is the time series model. Several related studies have been reported earlier by Box and Jenkins [80], then the finding continuously accepted by the researchers and shown that the results have improved significantly [15, 116-118]. The most fundamental time series models are the autoregressive (AR) model and the autoregressive moving

average (ARMA) model. In the AR model, the desired forecast load value is generally expressed as a linear combination of  $n$  previous load values and a noise term. Meanwhile, in the ARMA model, the desired forecast load value is generally expressed as a linear combination of  $n$  previous load values and  $m$  previous noise terms. However, the weather and other variables such as economic index, population index value, and etc, are not considered as a model inputs. Therefore, these methods could provide the practical model since other variables such as weather variables are quite inaccurate.

### ***2.2.2 Regression Model***

The building procedure of a regression model is to express the forecast load as a function of its influencing factors, such as previous load values, weather data and so on. Generally, the building procedure and the form of nonlinear regression models are much complex than linear ones [94, 119, 120]. In fact, heavy statistical analysis for model identification and parameter estimation is needed in the building process [119, 121].

### ***2.2.3 Expert System***

Generally for the expert system method, the system operator is treated as an expert person in the load forecast activities. The expert person must be very experienced in this particular area. A forecast model is developed for the prediction based on the experience of this expert person. There is no clear or direct algorithm form for the forecast model developed by this expert person. The related studies were suggested and proposed by [97, 98, 102, 122, 123].

### ***2.2.4 Neural Network***

Recently, neural network (NN) has been widely accepted in the area of the load forecast. The method of NN can develop a forecast model through a training process to a historic data. Such a training process enables the neural system to capture the complex and nonlinear load weather relationships that are not easily analyzed by

using conventional methods. The trained NN model can then be used to perform the task of load forecast. Based on the structure of the NN and the learning algorithm, various NN models for load forecast are suggested and proposed. For example, for the feed-forward NN networks with the back-propagation learning algorithm method can be found from [36, 62, 68], recurrent neural networks [32, 33, 124]. Several models developed by using hybrid method such as [68, 123, 125-127].

### **2.3 Summary**

Since in power systems the next days' power generation must be scheduled every day, day-ahead short-term load forecasting (STLF) is a necessary daily task for utilities system planning. Its accuracy affects the production operational cost and reliability of the system. Under forecast of STLF leads to insufficient reserve capacity preparation and, in turn, increases the operating cost by using expensive peaking units. On the other hand, over forecast of STLF leads to the unnecessarily large reserve capacity, which is also related to higher production operational cost. In spite of the numerous literatures on STLF, the research work in this area is still a challenge to the electrical engineering researchers because of its high complexity. The estimation of the future load demand with the help of historical data is still challenging until now. Especially, when solving for the future load demand for the holidays, days with extreme weather and other anomalous days. With the development of a new method of modified backward forward linear prediction (MFBLP) algorithms; it is expected that the STLF issues could be tackled and potentially will improve the forecast error results.



## CHAPTER 3

### STOCHASTIC MODELS FOR POWER LOAD FORECAST

In some applications, it is necessary to develop models for random processes. Examples include signals whose time evolution is affected or driven by random or unknown factors, as is the case for power load forecasting. Models for random processes differ from those for deterministic signals in the characteristics of the signal that is used as input to the system. Whereas for deterministic signals the input signal is usually a unit sample, for random process the input signal must be a random process. Typically, this input will be taken to be unit variance white noise.

#### 3.1 Autoregressive Moving Average (ARMA) Models

A time-series model that approximates many discrete-time stochastic processes encountered in practice is presented by the filter linear difference equation of complex coefficients, and it is given by

$$\begin{aligned}x(n) &= -\sum_{k=1}^p a_p(k)x(n-k) + \sum_{k=0}^q b_q(k)u(n-k) \\ &= \sum_{k=0}^{\infty} h(k)u(n-k)\end{aligned}\tag{3.1}$$

In which  $x(n)$  is the output sequence of a causal filter ( $h(k) = 0$  for  $k < 0$ ) that models the observed data and  $u(n)$  is an input driving white noise sequence. Eq. (3.1) determines the auto-regressive moving average (ARMA) model for the time-series  $x(n)$ .

The  $a_p(k)$  parameters form the autoregressive portion of the ARMA model. The  $b_q(k)$  parameters form the moving average portion of the ARMA model. Thus, a wide sense stationary ARMA( $p,q$ ) process may be generated by filtering the unit variance white noise  $u(n)$  with a causal linear shift-invariant filter having  $p$  poles and  $q$  zeros, and it is described as

$$H(z) = \frac{B_q(z)}{A_p(z)} = \frac{\sum_{k=0}^q b_q(k)z^{-k}}{1 + \sum_{k=1}^p a_p(k)z^{-k}} \quad (3.2)$$

Therefore, a random process  $x(n)$  may be modeled as an ARMA( $p,q$ ) process using the model shown in Figure 3.1, where  $u(n)$  is unit variance white noise.

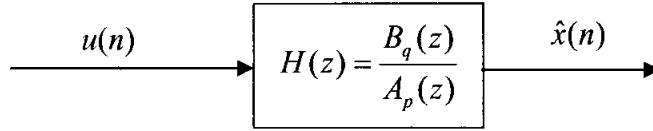


Figure 3.1 Modeling a random process  $x(n)$  as the response of a linear shift-invariant filter to unit variance white noise.

If Eq. (3.1) is multiplied by  $x^*(n-m)$  and the expectation taken, the result is known as Yule-Walker equation [74], given as,

$$r_x(k) + \sum_{l=1}^p a_p(l)r_x(k-l) = c_q(k) \quad (3.3)$$

Where  $r_x(k)$  is the autocorrelation sequence (ACS) of  $x(n)$  and the sequence  $c_q(k)$  is the convolution of  $b_q(k)$  and  $h^*(-k)$

$$c_q(k) = b_q(k) * h^*(-k) = \sum_{l=0}^{q-k} b_q(l+k)h^*(l) \quad (3.4)$$

Since  $h(n)$  is assumed to be causal, then  $c_q(k) = 0$  for  $k > q$  and the Yule-Walker equations for  $k > q$  are a function only of the coefficients  $a_p(k)$ ,

$$r_x(k) + \sum_{l=1}^p a_p(l)r_x(k-l) = 0 \quad ; \quad k > q \quad (3.5)$$

The auto-regressive parameters of an ARMA model are related by a set of linear equations to the autocorrelation sequence. Expressing (3.5) in matrix form for  $k = q+1, q+2, \dots, q+p$ , hence

$$\begin{bmatrix} r_x(q) & r_x(q+1) & \cdots & r_x(q-p+1) \\ r_x(q+1) & r_x(q) & \cdots & r_x(q-p+2) \\ \vdots & \vdots & \ddots & \vdots \\ r_x(q+p-1) & r_x(q+p-2) & \cdots & r_x(q) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} r_x(q+1) \\ r_x(q+2) \\ \vdots \\ r_x(q+p) \end{bmatrix} \quad (3.6)$$

Which is a set of  $p$  linear equations in the  $p$  unknowns,  $a_p(k)$ . These equations are referred to as the Modified Yule-Walker's equations.

Once the coefficients  $a_p(k)$  have been determined, the next step is to find the MA coefficients,  $b_q(k)$ . An MA( $q$ ) process may be generated by filtering unit variance white noise  $u(n)$  with an FIR filter [112, 117, 128] of order  $q$  as follows:

$$x(n) = \sum_{k=0}^q b_q(k)u(n-k) \quad (3.7)$$

The Yule-Walker equations relating the ACS to the filter coefficients  $b_q(k)$  are

$$r_x(k) = b_q(k) * b_q^*(-k) = \sum_{l=0}^{q-|k|} b_q(l+|k|)b_q^*(l) \quad (3.8)$$

Note that, unlike the case for an auto-regressive process, these equations are nonlinear in model coefficients  $b_q(k)$ . Therefore, even if the ACS were known

exactly, finding the coefficients  $b_q(k)$  may be difficult. Instead of attempting to solve the Yule-Walker equations directly, another approach is to perform a spectral factorization of the power spectrum  $P_x(z)$ . Specifically, since the autocorrelation of an MA( $q$ ) process is equal to zero for  $|k| > q$ , the power spectrum is a polynomial of the form

$$P_x(z) = \sum_{k=-q}^q r_x(k)z^{-k} = \sigma_0^2 Q(z)Q^*(1/z^*) \quad (3.9)$$

Where  $Q(z)$  is a minimum phase polynomial of a degree  $q$ .  $\sigma_0 = b_q(0)$  and  $Q(z)$  is the minimum phase version of  $B_q(z)$  that is formed by replacing each zero of  $b_q(z)$  that lies outside the unit circle with one that lies inside the unit circle at the conjugate reciprocal location [74]. Thus, given the autocorrelation sequence of MA( $q$ ) process, the model for  $x(n)$  could be determined.

From the autocorrelation sequence  $r_x(k)$  where for the polynomial  $P_x(z)$  and factor it into a product of a minimum phase polynomial,  $Q(z)$ , and a maximum phase polynomial  $Q^*(1/z^*)$ . The process  $x(n)$  may then be modeled as the input of the minimum phase FIR filter and the formulation is given by

$$H(z) = \sigma_0 Q(z) = \sigma_0 \sum_{k=0}^q q(k)z^{-k} \quad (3.10)$$

As an alternative to spectral factorization, a moving average model for a process  $x(n)$  may also be developed using Durbin's method [74]. This approach begins by finding a high-order all pole models  $A_p(z)$  for the moving average process. Then by considering the coefficients of the all-pole model  $a_p(k)$  to be a new "data set", the coefficients of a  $q$ th-order moving average model are determined by finding a  $q$ th-order all-pole model for the sequence  $a_p(k)$ . Once the high-order all-pole model for  $x(n)$  has been found, it is then necessary to estimate the MA coefficients  $b_q(k)$  from all-pole coefficients  $a_p(k)$ . Thus, the equation becomes

$$A_p(z) \approx \frac{1}{B_q(z)} = \frac{1}{b_q(0) + \sum_{k=1}^q b_q(k)z^{-k}} \quad (3.11)$$

Then  $1/B_q(z)$  represents a  $q$ th-order all-pole model for the “data”  $a(k)$ . The coefficients of the all-pole model for  $a(k)$  are taken as the coefficients of the moving average model. Typically, the model order  $p$  is chosen so that it is at least four times the order  $q$  of the moving average process [72].

### 3.2 The Autoregressive (AR) Models

A wide-sense stationary autoregressive process of order  $p$  is a special case of an ARMA( $p,q$ ) process in which  $q = 0$ . An AR( $p$ ) process may be generated by filtering unit variance white noise,  $u(n)$  with an all-pole filter of the form

$$H(z) = \frac{b_q(0)}{1 + \sum_{k=1}^p a_p(k)z^{-k}} \quad (3.12)$$

Just as with ARMA process, the autocorrelation sequence of an AR process satisfies the Yule-Walker equations

$$r_x(k) + \sum_{l=1}^p a_p(l)r_x(k-l) = |b_q(0)|^2 \delta(k) \quad ; \quad k \geq 0 \quad (3.13)$$

Writing these equations in a matrix form for  $k = 1, 2, \dots, p$ , using the conjugate symmetry of  $r_{xx}(k)$ , we have

$$\begin{bmatrix} r_x(0) & r_x^*(1) & r_x^*(2) & \cdots & r_x^*(p-1) \\ r_x(1) & r_x(0) & r_x^*(1) & \cdots & r_x^*(p-2) \\ r_x(2) & r_x(1) & r_x(0) & \cdots & r_x^*(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_x(p-1) & r_x(p-2) & r_x(p-3) & \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ a_p(2) \\ a_p(3) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} r_x(1) \\ r_x(2) \\ r_x(3) \\ \vdots \\ r_x(p) \end{bmatrix} \quad (2.14)$$

Therefore, given the autocorrelation  $r_x(k)$  for  $k=0,1, \dots, p$  we may solve (3.14) for the AR coefficients. These equations may be solved recursively using Levinson-Durbin Recursion [71, 72, 129] which led to a number of important discoveries including the lattice filter structure.

### 3.3 Lattice Methods for AR Modeling

A close relationship exists between a linear prediction filter and an AR process. If the random process  $x(n)$  is generated as an AR( $p$ ) process and the order of the linear predictor  $m = p$ , then the predictor coefficients will be identical to the AR parameters. This relationship is exploited by several algorithms in finding the AR coefficients through linear prediction [130].

Consider the forward linear prediction estimate

$$\hat{x}^f(n) = - \sum_{k=1}^p a_p^f(k) x(n-k) \quad (3.15)$$

of the sample  $x(n)$ , where  $a_p^f(k)$  is the forward linear prediction coefficients at time index,  $k$ . The hat  $\hat{\phantom{x}}$  is used to denote an estimate and the superscript  $f$  is used to denote that this is a forward estimate. The prediction is forward in the sense that the estimate at time index  $n$  is based on  $p$  samples indexed earlier in time. The complex forward linear prediction error is

$$e_p^f(n) = x(n) - \hat{x}_p^f(n) \quad (3.16)$$

The Eq. (3.16) has a real variance

$$\rho^f = \mathbf{E} \left\{ \left| e_p^f(n) \right|^2 \right\} \quad (3.17)$$

Where  $\mathbf{E} \{.\}$  denotes the expected value.

In similar way to forward prediction a backward linear prediction error estimate

$$\hat{x}_p^b(n) = - \sum_{k=1}^p a_p^b(k) x(n+k) \quad (3.18)$$

may also be formed, in which  $a_p^b(k)$  is the backward linear prediction coefficient at time index,  $k$ . A superscript  $b$  is used to tag elements associated with the backward linear prediction estimate. The prediction is backward in the sense that the estimate at time index  $n$  is based on  $m$  samples indexed later in time. The backward linear prediction error is

$$e_p^b(n) = x(n-m) - \hat{x}_p^b(n-m) \quad (3.19)$$

The Eq. (3.19) has the real variance of

$$\rho^b = \mathbf{E} \left\{ \left| e_p^b(n) \right|^2 \right\} \quad (3.20)$$

If the Levinson-Durbin recursion [71, 72, 131] is substituted for  $a_p^f(k)$  or  $a_p^b(k) = a_p^{f*}(k)$  in the Eq. (3.15) and (3.18) for the forward and backward linear prediction errors, then it is simple to see that

$$\begin{aligned}
e_{j+1}^f(n) &= e_j^f(n) + \Gamma_{j+1} e_j^b(n-1) \\
e_{j+1}^b(n) &= e_j^b(n-1) + \Gamma_{j+1}^* e_j^f(n)
\end{aligned} \tag{3.21}$$

Since the lattice filter provides an alternative parameterization of the all-pole filter, i.e., in terms of its reflection coefficients, formulating the all-pole signal modeling problem as one of finding the reflection coefficients that minimize some error may also be considered. In the following section two such lattice methods for signal modeling including Burg's method, and the modified covariance method are observed.

### 3.3.1 Burg's Method

Previously, Burg developed a method for spectrum estimation known as maximum entropy method [78]. As part of this method, which involves finding an all-pole model for the data, he proposed that the reflection coefficients be computed sequentially by minimizing the mean-square of the forward and backward prediction error [78, 132].

$$\varepsilon_j^{fb} = \varepsilon_j^f + \varepsilon_j^b = \mathbf{E} \left\{ \sum_{n=j}^N |e_j^f(n)|^2 + \sum_{n=j}^N |e_j^b(n)|^2 \right\} \tag{3.22}$$

Now, the value of the reflection coefficients  $\Gamma_j^{fb}$  may be found, which minimizes

$\varepsilon_j^{fb}$  by setting the derivatives of  $\varepsilon_j^{fb}$  with respect to  $(\Gamma_j^{fb})^*$  equal to zero as follows:



$$\begin{aligned}\frac{\partial}{\partial(\Gamma_j^{fb})^*} \mathcal{E}_j^{fb} &= \frac{\partial}{\partial(\Gamma_j^{fb})^*} \mathbf{E} \left[ \sum_{n=j}^N \left\{ |e_j^f(n)|^2 + |e_j^b(n)|^2 \right\} \right] \\ &= \mathbf{E} \left[ \sum_{n=j}^N \left\{ e_j^f(n) [e_{j-1}^b(n-1)]^* + [e_j^b(n)]^* e_{j-1}^f(n) \right\} \right] = 0\end{aligned}\quad (3.23)$$

Substituting the error update equations for  $e_j^f(n)$  and  $[e_j^b(n)]^*$ , which are similar to those given for  $e_{j+1}^f(n)$  and  $[e_{j+1}^b(n)]^*$  in (3.21), and solving for  $\Gamma_j^{fb}$  we find that the value of  $\Gamma_j^{fb}$  that minimizes  $\mathcal{E}_j^{fb}$  is

$$\Gamma_j^{fb} = - \frac{2 \sum_{n=j}^N e_{j-1}^f(n) [e_{j-1}^b(n-1)]^*}{\sum_{n=j}^N \left\{ |e_{j-1}^f(n)|^2 + |e_{j-1}^b(n-1)|^2 \right\}} \quad (3.24)$$

It is important to indicate that sequentially minimizing  $\mathcal{E}_p^{fb}$  by using Burg's method guarantee that the reflection coefficients are bounded by one in magnitude and thus, the AR model is stable.

### 3.3.2 The Modified Covariance Method

In the previous section, Burg recursion which finds the reflection coefficients for an AR model by sequentially minimizing the mean of the squared forward and backward prediction errors is described. In this section, the modified covariance method or forward-backward algorithm for AR signal modeling is observed. As with Burg algorithm, the modified covariance method minimizes the mean of the squares of the forward and backward prediction errors,

$$\mathcal{E}_p^{fb} = \mathcal{E}_p^f + \mathcal{E}_p^b \quad (3.25)$$

The difference, however, between the two approaches is that, in the modified covariance method, the minimization is not performed sequentially. In other words, for a  $p$ th-order model, the modified covariance method finds the set of reflection coefficients or, equivalently, the set of transversal filter coefficients  $a_p(k)$ , that

$$\text{minimize } \varepsilon_p^{fb} .$$

To find the filter coefficients that minimizes  $\varepsilon_p^{fb}$  the derivatives of  $\varepsilon_p^{fb}$  with respect to  $a_p^*(l)$  equal to zero for  $l = 1, 2, \dots, p$  are set. Since

$$e_p^f(n) = x(n) + \sum_{k=1}^p a_p^f(k)x(n-k) \quad (3.26)$$

and

$$e_p^b(n) = x(n-p) + \sum_{k=1}^p a_p^{f*}(k)x(n-p+k) \quad (3.27)$$

then

$$\begin{aligned} \frac{\partial \varepsilon_p^{fb}}{\partial a_p^*(l)} &= \mathbf{E} \left[ e_p^f(n) \frac{\partial [e_p^f(n)]^*}{\partial a_p^*(l)} + [e_p^b(n)]^* \frac{\partial e_p^b(n)}{\partial a_p^*(l)} \right] \\ &= \mathbf{E} \left[ e_p^f(n)x^*(n-l) + [e_p^b(n)]^* x(n-p+l) \right] = 0 \end{aligned} \quad (3.28)$$

Substituting Eq. (3.26) and (3.27) into (3.28) and simplifying, the normal equations for the modified covariance method are given by

$$\sum_{k=1}^p [r_x(l,k) + r_x(p-k, p-l)] a_p^f(k) = -[r_x(l,0) + r_x(p, p-l)]; \quad l = 1, \dots, p \quad (3.29)$$

where

$$r_x(l, k) = \sum_{n=p}^N x(n-k)x^*(n-l) \quad (3.30)$$

For the modified covariance error, the orthogonality condition in (3.28) to express

$\varepsilon_p^{fb}$  is used as follows

$$\varepsilon_p^{fb} = E \left[ e_p^f(n)x^*(n) + \left[ e_p^b(n) \right]^* x(n-p) \right] \quad (3.31)$$

Substituting the expression given in Eq. (3.26) and (3.27) for  $e_p^f(n)$  and  $e_p^b(n)$  and simplifying, gives

$$\varepsilon_p^{fb} = r_x(0,0) + r_x(p,p) + \sum_{k=1}^p a(k) [r_x(0,k) + r_x(p,p-k)] \quad (3.32)$$

One of the properties of the modified covariance method is that the reflection coefficients are not guaranteed to be less than one in magnitude. As a result, it is possible for this method to produce an unstable model.

### 3.4 Box-Jenkins ARIMA Method

The autoregressive integrated moving average (ARIMA) or often called the Box-Jenkins method is a univariate approach which is built on the premise that knowledge of past values of a time series is sufficient to make forecasts of the variable. Briefly, Box-Jenkins method involves the following steps:

Step one is a model identification one, involves the comparison of estimated autocorrelation functions (ACF) and partial autocorrelation functions (PACF) of known ARIMA processes. Given a class of ARIMA models from step one; their

parameters are to be estimated from the historical series using nonlinear least squares. In step three a diagnostic checks are applied to determine any possible inadequacies in the model, and the process is repeated if any are found. Finally, having arrived at an adequate model, "optimal" forecasts are generated by recursive calculation. This algorithm is given in details in [21, 80-82, 95, 116, 133].

Generally speaking, Box-Jenkins' ARMA method can only be used with stationary time series. In practice many time series are non-stationary including the power load data series. Consequently, one possible way of handling non-stationary series is to apply differencing to the data, so as to make them stationary. The first differences, namely  $x(n) - x(n-1) = (1-B)x(n)$ , may themselves be differenced to give second differences and so on. The  $d$ th differences may be written as  $(1-B)^d x(n)$ , where  $B$  is backward shift operator, such that  $Bx(n) = x(n-1)$ . If the raw data series is differenced  $d$  times before fitting an ARMA ( $p,q$ ) process, the model is said to be an ARIMA ( $p,d,q$ ) process; where the letter 'I' in the acronym stands for integrated and  $d$  denotes the number of differences. However, Box-Jenkins' ARMA (ARIMA) will be revisited in Chapter 4 for more detailed explanation.

### 3.5 Conclusion

In this chapter, the foundations of stochastic models for data series are laid out. The auto regressive moving average (ARMA) as a general model for time-series, is discussed. The major drawbacks of ARMA which mainly come from the necessity for nonlinear solutions of the moving average part, are outlined. Durbin's ARMA and Box-Jenkins ARMA are explained as possible solutions to ARMA coefficients. In the second part of this chapter, the autoregressive model is outlined and its relationship with the lattice structure is explained. Burg is thoroughly outlined, since it is widely used for power load demand time-series and it will be used later for the comparison with the proposed algorithm. The modified covariance as AR-based algorithm is also explained and its form of least squares solution to the filter coefficients is outlined. This will help deriving the proposed algorithm in Chapter 4.

## CHAPTER 4

### A MODIFIED FORWARD BACKWARD LINEAR PREDICTION

#### (MFBLP) ALGORITHM

##### 4.1 Introduction

One of the short term load forecast (STLF) methods that received significant attention in literature is the autoregressive-moving average (ARMA). Because of the nonlinear nature required for ARMA's algorithms, all existing solutions to this problem appear to suffer from one or more drawbacks, as explained briefly in the following:

- Several methods may end up in a hard failure mode. This means that the identification algorithm may return an invalid model or the algorithm cannot be carried out to completion because, as a result of a step during its execution, a parameter set outside the class of permissible ones arises for which no provisions have been adopted. This may happen for moment fitting procedures, as well as the methods that first estimate the AR parameters and then the MA.
- Maximum-likelihood methods that depend on search over the parameter space involve significant computations and are not guaranteed to converge, or they may converge to the wrong solution.
- Finally, some methods may be inaccurate (e. g., significantly biased) in finite

samples. This is the case of with Durbin's two stage least-squares method [71-73], and with the approximate subspace methods that enforce positivity of the estimated MA spectrum based on [74].

Though the all-pole models have less degree of freedom than ARMA, they have attracted the attention for their linear nature and their efficient solution to STLF problem especially in short data records. However, in long data records, as it is the case with the most of power loads time series, the required number of all-pole coefficients dramatically increases making most of the known algorithms less efficient in estimating their values and accordingly less accurate in modeling the power load data.

In this chapter an AR algorithm designed for long data records is proposed. The algorithm divides the data record into segments and searches for AR coefficients that simultaneously model all of them with least means squared errors.

## **4.2 Modified Forward-Backward Linear Prediction (MFBLP) Algorithm**

Assume the  $m$ -points data sequences  $x(1), x(2), \dots, x(m)$  are to be used to estimate the  $p$ -th AR parameters. Since, with AR algorithms the order of the model is proportional to the length of data record [73, 134, 135]. In order to avoid using large orders with long data records (as in the thesis, three years hourly data), it is considered that the segmentation of the  $m$ -points data sequence into  $Q$  segments of  $N$  samples each.

Assume one segment of data out of the available  $Q$  segments. Because forward and backward linear predictions have similar statistical information [73], it seems reasonable to combine the linear prediction error statistics of both directions in order to generate more error points. The net result should be an improved estimate of the auto-regressive parameters.

Therefore, the matrix form of (N-p) forward and the (N-p) backward linear prediction samples representing the Q segments and N samples are formulated by the following steps of equations, and it is given by

$$\mathbf{D}_p^f(q) = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(m-2) & x(m-3) & \cdots & x(m-p-1) \end{bmatrix} \quad (4.1)$$

$$\mathbf{D}_p^f(q) = [x_0^f(q) \ x_1^f(q) \ \cdots \ x_{m-p-1}^f(q)]^T \quad (4.2)$$

Where

$$\mathbf{z}_k^f(q) = [x(p+k-1) \ x(p+k-2) \ \cdots \ x(k)]^T \quad (4.3)$$

Where  $k = 0, 1, 2, \dots, m-p-1$ .

The linear predicted array outputs corresponding to the forward data matrix  $\mathbf{D}_p^f(q)$ , can be described as

$$\mathbf{w}_p^f(q) = [\hat{x}(p) \ \hat{x}(p+1) \ \cdots \ \hat{x}(m-1)]^T \quad (4.4)$$

The similar approach from the above forward linear prediction, the  $(m-p)$  sub vectors in the backward direction is drawn as follows

$$\mathbf{D}_p^b(q) = \begin{bmatrix} x^*(p-1) & x^*(p-2) & \cdots & x^*(0) \\ x^*(p) & x^*(p-1) & \cdots & x^*(1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-2) & x^*(m-3) & \cdots & x^*(m-p-1) \end{bmatrix} \quad (4.5)$$

$$\mathbf{D}_p^b(q) = [x_0^b(q) x_1^b(q) \cdots x_{m-p-1}^b(q)]^T \quad (4.6)$$

Where

$$x_k^b(t) = [x^*(k+1) x^*(k+2) \cdots x^*(k+L)]^T \quad (4.7)$$

$k = 0, 1, 2, \dots, m-p-1$ .

The predicted array outputs corresponding to the backward data matrix  $\mathbf{D}_p^b(t)$ , are given as

$$w_p^b(t) = [\hat{x}^*(0) \hat{x}^*(1) \cdots \hat{x}^*(m-p-1)]^T \quad (4.8)$$

Therefore, with the combination of forward and backward linear predicted algorithm the MFBLP data matrix can be described as



$$\mathbf{D}^{fb} = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(p-2) & x(p-3) & \cdots & x(m-p-1) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-p) & x^*(m-p+1) & \cdots & x^*(m-1) \end{bmatrix} \quad (4.9)$$

Assume that  $w^{fb}$  is the desired response at the predictor output of the  $D^{fb}$ , which can be defined as

$$\mathbf{w}^{fb} = [x(p) x(p+1) \cdots x(m-1) \hat{x}^*(1) \hat{x}^*(2) \cdots \hat{x}^*(m-p-1)]^T \quad (4.10)$$

With the  $\mathbf{f}$ , matrix for the MFBLP prediction coefficients, where  $\mathbf{f}$ , is given by

$$\mathbf{f} = [a_p^{fb}(1) a_p^{fb}(2) \cdots a_p^{fb}(p)]^T \quad (4.11)$$

Hence the MFBLP model can be written in matrix form as

$$\mathbf{D}^{fb} \mathbf{f}_p^{fb} = \mathbf{w}^{fb} \quad (4.12)$$

For the simplicity, the Eq. (4.12) can be reduced to

$$\mathbf{D} \mathbf{f} = \mathbf{w} \quad (4.13)$$

Where  $\mathbf{D}$  is the data series,  $\mathbf{f}$  is the predicted coefficients and  $\mathbf{w}$  is the predicted response (signal).

The (N-p) forward and the (N-p) backward linear prediction formulation in Eq. (4.13) associated with the Q-segments of data series can be redefined as

$$\mathbf{D}_q \mathbf{f} = \mathbf{w}_q \quad (3.14)$$

Where the  $2 \times (N - p)$  forward-backward linear prediction data matrix is given by

$$\mathbf{D}_q = \begin{bmatrix} x(p-1) & x(p-2) & \cdots & x(0) \\ x(p) & x(p-1) & \cdots & x(1) \\ \vdots & \vdots & \cdots & \vdots \\ x(p-2) & x(p-3) & \cdots & x(m-p-1) \\ x^*(1) & x^*(2) & \cdots & x^*(p) \\ x^*(2) & x^*(3) & \cdots & x^*(p+1) \\ \vdots & \vdots & \cdots & \vdots \\ x^*(m-p) & x^*(m-p+1) & \cdots & x^*(m-1) \end{bmatrix} \quad (3.15)$$

Let  $\mathbf{w}$  denotes the desired response at the predictor output, given by

$$\mathbf{w}_q = [x(p) \ x(p+1) \ \cdots \ x(m-1) \ \hat{x}^*(1) \ \hat{x}^*(2) \ \cdots \ \hat{x}^*(m-p-1)]^T \quad (3.16)$$

Therefore, the coefficients of MFBLP algorithm is described as

$$\mathbf{f} = [a_p(1) \ a_p(2) \ \cdots \ a_p(p)]^T \quad (3.17)$$

By forming the data matrix  $\mathbf{D}_q$  in corresponds to each data segment, Q, and arranging the resultant matrices in the following form

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \vdots \\ \mathbf{D}_Q \end{bmatrix} \quad (3.18)$$

The corresponding predicted vector to matrix  $\mathbf{D}$  is defined as

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_Q \end{bmatrix} \quad (3.19)$$

Hence, the Eq. (4.17), (4.18) and (4.19) can be rewritten and summarized as,

$$\mathbf{D} \mathbf{f} = \mathbf{w} \quad (3.20)$$

where  $\mathbf{f} = [f_L(1) f_L(2) \dots f_L(L)]^T$ .

A well known criterion called the Least-Squares, will be used to obtain a solution to Eq. (4.20) for the predictor coefficient vector  $\mathbf{f}$ . This solution will guarantee the minimum sum of squared values of the predicted errors [51, 136, 137]. Since the number of the predicted values has been significantly increased by segmentation of the data, then it is expected that the least-squares solution for the predictor coefficients will provide effective solution to power load forecast.

### 4.3 The Least-Squares Solution

The general least squares will be used here to obtain a solution to Eq. (4.20) for the predictor coefficient vector  $\mathbf{f}$ . This solution will guarantee the minimum sum of the squared values of the predicted errors.

According to the Eq. (4.20), the linear predicted errors or the residuals of estimation are given by

$$\mathbf{e} = \mathbf{w} - \mathbf{D}\mathbf{f} \quad (4.21)$$

and the sum of squared errors, is

$$\zeta = \mathbf{e}^H \mathbf{e} \quad (4.22)$$

By substituting Eq. (4.20) into Eq. (4.22), the dependence of the sum of squared errors on the predictor coefficients, can be expressed as follows

$$\zeta = \mathbf{w}^H \mathbf{w} - \mathbf{w}^H \mathbf{D}\mathbf{f} - \mathbf{f}^H \mathbf{D}^H \mathbf{w} + \mathbf{f}^H \mathbf{D}^H \mathbf{D}\mathbf{f} \quad (4.23)$$

Now, by differentiating Eq. (4.23) with respect to the  $\mathbf{f}$ , the gradient vector can be expressed as

$$\begin{aligned} \frac{\partial \zeta}{\partial \mathbf{f}} &= -2\mathbf{D}^H \mathbf{w} + 2\mathbf{D}^H \mathbf{D}\mathbf{f} = 0 \\ &= -2\mathbf{D}^H \mathbf{e} \end{aligned} \quad (4.24)$$

As it is clear the sum of squared errors reaches its minimum value when the gradient vector is zero. Then from Eq. (4.24), the formulation can immediately deduce to

$$\mathbf{D}^H \mathbf{D}\mathbf{f} = \mathbf{D}^H \mathbf{w} \quad (4.25)$$

Thus, the predictor coefficients that give the least squares errors are obtained as a solution to Eq. (4.25). However, this solution is unique only when the matrix  $\mathbf{D}$  is full rank. When this condition is satisfied the matrix  $\mathbf{D}^H\mathbf{D}$  is nonsingular and the solution is unique, given as

$$\mathbf{f} = (\mathbf{D}^H\mathbf{D})^{-1}\mathbf{D}^H\mathbf{w} \quad (4.26)$$

Thus the predictor coefficients that give the least squares errors are obtained as a solution to Eq. (4.26). However, this solution is unique only when the nullity of the matrix  $\mathbf{D}$  is zero [138]. The nullity of a matrix denoted as  $\text{null}(\cdot)$  is defined as the dimension of the matrix null space. In other words, the least-squares solution is unique when the matrix  $\mathbf{D}$  is of full rank. When this condition is satisfied, the  $p$ -by- $p$  matrix  $\mathbf{D}^H\mathbf{D}$  is nonsingular and the solution is unique, given as

$$\begin{aligned} \mathbf{f} &= (\mathbf{D}^H\mathbf{D})^{-1}\mathbf{D}^H\mathbf{w} \\ \mathbf{f} &= \mathbf{D}^\# \mathbf{w} \end{aligned} \quad (4.27)$$

Where  $\mathbf{D}^\#$  is called the pseudo-inverse of the matrix  $\mathbf{D}$ , and it is given by

$$\mathbf{D}^\# = (\mathbf{D}^H\mathbf{D})^{-1}\mathbf{D}^H \quad (4.28)$$

In some particular conditions, it may be facing with the data matrix  $\mathbf{D}$  that has linearly dependent columns; meaning that there is no longer  $\text{null}(\mathbf{D}) = 0$ . Consequently, then it will create a new situation where the decision has to be made; on which of an infinite number of possible solutions to work with for the described least squares solution. This issue can be solved by developing a general definition of a pseudo-inverse that guarantees a unique least-squares solution even with  $\text{null}(\mathbf{D}) \neq 0$  [138]. The general pseudo-inverse solution is unique in that it is the only solution that satisfies two requirements:

- It produces the minimum prediction errors or estimation residuals, and
- It has the smallest Euclidean norm possible.

#### 4.4 Summary of the MFBLP Algorithm

A brief summary of the MFBLP algorithm applied to power load demand data is as follows:

1. The power load demand data  $x(1), x(2), \dots, x(n)$  is formed.
2. The data series is segmented into  $Q$  segments and  $N$  number of samples for each segments as indicated in Eq. (4.18).
3. From the data matrix, the optimum number for  $Q$  and  $N$  are obtained.
4. The estimation data matrix (in-sample data) is then estimated.
5. Finally, the value of  $f_l$  are calculated using Eq. (4.21).

#### 4.5 The Optimum Length of the Data Segments and Predictor Order

In this section we will try to find the best values for  $Q$ ,  $N$ , and  $L$ . Three years data load demand record collected by NEMMCO in NSW, Australia, between the beginning of 2005 and the end of 2007, is used for our study. The first two years hourly data (17520 samples) are used for model extraction and the remaining year of data is used for model validation. The mean average peak error (MAPE) is used as a metric indicate the accuracy of the model in predicting data.

In the first experiment, the two years 17520 data samples are segmented into  $Q$  segments of different lengths  $N$ . The different arrangements of  $Q$  and  $N$  are shown in Table 4.1.

Table 4.1 Q and N for the NSW two-year hourly load demand data

	Q	N
Segmentation 1	2	8760
Segmentation 2	5	3504
Segmentation 3	10	1752
Segmentation 4	24	730
Segmentation 5	48	365

Next, the predictor order (L) is varied from 0.1N to 0.25N and the best order of the MFBLP predictor is defined. As indicator of performance, the MAPE value is calculated for each L over the validation year of data. Since the MAPE values are function of the leading forecast time, they are calculated with three different leading times over the validation year of data:

- 1) 1 hour ahead,
- 2) one day ahead, and
- 3) one week ahead.

The results are included in Tables 4.2, 4.3, and 4.4. Table 4.2 shows the MAPE values of MFBLP algorithm as a function of L in case of 1 hour ahead forecast. The MAPE values are calculated from  $350 \times 24$  data samples over the validation year.

Table 4.2 The MAPE values of the MFBLP as a function of L in 1-hour ahead forecast

Segmentation 1	1.21	1.16	1.05	0.84
Segmentation 2	1.22	1.19	1.07	0.88
Segmentation 3	1.21	1.18	1.07	0.89
Segmentation 4	1.20	1.18	1.05	0.85
Segmentation 5	1.20	1.19	1.05	0.83

In Table 4.3, we have the MAPE values calculated with one day leading time (24 hours ahead forecast) over the 350 days of the validation year.

Table 4.3 The MAPE values of the MFBLP as a function of L in 24 hours-ahead forecast

Segmentation 1	3.41	3.33	3.26	3.28
Segmentation 2	3.56	3.42	3.40	3.39
Segmentation 3	3.55	3.41	3.39	3.36
Segmentation 4	3.40	3.32	3.25	3.20
Segmentation 5	3.40	3.30	3.24	3.19

The results in Table 4.4 shows MAPE values of the MFBLP calculated with one week leading time (168 hours ahead) over the 50 weeks of the validation year.



Table 4.4 The MAPE values of the MFBLP as a function of L in 168 hours-ahead forecast

Segmentation 1	4.79	4.70	4.60	4.48
Segmentation 2	4.90	4.81	4.74	4.51
Segmentation 3	4.91	4.83	4.75	4.51
Segmentation 4	4.73	4.68	4.62	4.42
Segmentation 5	4.72	4.66	4.62	4.40

The results in Table 4.2 to 4.4 indicate clearly that the MFBLP is reaching its best performance when  $L = 0.25N$ . Now, if we look into the  $0.25N$  columns in Tables 4.2, 4.3, and 4.4, we can easily find that the MFBLP is achieving its best performance with the 5th segmentation, namely when  $Q = 48$  and  $N = 365$ . Having obtained the best values for  $Q$  and  $N$  at  $L = 0.25N$ , would the MFBLP algorithm show any further improvement with increased  $L$ ? In order to answer this question,  $L$  is varied over the range from  $0.1N$  to  $0.65N$  and the MAPE values are calculated and depicted in Figure 4.1.

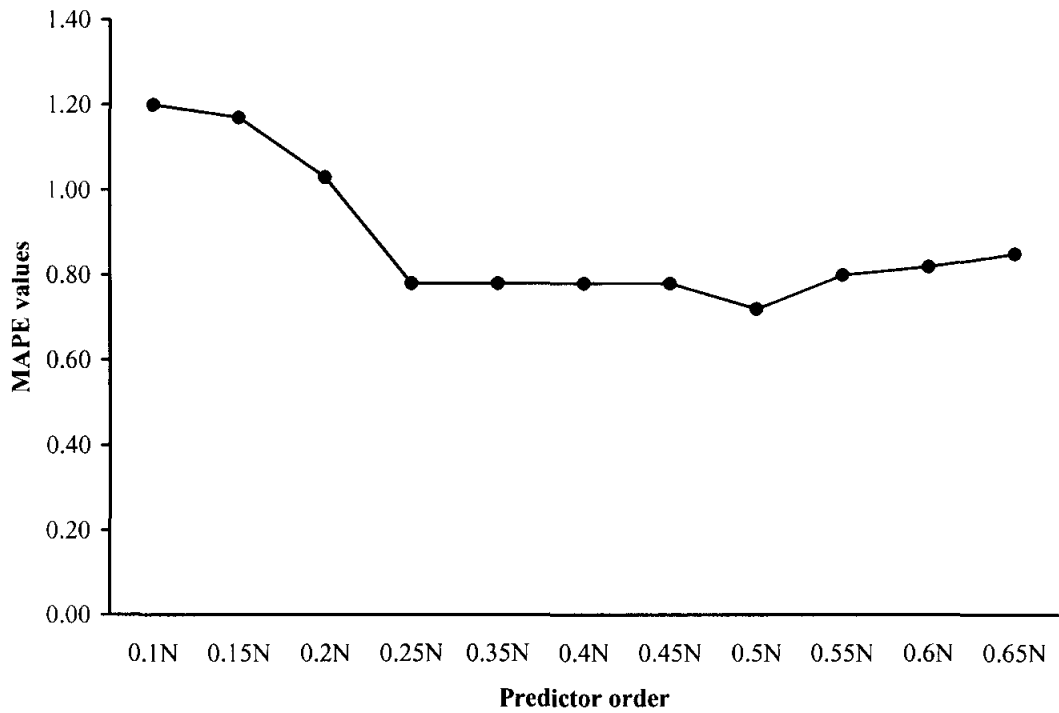


Figure 4.1 MAPE values of the MFBLP algorithm as a function of predictor order

Figure 4.1 shows clearly that the MFBLP is achieving its best performance for  $L = 0.25N$ . For higher values than  $L = 0.25N$ , the algorithm is showing a relatively constant behavior till  $L = 0.55N$ , where it starts to show some signs of deterioration with increased  $L$ . Thus, we chose  $L = 0.25N$  as the best value for the predictor order. In the second experiment, the computational time of the MFBLP is found for the different segmentation schemes. The algorithm is written in MATLAB 7.2 and run on Pentium 4 machine.  $L$  is considered to be  $0.25N$  and  $N = 17520$  samples (two years). The results are shown in Table 4.5.

Table 4.5 The computational time of the MFBLP algorithm with  $L = 0.25N$  and  $N = 17520$  samples as a function of the different segmentations

Segmentation	Computational Time
Segmentation 1	21
Segmentation 2	13
Segmentation 3	10
Segmentation 4	5
Segmentation 5	3

The results in Table 4.5 show clearly that Segmentation-5 is giving the least computational time for the MFBLP algorithm. This will add another important factor to the smallest MAPE value which obtained in the first experiment, in order to consider Segmentation-5 as the best one for the MFBLP algorithm.

Based on the obtained results in the two above experiments it is found that the MFBLP is achieving its optimum performance in terms of MAPE and computational time with  $Q = 48$  ( in case of 17520 data samples) and  $L = 0.25N$ .

#### 4.6 Conclusion

In this chapter, the proposed MFBLP algorithm is thoroughly described as a solution to power load forecast problem, and its optimum parameters are experimentally obtained. The algorithm is based on the segmentation of the power load demand data samples into  $Q$  segments and finding the forward back linear prediction data matrix for each segment. From the  $Q$  data matrices an overall matrix is obtained and the least squares criterion is used to find the predictor coefficients. Three years data load demand record collected by NEMMCO in NSW, Australia, between the beginning of 2005 and the end of 2007, is used to find the best values for the number of

segments ( $Q$ ) and predictor order ( $L$ ) as a function of segment length ( $N$ ). Two metrics are used as performance indicators of the MFBLP, which are the MAPE and the computational time. The results shows that the best value of  $Q$  with two years data samples (17520 samples) is 48 and the best value for the predictor order as a function of segment length is  $0.25N$ .

# **CHAPTER 5**

## **POWER LOAD TIME SERIES: CHARACTERISTICS AND PROCESSING**

### **5.1 Introduction**

In general, power load time series is a complex type of signals. In this chapter we address the major characteristics of the power load time series concentrating on the periodical or seasonal behavior of the signal. The redundancy issue is also addressed and the differencing techniques to reduce redundancy in the signal are discussed. Real power load demand time series are collected by the National Electricity Market Management Company Ltd (NEMMCO) in four Australian states and used for this study. The data is an hourly one collected over the duration of three years in the regions of New South Wales (NSW), Queensland (QLD), South Australia (SA) and Victoria (VIC).

### **5.2 Analyzing the Seasonal Behavior of the Power Load Time Series**

In this section, the daily, the weekly, and the seasonal periodic behaviors of the hourly power load time series of the four Australian states, are addressed and thoroughly analyzed. Two years of hourly data collected over the years 2005 and 2006 is used for this study.

### 5.2.1 New South Wales (NSW)

In order to show the seasonal periodic behavior of NSW power load time series, the two years data is shown Figures 5.1.

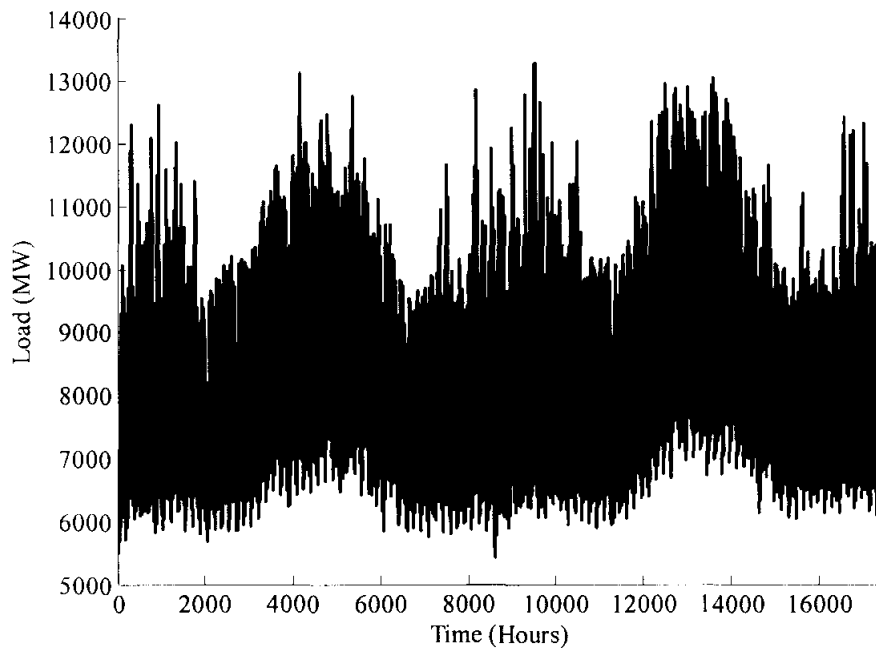


Figure 5.1 Two years of hourly power load demand of NSW

To show the weekly and daily periodic patterns, respectively, three weeks of data are depicted in Figure 5.2 and seven days in Figure 5.3.

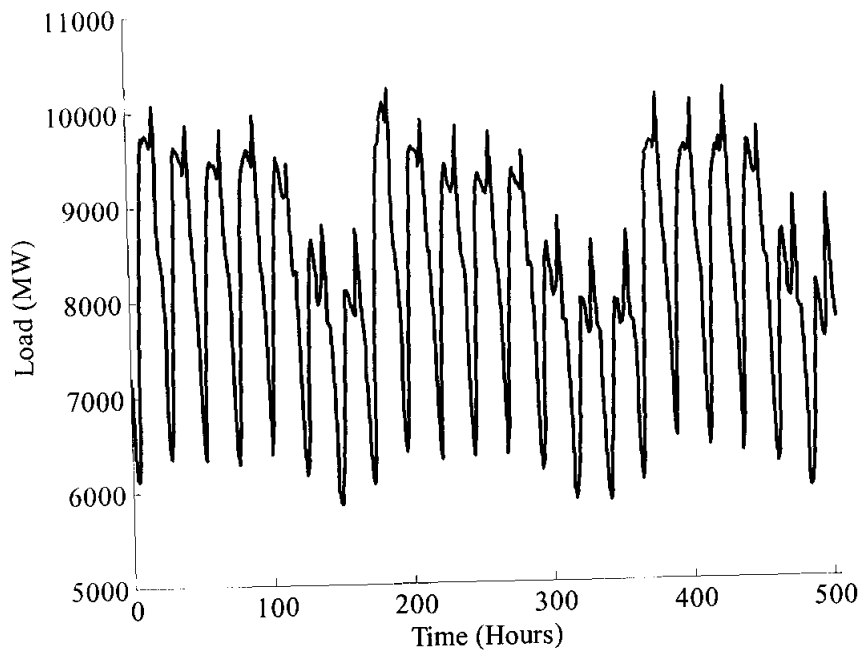


Figure 5.2 Three weeks of hourly power load demand of NSW

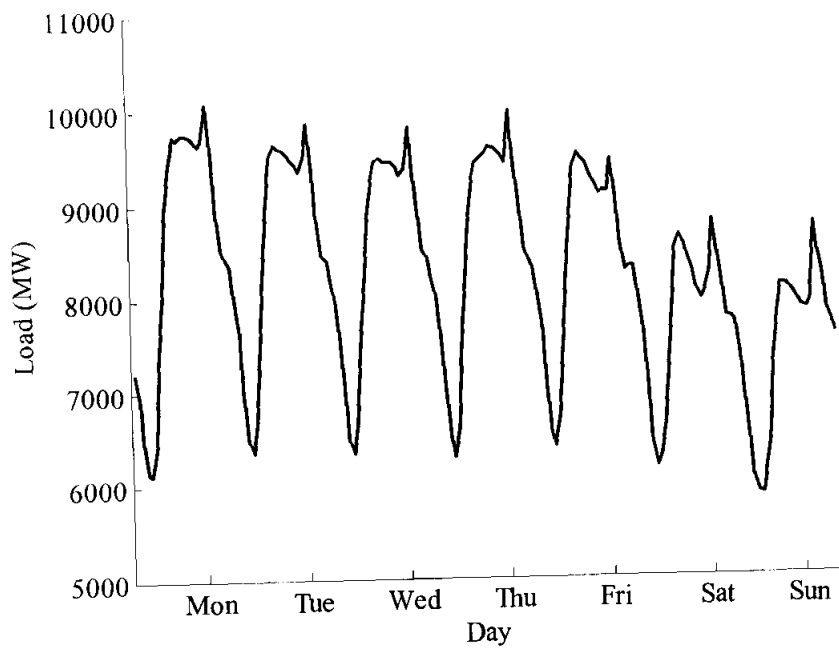


Figure 5.3 Seven days of hourly power load demand of NSW

As shown in Figure 5.1, 5.2 and 5.3, the power load time series of NSW exhibits well-defined daily, weekly and seasonal patterns. Figure 5.1 shows different dynamic range for the summer and winter patterns from fall and spring. The reason is directly related to the higher rates of power consumption, in terms of heating and cooling, during winter and summer than in spring and fall. Figure 5.1 also shows that the base load is around 6000MW and the peak load is about 13000MW. This means that the dynamic range of NSW power load demand is approximately 7000MW, indicating high level of fluctuation.

### 5.2.2 Queensland (QLD)

In similar way to above NSW results, Figures 5.4, 5.5 and 5.6 shows the different forms of seasonality in QLD power load data series.

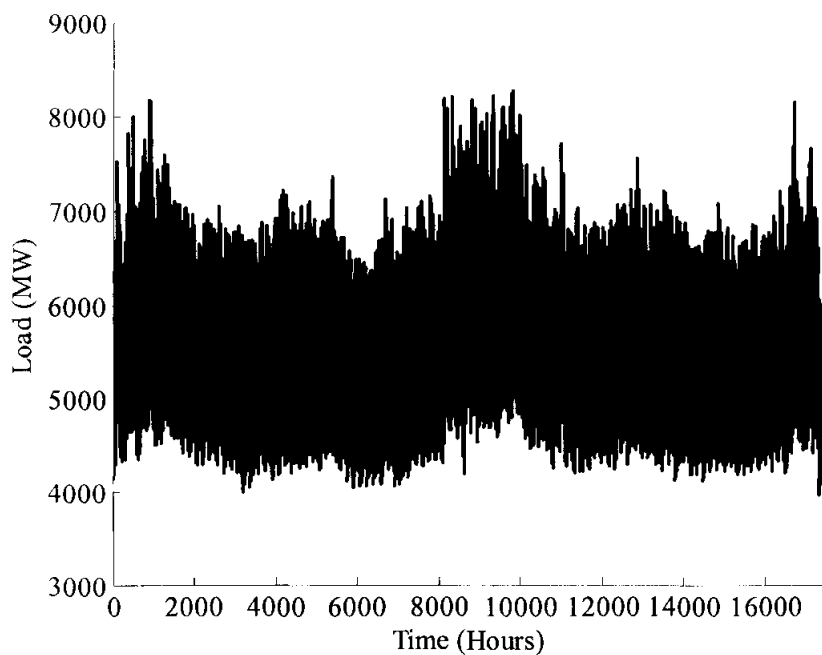


Figure 5.4 Two years of hourly power load demand of QLD



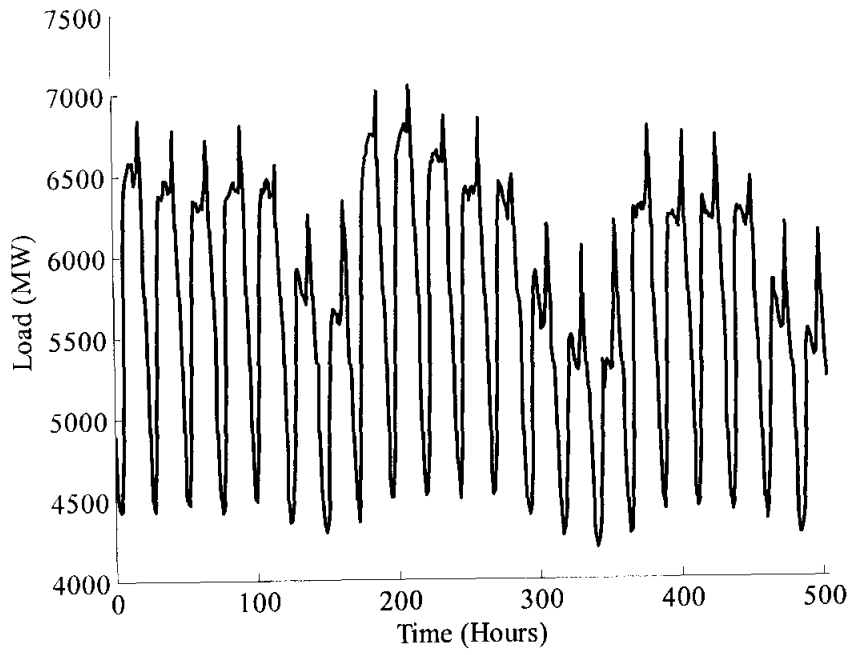


Figure 5.5 Three weeks of hourly power load demand of QLD

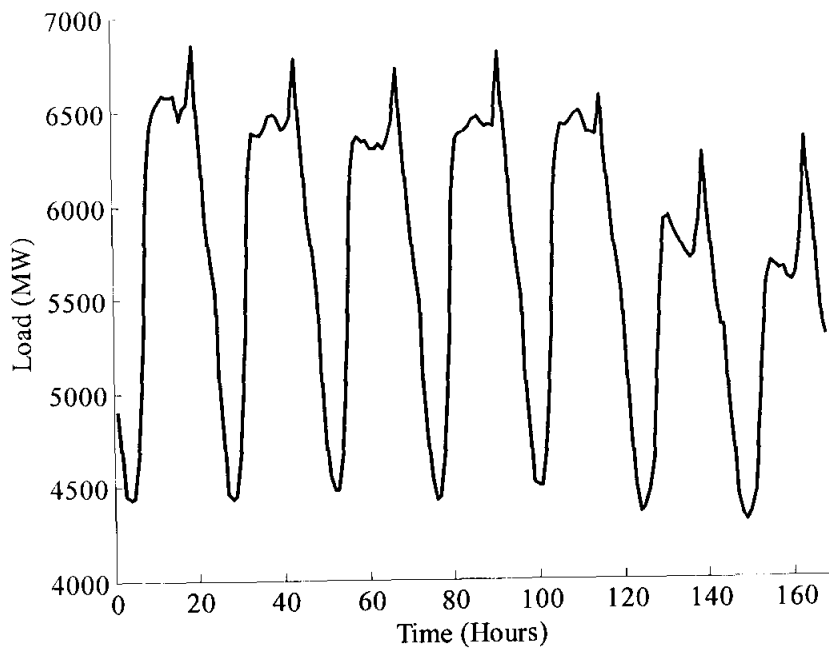


Figure 5.6 Seven days of hourly power load demand of QLD

Figure 5.4 shows the seasonal behavior of the QLD power load data series. Because of the less extreme weather in comparison to NSW, the QLD time series is showing similar behavior for winter, spring, and fall and slightly higher pattern for summer. This makes the dynamic range of QLD time series less than NSW. From Figure 5.4, it is evident that the base load is around 4500MW and the peak load is almost 8000MW, making the dynamic range of approximately 3500MW. The shown daily and weekly pattern in Figures 5.5 and 5.6, respectively, are similar to NSW. The patterns are uniform in their periodic behavior with slight changes over the weekends.

### 5.2.3 South Australia (SA)

Figures 5.7, 5.8 and 5.9 show, respectively, the power demand over two years, three weeks and seven days.

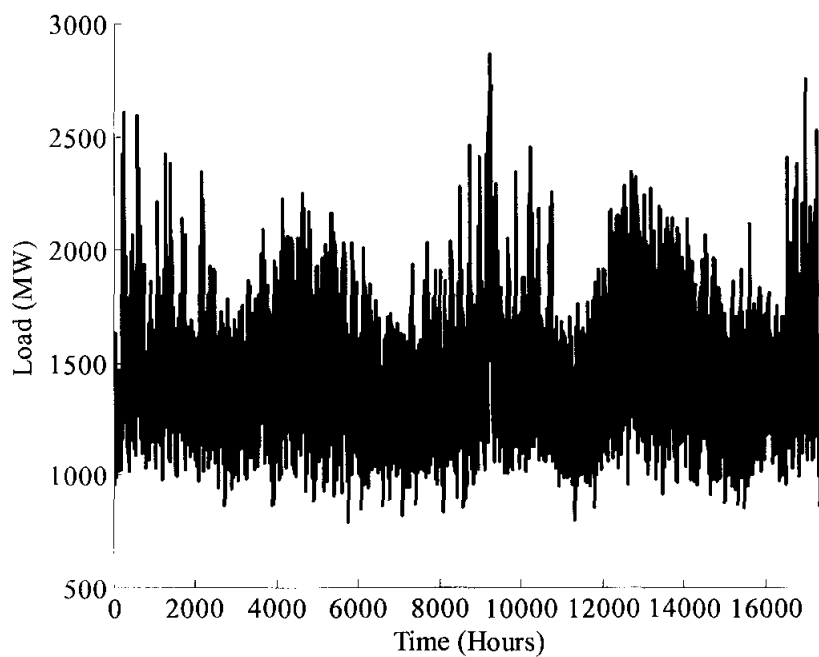


Figure 5.7 Two years of hourly power load demand of SA

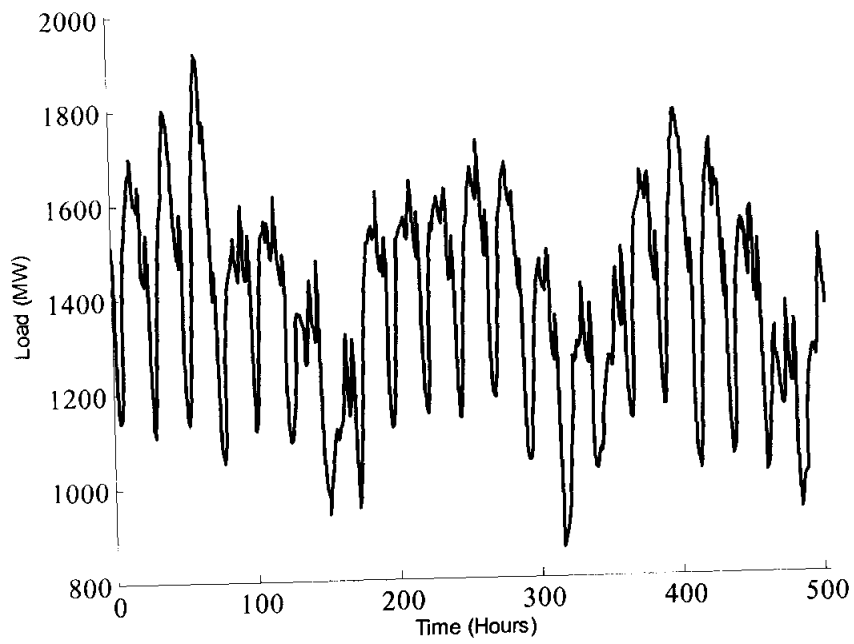


Figure 5.8 Three weeks of hourly power load demand of SA

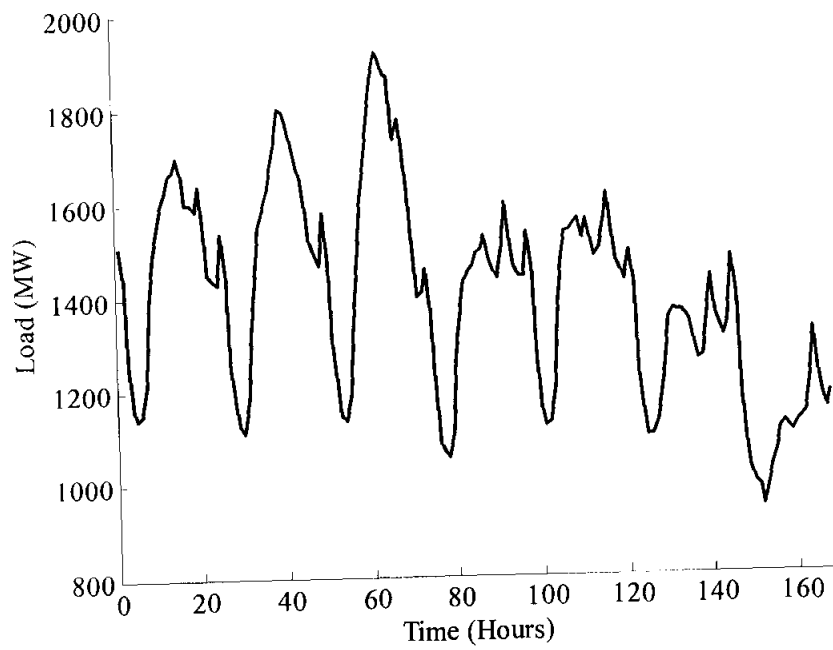


Figure 5.9 Seven days of hourly power load demand of SA

Figure 5.7 shows clearly the impact of the extreme weather in SA on the power load demand. The seasonal behavior of the data is clear but with considerable amount of irregularities, making load forecast in SA much more difficult than in NSW and QLD. This high level of seasonal fluctuation, find its way to the weekly patterns in Figure 5.8 through less similarity factor over the considered three weeks. Figure 5.9 shows that the irregularities with SA data are not confined to the weekend but extend through the weekdays.

#### 5.2.4 Victoria (VIC)

Figures 5.10, 5.11 and 5.12 show the power load demand in VIC for the duration of two years, three weeks and seven days, respectively.

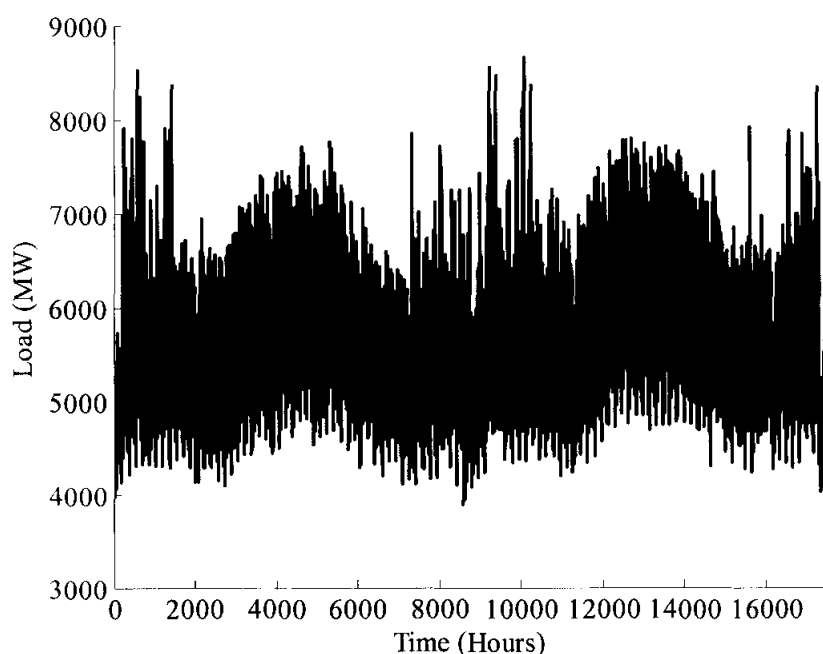


Figure 5.10 Two years of hourly power load demand of VIC

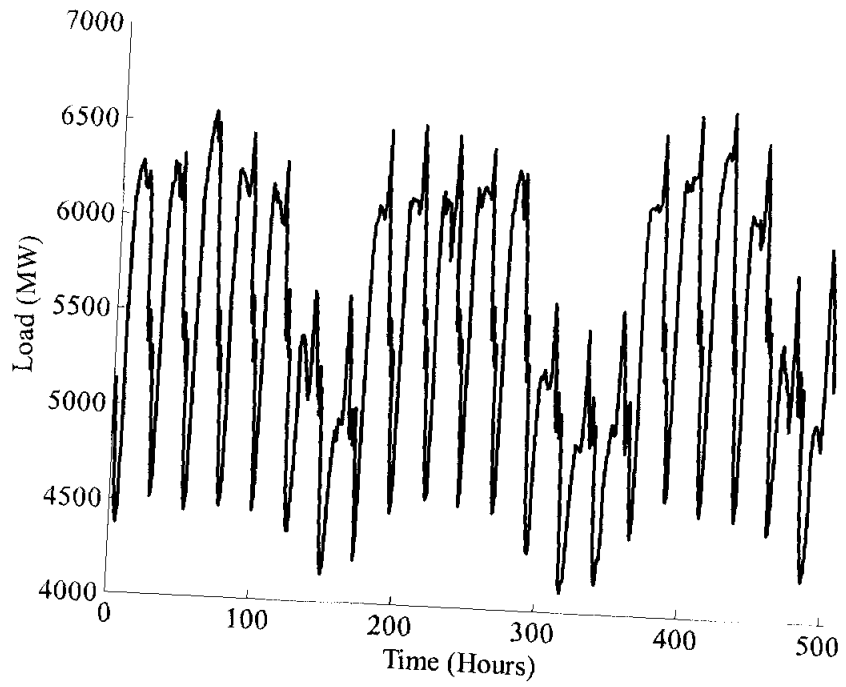


Figure 5.11 Three weeks of hourly power load demand of VIC

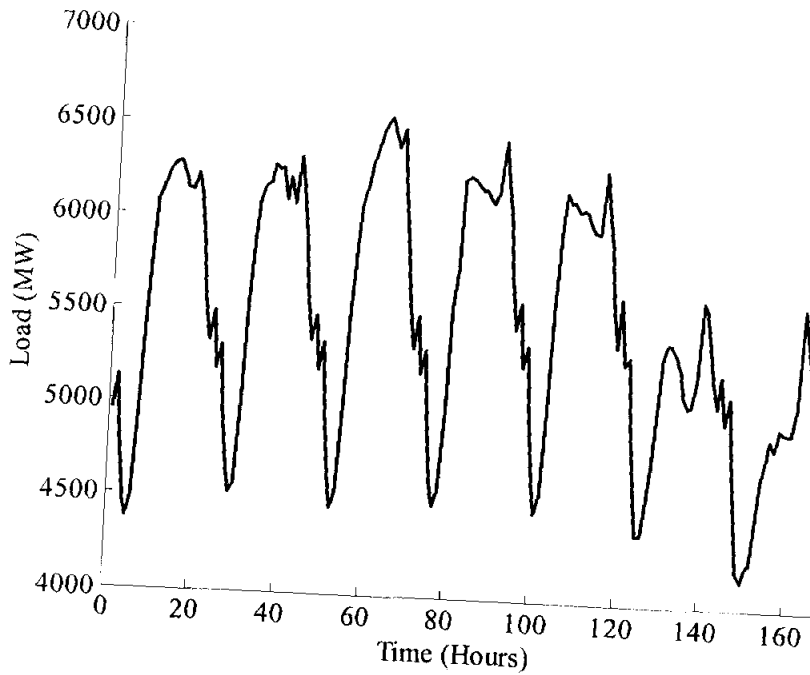


Figure 5.12 Seven days of hourly power load demand of VIC

As shown in Figure 5.10, 5.11 and 5.12, the power load time series for VIC is showing annual, weekly, and daily periodic behavior in similar way to the other states. The periodic patterns are well-defined except for some irregularities during summer time. The power load demand in VIC is almost similar to the QLD and the dynamic range is approximately 4500 MW.

### 5.3 Demand Curves Affected by the Anomalous Day

During anomalous days the majority of industries and commercial activities are significantly reduced, if not totally ceased. This produces large drop in power consumption which may initiates different forms of variation in the load demand time series. For example, the New Year day initiates a particular pattern of variation over the first week of the year different from the ordinary weeks. The same can be said about Australia day, Anzac day and Queen’s Birthday.

Figures 5.13 to 5.16 illustrate a few examples of the demand curves during anomalous days.

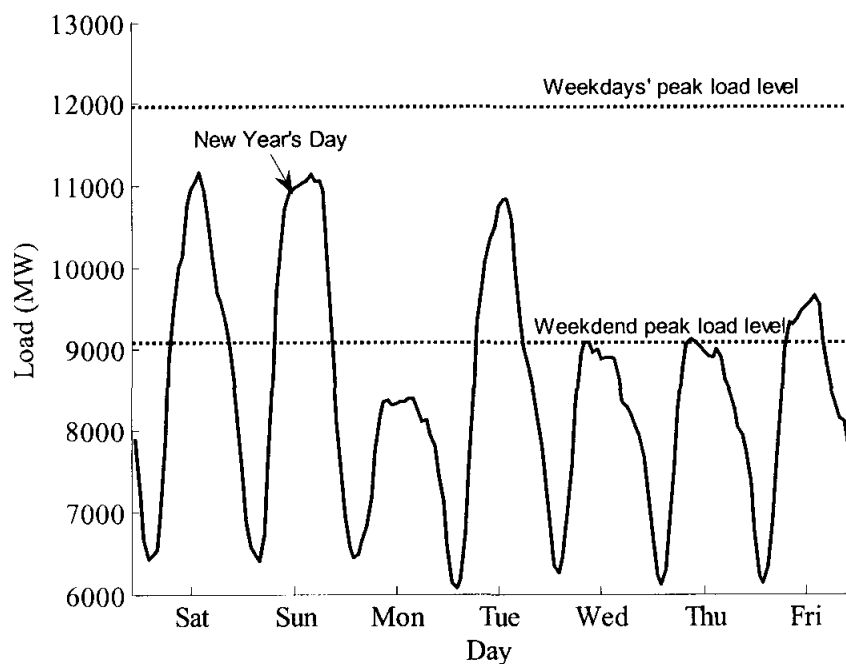


Figure 5.13 The load demand over the days of the week of year 2006 in NSW, initiated by the New Year day

It is clear from Figure 5.13, that the New Year day affects the power load demand not only over the following days, but also over the preceding days. For instance, the day before the New Year day shows power load of 11000 MW, which exceeds the average value of that day by approximately, 2000MW. It is also clear that the effect of this anomalous day propagates through the following days generating totally different pattern of variation from the rest of the year.

The special pattern of variation, generated by the Anzac day over one week is shown in Figure 5.14.

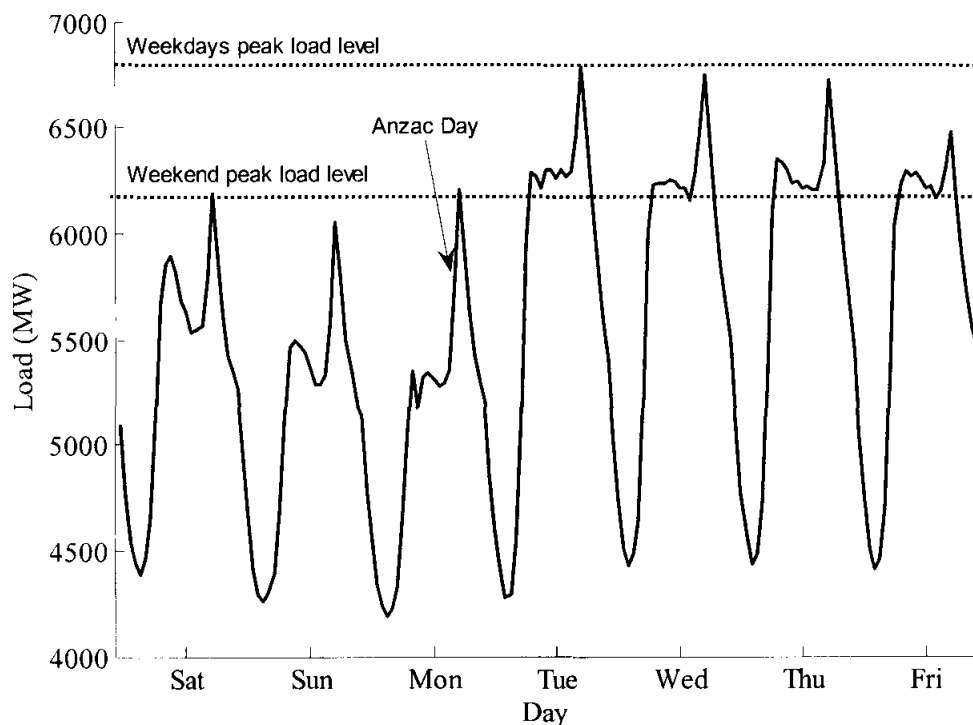


Figure 5.14 The initiated load demand by Anzac day in QLD over one week in April 2005

Figure 5.14 shows clearly the different pattern from ordinary weeks, because of Anzac day. In similar way to New Year day, the effect of the Anzac day is propagating backward and forward affecting the load demand in both sides. The weekly patterns of other anomalous days, Australia day and Queen's Birthday, are shown, respectively, in Figures 5.15, and 5.16.

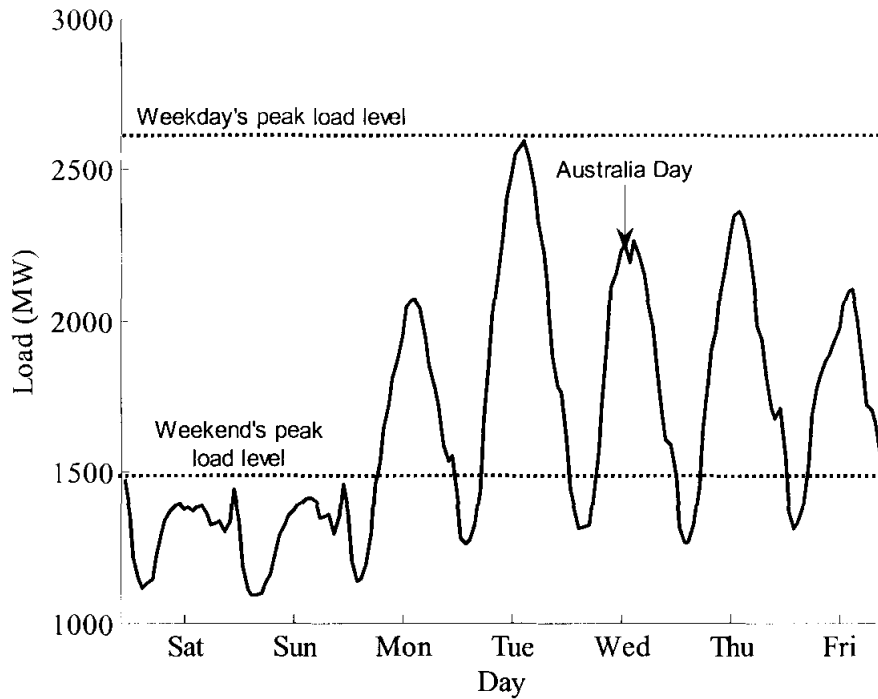


Figure 5.15 The initiated load demand by Australia day in SA over one week in January 2005

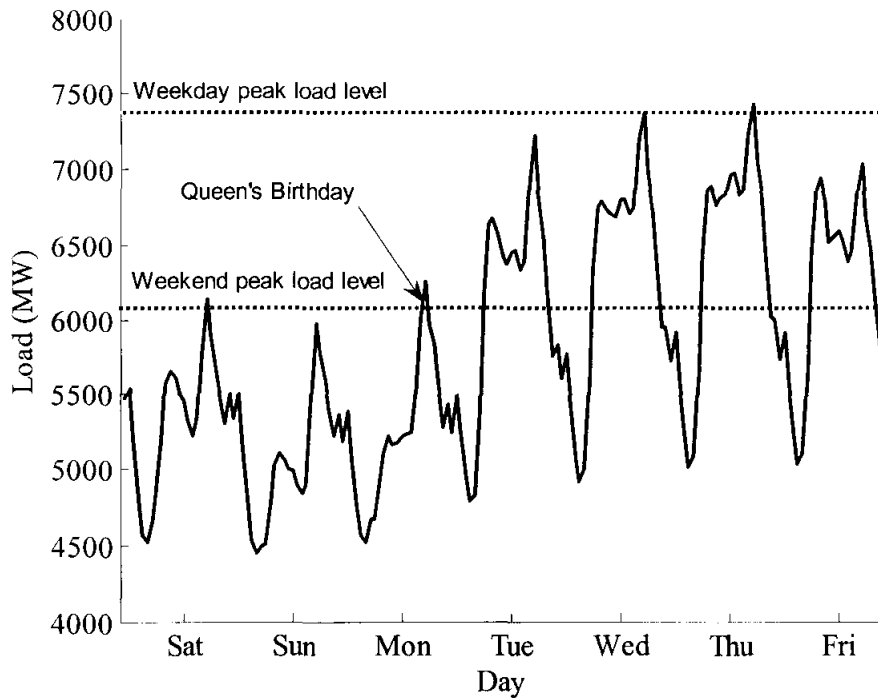


Figure 5.16 The VIC load demand curves for one week during the Queen's Birthday in June 2005



## 5.4 Data Processing for Redundancy Reduction

Due to the daily and weekly seasonal behaviors of the power load demand, the data is showing great amount of interdependency, which makes it highly redundant. Since, the building up suitable model that fits the data record is in the process and, the redundancy from the data before constructing the model is removed in order to avoid model overload. Many techniques for redundancy mitigation in time series were proposed over the last few decades, with different degrees of efficiency [13, 14, 37, 80, 139-141]. However, one of the best and most used techniques is the differencing [1, 80, 81].

In this section, the differencing technique is outlined through Box-Jenkins ARIMA and the seasonal ARIMA (SARIMA) algorithms.

### 5.4.1 Box-Jenkins ARIMA Algorithm

Consider the power load time-series  $x(1), x(2), \dots, x(n)$ . The first differencing ( $I$ -th difference) of the data is given as

$$x(n) - x(n - k) \tag{5.1}$$

Where  $k$  is a lag parameter.

The Eq. (5.1) may further difference in order to produce the  $d$  order of differencing ( $d$ -th difference). The Equation (5.1) may rewritten as

$$x(n) - x(n - k) = (1 - B)x(n) \tag{5.2}$$

Where  $B$  is backward shift operator, such that  $Bx(n) = x(n - 1)$ .

The  $d$ -th differences may be written as  $(1 - B)^d x(n)$ , where  $k = 1$ . If the original raw data series is differenced  $d$  times before fitting an ARMA  $(p, q)$  process, then the model for the original un-differenced series is said to be an ARIMA  $(p, d, q)$  process where  $d$  denotes the number of differences taken.

Generally, the ARMA  $(p, q)$  may be written as

$$\phi(B)x(n) = \theta(B)a(n) \quad (5.3)$$

Where  $\phi(B), \theta(B)$  are the polynomials in  $B$  of finite order  $p, q$ , respectively. From Eq. (5.2) and (5.3), it is generalized that an ARIMA model can be represented by

$$\phi(B)(1 - B)^d x(n) = \theta(B)a(n) \quad (5.4)$$

When fitting AR and MA models, the main difficulty is assessing the order of the process, rather than estimating the coefficients. Moreover, with ARIMA models, there is an additional problem in choosing the required order of differencing such as the value of  $d$ . Some formal procedures are available, including testing for the presence of a unit root [142-144]. However, many literatures simply difference the series until the autocorrelation value declines to zero. First-order differencing is usually adequate for non-seasonal series, albeit the second-order differencing is occasionally needed. Once the redundancy in the data series has been removed, an ARMA model can be fitted to the differenced data in the usual manner.

#### **5.4.2 Box-Jenkins SARIMA processes**

The seasonal ARIMA abbreviated SARIMA model can be obtained as a generalization of Eq. (5.4).

Suppose the seasonal period is  $s$ . Then we may denote  $B^s$  as the operator such that

$$x(n) - x(n - s) = (1 - B^s)x(n) \quad (5.5)$$

A seasonal autoregressive term, for example, is one where  $x(n)$  depends linearly on  $x_{n-s}$ . A SARIMA model with non-seasonal terms of order  $(p, d, q)$  and seasonal terms of order  $(P, D, Q)$  is abbreviated as SARIMA  $(p, d, q) \times (P, D, Q)_s$  model and can be represented by

$$\phi_1(B)\phi_2(B^s)(1 - B)^d(1 - B^s)^D x(n) = \theta_1(B)\theta_2(B^s)a(n) \quad (5.6)$$

where  $\phi, \theta$  are the polynomials in  $B^s$  of order  $P, Q$ , respectively.

In order to remove the redundancy and the seasonal in the data series, we apply first order seasonal differencing technique. The seasonal parameters are  $S1 = 24$  and  $S2 = 168$ . Thus the seasonal first and second differencing can be written as

$$x(n) - x(n - 24) = (1 - B^{24})x(n) \quad (5.7)$$

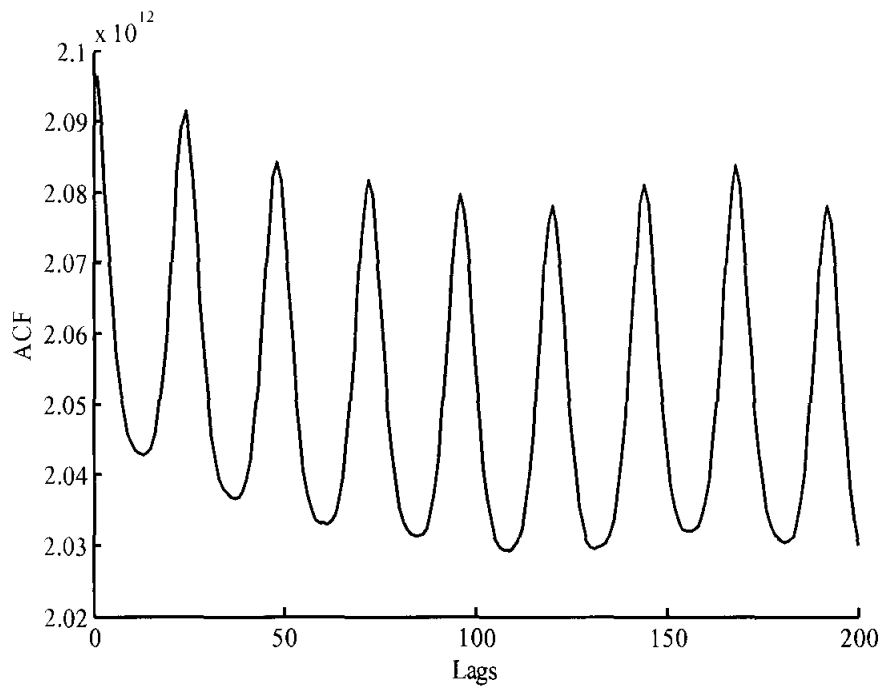
and,

$$x(n) - x(n - 168) = (1 - B^{168})x(n) \quad (5.8)$$

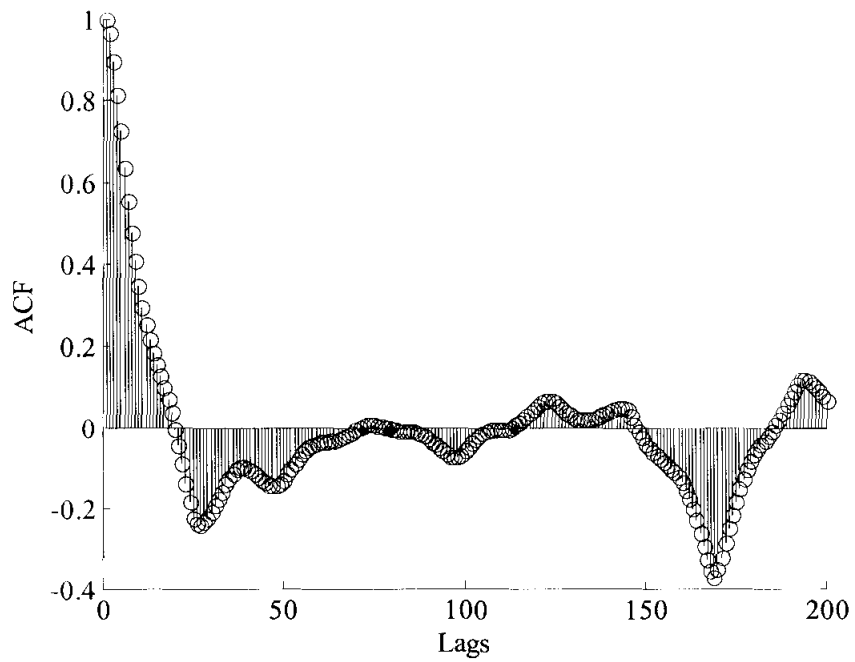
Since the data series need to apply for double differencing, thus the SARIMA model is now multiplicative. Hence, we can represent the multiplicative SARIMA as

$$\phi_1(1 - B^{24})\phi_2(1 - B^{168})x(n) = \theta_1(1 + B^{24})\theta_2(1 + B^{168})a(n) \quad (5.9)$$

The autocorrelation functions of the raw time series of the four Australian states and the autocorrelation functions of the processed data according to Eq. 5.9, are shown in Figures 5.17, 5.18, 5.19, and 5.20.

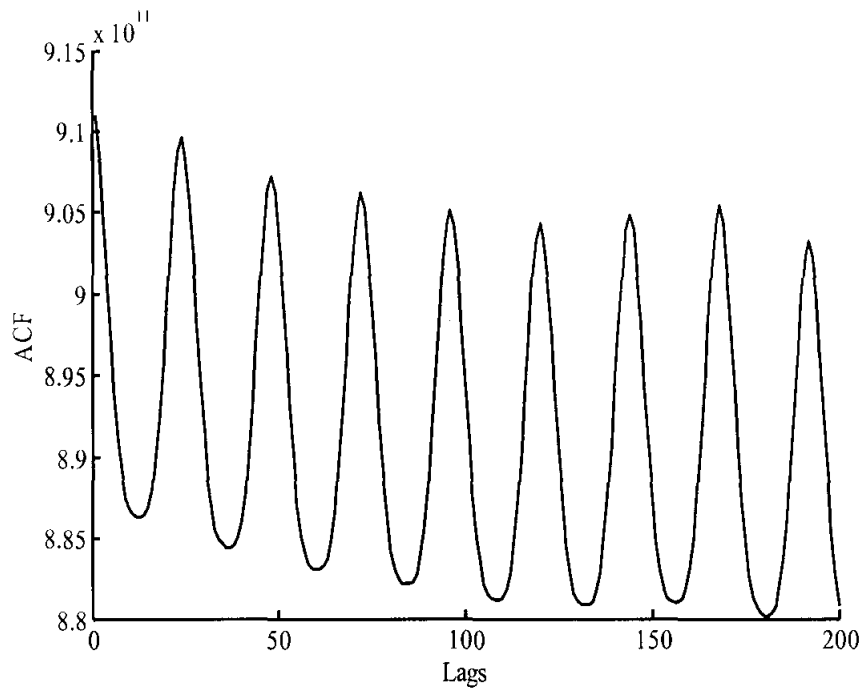


(a)

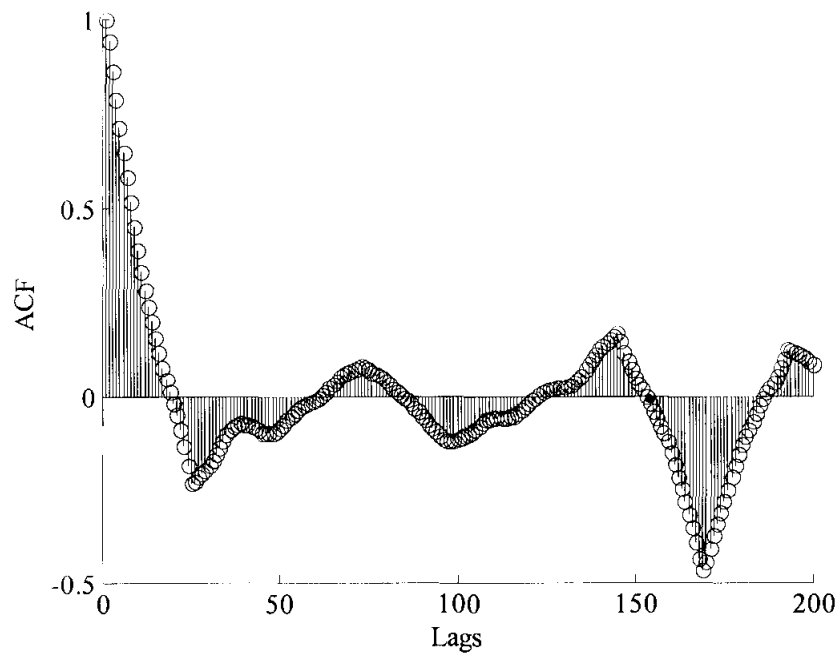


(b)

Figure 5.17 The ACF of NSW time series: a) ACF of raw data, b) ACF of processed data

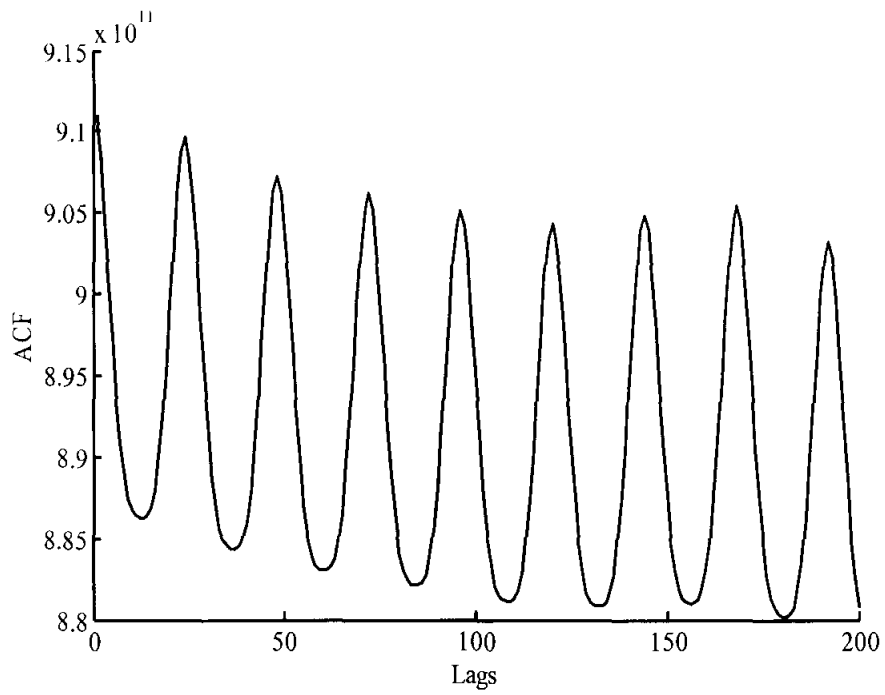


(a)

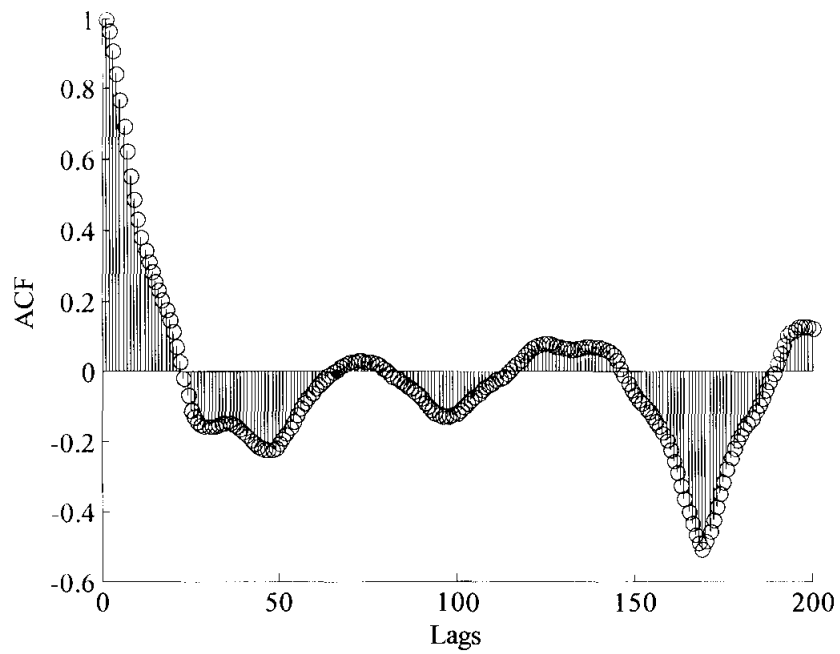


(b)

Figure 5.18 The ACF of QLD time series: a) ACF of raw data, b) ACF of processed data

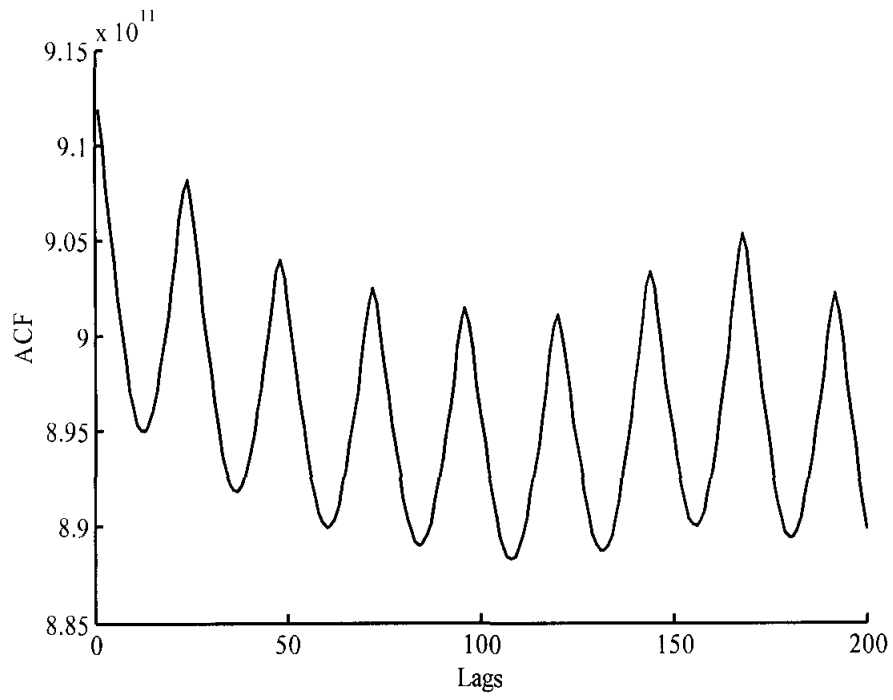


(a)

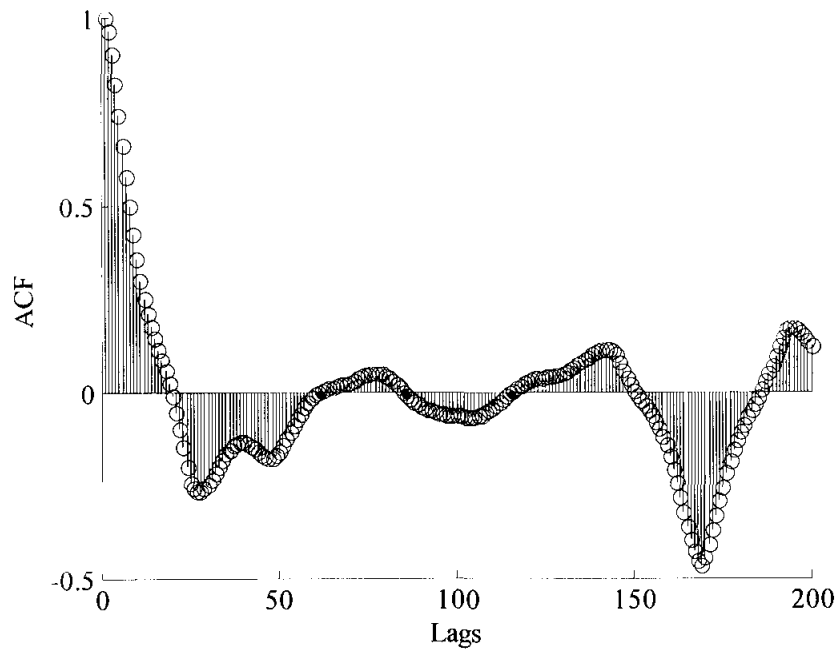


(b)

Figure 5.19 The ACF of SA time series: a) ACF of raw data,  
b) ACF of processed data



(a)



(b)

Figure 5.20 The ACF of VIC time series: a) ACF of raw data, b) ACF of processed data

It is quite clear from the figures above that processing the power load time series according to Eq. (5.9) will significantly reduce the level of redundancy in the signal. The fast decaying ACF in the figures above indicates the adequateness of the differencing order in Eq. (5.9) in removing most of the redundancy in the original data of the four Australian states. This makes the data in the required form to be used with SARIMA algorithm.

In addition to the seasonal differencing in Eq. (5.9), the implementation of non-seasonal difference may lead to some improvement in ACF. The additional non-seasonal differencing order to Eq. (5.9) is suggested by different studies [81, 82].

The implementation of one non-seasonal difference with  $s = 1$  into Eq. (5.9), makes the multiplicative SARIMA of the form

$$\phi_1(1 - B)\phi_2(1 - B^{24})\phi_3(1 - B^{168})x(n) = \theta_1(1 + B)\theta_2(1 + B^{24})\theta_3(1 + B^{168})a(n) \quad (5.10)$$

Accordingly, when fitting SARIMA models, one must first choose suitable values for the seasonal and non-seasonal differencing schemes.



## 5.5 Summary

The seasonal behavior of the power load data series in NSW, QLD, SA, and VIC is addressed and extensively analyzed. The four data series are showing daily, weekly and annual patterns. The daily and the weekly patterns are clearly defined in the four states, whereas the annual pattern is highly dependent on the seasonal changes in the state. The impact of the anomalous days on the weekly daily and weekly patterns is also studied.

The amount of redundancy in the considered time series is also addressed and the differencing technique as efficient way to reduce it is discussed. In the last part of this chapter, SARIMA algorithm based on seasonal and non-seasonal differencing schemes is introduced and made ready for the comparison with the proposed MFBLP algorithm in Chapter 6.

## CHAPTER 6

### RESULTS AND DISCUSSION

To verify the performance of the MFBLP as a solution to STLF problem, the performance is obtained and compared with other parametric techniques, like Burg, and Box-Jenkins SARIMA and with nonparametric techniques, like artificial neural network (ANN). The optimum parameters for the MFBLP, Burg, Box-Jenkins and ANN are obtained and it is discussed in the section 6.3.

Three years data load demand record collected by NEMMCO in four Australian states, NSW, QLD, SA, and VIC, between the beginning of 2005 and the end of 2007, is used for our study. The first two years hourly data (17520 samples) are used for model extraction with the three parametric techniques and for learning phase for the ANN. The remaining one year of data is used for validation. The three years data is used without any form of manipulation, no exceptions, no corrections for anomalous day, vacation periods or cold snaps or unexpected events or weather. In addition, there are no supplementary subjective adjustments by the control operator or any other expert.

The algorithms are run with raw data and with differenced (differencing) data. The parametric algorithms and the ANN technique are all written in MATLAB 7.2 and run on Pentium 4 machine.

## 6.1 Metrics of Performance

Before the study of the performance of the different algorithms and techniques is initiated, it is necessary to define carefully the used metrics of performance. There are many statistical techniques that describe how well a model fits a given sample of data or how well it extrapolates the data. The first type of error is called the residue error of the model or the goodness-of-fit and the second type is called the extrapolation error or out-of-sample. The commonly used metric to indicate these different types of error is the mean-squared error or mean absolute percentage error (MAPE) [39, 145, 146].

The error in fitting data sample  $y(n)$  or forecasting it, is given by

$$e(n) = y(n) - \hat{y}(n) \quad (6.1)$$

where  $\hat{y}(n)$  is the forecast or fitting of  $y(n)$ . The mean error (ME) value of the function in Eq. (6.1) calculated from  $N$  samples, is given as

$$ME = \frac{1}{N} \sum_{n=1}^N e(n) \quad (6.2)$$

The mean absolute error (MAE) is given by

$$MAE = \frac{1}{N} \sum_{n=1}^N |e(n)| \quad (6.3)$$

The mean squared error (MSE) is given by

$$MSE = \frac{1}{N} \sum_{n=1}^N |e(n)|^2 \quad (6.4)$$

The mean forecast error in Eq. (6.2) is sometimes called the bias of the estimator and if it is of zero value the estimator is called unbiased. On the other hand, the mean squared error in Eq. (6.4) is called the variance of the algorithm. The smaller the bias and the variance values of estimator, the better is its performance. However, estimators with zero variance when  $N$  approaches infinity is called consistent estimators and those of nonzero variance for infinite number of samples are called inconsistent [147, 148].

The relative forecast error in percentage is given by

$$\begin{aligned}\tilde{e}(n) &= \left( \frac{y(n) - \hat{y}(n)}{y(n)} \right) \times 100\% \\ &= \frac{e(n)}{y(n)} \times 100\%\end{aligned}\tag{6.7}$$

The mean percent forecast error (MPE) is given by

$$MPE = \frac{1}{N} \sum_{n=1}^N \tilde{e}(n)\tag{6.8}$$

Accordingly, the mean absolute percent forecast error (MAPE) is given as

$$MAPE = \frac{1}{N} \sum_{n=1}^N |\tilde{e}(n)|\tag{6.9}$$

Practically, the MAPE values are widely used in the literature of power load forecast to indicate the performance of the different techniques [115, 149, 69, 88, 150, 151]. In this chapter we continue with this trend and use MAPE values for performance assessment.

## 6.2 Forecast Horizons of Forecast

Forecast horizon is the period for which a forecast is conducted. The forecast horizon is also called as forecast period or forecast steps ahead. In the chapter and in similar way to chapter 2, three forecast horizons are considered, namely 1 hour, 24 hours and 168 hours. Commonly, this span of horizon is referred to as short term load forecast. In the most of forecast studies, researchers addressed the first two horizons [68, 81, 116] of forecast and some of them extended towards 48 hours [82, 95]. In this study, the horizon of forecast from one or two days to one-week or 168 hours forecast is extended. This would benefit more the utilities' planner in arranging the power generation optimally. In practice the load system planner uses these forecast values in advance to plan for power plants output and to avoid unstable system at the load distribution site [45, 64, 150].

In the first part of this study, the performance of the proposed MFBLP with the four Australian data series is explored. The performance of the MFBLP is assessed with raw and differenced data and the algorithm is implemented with the shown parameters in Table 6.1.

Table 6.1 Data specification and the MFBLP parameters

Estimation data	17520 (two years) hourly samples
Validation data	8760 (one year) hourly samples
Number of segments (Q)	48
Length of the segment (N)	365
Order of the filter (L)	$0.25 \times 365 \approx 90$

Table 6.1 depicts the data specification and the proposed MFBLP parameters. The estimation data are applied to determine the model coefficients. Number of data samples is about two years which are 17520 hourly data samples. The data are gathered between January 2, 2005 and December 31, 2006. Validation data samples are considered unknown to the model. There is no attempt to smooth the data that has been done except the data differencing. The validation data are set to be out-of-sample and the number of data samples is about one year. The validation data are also hourly data samples which was collected between January 1, 2007 and December 31, 2007. Meanwhile, number of segments,  $Q$ , length of the segments,  $N$ , and order of the filter,  $L$ , are set to be 48, 365 and 90 respectively.

### **6.3 The Performance of the MFBLP Algorithm with Raw and Differenced Data**

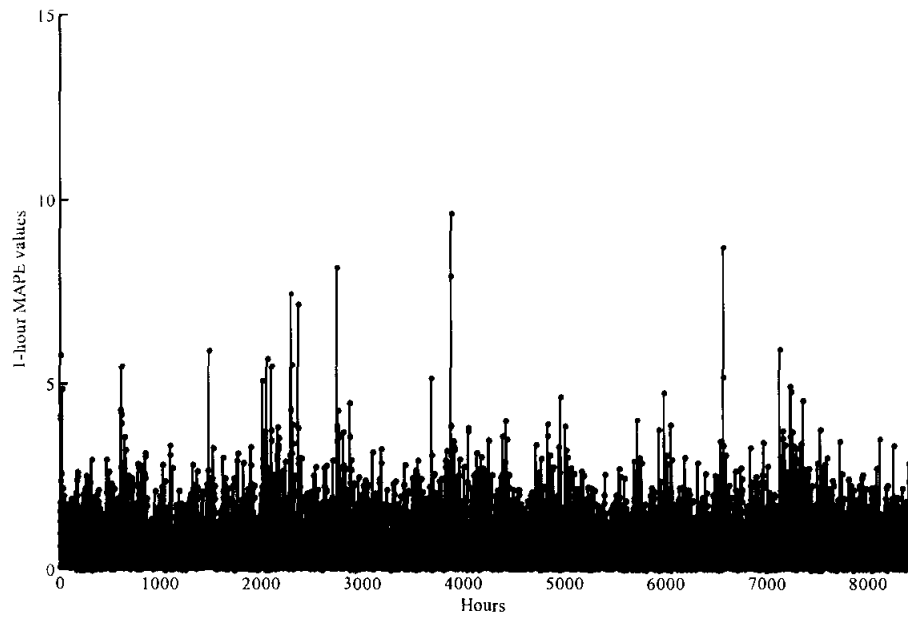
As indicated previously, the difference values between the predicted future data sample and its actual value is referred to as forecast error. The MAPE metric estimated from the predicted data samples over the horizon of prediction is used to indicate the accuracy of the forecast technique.

In the first part of this experiment the four states raw data are implemented with the MFBLP algorithm. Three forecast horizons 1, 24 and 168-hours are considered and the predicted samples for each horizon are calculated and validated over the entire year of 2007. The error is calculated by comparing the actual load data and the predicted one over the different spans of forecast.

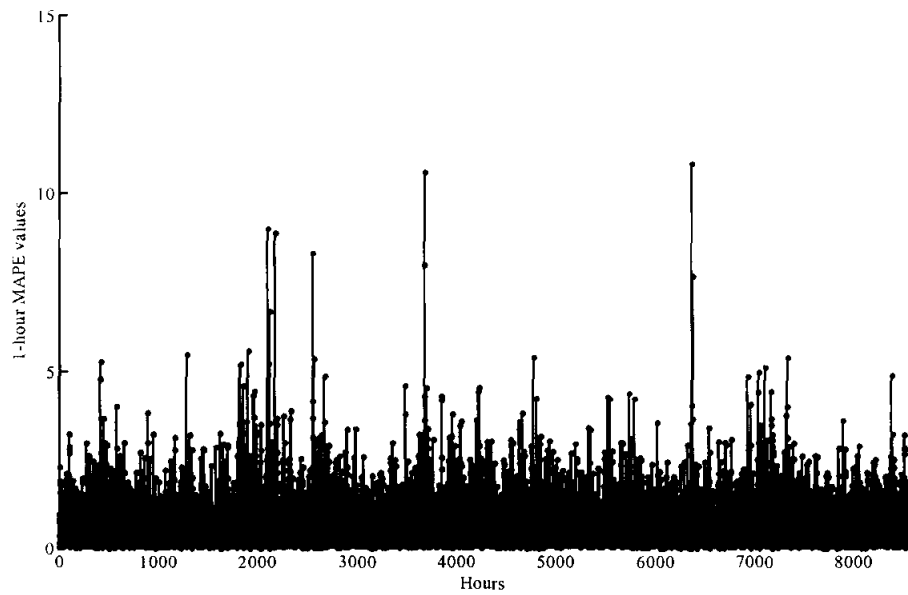
In the second part of this experiment, the performance of the proposed algorithm is investigated with the time differenced data. Data differencing schemes of order 1 ( $d = 1$ ) with 1-hour ahead forecast and of order 24 ( $d = 24$ ) with one-day ahead forecast and of order 168 ( $d = 168$ ) with one-week ahead forecast.

To make it easy to compare, the results for the raw and differenced data are combined together in the same graphs and tables. First, the MFBLP is run for 1-hour forecast with the raw and the differenced data of NSW, QLD, SA, and VIC. The percentage error samples over the validation year for the four Australian states are shown in Figures 6.1, 6.2, 6.3, and 6.4. The 1-hour MAPE value, defined as MAPE value estimated from one data sample, is used to indicate the performance in case of 1-hour ahead forecast.

It is clear from Figure 6.1(a) that the MFBLP manages to forecast the load over the majority of the hours of the validation year with considerably low error ( $\approx 1.0\%$ ). It is also clear that with some data points the forecast is poor so the error is high. This is mainly because of the anomalous days. Figure 6.1(b) shows that the differencing is not of significant effect on the MFBLP performance. The reason is that the forward and backward sub vectors used in the data matrix  $D$ , act as data decorrelator removing considerable part of its redundancy [73].



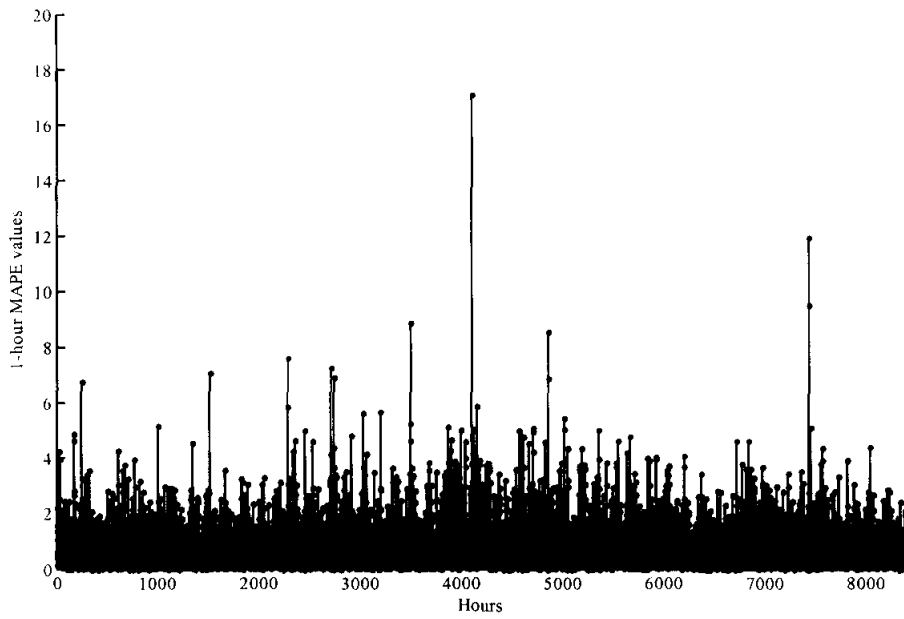
(a)



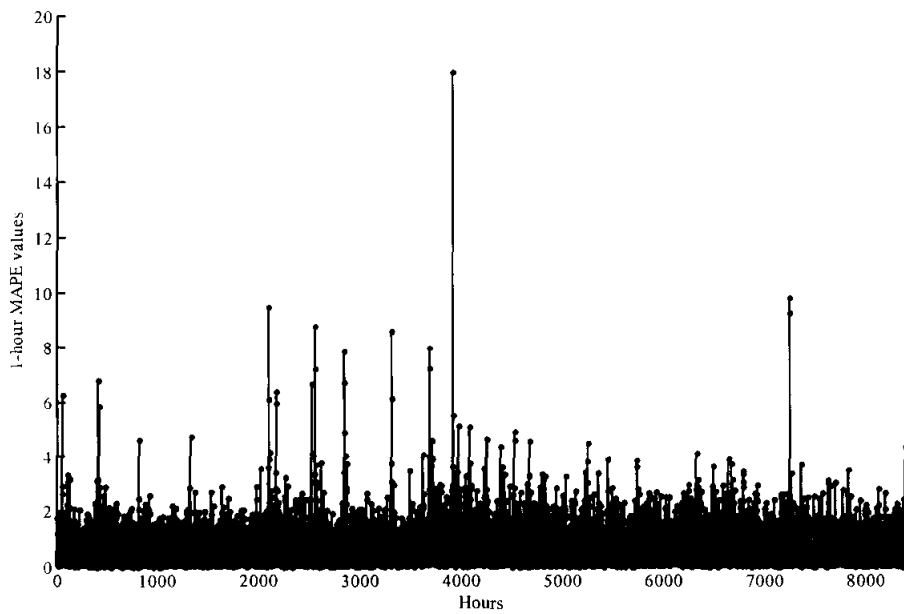
(b)

Figure 6.1 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for NSW. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$





(a)

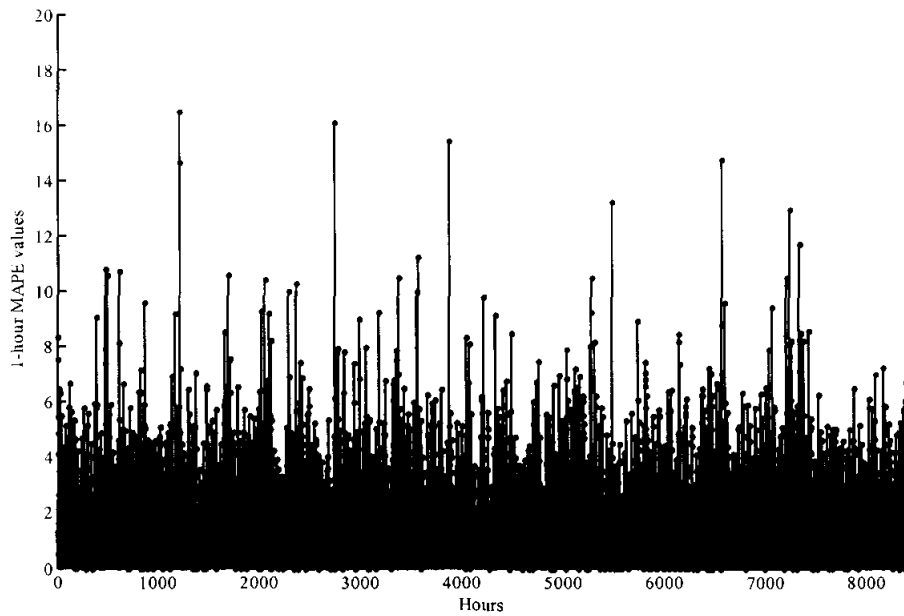


(b)

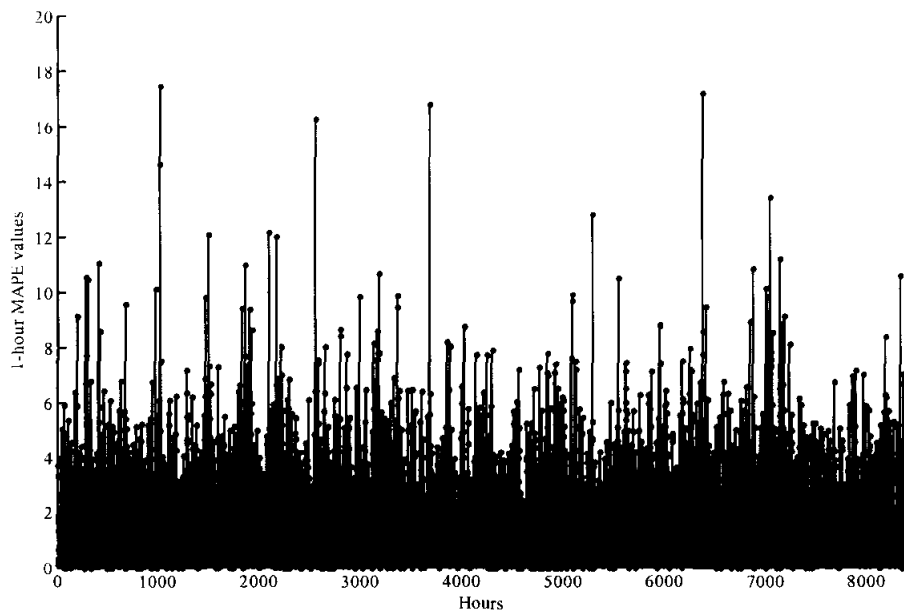
Figure 6.2 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for QLD. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

It is quite clear from Figure 6.2, that the MFBLP is performing better with QLD time series than in NSW. The reason is the less harsh weather in QLD, which leads to less fluctuating data.

Figure 6.3(a) shows that the extreme weather in SA, makes the forecast error of the MFBLP higher than in NSW and QLD ( $\approx 1.5\%$ ). The little effect of the differencing scheme on the performance of the MFBLP algorithm is also evident in Figure 6.3(b).

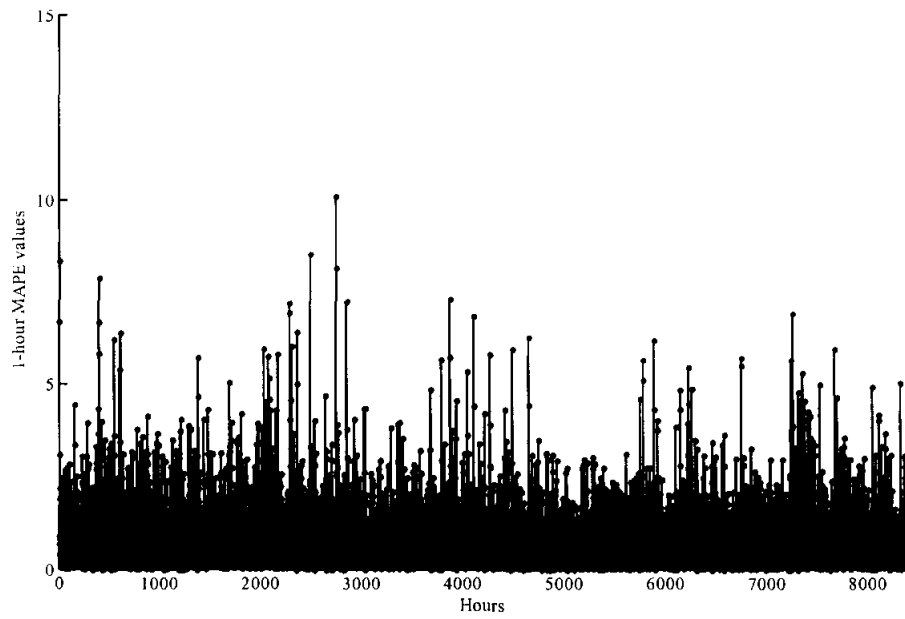


(a)

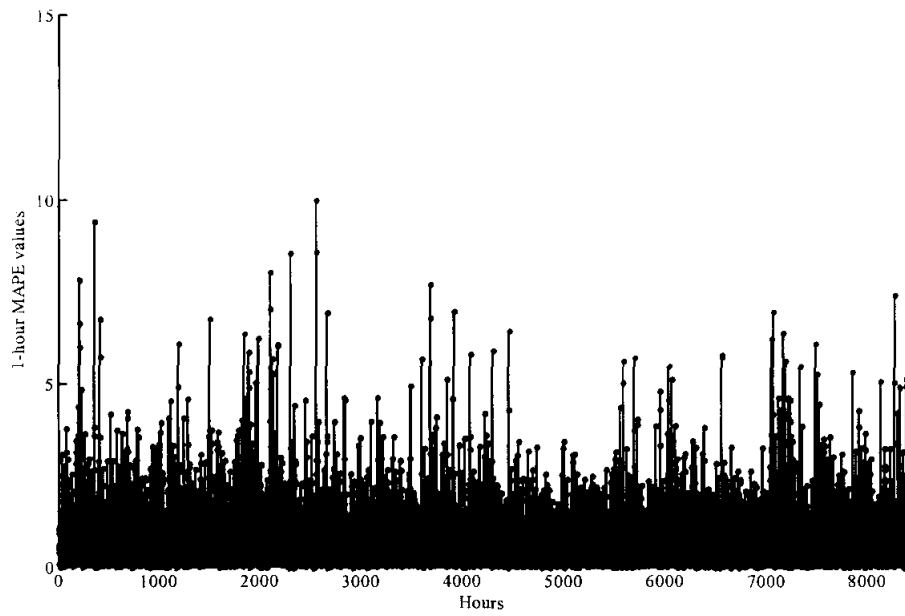


(b)

Figure 6.3 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for SA. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$



(a)



(b)

Figure 6.4 The MAPE values of the MFBLP algorithm in 1-hour forecast horizon for VIC. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

Figure 6.4 shows that the proposed algorithm is capable of forecasting the 1-hour ahead in VIC with reasonable 1-hour MAPE values ( $\approx 1.0\%$ ).

To gain more insight into the results, the overall MAPE value for the proposed algorithm is calculated from the 8760 samples over the year of validation. The results for the four data series are shown in Table 6.2 for the raw and the differenced data.

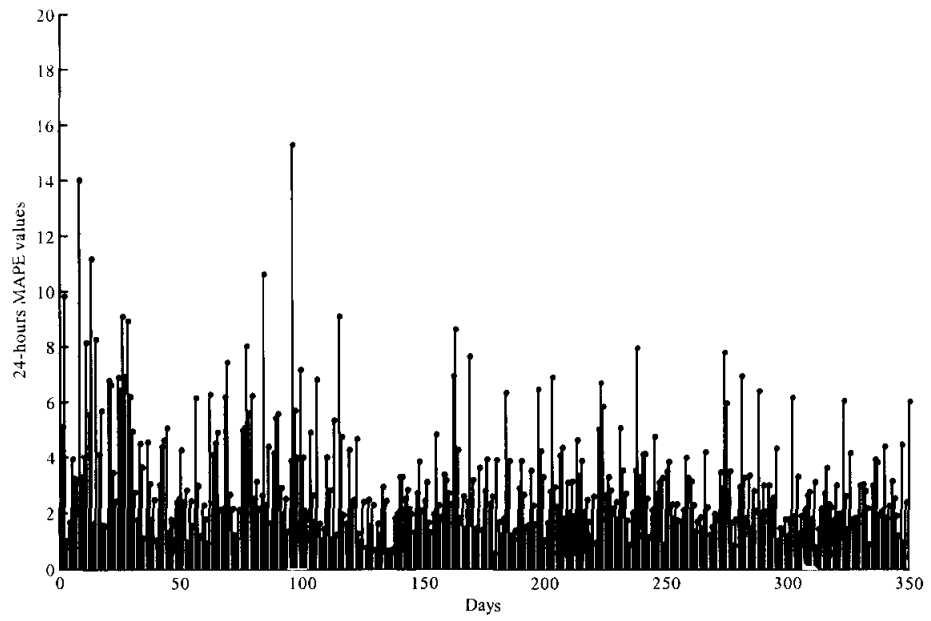
Table 6.2 The MAPE values of the MFBLP in 1-hour forecast scheme

Type of data	MAPE values			
Raw data	0.85	0.78	1.42	0.82
Differenced data	0.83	0.75	1.35	0.79

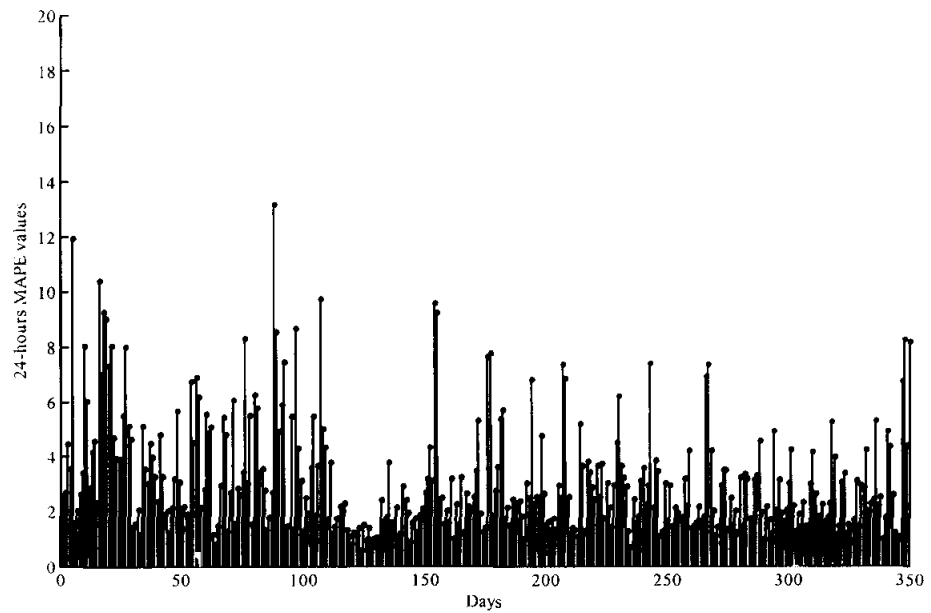
The results in Table 6.2 indicate clearly the accuracy of the proposed MFBLP algorithm in forecasting the data samples. The results also show great amount of independency of the MFBLP from the nature of the power load time series. As indicated before, differencing has no considerable effect on the proposed algorithm. Table 6.2 reflects this fact through approximately similar MAPE values between the raw and the differenced data.

In the second step of this experiment, the forecast level of difficulty is enhanced by extending the forecast horizon from 1-hour to 24-hours. The time differencing scheme of the same order number  $d = 1$ ,  $d = 24$ , and  $d = 168$  is performed. Figures 6.5, 6.6, 6.7, and 6.8 are showing the performance of the MFBLP over the days of the validation year for both raw and differenced data. MAPE values estimated from 24 error samples are used to indicate the performance. This form of MAPE is referred to as 24-MAPE values, and given as

$$24 - MAPE = \frac{1}{24} \sum_{n=1}^{24} |\tilde{e}(n)| \quad (6.10)$$

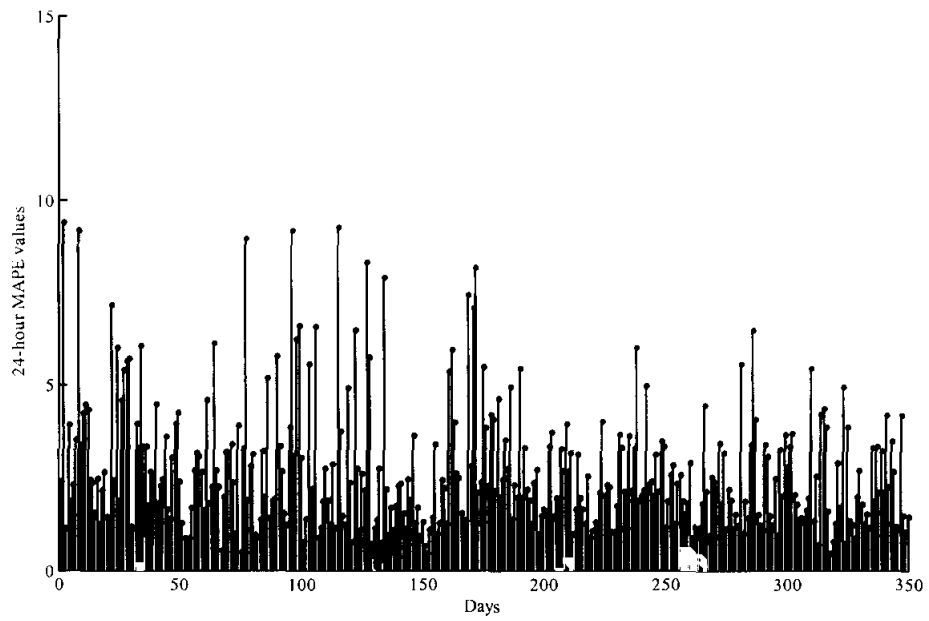


(a)

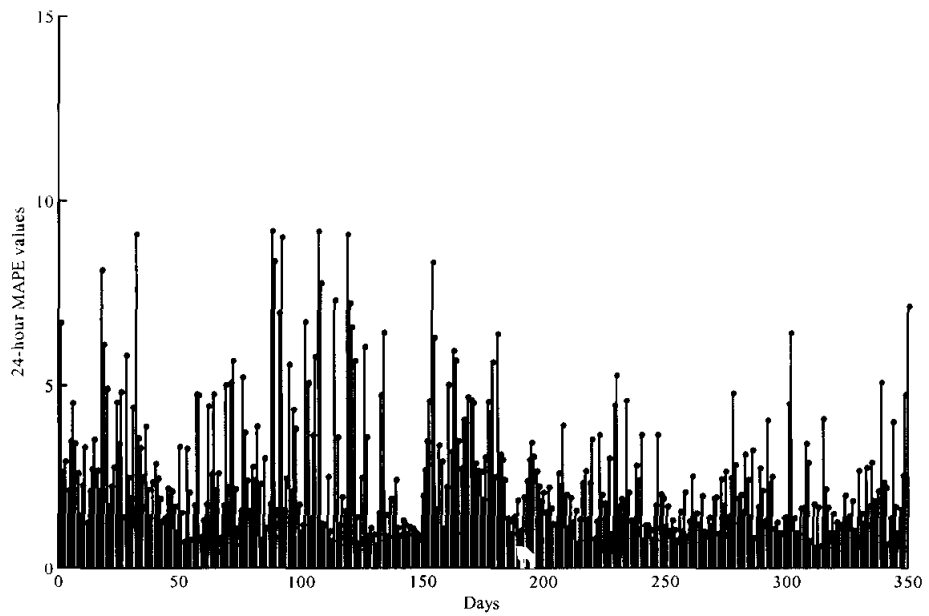


(b)

Figure 6.5 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for NSW. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$



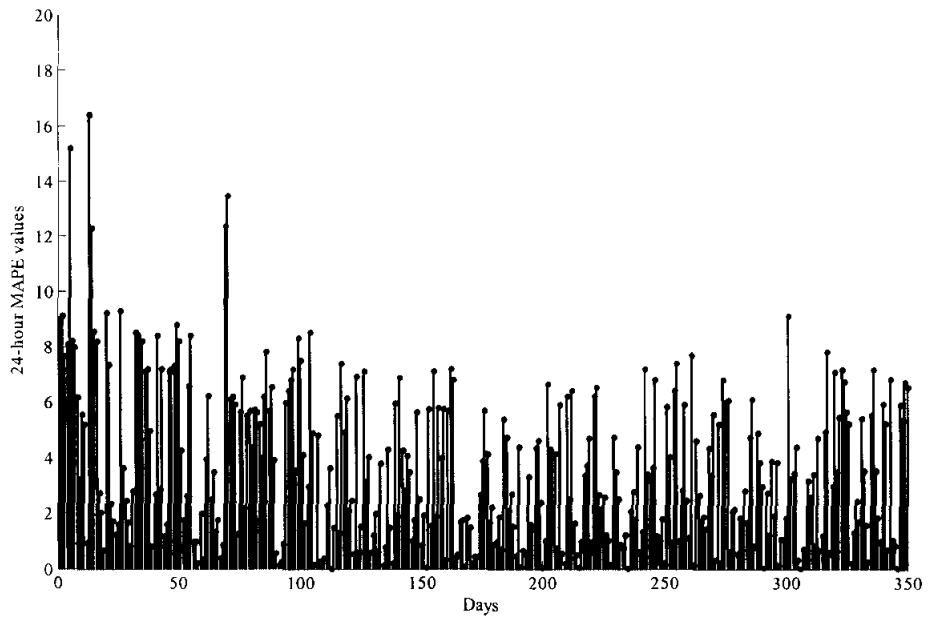
(a)



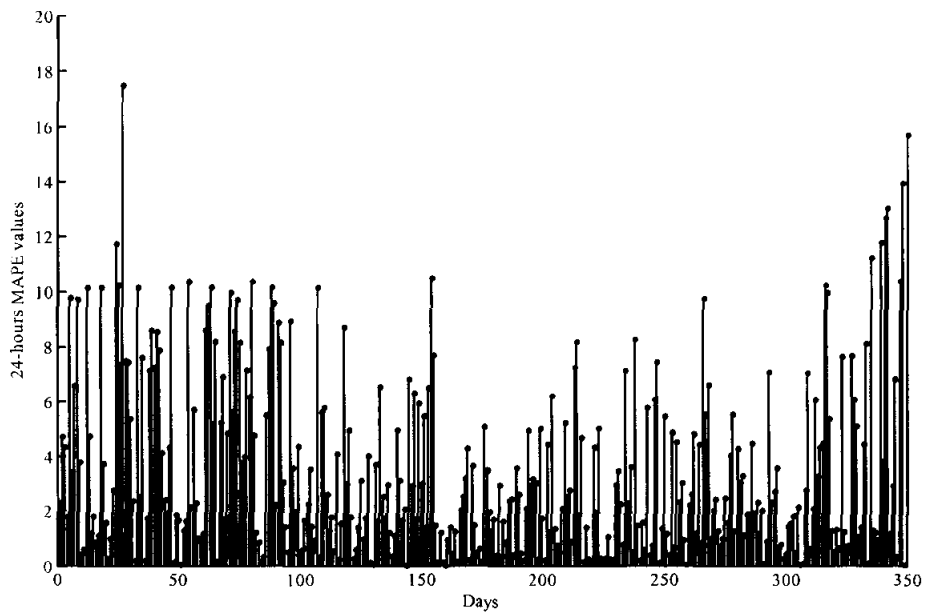
(b)

Figure 6.6 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for QLD. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$



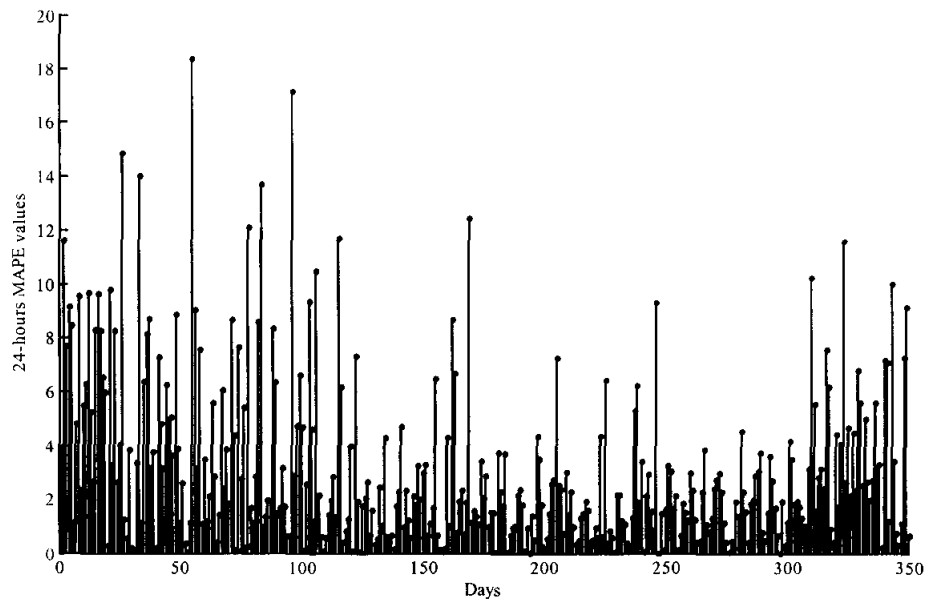


(a)

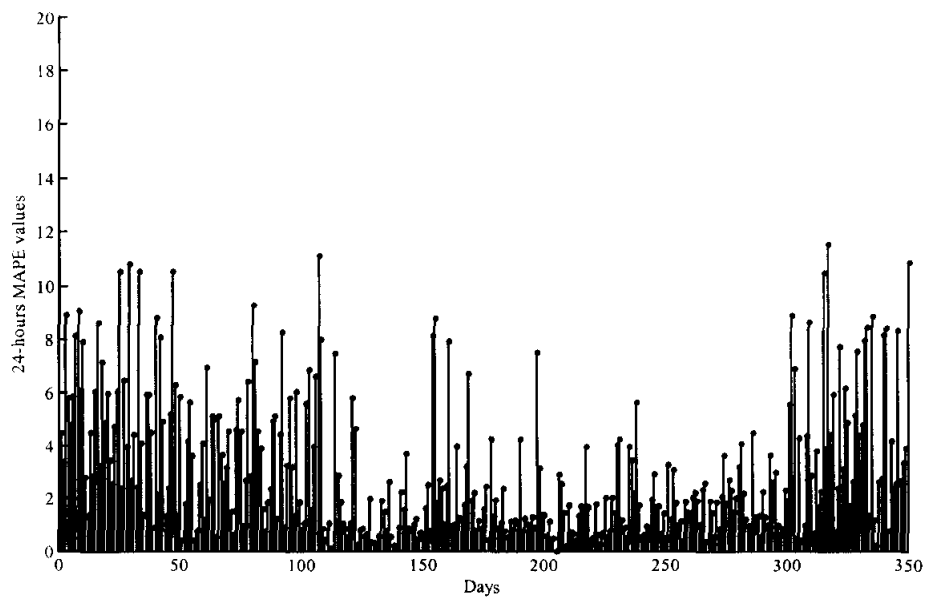


(b)

Figure 6.7 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for SA. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$



(a)



(b)

Figure 6.8 The MAPE values of the MFBLP algorithm in 24-hours forecast horizon for VIC. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

The results in Figures 6.5, 6.6, 6.7 and 6.8 indicate the capability of the proposed algorithm to extend its good estimation towards one-day ahead forecast. The results shows that the MFBLP has managed to keep its 24-hours MAPE values, in average, below 3.5 % for the raw data and below 3.2 for the differenced data. This fact is verified through the calculation of MAPE values over the days of the year. The results are shown in Table 6.3 for the different data series.

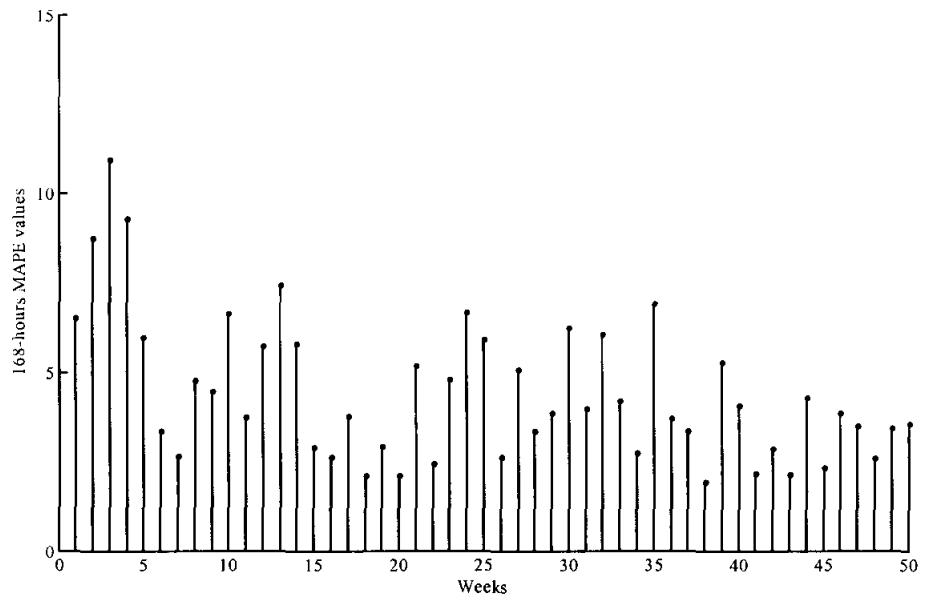
Table 6.3 The MAPE values of the MFBLP in 24-hours forecast scheme

Type of data	MAPE values			
Raw	3.2	3.03	3.36	3.07
Differenced data	3.13	2.95	3.32	3.01

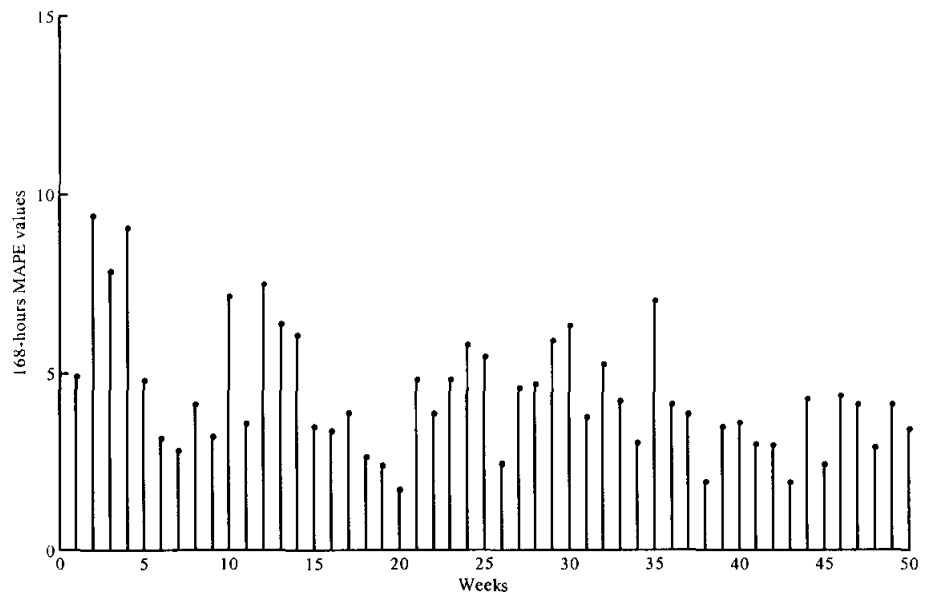
In the third part of this experiment, the forecast horizon to 168 hours or one week ahead is extended. The performance is indicated through the calculation of the MAPE values over the 168 hours of each week. This form of MAPE is referred as 168-MAPE values and it is obtained as

$$168 - MAPE = \frac{1}{168} \sum_{n=1}^{168} |\tilde{e}(n)| \quad (6.11)$$

The 168-MAPE values calculated over the 50 weeks of the year are shown in Figures 6.9, 6.10, 6.11, and 6.12 for both raw and differenced data.

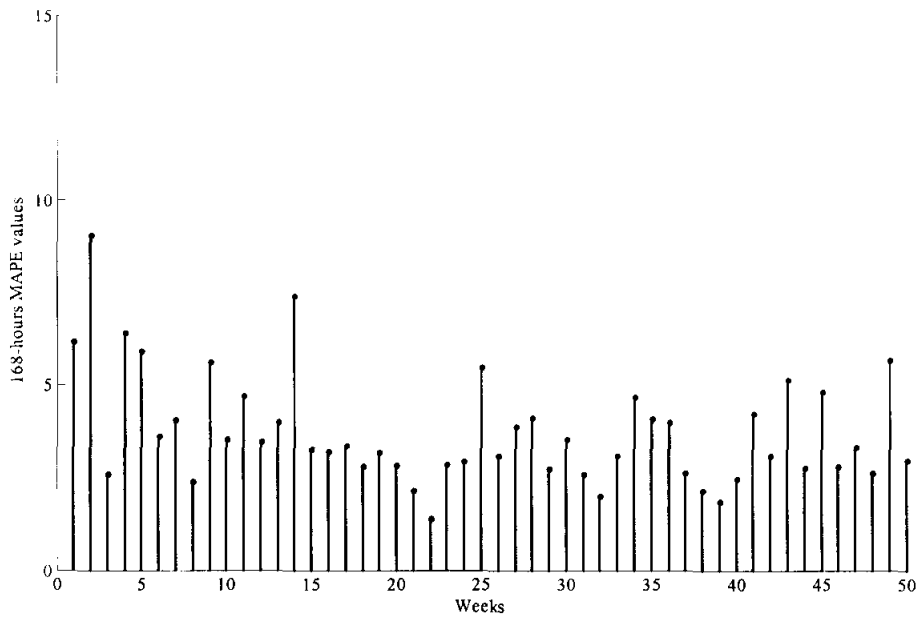


(a)

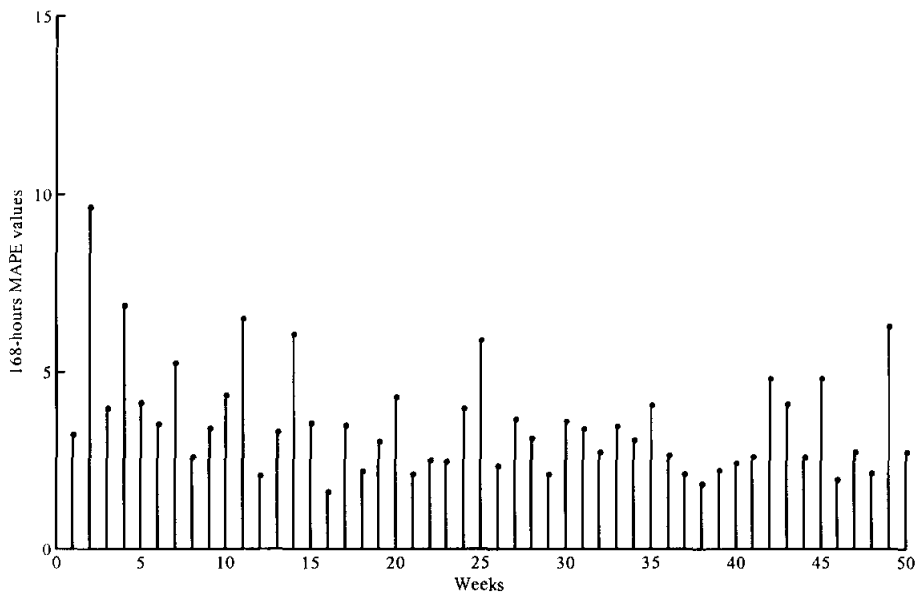


(b)

Figure 6.9 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for NSW. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

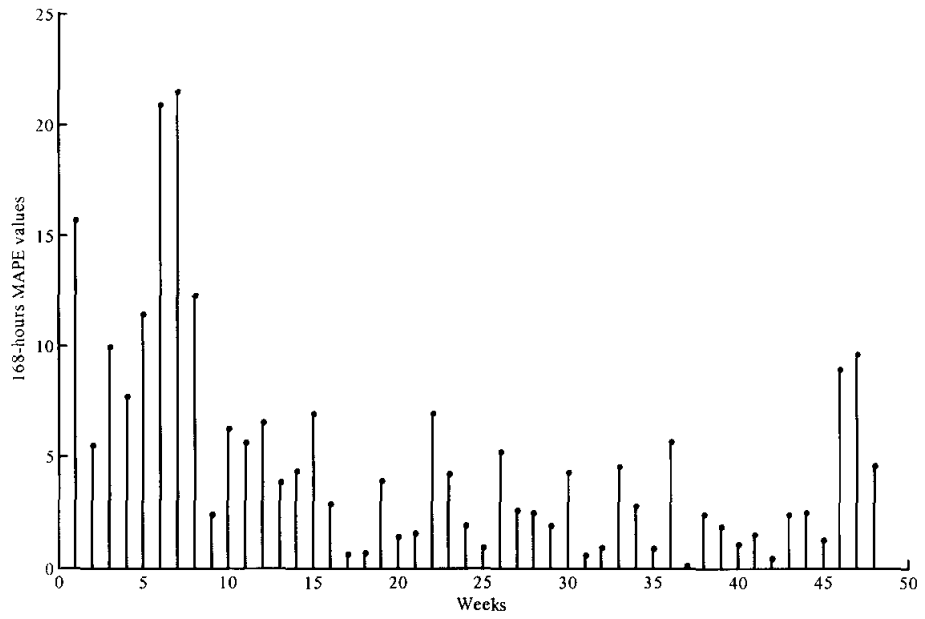


(a)

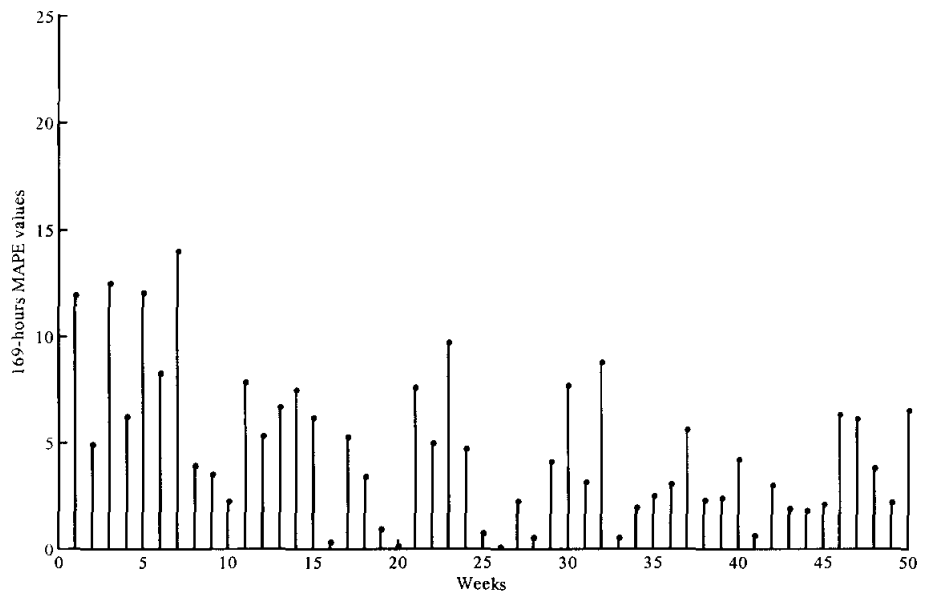


(b)

Figure 6.10 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for QLD. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

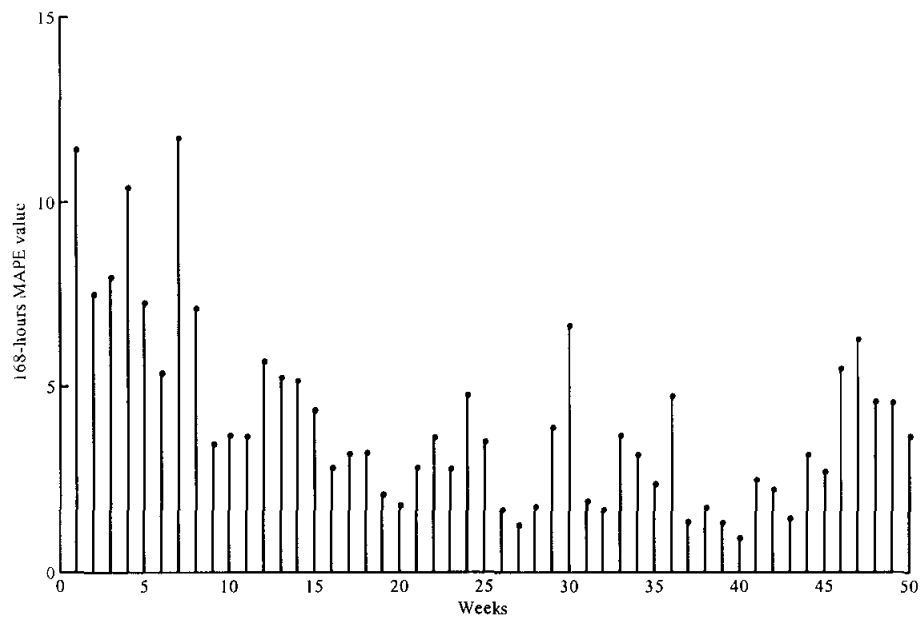


(a)

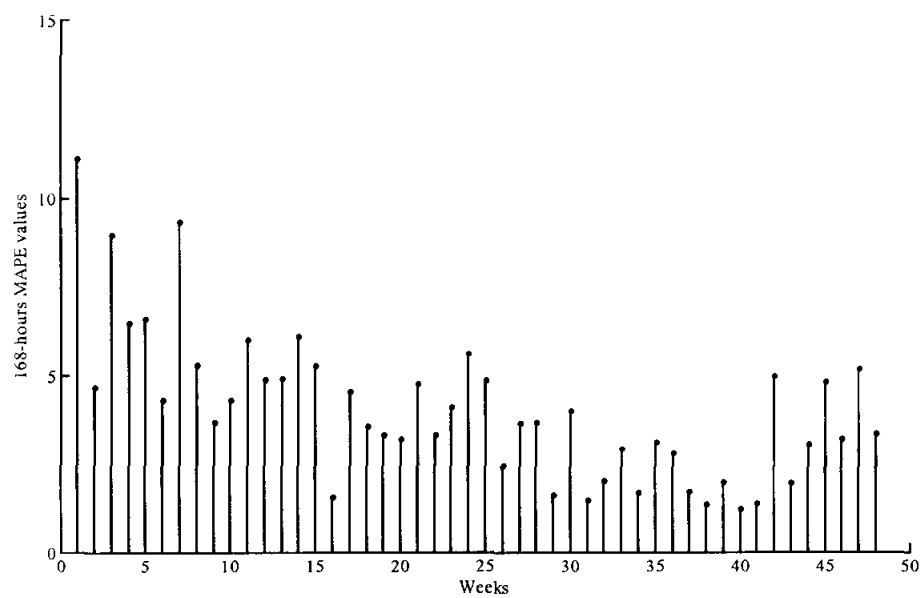


(b)

Figure 6.11 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for SA. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$



(a)



(b)

Figure 6.12 The MAPE values of the MFBLP algorithm in 168-hours forecast horizon for VIC. a) Raw data, b) Differenced data with  $d = 1$ ,  $d = 24$ , and  $d = 168$

The results in Figures 6.9, 6.10, 6.11, and 6.12 show the reliable performance of the MFBLP algorithm in one-week ahead forecast. The algorithm 168-MAPE values are kept below 5% over the most of the 50 weeks of the validation year. As indicated before, differencing has no big impact on the performance of the MFBLP. This is again, because of the implicit decorrelation effect of the forward and backward subsectors, used in the data matrix **D**.

To gain more insight into the overall performances, we find the MAPE values over the 50 weeks. The results are shown for the four data series in Table 6.4.

Table 6.4 The MAPE values of the MFBLP in 168-hours forecast scheme

Type of data	MAPE values			
Raw	4.44	3.75	4.78	4.12
Differenced data	4.42	3.58	4.69	4.05

The MAPE values in Table 6.4 are another indication of the acceptable performance of the proposed MFBLP algorithm in solving the STLTF in one-week horizon of forecast.



In the second experiment, the performance of proposed MFBLP algorithm is compared with Burg and SARIMA as parametric techniques and with the ANN as nonparametric technique.

For Burg's method, it is implemented with order of  $L = 0.2N$ , as it is recommended by the most of literatures in area of forecasting [77, 135]. Box-Jenkins SARIMA model is implemented with the parameters and estimated using the generalized squares method. The values of the estimated parameters are shown in Table 6.5.

Table 6.5 Estimated parameters coefficients for Box-Jenkins SARIMA model

$\phi_1 = 0.752$	$\phi_2 = 0.026$	$\phi_3 = 0.027$
$\theta_1 = 0.938$	$\theta_2 = 0.611$	$\theta_3 = 0.826$

Based on the values given in Table 6.5 the Box-Jenkins SARIMA model becomes

$$\begin{aligned}
 &0.752(1 - B)0.026(1 - B^{24})0.027(1 - B^{168})x(n) = \\
 &0.938(1 - B)0.611(1 - B^{24})0.826(1 - B^{168})a(n)
 \end{aligned}
 \tag{6.12}$$

The ANN as representative of nonparametric techniques is implemented and briefly describe as follows:

For an overview of this field, we refer to, for example [152]. ANN is usually used as nonlinear function predictors. They map an input space for the present and past values of the time series onto an output space or future values. The ANN in the thesis will be of the following feed-forward type.

$$\begin{aligned}
 \hat{x}_n &= f(x_{n-1}, \dots, x_{n-m}) \\
 &= b_0 + \sum_{i=1}^t b_i \tanh \left( a_{i0} + \sum_{j \geq 1} a_{ij} x_{n-j} \right)
 \end{aligned}
 \tag{6.13}$$

Where  $\tanh$  is the tangent hyperbolicus function. It is a nonlinear transformation, with a sigmoid shape. The inputs are  $x_{n-j}$ , with  $j$  running over an index set of not necessarily sequential positive integers ( $n$  denotes the largest lag in the model). These inputs form the so-called input layer. In the second layer, which is referred to as the hidden layer, there are  $m$  nonlinear processing units. These units transform the inputs, by means of the multiplicative weights  $a_{ij}$ , the additive weights  $a_{i0}$  and the sigmoid functions. A weighted sum, with weights  $b_i$ , over the outputs of these hidden units plus a shift,  $b_0$ , produces the final output. The network parameters are estimated by minimizing the error function

$$\sum_{n=m+1}^M [f(x_{n-1}, \dots, x_{n-m}) - x_n]^2 \quad (6.14)$$

where  $N$  is the number of elements in the estimation set. This error function takes all input vectors of the estimation set into account. It takes, of course, several passages over the estimation set to obtain reasonable values for the parameters  $a_{ij}$  and  $b_i$ . More discussion of the method can be referred from the Appendix C.

In the first part of this experiment, the MFBLP is compared with Burg, SARIMA, and ANN using the raw data of the four Australian states. The results are shown in terms of MAPE values in Tables 6.6, 6.7, 6.8, and 6.9.

Table 6.6 The performance of the different techniques with the raw data of NSW

Forecast horizon	MAPE			
1-hour	0.85	1.20	2.80	0.92
24-hours	3.07	3.42	6.25	4.82
168-hours	4.44	4.80	12.28	7.56

Table 6.7 The performance of the different techniques with the raw data of QLD

Forecast horizon	MAPE			
1-hour	0.78	1.12	2.72	0.89
24-hours	3.03	3.31	6.15	4.72
168-hours	3.75	3.95	11.78	7.43

Table 6.8 The performance of the different techniques with the raw data of SA

Forecast horizon	MAPE			
1-hour	1.42	1.65	3.23	1.86
24-hours	3.36	3.85	6.87	4.97
168-hours	4.78	5.35	12.98	7.96

Table 6.9 The performance of the different techniques with the raw data of VIC

Forecast horizon	MAPE			
1-hour	0.82	1.18	2.68	0.91
24-hours	3.07	3.38	6.15	4.81
168-hours	4.12	4.33	12.15	7.42

The results in Tables 6.6, 6.7, 6.8, and 6.9, show clearly the better performance of the proposed MFBLP algorithm to Burg, SARIMA, and ANN with the raw data of the four Australian states over the three horizons of forecast. With the four time series and over the three horizons of forecast, the proposed MFLP algorithm is showing the smallest MAPE values among the considered techniques. With the 1-hour leading time forecast, the ANN shows comparable results to the MFBLP in the four states followed by Burg and SARIMA. But, with the 24-hours and 168-hours of forecast horizons, the ANN deteriorates faster than Burg which

occupies the second place after the proposed MFBLP. SARIMA is performing the worst especially with the third range of forecast, where it shows considerably high values of MAPE. This is mainly because of the seasonality of the raw data in the four Australian states.

In the second part of this experiment, the performances of the different algorithms are tested with differenced data. Three differencing schemes of  $d1 = 1$ ,  $d2 = 24$ , and  $d3 = 168$  are sequentially implemented on the four time series of the data in order to mitigate, respectively, the hourly, the daily, and the weekly patterns. The results are shown in Tables 6.10, 6.11, 6.12, and 6.13.

Table 6.10 The performance of the different techniques with the differenced (differencing) data of NSW

Forecast horizon	MAPE			
1-hour	0.83	0.87	1.10	0.88
24 hours	3.13	3.32	3.48	3.65
168-hours	4.42	4.68	5.12	6.26

Table 6.11 The performance of the different techniques with the differenced (differencing) data of QLD

Forecast horizon	MAPE			
1-hour	0.75	0.81	0.92	0.88
24-hours	2.95	3.27	3.39	3.61
168-hours	3.58	3.75	5.09	6.21

Table 6.12 The performance of the different techniques with the differenced (differencing) data of SA

Forecast horizon	MAPE			
1-hour	0.91	1.01	1.15	1.14
24-hours	3.32	3.42	3.51	3.71
168-hours	4.69	4.86	5.37	6.38

Table 6.13 The performance of the different techniques with the differenced (differencing) data of VIC

Forecast horizon	MAPE			
1-hour	0.79	0.83	0.95	0.94
24-hours	3.01	3.30	3.43	3.68
168-hours	4.05	4.23	5.14	6.25

It is obvious from the results in Tables 6.10, 6.11, 6.12, and 6.13 that reducing the hourly, the daily, and weekly patterns through sequential differencing, has resulted in better performance by the four considered techniques. However, the degree of improvement varies from one technique to another. The main beneficiary of the data differencing scheme is the SARIMA, which achieves about 50% improvement in its performance and instead of occupying the last place in this comparison it moves forward to the third place leading the ANN technique. The MFBLP and Burg algorithms and to some extent the ANN technique, benefit from data differencing by about 10% over their performance with the raw data. This simply means that data differencing is not as important to the MFBLP, Burg, and the ANN as it is to SARIMA. This is because SARIMA is derived based on underlying principal of pattern free data. It is also worth mentioning that with differenced data the considered algorithms are showing more independent behavior of the degree of fluctuation in the original data. This is evident from their relatively similar MAPE values for the four Australian data series.

In the third experiment, the algorithms are tested with seasonally sorted data. The reason is to mitigate the seasonal pattern of the data series. To achieve this goal the four Australian power load demand time series, namely, NSW, QLD, SA, and VIC, are sorted according to the seasons of the Australian year, as shown in Table 6.14.

Table 6.14 The date for the seasons in Australia

Seasons	Duration of time
Summer	December 23 to March 22
Fall	March 23 to June 22
Winter	June 23 to September 22
Spring	September 23 to Decemeber 22

The 17520 hourly data samples over the years 2005 and 2007 (modeling years) are sorted into four groups of seasonal data, each of two similar seasons. This means that the summer group data samples are composed of summer 2005 and summer 2006. The group data samples are the same for the other seasons. By simple calculation we find the number of hourly data samples used to extract the coefficient of the different parametric algorithm and also to train the ANN. The calculation of hourly data samples are as follows:

$$(90 \text{ days} + 90 \text{ days}) \times 24 \text{ hours} = 4320 \text{ samples}$$

The four groups of data over the years 2005 and 2006 are used to forecast their counterpart groups in the validation year 2007. For instance, the 4320 data samples of the summer of 2005 and the summer of 2006 are used to forecast the data over the summer of the validation year 2007. The same applies to other seasonal data groups. The MFBLP is implemented with  $Q = 10$ ,  $N = 432$ , and  $L = 108$ . Burg and SARIMA, are implemented with their optimum order [77, 82, 95, 114].



In this experiment, the MFBLP is compared with Burg, SARIMA, and ANN using the four groups of raw seasonal data. The results for the seasons forecasting of NSW are shown in terms of MAPE values in Tables 6.15, 6.16, 6.17, and 6.18.

Table 6.15 The performance of the different techniques with the summer data of NSW

Forecast horizon	MAPE			
1-hour	1.05	1.32	2.98	2.54
24-hours	3.35	3.65	6.63	6.52
168-hours	4.65	4.95	14.21	13.24

Table 6.16 The performance of the different techniques with the fall data of NSW

Forecast horizon	MAPE			
1-hour	0.78	1.23	2.75	2.37
24-hours	3.05	3.21	6.20	6.12
168-hours	4.41	4.62	12.10	12.03

Table 6.17 The performance of the different techniques with the winter data of NSW

Forecast horizon	MAPE			
1-hour	0.88	1.30	2.92	2.86
24-hours	3.15	3.35	6.50	6.58
168-hours	4.58	4.78	13.58	13.21

Table 6.18 The performance of the different techniques with the spring data of NSW

Forecast horizon	MAPE			
1-hour	0.76	1.28	2.68	2.51
24 hours	3.02	3.25	6.25	6.12
168-hours	4.38	4.65	10.20	10.08

Next, in Tables 6.19, 6.20, 6.21, and 6.22 the performance of the different techniques with seasonally sorted QLD data.

Table 6.19 The performance of the different techniques with the summer data of QLD

Forecast horizon	MAPE			
1-hour	0.89	1.22	2.87	2.51
24-hours	3.25	3.52	6.52	6.45
168-hours	3.98	4.22	13.86	13.11

Table 6.20 The performance of the different techniques with the fall data of QLD

Forecast horizon	MAPE			
1-hour	0.78	1.23	2.62	2.59
24-hours	3.05	3.18	6.08	6.02
168-hours	3.72	3.85	11.97	11.21

Table 6.21 The performance of the different techniques with the winter data of QLD

Forecast horizon	MAPE			
1-hour	0.82	0.98	2.81	2.52
24-hours	3.17	3.28	6.48	6.32
168-hours	3.92	4.12	13.15	12.95

Table 6.22 The performance of the different techniques with the spring data of QLD

Forecast horizon	MAPE			
1-hour	0.77	0.96	2.53	2.45
24-hours	3.03	3.18	6.01	5.88
168-hours	3.70	3.87	9.98	9.78

Then, the experiments are done for the season's data samples of SA. The results for the seasonally sorted SA data are shown in Tables 6.23, 6.24, 6.25, and 6.26

Table 6.23 The performance of the different techniques with the summer data of SA

Forecast horizon	MAPE			
1-hour	1.21	1.41	3.02	3.01
24-hours	3.87	3.92	6.74	6.69
168-hours	5.96	5.98	14.52	14.45

Table 6.24 The performance of the different techniques with the fall data of SA

Forecast horizon	MAPE			
1-hour	1.01	1.35	2.84	2.78
24-hours	3.55	3.68	6.23	6.21
168-hours	5.12	5.18	12.15	12.11

Table 6.25 The performance of the different techniques with the winter data of SA

Forecast horizon	MAPE			
1-hour	1.19	1.30	2.98	2.92
24-hours	3.75	3.85	6.68	6.55
168-hours	5.85	5.92	13.98	13.52

Table 6.26 The performance of the different techniques with the spring data of SA

Forecast horizon	MAPE			
1-hour	0.98	1.12	2.68	2.59
24-hours	3.42	3.55	6.25	6.21
168-hours	5.08	5.12	10.20	10.12

Subsequently, the experiments are done for the season's data samples of VIC. The forecast results for the seasonally sorted VIC data are shown in Tables 6.27, 6.28, 6.29, and 6.30.

Table 6.27 The performance of the different techniques with the summer data of VIC

Forecast horizon	MAPE			
1-hour	0.97	1.27	2.93	2.53
24-hours	3.30	3.59	6.58	6.49
168-hours	4.32	4.59	14.04	13.18

Table 6.28 The performance of the different techniques with the fall data of VIC

Forecast horizon	MAPE			
1-hour	0.78	1.23	2.69	2.48
24-hours	3.05	3.20	6.14	6.07
168-hours	4.07	4.24	12.04	11.62

Table 6.29 The performance of the different techniques with the winter data of VIC

Forecast horizon	MAPE			
1-hour	0.85	1.14	2.87	2.69
24-hours	3.16	3.32	6.49	6.45
168-hours	4.25	4.45	13.37	13.08

Table 6.30 The performance of the different techniques with the spring data of VIC

Forecast horizon	MAPE			
1-hour	0.77	1.12	2.61	2.48
24-hours	3.03	3.22	6.13	6.00
168-hours	4.04	4.26	10.09	9.93

Tables from 6.15 to 6.30 illustrate the MAPE values for the different algorithms with seasonally sorted data. Generally speaking, the MFBLP and Burg are showing small improvement or deterioration over their MAPE values with unsorted data. The reason is the little data samples used to find the parameters of the autoregressive



model (only 4320) and accordingly the less used predictor order. This clearly indicates that if we had more data than two years, the MFBLP and Burg would have come with better results for the sorted data over the unsorted. The improvement achieved by SARIMA with the sorted data is far less than MFBLP and Burg. The reason is the insufficient data used to estimate the parameters of the autoregressive and the moving average of SARIMA. Since the number of SARIMA coefficients is more than MFBLP and Burg, the data shortage because of seasonal sorting is much severe with SARIMA. ANN is showing deterioration in its performance with the sorted data. The reason is insufficient data for the learning process of the neural network. In general, sorting the power load data will lead to improvement in the performance of the forecast techniques, if sufficient data is used.

#### **6.4 Conclusion**

In this chapter, the results for the forecast models are presented. The forecast performances of the MFBLP, Burg, Box-Jenkins SARIMA, and ANN are investigated. Two-year hourly data collected from the Australian states, NSW, QLD, SA, and VIC are used to estimate the autoregressive model coefficients of the MFBLP and Burg algorithms and the autoregressive-moving average models for SARIMA. The same data is used to train the neural network. One-year data is used for validation.

In the first experiment, the algorithms are run with the raw two-year modeling data. The results indicate better performance by the proposed technique MFBLP over the rest. The algorithm shows less MAPE values maintained over the four power load time-series of the Australian states.

In the second experiment, the performances are tested with differenced data in order to mitigate the seasonality. The results show clearly that the beneficiary of the differencing scheme is SARIMA. The MFBLP, Burg and ANN, are not showing that amount of dependency on the differencing scheme as SARIMA. The reason in

case of the MFBLP algorithm is the implicit decorrelation effect of using forward-backward data subsectors in constructing the data matrix  $\mathbf{D}$ .

In the third experiment, the performances are tested with seasonally sorted data. The results show, no significant improvement for all algorithms. The reason is insufficient data to extract the autoregressive model for MFBLP and Burg algorithms and the autoregressive-moving average coefficients for SARIMA.

## CHAPTER 7

### CONCLUSIONS AND RECOMMENDATIONS

#### 7.1 Conclusions

A novel algorithm, MFBLP, for autoregressive model estimation is proposed as a solution to STLF problem. The proposed algorithm divides the power load long data record into  $Q$  shorter segments of  $N$  length each, then uses the least squares criterion to solve for the best AR coefficients. Three years of load demand records collected by NEMMCO-Australia in NSW, QLD, SA, and VIC, between the beginning of 2005 and the end of 2007, were used to find the optimum parameters of the MFBLP, and to verify its solution to the STLF problem. The results show that the best value of  $Q$  with two years data samples (17520 samples) is 48 and the best value for the predictor order ( $L$ ) as a function of segment length ( $N$ ) is  $0.25N$ .

In order to verify the performance of the MFBLP algorithm as a solution to the STLF problem, it was compared with other parametric-based algorithms, like Burg and SARIMA, and with non parametric technique, like ANN. Hourly data of a two year period collected from the Australian states, NSW, QLD, SA, and VIC were used to estimate the autoregressive model coefficients of the MFBLP and Burg algorithms, and the autoregressive moving average models for SARIMA. The same data were used to train the neural network. One-year data were used for validation.

In the first experiment, the algorithms were tested with raw data. The results indicate better performance by the proposed technique MFBLP over Burg, SARIMA and

ANN. The algorithm shows less MAPE values maintained over the four power load time-series of the Australian states. The performance of the MFBLP was verified with three horizons of forecast, namely, 1-hour, 24-hour and 168-hour, and was found to give a better forecast than the other algorithms.

In the second experiment, the performances were tested with differenced data in order to mitigate the seasonality. The results show that the MFBLP was not benefiting too much out of data differencing. The reason is the implicit decorrelation effect of using forward-backward data subsectors in constructing the data matrix  $\mathbf{D}$  of the proposed MFBLP.

In the third experiment, the performances were tested with seasonally sorted data. The results show no significant improvement for all algorithms. The reason is due to insufficient data to extract the autoregressive model for MFBLP and Burg algorithms, and the autoregressive moving average coefficients for SARIMA.

This establishes the MFBLP as one of the best known algorithms for short term power load forecast. The improvement over other widely used algorithms in terms of few percent less in MAPE values could save large amount of money in operating power stations worldwide.

## **7.2 Summary of the Main Contribution**

The main contribution of the research reported in this thesis is the development of the proposed method that is applicable for both raw and differencing data, to forecast the future power demand. It has been proven that from the experimental works, the MFBLP algorithm shows outstanding results. The method is assessed by applying the Australian states power demand data to the experimental works. Subsequently, ascertains the forecast results were ascertained from the proposed method. All states forecast results obtained are reasonably low errors. The experimental forecast results are also in comparison to earlier widely used

methods in STLF. From the comparison, MFBLP algorithm performs significantly better than the other methods.

The development of the MFBLP method is very important in finding solution to the STLF problem. It provides a new approach in forecasting future load demand. One of the advantages of the MFBLP algorithm is such that it only requires single input to implement the forecast model. Furthermore, the algorithm is able to solve forecast problems by using the raw data.

In the thesis, an alternative method to STLF problem has been successfully developed. The MFBLP algorithm has shown a reasonably good forecast result. The overall forecast MAPE values are reduced to some degree in comparison to earlier widely used methods. Hence, in the long-run the improvement of MAPE can contribute in saving the operational cost of the utilities. As stated in the literature [2] that 1% of the forecast accuracy constitutes to 10 million pounds in operating cost per annum.

### **7.3 Recommendations for Future Work**

The most significant areas for further research are listed as the followings:

- i. The proposed MFBLP algorithm is only tested with Australian data, and it is required to test it with other utility load demand systems of different nature.
- ii. Because of its efficiency in modeling the data with autoregressive model, the MFBLP can be used to extract the AR parameters of ARMA model. This will enhance the ARMA-based STLF solutions in terms of lower MAPE values.
- iii. Performance evaluation of the MFBLP algorithms with longer forecast horizon of than 168 hours is needed. This form of forecast is called long term load forecast (LTLF), and is useful for planning of utilities infrastructure

such as building a new power plant or purchasing a new generator to cater for the increase of power load demand.

- iv. The performance of the MFBLP was tested with raw data without any form of data filtering or classification in terms of weekends, public holidays or festival days. It is expected that data filtering will mitigate the variation in the data and makes forecast more accurate. Thus, the performance of the MFBLP with filtered data is required.

## REFERENCES

- [1] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, pp. 1558-1573, 1987.
- [2] D. W. Bunn and E. D. Farmer, *Comparative Models for Electrical Load Forecasting*. United States: John Wiley and Sons Inc., New York, NY, 1985 Jan 01.
- [3] E. Doveh, P. Feigin, D. Greig and L. Hyams, "Experience with FNN models for medium term power demand predictions," *Power Systems, IEEE Transactions on*, vol. 14, pp. 538-546, 1999.
- [4] D. Asber, S. Lefebvre, M. Saad and C. Desbiens, "Modeling of Distribution Loads for Short and Medium-Term Load Forecasting," *Power Engineering Society General Meeting, 2007. IEEE*, pp. 1-5, 2007.
- [5] M. S. Kandil, S. M. El-Debeiky and N. E. Hasanien, "Overview and comparison of long-term forecasting techniques for a fast developing utility: part I," *Electric Power Systems Research*, vol. 58, pp. 11-17, 5/21, 2001.
- [6] M. S. Kandil, S. M. El-Debeiky and N. E. Hasanien, "Long-Term Load Forecasting for Fast-Developing Utility Using a Knowledge-Based Expert System," *Power Engineering Review, IEEE*, vol. 22, pp. 78-78, 2002.
- [7] Qia Ding, "Long-Term Load Forecast Using Decision Tree Method," *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, pp. 1541-1543, 2006.
- [8] S. V. Bajay, "Long-term electricity demand forecasting models: A review of methodologies," *Electric Power Systems Research*, vol. 6, pp. 243-257, 12, 1983.
- [9] H. A. Amarawickrama and L. C. Hunt, "Electricity demand for Sri Lanka: A time series analysis," *Energy*, vol. 33, pp. 724-739, 5, 2008.
- [10] J. Toyoda, Mo-Shing Chen and Y. Inoue, "An Application of State Estimation to Short-Term Load Forecasting, Part I: Forecasting Modeling," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, pp. 1678-1682, 1970.
- [11] T. N. Goh, H. L. Ong and Y. O. Lee, "A study on the accuracy of time series modeling of daily power demand," *Electric Power Systems Research*, vol. 9, pp. 73-77, 6, 1985.
- [12] L. D. Paarmann and M. D. Najjar, "Adaptive online load forecasting via time series modeling," *Electric Power Systems Research*, vol. 32, pp. 219-225, 3, 1995.

- [13] R. E. Abdel-Aal and A. Z. Al-Garni, "Forecasting monthly electric energy consumption in eastern Saudi Arabia using univariate time-series analysis," *Energy*, vol. 22, pp. 1059-1069, 11, 1997.
- [14] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *Power Systems, IEEE Transactions on*, vol. 16, pp. 798-805, 2001.
- [15] M. Espinoza, C. Joye, R. Belmans and B. DeMoor, "Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1622-1630, 2005.
- [16] C. C. Holt. (1957, Forecasting seasonals and trends by exponentially weighted moving averages. *Office of Naval Research Memorandum no. 52, Carnegie Institute of Technology* .
- [17] P. R. Winters. (1960, Forecasting sales by exponentially weighted moving averages. *Management Science*, 6 pp. 235-239.
- [18] W. R. Christiaanse, "Short-Term Load Forecasting Using General Exponential Smoothing," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-90, pp. 900-911, 1971.
- [19] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *Power Systems, IEEE Transactions on*, vol. 4, pp. 1484-1491, 1989.
- [20] E. S. Gardner, "Exponential Smoothing: The state of the art-part II," *International J. Forecasting*, vol. 22, pp. 637-666, 2006.
- [21] J. W. Taylor, "Sort-term electricity demand forecasting using double seasonal smoothing," *Journal of the Operational Research Society*, pp. 799-805, 2003.
- [22] A. Khotanzad, R. Afkhani-Rohani, Tsun-Liang Lu, A. Abaye, M. Davis and D. J. Maratukulam, "ANNSTLF-a neural-network-based electric load forecasting system," *Neural Networks, IEEE Transactions on*, vol. 8, pp. 835-846, 1997.
- [23] A. Khotanzad, R. Afkhani-Rohani and D. Maratukulam, "ANNSTLF-Artificial Neural Network Short-Term Load Forecaster generation three," *Power Systems, IEEE Transactions on*, vol. 13, pp. 1413-1422, 1998.
- [24] I. Drezga and S. Rahman, "Input variable selection for ANN-based short-term load forecasting," *Power Systems, IEEE Transactions on*, vol. 13, pp. 1238-1244, 1998.
- [25] I. Drezga and S. Rahman, "Short-term load forecasting with local ANN predictors," *Power Systems, IEEE Transactions on*, vol. 14, pp. 844-850, 1999.



- [26] G. J. Tsekouras, N. D. Hatziargyriou and E. N. Dialynas, "An optimized adaptive neural network for annual midterm energy forecasting," *Power Systems, IEEE Transactions on*, vol. 21, pp. 385-391, 2006.
- [27] P. Mandal, T. Senjyu, N. Urasaki and T. Funabashi, "A neural network based several-hour-ahead electric load forecasting using similar days approach," *International Journal of Electrical Power & Energy Systems*, vol. 28, pp. 367-373, 7, 2006.
- [28] N. Kandil, R. Wamkeue, M. Saad and S. Georges, "An efficient approach for short term load forecasting using artificial neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 28, pp. 525-530, 10, 2006.
- [29] K. Methaprayoon, Wei-Jen Lee, S. Rasmiddatta, J. R. Liao and R. J. Ross, "Multistage Artificial Neural Network Short-Term Load Forecasting Engine With Front-End Weather Forecast," *Industry Applications, IEEE Transactions on*, vol. 43, pp. 1410-1416, 2007.
- [30] V. H. Ferreira and A. P. Alves da Silva, "Toward Estimating Autonomous Neural Network-Based Electric Load Forecasters," *Power Systems, IEEE Transactions on*, vol. 22, pp. 1554-1562, 2007.
- [31] Zhang Yun, Zhou Quan, Sun Caixin, Lei Shaolan, Liu Yuming and Song Yang, "RBF Neural Network and ANFIS-Based Short-Term Load Forecasting Approach in Real-Time Price Environment," *Power Systems, IEEE Transactions on*, vol. 23, pp. 853-858, 2008.
- [32] T. Senjyu, P. Mandal, K. Uezato and T. Funabashi, "Next day load curve forecasting using recurrent neural network structure," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 151, pp. 388-394, 2004.
- [33] K. Kalaitzakis, G. S. Stavrakakis and E. M. Anagnostakis, "Short-term load forecasting based on artificial neural networks parallel implementation," *Electric Power Systems Research*, vol. 63, pp. 185-196, 10/28, 2002.
- [34] C. Kim, I. Yu and Y. H. Song, "Kohonen neural network and wavelet transform based approach to short-term load forecasting," *Electric Power Systems Research*, vol. 63, pp. 169-176, 10/28, 2002.
- [35] S. Osowski and K. Siwek, "Regularisation of neural networks for improved load forecasting in the power system," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 149, pp. 340-344, 2002.
- [36] L. M. Saini and M. K. Soni, "Artificial neural network based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 149, pp. 578-584, 2002.
- [37] J. Chen, W. Wang and C. Huang, "Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting," *Electric Power Systems Research*, vol. 34, pp. 187-196, 9, 1995.

- [38] Shyh-Jier Huang and Kuang-Rong Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *Power Systems, IEEE Transactions on*, vol. 18, pp. 673-679, 2003.
- [39] Feng Ding, Yang Shi and Tongwen Chen, "Performance analysis of estimation algorithms of nonstationary ARMA processes," *Signal Processing, IEEE Transactions on*, vol. 54, pp. 1041-1053, 2006.
- [40] M. Jachan, G. Matz and F. Hlawatsch, "Time-Frequency ARMA Models and Parameter Estimators for Underspread Nonstationary Random Processes," *Signal Processing, IEEE Transactions on*, vol. 55, pp. 4366-4381, 2007.
- [41] S. S. Pappas, L. Ekonomou, D. C. Karamousantas, G. E. Chatzarakis, S. K. Katsikas and P. Liatsis, "Electricity demand loads modeling using Autoregressive Moving Average (ARMA) models," *Energy*, vol. 33, pp. 1353-1360, 9, 2008.
- [42] H. M. Al-Hamadi and S. A. Soliman, "Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model," *Electric Power Systems Research*, vol. 68, pp. 47-59, 1, 2004.
- [43] H. M. Al-Hamadi and S. A. Soliman, "Fuzzy short-term electric load forecasting using Kalman filter," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 153, pp. 217-227, 2006.
- [44] H. El-Sherief and Y. Abdel-Magid, "An efficient on-line load-modeling algorithm for short-term forecasting of interconnected power systems," *Automatic Control, IEEE Transactions on*, vol. 29, pp. 190-192, 1984.
- [45] S. Chentur Pandian, K. Duraiswamy, C. Christober Asir Rajan and N. Kanagaraj, "Fuzzy approach for short term load forecasting," *Electric Power Systems Research*, vol. 76, pp. 541-548, 4, 2006.
- [46] H. -. Wu and C. Lu, "Automatic fuzzy model identification for short-term load forecast," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 146, pp. 477-482, 1999.
- [47] J. Nazarko and W. Zalewski, "The fuzzy regression approach to peak load estimation in power distribution systems," *Power Systems, IEEE Transactions on*, vol. 14, pp. 809-814, 1999.
- [48] P. A. Mastorocostas, J. B. Theocharis and A. G. Bakirtzis, "Fuzzy modeling for short term load forecasting using the orthogonal least squares method," *Power Systems, IEEE Transactions on*, vol. 14, pp. 29-36, 1999.
- [49] R. -. Liang and C. -. Cheng, "Combined regression-fuzzy approach for short-term load forecasting," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 147, pp. 261-266, 2000.
- [50] A. M. Al-Kandari, S. A. Soliman and M. E. El-Hawary, "Fuzzy short-term electric load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 26, pp. 111-122, 2, 2004.

- [51] G. Li, C. - . Liu, C. Mattson and J. Lawarree, "Day-Ahead Electricity Price Forecasting in a Grid Environment," *Power Systems, IEEE Transactions on*, vol. 22, pp. 266-274, 2007.
- [52] S. E. Papadakis, J. B. Theocharis, S. J. Kiartzis and A. G. Bakirtzis, "A novel approach to short-term load forecasting using fuzzy neural networks," *Power Systems, IEEE Transactions on*, vol. 13, pp. 480-492, 1998.
- [53] D. Srinivasan, S. S. Tan, C. S. Chang and E. K. Chan, "Practical implementation of a hybrid fuzzy neural network for one-day-ahead load forecasting," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 145, pp. 687-692, 1998.
- [54] P. K. Dash, A. C. Liew and S. Rahman, "Fuzzy neural network and fuzzy expert system for load forecasting," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 143, pp. 106-114, 1996.
- [55] A. G. Bakirtzis, J. B. Theocharis, S. J. Kiartzis and K. J. Satsios, "Short term load forecasting using fuzzy neural networks," *Power Systems, IEEE Transactions on*, vol. 10, pp. 1518-1524, 1995.
- [56] P. K. Dash, A. C. Liew and S. Rahman, "Peak load forecasting using a fuzzy neural network," *Electric Power Systems Research*, vol. 32, pp. 19-23, 1, 1995.
- [57] P. K. Dash, G. Ramakrishna, A. C. Liew and S. Rahman, "Fuzzy neural networks for time-series forecasting of electric load," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 142, pp. 535-544, 1995.
- [58] 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," *Evolutionary Computation, IEEE Transactions on*, vol. 10, pp. 330-340, 2006.
- [59] G. Liao and T. Tsao, "Application of fuzzy neural networks and artificial intelligence for load forecasting," *Electric Power Systems Research*, vol. 70, pp. 237-244, 8, 2004.
- [60] S. H. Ling, F. H. F. Leung, H. K. Lam and P. K. S. Tam, "Short-term electric load forecasting based on a neural fuzzy network," *Industrial Electronics, IEEE Transactions on*, vol. 50, pp. 1305-1316, 2003.
- [61] A. Khotanzad, Enwang Zhou and H. Elragal, "A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment," *Power Systems, IEEE Transactions on*, vol. 17, pp. 1273-1282, 2002.
- [62] S. H. Ling, F. H. F. Leung, H. K. Lam, Yim-Shu Lee and P. K. S. Tam, "A novel genetic-algorithm-based neural network for short-term load forecasting," *Industrial Electronics, IEEE Transactions on*, vol. 50, pp. 793-799, 2003.

- [63] T. Maifeld and G. Sheblé, "Short-term load forecasting by a neural network and a refined genetic algorithm," *Electric Power Systems Research*, vol. 31, pp. 147-152, 12, 1994.
- [64] A. A. El Desouky, R. Aggarwal, M. M. Elkateb and F. Li, "Advanced hybrid genetic algorithm for short-term generation scheduling," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 148, pp. 511-517, 2001.
- [65] A. P. A. da Silva and L. S. Moulin, "Confidence intervals for neural network based short-term load forecasting," *Power Systems, IEEE Transactions on*, vol. 15, pp. 1191-1196, 2000.
- [66] Hong-Tzer Yang and Chao-Ming Huang, "A new short-term load forecasting approach using self-organizing fuzzy ARMAX models," *Power Systems, IEEE Transactions on*, vol. 13, pp. 217-225, 1998.
- [67] L. A. Aguirre, D. D. Rodrigues, S. T. Lima and C. B. Martinez, "Dynamical prediction and pattern mapping in short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 30, pp. 73-82, 1, 2008.
- [68] N. Amjady, "Short-Term Bus Load Forecasting of Power Systems by a New Hybrid Method," *Power Systems, IEEE Transactions on*, vol. 22, pp. 333-341, 2007.
- [69] P. Mandal, T. Senjyu, N. Urasaki, T. Funabashi and A. K. Srivastava, "A Novel Approach to Forecast Electricity Price for PJM Using Neural Network and Similar Days Method," *Power Systems, IEEE Transactions on*, vol. 22, pp. 2058-2065, 2007.
- [70] J. N. Fidalgo and J. A. P. Lopes, "Load forecasting performance enhancement when facing anomalous events," *Power Systems, IEEE Transactions on*, vol. 20, pp. 408-415, 2005.
- [71] J. Durbin, "The fitting of time series models," *Rev. Internet. Statist., Vol. 23, Pp. 233-244, 1960*, 1960.
- [72] P. M. T. Broersen, "Autoregressive model orders for Durbin's MA and ARMA estimators," *Signal Processing, IEEE Transactions on*, vol. 48, pp. 2454-2457, 2000.
- [73] N. S. Kamel, "High resolution sonar signal processing for detection of arrivals estimation," June 1993.
- [74] M. H. Hayes, *Statistical Digital Signal Processing and Modeling* Wiley, 1996.
- [75] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. New Jersey: Prentice Hall, 1997.
- [76] A. Swami and J. M. Mendel, "ARMA parameter estimation using only output cumulants," *IEEE Trans. Acoustic, Speech, Signal Processing*, vol. 38, pp. 1257-1265, 1990.

- [77] N. Kamel and Z. Baharudin, "Short term load forecast using Burg autoregressive technique," *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, pp. 912-916, 2007.
- [78] J. Burg, "Maximum entropy spectral analysis," *Proc. 37th Meeting of Soc. of Exploration Geophysicists, Oklahoma City*, October 1967.
- [79] Z. Baharudin and N. S. Kamel, "AR modified covariance in one week ahead short term load forecast," *IASTED Proceeding Power and Energy Systems*, vol. 606-070, 2008.
- [80] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, *Time Series Analysis: Forecasting & Control (3rd Edition)*. New York: Prentice Hall; 3rd edition (February 28, 1994), 1994.
- [81] G. A. Darbellay and M. Slama, "Forecasting the short-term demand for electricity: Do neural networks stand a better chance?" *International Journal of Forecasting*, vol. 16, pp. 71-83, 0, 2000.
- [82] J. W. Taylor, L. M. de Menezes and P. E. McSharry, "A comparison of univariate methods for forecasting electricity demand up to a day ahead," *International J. Forecasting*, vol. 22, pp. 1-16, 2006.
- [83] H. S. Hippert, C. E. Pedreira and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *Power Systems, IEEE Transactions on*, vol. 16, pp. 44-55, 2001.
- [84] J. W. Taylor and R. Buizza, "Neural network load forecasting with weather ensemble predictions," *Power Systems, IEEE Transactions on*, vol. 17, pp. 626-632, 2002.
- [85] R. Billinton and Dange Huang, "Effects of Load Forecast Uncertainty on Bulk Electric System Reliability Evaluation," *Power Systems, IEEE Transactions on*, vol. 23, pp. 418-425, 2008.
- [86] Chao-Ming Huang, Chi-Jen Huang and Ming-Li Wang, "A particle swarm optimization to identifying the ARMAX model for short-term load forecasting," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1126-1133, 2005.
- [87] S. Mirasgedis, Y. Sarafidis, E. Georgopoulou, D. P. Lalas, M. Moschovits, F. Karagiannis and D. Papakonstantinou, "Models for mid-term electricity demand forecasting incorporating weather influences," *Energy*, vol. 31, pp. 208-227, 0, 2006.
- [88] Ching-Lai Hor, S. J. Watson and S. Majithia, "Analyzing the impact of weather variables on monthly electricity demand," *Power Systems, IEEE Transactions on*, vol. 20, pp. 2078-2085, 2005.
- [89] F. Galiana, E. Handschin and A. Fiechter, "Identification of stochastic electric load models from physical data," *Automatic Control, IEEE Transactions on*, vol. 19, pp. 887-893, 1974.

- [90] J. A. Cadzow, "ARMA Time Series Modeling: an Effective Method," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-19, pp. 49-58, 1983.
- [91] Kyung-Bin Song, Young-Sik Baek, Dug Hun Hong and G. Jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," *Power Systems, IEEE Transactions on*, vol. 20, pp. 96-101, 2005.
- [92] W. Charytoniuk, M. S. Chen and P. Van Olinda, "Nonparametric regression based short-term load forecasting," *Power Systems, IEEE Transactions on*, vol. 13, pp. 725-730, 1998.
- [93] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *Power Systems, IEEE Transactions on*, vol. 5, pp. 1535-1547, 1990.
- [94] T. Haida and S. Muto, "Regression based peak load forecasting using a transformation technique," *Power Systems, IEEE Transactions on*, vol. 9, pp. 1788-1794, 1994.
- [95] J. W. Taylor and P. E. McSharry, "Short-Term Load Forecasting Methods: An Evaluation Based on European Data," *Power Systems, IEEE Transactions on*, vol. 22, pp. 2213-2219, 2007.
- [96] A. C. Ferguson, C. H. Davis and J. E. Cavanaugh, "An autoregressive model for analysis of ice sheet elevation change time series," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, pp. 2426-2436, 2004.
- [97] M. S. Kandil, S. M. El-Debeiky and N. E. Hasanien, "Long-term load forecasting for fast developing utility using a knowledge-based expert system," *Power Systems, IEEE Transactions on*, vol. 17, pp. 491-496, 2002.
- [98] M. S. Kandil, S. M. El-Debeiky and N. E. Hasanien, "The implementation of long-term forecasting strategies using a knowledge-based expert system: part-II," *Electric Power Systems Research*, vol. 58, pp. 19-25, 5/21, 2001.
- [99] D. Srinivasan, Swee Sien Tan, C. S. Cheng and Eng Kiat Chan, "Parallel neural network-fuzzy expert system strategy for short-term load forecasting: system implementation and performance evaluation," *Power Systems, IEEE Transactions on*, vol. 14, pp. 1100-1106, 1999.
- [100] Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang and Sung-Hak Kim, "Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems," *Power Systems, IEEE Transactions on*, vol. 10, pp. 1534-1539, 1995.
- [101] S. Rahman, "Formulation and analysis of a rule-based short-term load forecasting algorithm," *Proceedings of the IEEE*, vol. 78, pp. 805-816, 1990.

- [102] 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," *Power Systems, IEEE Transactions on*, vol. 21, pp. 869-876, 2006.
- [103] R. Campo and P. Ruiz, "Adaptive Weather-Sensitive Short Term Load Forecast," *Power Systems, IEEE Transactions on*, vol. 2, pp. 592-598, 1987.
- [104] Bo-Juen Chen, Ming-Wei Chang and Chih-Jen lin, "Load forecasting using support vector Machines: a study on EUNITE competition 2001," *Power Systems, IEEE Transactions on*, vol. 19, pp. 1821-1830, 2004.
- [105] S. Rahman and G. Shrestha, "A priority vector based technique for load forecasting," *Power Systems, IEEE Transactions on*, vol. 6, pp. 1459-1465, 1991.
- [106] H. S. Hippert and C. E. Pedreira, "Estimating temperature profiles for short-term load forecasting: neural networks compared to linear models," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 151, pp. 543-547, 2004.
- [107] Z. Baharudin, R. Ibrahim and N. Saad, "Fuzzy logic approach to short term load forecast," *Proceeding of the 2nd ICAIET, Kota Kinabalu, Malaysia*, 2004.
- [108] T. Senjyu, S. Higa and K. Uezato, "Future load curve shaping based on similarity using fuzzy logic approach," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 145, pp. 375-380, 1998.
- [109] M. -. Chow and H. Tram, "Application of fuzzy logic technology for spatial load forecasting," *Power Systems, IEEE Transactions on*, vol. 12, pp. 1360-1366, 1997.
- [110] R. Liang and C. Cheng, "Short-term load forecasting by a neuro-fuzzy based approach," *International Journal of Electrical Power & Energy Systems*, vol. 24, pp. 103-111, 2, 2002.
- [111] M. Tamimi and R. Egbert, "Short term electric load forecasting via fuzzy neural collaboration," *Electric Power Systems Research*, vol. 56, pp. 243-248, 12/1, 2000.
- [112] A. A. El Desouky and M. M. Elkateb, "Hybrid adaptive techniques for electric-load forecast using ANN and ARIMA," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 147, pp. 213-217, 2000.
- [113] M. El-Telbany and F. El-Karmi, "Short-term forecasting of Jordanian electricity demand using particle swarm optimization," *Electric Power Systems Research*, vol. 78, pp. 425-433, 3, 2008.
- [114] Z. Baharudin and N. S. Kamel, "One week ahead short term load forecasting ," *IASTED Proceeding European Power and Energy Systems*, vol. 582-117, 2007.
- [115] M. Espinoza, J. A. K. Suykens, R. Belmans and B. De Moor, "Electric Load Forecasting," *Control Systems Magazine, IEEE*, vol. 27, pp. 43-57, 2007.

- [116] N. Amjady, "Short-term hourly load forecasting using time-series modeling with peak load estimation capability," *Power Systems, IEEE Transactions on*, vol. 16, pp. 498-505, 2001.
- [117] P. M. T. Broersen and S. de Waele, "Automatic identification of time-series models from long autoregressive models," *Instrumentation and Measurement, IEEE Transactions on*, vol. 54, pp. 1862-1868, 2005.
- [118] P. M. T. Broersen, "ARMAseI for Detection and Correction of Outliers in Univariate Stochastic Data," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, pp. 446-453, 2008.
- [119] G. J. Tsekouras, E. N. Dialynas, N. D. Hatziargyriou and S. Kavatza, "A non-linear multivariable regression model for midterm energy forecasting of power systems," *Electric Power Systems Research*, vol. 77, pp. 1560-1568, 10, 2007.
- [120] J. -. Le Caillec, "Study of the SAR signature of internal waves by nonlinear parametric autoregressive models," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 44, pp. 148-158, 2006.
- [121] J. Yang and J. Stenzel, "Short-term load forecasting with increment regression tree," *Electric Power Systems Research*, vol. 76, pp. 880-888, 6, 2006.
- [122] A. J. R. Reis and A. P. A. da Silva, "Feature extraction via multiresolution analysis for short-term load forecasting," *Power Systems, IEEE Transactions on*, vol. 20, pp. 189-198, 2005.
- [123] T. Senjyu, P. Mandal, K. Uezato and T. Funabashi, "Next day load curve forecasting using hybrid correction method," *Power Systems, IEEE Transactions on*, vol. 20, pp. 102-109, 2005.
- [124] B. Kermanshahi and H. Iwamiya, "Up to year 2020 load forecasting using neural nets," *International Journal of Electrical Power & Energy Systems*, vol. 24, pp. 789-797, 11, 2002.
- [125] O. A. S. Carpinteiro, R. C. Leme, A. C. Z. de Souza, C. A. M. Pinheiro and E. M. Moreira, "Long-term load forecasting via a hierarchical neural model with time integrators," *Electric Power Systems Research*, vol. 77, pp. 371-378, 3, 2007.
- [126] S. Fan and L. Chen, "Short-term load forecasting based on an adaptive hybrid method," *Power Systems, IEEE Transactions on*, vol. 21, pp. 392-401, 2006.
- [127] A. Sfetsos, "Short-term load forecasting with a hybrid clustering algorithm," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 150, pp. 257-262, 2003.
- [128] S. Sargunraj, D. P. S. Gupta and S. Devi, "Short-term load forecasting for demand side management," *Generation, Transmission and Distribution, IEE Proceedings-*, vol. 144, pp. 68-74, 1997.



- [129] A. B. Frakt, H. Lev-Ari and A. S. Willsky, "A generalized Levinson algorithm for covariance extension with application to multiscale autoregressive modeling," *Information Theory, IEEE Transactions on*, vol. 49, pp. 411-424, 2003.
- [130] S. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall, Inc., EnglewoodCliffs, N.J. 1987.
- [131] P. M. T. Broersen and S. de Waele, "Windowed periodograms and moving average models," *Decision and Control, 2000. Proceedings of the 39th IEEE Conference on*, vol. 3, pp. 2706-2709 vol.3, 2000.
- [132] J. Burg, "Maximum entropy spectral analysis," *Ph.D. Dissertation, Stanford University, CA*, May 1975.
- [133] G. E. P. Box and G. C. Reinsel, *Time Series Analysis Forecasting and Control*. San Francisco: Holden Day, 1976.
- [134] S. de Waele and P. M. T. Broersen, "Order selection for vector autoregressive models," *Signal Processing, IEEE Transactions on*, vol. 51, pp. 427-433, 2003.
- [135] P. M. T. Broersen, "Let the Data Speak for Themselves," *Instrumentation and Measurement, IEEE Transactions on*, vol. 56, pp. 814-823, 2007.
- [136] G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Review of linear least squares theory," in *Time Series Analysis Forecasting and Control*, Third Edition ed. Anonymous Prentice Hall, 1994, pp. 286-304.
- [137] M. H. Hayes, "The least squares method," in *Statistical Digital Signal Processing and Modeling*, First ed. Anonymous USA: John Wiley & Sons, Inc., 1996, pp. 131-133.
- [138] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice Hall Inc, Englewood Cliff, 1986.
- [139] S. Vemuri, Wen Liang Huang and D. J. Nelson, "On-Line Algorithms for Forecasting Hourly Loads of an Electric Utility," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-100, pp. 3775-3784, 1981.
- [140] M. T. Hagan and S. M. Behr, "The Time Series Approach to Short Term Load Forecasting," *Power Systems, IEEE Transactions on*, vol. 2, pp. 785-791, 1987.
- [141] G. A. N. Mbamalu and M. E. El-Hawary, "Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation," *Power Systems, IEEE Transactions on*, vol. 8, pp. 343-348, 1993.
- [142] D. N. Dejong and C. H. Whiteman, "Unit roots in U.S. macroeconomic time series: A survey of classical and Bayesian perspectives. In *New directions in Time Series Analysis: Part II*," pp. 43-59, 1993.
- [143] J. D. Hamilton, *Time Series Analysis*. Princeton, NJ: Princeton Univ. Press., 1994.

- [144] J. H. Stock, *Unit Roots, Structural Breaks and Trends. in Hand Book of Econometrics*. Amsterdam: Elsevier, .
- [145] L. Weruaga, "All-Pole Estimation in Spectral Domain," *Signal Processing, IEEE Transactions on*, vol. 55, pp. 4821-4830, 2007.
- [146] M. Athineos and D. P. W. Ellis, "Autoregressive Modeling of Temporal Envelopes," *Signal Processing, IEEE Transactions on*, vol. 55, pp. 5237-5245, 2007.
- [147] H. L. Willis, "Load Forecasting for Distribution Planning-Error and Impact on Design," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-102, pp. 675-686, 1983.
- [148] S. de Waele and P. M. T. Broersen, "Error measures for resampled irregular data," *Instrumentation and Measurement, IEEE Transactions on*, vol. 49, pp. 216-222, 2000.
- [149] H. Akaike, "A new look at the statistical model identification," *Automatic Control, IEEE Transactions on*, vol. 19, pp. 716-723, 1974.
- [150] B. F. Hobbs, S. Jitprapaikularn, S. Konda, V. Chankong, K. A. Loparo and D. J. Maratukulam, "Analysis of the value for unit commitment of improved load forecasts," *Power Systems, IEEE Transactions on*, vol. 14, pp. 1342-1348, 1999.
- [151] R. E. Abdel-Aal, "Short-term hourly load forecasting using abductive networks," *Power Systems, IEEE Transactions on*, vol. 19, pp. 164-173, 2004.
- [152] C. M. Bishop, "Neural Networks for Pattern Recognition," *Oxford University Press, Oxford*, 1995.
- [153] G. Zhang, B. E. Patuwo and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art," *International Journal of Forecasting*, vol. 14,35-62, 1998.

## APPENDIX A

### DATA FILES

This appendix provides a description of data files that have been used in the simulation works. There are four set of data series and the filename data in the extension of .mat are depicted in Table A.1.

Table A.1 The data filename that are used to upload all data at the MATLAB workspace.

Region	Filename
NSW	nsw24odd_05to07.mat
QLD	qld24odd_05to07.mat
SA	sa24odd_05to07.mat
VIC	vic24odd_05to07.mat

In the Table A.2 depicts the details of the main data file that has been described in Table A.1. Each of the data file in the Table A.1 consist of three sets of data.

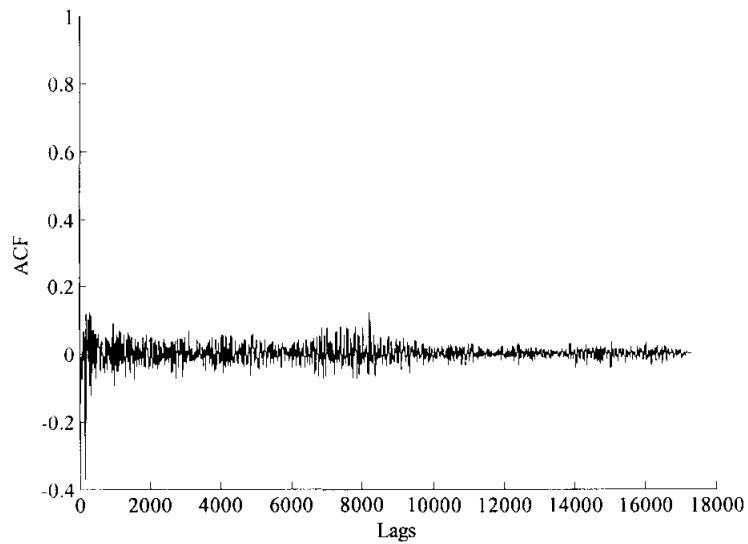
Table A.2 Three sets of data series that are applied in the simulation works.

Filename	Total Data	Data Size	Estimation Data	Data Size	Validation Data	Data Size
nsw24odd_05to07.mat	nsw24odd	26280x1	nsw24odd_series1	17520x1	nsw24odd_07	8400x1
qld24odd_05to07.mat	nsw24odd	26280x1	qld24odd_series1	17520x1	qld24odd_07	8400x1
sa24odd_05to07.mat	nsw24odd	26280x1	sa24odd_series1	17520x1	sa24odd_07	8400x1
vic24odd_05to07.mat	nsw24odd	26280x1	vic24odd_series1	17520x1	vic24odd_07	8400x1

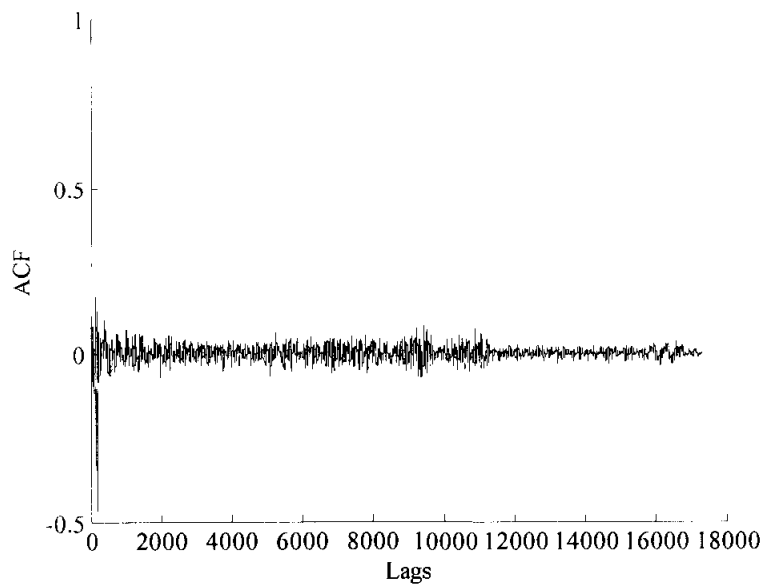
## APPENDIX B

### AUTOCORRELATION FUNCTION (ACF) PLOTS

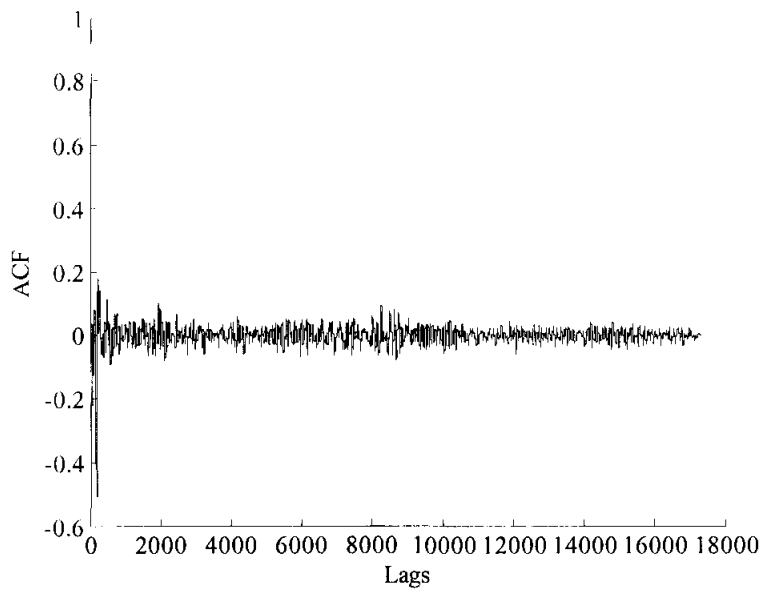
B.1 Autocorrelation Function (ACF) plots for the Eq. (4.9)



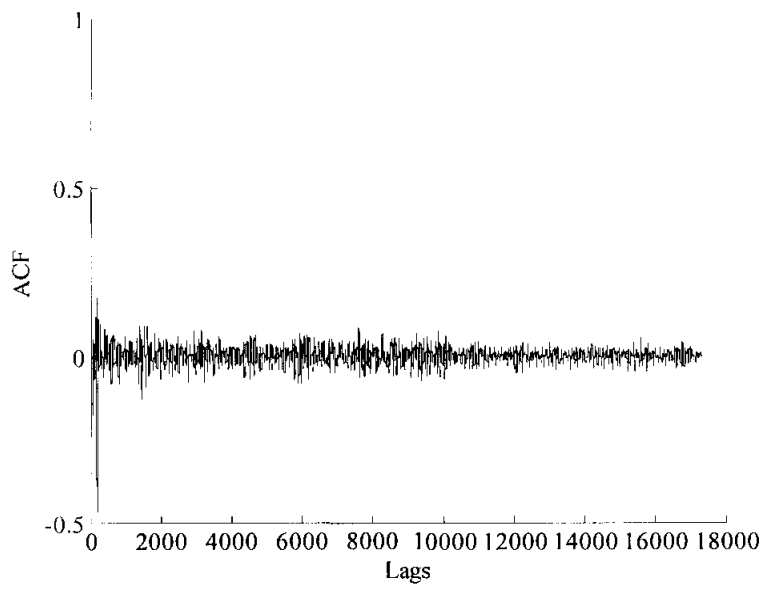
The ACF plots for the NSW data series



The ACF plots for the QLD data series

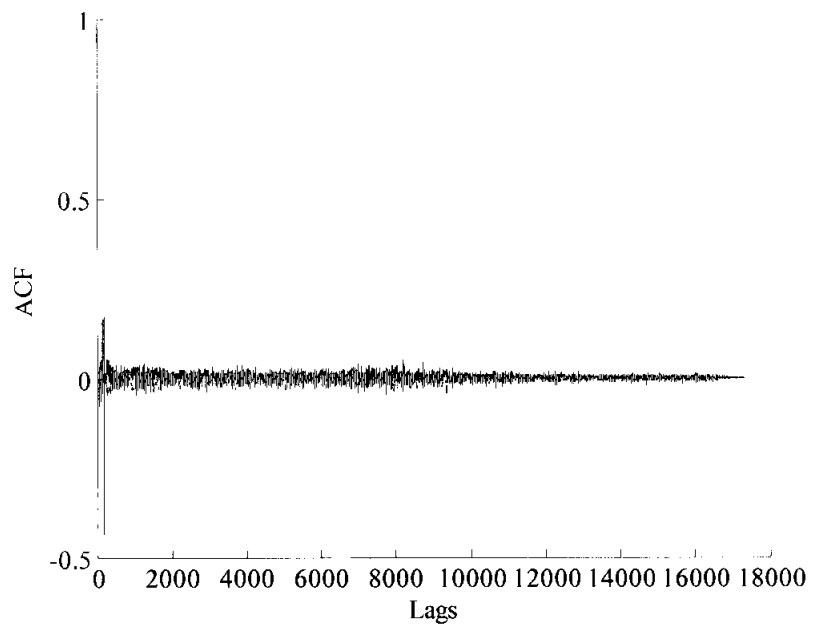


The ACF plots for the SA data series

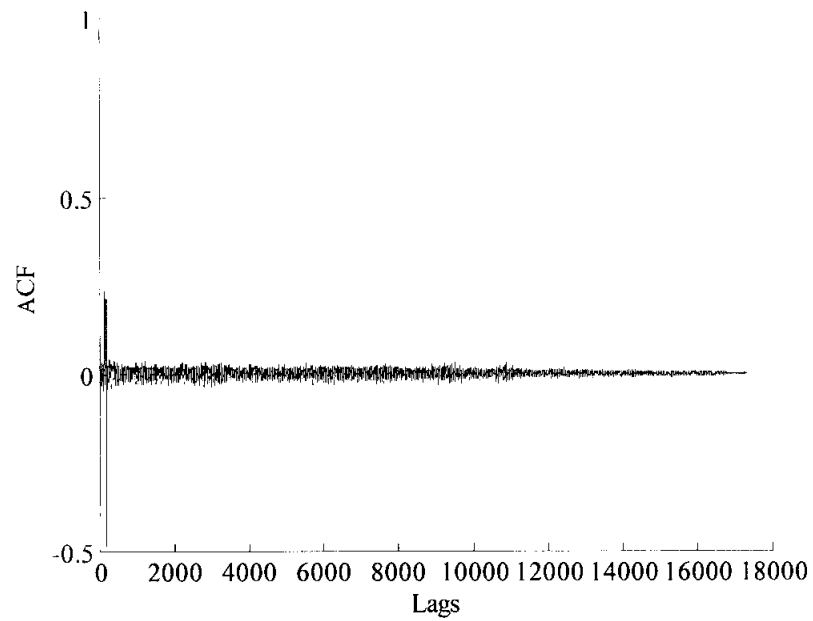


The ACF plots for the VIC data series

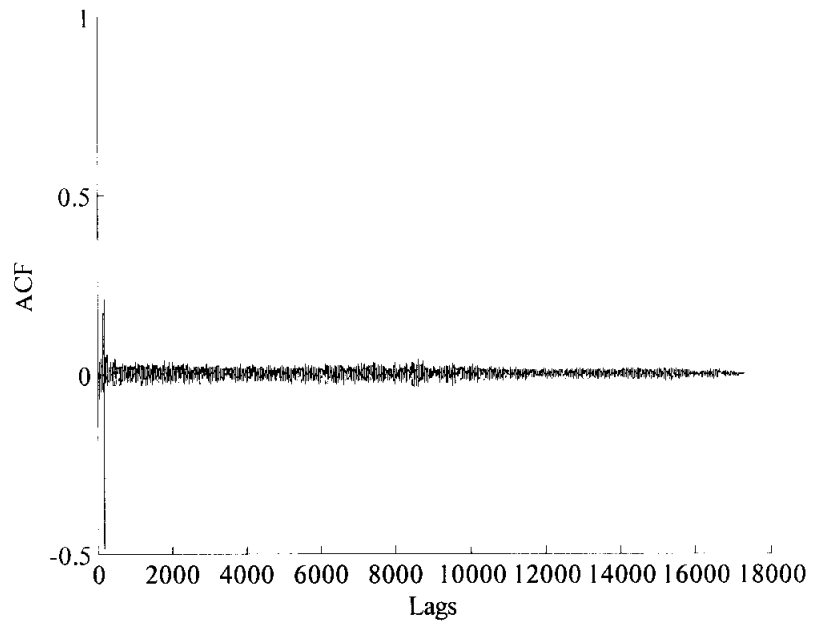
B.2 Autocorrelation Function (ACF) plots for the Eq. (4.10)



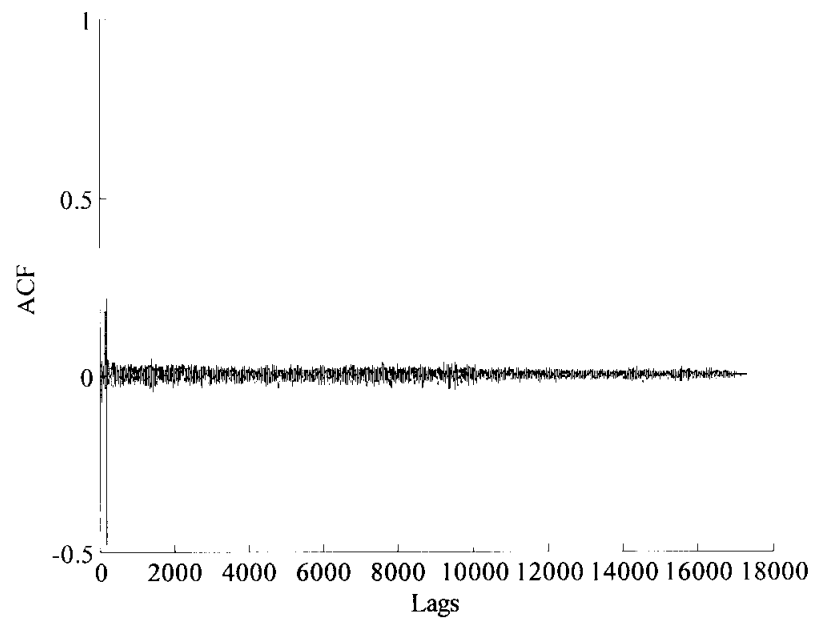
The ACF plots for the NSW data series



The ACF plots for the QLD data series

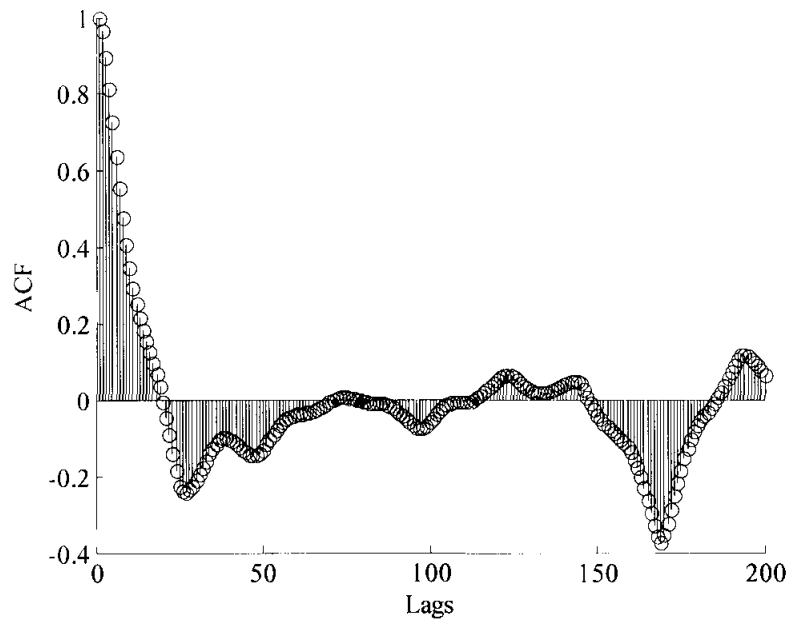


The ACF plots for the SA data series

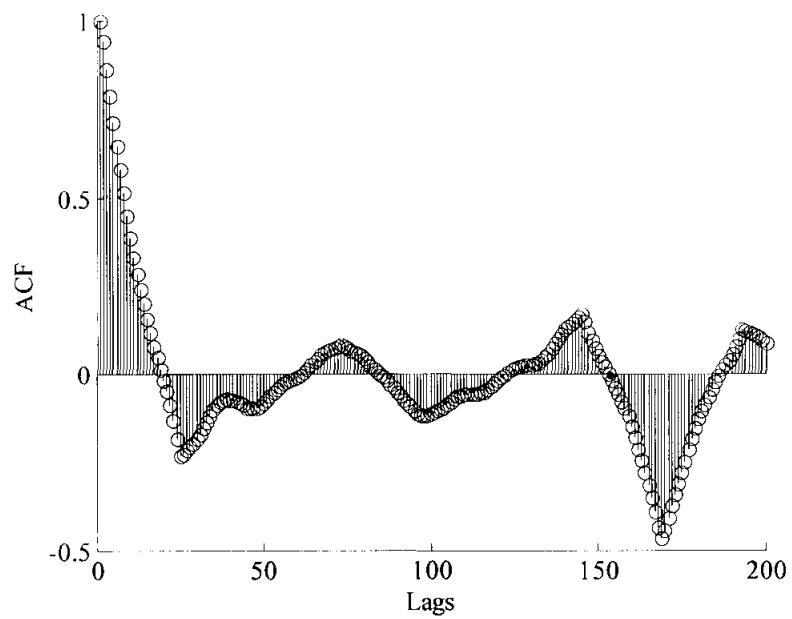


The ACF plots for the VIC data series

B.3 The first 200 samples Autocorrelation Function (ACF) plots for the Eq. (4.10)

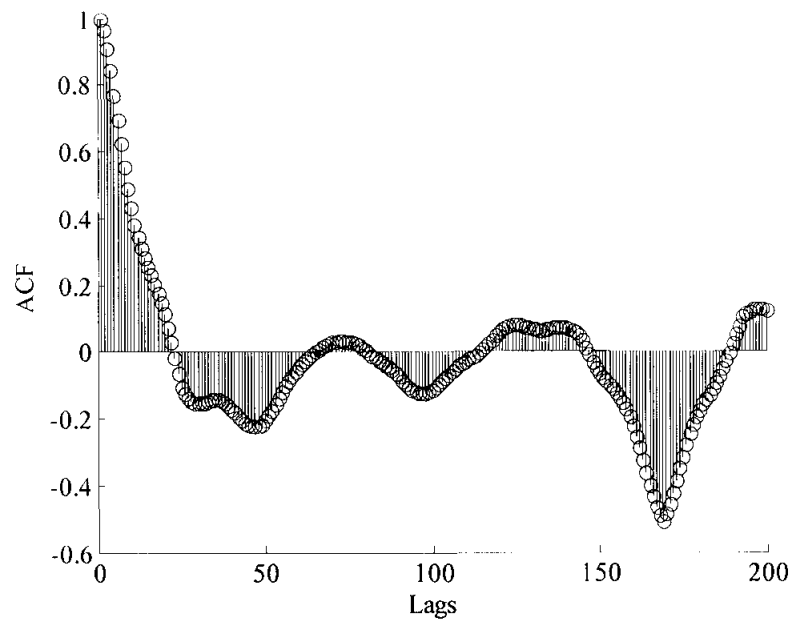


The first 200 samples ACF plots of the NSW data series

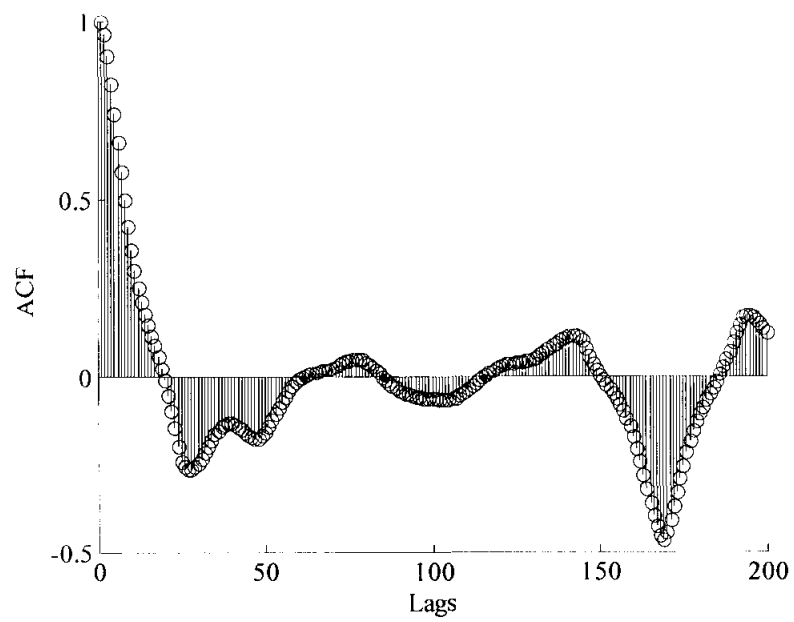


The first 200 samples ACF plots of the QLD data series





The first 200 samples ACF plots of the SA data series



The first 200 samples ACF plots of the VIC data series

## APPENDIX C

### ARTIFICIAL NEURAL NETWORK (ANN)

For an overview of this field, we refer to, for example [152]. A recent review of their use in forecasting is that of [153]. In time series analysis, neural networks are usually used as nonlinear function predictors. They map an input space for the present and past values of the time series onto an output space or future values. The ANN in the thesis will be of the following feed-forward type:

$$\begin{aligned}\hat{x}_n &= f(x_{n-1}, \dots, x_{n-m}) \\ &= b_0 + \sum_{i=1}^l b_i \tanh\left(a_{i0} + \sum_{j \geq 1} a_{ij} x_{n-j}\right)\end{aligned}\tag{C.1}$$

Where  $\tanh$  is the tangent hyperbolicus function. It is a nonlinear transformation, with a sigmoid shape. The inputs are  $x_{n-j}$ , with  $j$  running over an index set of not necessarily sequential positive integers ( $n$  denotes the largest lag in the model). These inputs form the so-called input layer. In the second layer, which is referred to as the hidden layer, there are  $m$  nonlinear processing units. These units transform the inputs, by means of the multiplicative weights  $a_{ij}$ , the additive weights  $a_{i0}$  and the sigmoid functions. A weighted sum, with weights  $b_i$ , over the outputs of these hidden units plus a shift,  $b_0$ , produces the final output. The network parameters are estimated by minimizing the error function

$$\sum_{n=m+1}^M [f(x_{n-1}, \dots, x_{n-m}) - x_n]^2\tag{C.2}$$

where  $N$  is the number of elements in the estimation set. This error function takes all input vectors of the estimation set into account. It takes, of course, several passages over the estimation set to obtain reasonable values for the parameters  $a_{ij}$  and  $b_i$ .

## APPENDIX D

### MATLAB FILES

#### D.1 Raw data

```

tic
clear
clc
%Setting Parameter, Q,m.
Q=;   m=%<=====SPECIFIED   SEGMENTATION   SIZE,
Q*M=17520
mm=m;
load ;%Load data
LF_168step=[];
AL_all=[];
res=[];
UPdL=[];
result=[];
start_pred=input('Enter start week/day/hour number==> ');
end_pred=input('Enter end week/day/hour number==> ');
L=input('Enter # of order==> ');%Order of coefficients
nsw24odd_series1=nsw24odd_series1';
%w=50,168(168 steps ahead): w=350,24(24 steps ahead): w=8400,1(1 step ahead)
step_ahead=%<===== CHANGE HERE
for w=start_pred:1:end_pred%<===== 50 for 168: 350 for 24: 8400 for 1.
    tic
    ssize_start=1+(step_ahead*(w-1));
    ssize_end=17520+(step_ahead*(w-1));
    h1=nsw24odd_series1(ssize_start:ssize_end);
    y=reshape(h1,m,Q)';%<=====SEGMENTATION DATA Q=, M=,
    u = m-L;

```

```

%%ESTIMATION
D = zeros(2*(m-L),L);%Preallocate to improve speed
dfb = zeros(2*(m-L),1);
% disp(['week no ' num2str(w) '==>' ]);%<===== DISPLAY CURRENT WEEK #
disp(['day no ' num2str(w) '==>' ]);%<===== DISPLAY CURRENT DAY #
%disp(['hour no ' num2str(w) '==>' ]);%<===== DISPLAY CURRENT HOUR #
for q = 1:Q
    y1 = y(q,:);
    for r = 1:u,
        for c = 1:L,
            Df(r,c) = y1(1,L-c+r);
            df    = y1(L+1:m).';
            Db(r,c) = conj(y1(1,c+r));
            db    = y1(1:m-L).';
        end
        Dfb = [Df;Db];
        fb = -[df;db];
    end
    D = [D;Dfb];
    dfb = [dfb;fb];
end
D(1:2*(m-L),:)=[];
dfb(1:2*(m-L),:)= [];
f = pinv(D)*dfb;
f = [1;conj(f)];
f(1)=[];
sample_load=(Q*m)+(step_ahead*(w-1));%2*365*24
ss=(sample_load+step_ahead)+(step_ahead*(w-1)+1);%2*365*24
SL=nsw24odd_series1(ssize_start:sample_load);
file=nsw24odd_series1;%Call file for validation with actual load

```

```

estimation_time=toc
%=====FORECASTING=====
for step=1:step_ahead;%<==== STEP AHEAD/FORECAST HORIZONS 1/24/168.
    s=(Q*m)+step-1;
    PdL=SL(s:-1:(s-L+1))*-f;%
    PdL=round(PdL);%
    SL=[SL(1:(Q*m)+step-1) PdL];
    %pause
    %UPdL=SL;%updated predicted diff load
    res=[res PdL];
end
%result=[result;res];
AL_168=nsw24odd_series1(ssize_end+1:ssize_end+(step_ahead));
%Display actual load
AL_all = [AL_all;AL_168];
end
AL_all = AL_all';
AL=reshape(AL_all,1,(step_ahead*(w-start_pred+1)));
er=AL-res;
err=(abs(AL-res)./AL)*100;
err_mape_all=mean(err);
toc
Time=toc

```

## D.2 Differencing Data

```
tic
clear
clc
%Setting Parameter, Q,m.
Q=;   m=%<=====SPECIFIED   SEGMENTATION   SIZE,
Q*M=17520
mm=m;
load ;%Load data
LF_168step=[];
AL_all=[];
res=[];
UPdL=[];
result=[];
start_pred=input('Enter start week/day/hour number==> ');
end_pred=input('Enter end week/day/hour number==> ');
L=input('Enter # of order==> ');%Order of coefficients
nsw24odd_series1=nsw24odd_series1';
%w=50,168(168 steps ahead): w=350,24(24 steps ahead): w=8400,1(1 step ahead)
step_ahead=%<===== CHANGE HERE
for w=start_pred:1:end_pred%<===== 50 for 168: 350 for 24: 8400 for 1.
    tic
    ssize_start=1+(step_ahead*(w-1));
    ssize_end=17520+(step_ahead*(w-1));
    h1=d1d24d168(ssize_start:ssize_end);
    y=reshape(h1,m,Q)';%<=====SEGMENTATION DATA Q=, M=,
    u = m-L;
    %%ESTIMATION=====
    D = zeros(2*(m-L),L);%Preallocate to improve speed
    dfb = zeros(2*(m-L),1);
```

```

%disp(['week no ' num2str(w) '==>' ]);%<== DISPLAY CURRENT WEEK #
%disp(['day no ' num2str(w) '==>' ]);%<===== DISPLAY CURRENT DAY #
disp(['hour no ' num2str(w) '==>' ]);%<===== DISPLAY CURRENT HOUR #
for q = 1:Q
    y1 = y(q,:);
    for r = 1:u,
        for c = 1:L,
            Df(r,c) = y1(1,L-c+r);
            df    = y1(L+1:m).';
            Db(r,c) = conj(y1(1,c+r));
            db    = y1(1:m-L).';
        end
        Dfb = [Df;Db];
        fb = -[df;db];
    end
    D = [D;Dfb];
    dfb = [dfb;fb];
end
D(1:2*(m-L),:)=[];
dfb(1:2*(m-L),:) = [];
f = pinv(D)*dfb;
f = [1;conj(f)];
f(1)=[];
sample_load=(Q*m)+(step_ahead*(w-1));%2*365*24
ss=(sample_load+step_ahead)+(step_ahead*(w-1)+1);%2*365*24
SL=d1d24d168(ssize_start:sample_load);
file=nsw24odd_series1;%Call file for validation with actual load
estimation_time=toc
%=====FORECASTING=====
for step=1:step_ahead;%<== STEP AHEAD/FORECAST HORIZONS 1/24/168.
    s=(Q*m)+step-1;
    PdL=SL(s:-1:(s-L+1))*-f;%

```

```

PdL=round(PdL);%
SL=[SL(1:(Q*m)+step-1) PdL];
res=[res PdL];
end
AL_168=nsw24odd_series1(ssize_end+1+193:ssize_end+193+(step_ahead));
%Display actual load=====
AL_all = [AL_all;AL_168];
end
res2=res';
AL_all = AL_all';
AL=reshape(AL_all,1,(step_ahead*(w-start_pred+1)));
AL=AL';
res_fload=[];
toc
Time=toc

```