UNIVERSITY OF ALBERTA

# A STUDY OF PARAMETER ESTIMATION TECHNIQUES IN SHORT-TERM LOAD FORECASTING

BY

© SHASHI PERSAUD

A Thesis submitted to the Faculty of Graduate Studies and research in partial fulfilment of the requirements for the degree of MASTER OF SCIENCE.

DEPARTMENT OF ELECTRICAL ENGINEERING

Edmonton, Alberta

FALL 1991

ISBN 0-315-70207-9

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **SHASHI PERSAUD**

TITLE OF THESIS: **A STUDY OF PARAMETER ESTIMATION TECHNIQUES IN SHORT-TERM LOAD FORECASTING**

DEGREE: **MASTER OF SCIENCE**

YEAR THIS DEGREE GRANTED: **1991**



Shashi Persaud

24 LEONORA
EAST HOUSING SCHEME
WEST COAST DEMERARA
GUYANA

Date: 16th July 1991

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **A STUDY OF PARAMETER ESTIMATION TECHNIQUES IN SHORT-TERM LOAD FORECASTING** submitted by **SHASHI PERSAUD** in partial fulfilment of the requirements for the degree of **MASTER OF SCIENCE**.

G.S.Christensen

D.H.Kelly

A.E.Peterson

Date: July 10, 1991.

To My Parents

and

To My Wife And Son

## ABSTRACT

In short-term forecasting, the future load on a power system is predicted by extrapolating a predetermined relationship between the load and its influential variables. Determination of this relationship, involves modelling the load as mathematical function of its influential variables, and estimating the coefficients of the model through the use of an efficient parameter estimation technique.

Parameter estimation techniques in short-term load forecasting, can be classified as either static or dynamic, and are generally based on either the least squares or the least absolute value error minimisation criterion. In this thesis, a comparative study of these techniques are presented.

First, the application of the least squares, linear programming and a new least absolute value estimation technique to off-line forecasting are comparatively investigated. Next, the effectiveness of the Kalman filter, a recently developed weighted least absolute value filter and an adaptive general exponential smoothing algorithm, as dynamic on-line forecasting methods are compared.

From the results presented, it will be seen that the new least absolute value technique offers the best choice of a static estimator in all cases, whilst the weighted least absolute value filter is a comparable alternative to the popular Kalman filter.

ACKNOWLEDGEMENT

My sincerest gratitude is extended to Professor G.S.Christensen for the efforts he has undertaken, to make it possible for me to attend this fine University. I would also like to thank Professor Christensen for his help, the many advice and for the patience and understanding has shown during my course of studies.

I would like to express my deepest appreciation to Dr. S.A.Soliman, for the time and effort that he has spent in making me a better researcher. I would also like to thank Dr.Soliman for his many helpful advice and for his frequent invigorating discussions.

I am grateful to Mr.K.El-Naggar, for the many advice and for the time and effort he has spent helping me formulate and debug my many Fortran programs. My thanks are also extended to the Edmonton Power Company and to Mr.T.Thompson for making available the data used in this study.

I would like to extend my thanks to the University of Guyana for offering me this opportunity to continue my studies. Thanks are also due to the Mr. and Mrs. Roopchand Mangru and son Virendra of Leonora, Guyana for making it possible for me to accept this offer.

Finally I would like to thank my wife, son and family for their patience, understanding and support during these past two years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

# CHAPTER I

## INTRODUCTION

In most countries the world over, economic development is directly dependent on the state of available electric energy, especially since most industries depend almost entirely on its use. Such a source of continous, cheap and reliable energy is of foremost economic importance.

Electrical load forecasting is one of the more important strategies used to ensure that the energy supplied by a utility has these qualities. To this end most utilities maintain a staff of trained personnel to carry out this specialised function.

Load forecasting is basically the science or art of predicting the future load on a given system, for a specified period of time ahead. These predictions may be just for a fraction of an hour ahead, or as much as fifty years into the future.

Researchers on the subject have categorized load forecasting into three general subject areas, namely [1,2,3,4]

(a) Long range forecasting which is used to predict loads as distant as fifty years ahead, so that expansion planning can be facilitated.

(b) Medium range forecasting which is used to predict weekly, monthly and yearly peak loads up to ten years ahead, so that efficient operational planning can be carried out, and

(c) Short range forecasting which is used to predict loads up to a week ahead, so that day to day running and dispatching costs can be minimised.

In all categories of load forecasting, load models are developed to mathematically represent the relationship between the load and influential variables such as time, weather, economic factors etc.. The precise relationship between the load and these variables is usually determined through statistical analysis of previous observations.

Correllation analyses are performed to identify suspected influential factors and statistical tests are used to determine their role in the load model. Once the mathematical model is constructed, parameter estimation techniques are then used to determine the coefficients of the terms in the model.

Forecasts can then be made by extrapolating this relationship to the required lead time ahead, given that the corresponding values of influential variables are available or predictable. Since factors such as weather and economic indices are increasingly difficult to accurately predict for longer lead times ahead, the greater the lead time, the less accurate the prediction is likely to be.

The final accuracy of any forecast is thus dependent on the load model employed, the accuracy of predicted variables and the coefficients assigned by the relevant estimation technique. Since different methods of estimation will result in differing values of estimated parameters, it follows that the resulting forecasts will differ in predictive accuracy.

In general, parameter estimation algorithms used in load forecasting have been limited to those based on the least squares minimisation criterion, even though estimation theory indicates that algorithms based on the least absolute value criterion are a viable alternative [5].

This can be attributed to the fact that the least squares principle is quite easy to implement and resultsin efficient and robust estimation algorithms. The least absolute value principle on the other hand, usually requires the use of iterative linear programming methods and as such resulting algorithms tend to be computationally inefficient and not as easy to implement as their least squares based counterparts.

This has been the case until quite recently, when a new non-iterative short cut to least absolute value solutions has opened up new possibilities for least absolute value based algorithms [5]. This new technique claims to be able to produce estimates with efficiencies not far removed from those produced by least squares methods.

In this thesis the primary objective is to compare the accuracy of forecasts made via the least absolute criterion to those obtained via the least squares principle. Also to be compared, is the performance of the new least absolute value technique to that of conventional linear programming methods. The final aim here is to determine the suitability of the new method as a forecasting tool.

The load forecasting problem will be restricted to the short-term case primarily because it is beyond the scope of this thesis to develop load models covering all aspects of load forecasting. It is expected however, that comparisons made will hold true for most aspects of load forecasting, especially since the role of parameter estimation in forecasting is essentially similar

## 1.1: Outline of thesis

The thesis begins in chapter II by introducing the short-term load forecasting problem. Here the general characteristics of load behaviour and the short term modelling techniques applicable to such loads are reviewed.

In chapter III the static parameter estimation problem is presented followed by the theory and development the least squares and least absolute value methods. Chapter IV on the other hand deals with the dynamic estimation problem.

Here the dynamic estimation problem is presented followed by the derivations of the Kalman and a recently developed least absolute value based filter. Also included in this chapter is the formulation of an adaptive general exponential smoothing algorithm that will feature in on-line simulations.

Load models for on and off-line simulations are developed in chapter V. Three off-line models and a single on-line model are used to simulate the responses of the static and dynamic estimation algorithms respectively, and the results are presented in chapter VI. Finally conclusions are drawn and recommendations made in chapter VII.

# CHAPTER II

## SHORT-TERM LOAD FORECASTING

In this chapter, the aspects of short term load forecasting relevant to this thesis are reviewed.

The characteristics of expected load behaviour and the state of the art short-range forecasting techniques are covered in general. Finally the chapter concludes by examining the basic criteria that are used to gauge the design and performance of any short-term load predictor.

### 2.1: Introduction to short-term load forecasting

Short-term load forecasting is an integral part of power system operation that is essential for an inexpensive supply of reliable electric energy. It is used to predict load demands up to a week ahead so that the day to day operation of a power system can be efficiently planned and operating costs minimised.

Short-term forecasting can be subclassified into two general categories, namely on and off-line forecasting. This categorisation, as the names suggest, stem from the areas of application of the load predictors.

Off-line predictors are primarily applied to the scheduling of the large generating units whose "start up" time may vary from a few hours ahead to a few days ahead. The scheduling process is termed unit commitment and ensures that there is sufficient operating generation capacity to meet the variable load demand with specified reliability [6].

Incorrect scheduling by way of poor forecasting can result in excessive daily operational cost as it may necessitate the use of costly quick start units in the event of underscheduling, or alternatively result in the uneconomic operation of large generating units, in the case of overscheduling [6].

In the minute to minute operation of a power system, the economic dispatching of load to the various generating units that make up the generating mix, depends upon calculations minimising a cost function that is subjective to the characteristics of the units. These calculations are based on values of load demand predicted a few hours in advance, and as such the optimum generating mix is dependent upon the accuracy of the on-line forecasts.

Accurate short-term load predictors have long been recognised as a basic necessity for the optimum economic operation of power systems and this has led to many accurate forecasting models being developed. While it is recognised that further sophistication of load models will only result in marginal increases in predictive accuracy, there is much to be gained economically as incremental savings integrated over a year can lead to a significant reduction in capital expenditure on the balance sheets. This is especially important since the cost of electricity is regulated by local government and competition from non-electric energy sources, for example natural gas for heating, cooking, etc., may threaten market share.

## 2.2: Load characteristics

A prerequisite to the development of an accurate load foreasting model, is an in depth understanding of the characteristics of the load to be modelled. This knowledge of load behaviour is gained from experience with the load and through statistical analyses of past load data. Utilities with similar climatic and economic enviroments usually experience similar load behaviour and load models developed for one utility can usually be modified to suit another.

A review of the literature on short-term load modelling indicates that the load supplied by a power system is dynamic in nature and directly reflects the activities and conditions in the surrounding environment. Researchers have found that this load can be separated into a standard or base load, a weather dependent load and a residual load [7,8]. In the following sections the characteristics of each of these load components are reviewed in turn.

### 2.2.1: Standard load

The standard or base load is the largest component of total system load and results directly from the business and economic conditions of the load environment. The base load usually accounts for about ninety percent of total load and can be spectrally decomposed into four distinct components , namely [9,10,11,12,13]

(a) A long-term load that reflects the economic growth of the industry and is usually directly proportional to the growth of the national economy.

(b) A seasonal component that results from the changes in electricity demand from one season to another. In North America this load pattern is characterised by midwinter and midsummer peaks interspaced by troughs occuring during the central spring and autumn seasons.

(c) A weekly load cycle that results from one day of the week being characteristically different the from others. Weekly business cycles and repetitive local activities are the main reasons for this aspect of load behaviour that is characterised by relatively constant midweek demands and smaller weekend loads.

(d) A daily load cycle that result from the basic daily similarity of consumer activities. Low early morning demand peaking to a mid-afternoon high usually characterises this load cycle.

### 2.2.2: Weather dependent load

Load forecasters have always been aware that weather contributes significantly to the dynamics of load demand, and researchers over the years have tried to determine viable relationships between the two so that accurate load models could be developed [7,8,14,15,16].

The effects of weather on load is usually modelled by expressing the load as a linear regression of explanatory meteorological factors such as temperature, windspeed, humidity etc. While it is recognised that an extremely wide variety of explanatory weather variables is required to totally represent the effects of weather, studies have shown that a few basic meteorological factors usually account for most of theweather dependent load.

The specific weather variables that are normally used to model weather dependent load are dry bulb temperature, windspeed, humidity and daylight illumination. Daylight illumination is usually the least significant of these weather variables and since metering is difficult and costly, it is usually omitted from most models.

The general effects of these weather variables on load are summarised below, under their respective headings.

Temperature

In most load enviroments, dry bulb temperature is the most significant weather variable and usually accounts for the largest percentage of weather dependent load. Deviations of temperatures from the norm can result in major changes in the load pattern. These changes however, do not occur immediately, but are rather delayed due to thermal storage in buildings [7].

The effects of temperature are not uniform and can only be specified for a short period of time. This is seen especially at night when changes in temperature have less effect on the load than equivalent changes during the day. In general, the effects of temperature are to increase the heating load when the temperature is below room temperature and to increase the cooling load when it is above.

Temperature effects are usually modelled by considering the load to be a function of the effective temperature or temperature deviation, rather than the actual temperature. This stems from the realisation that the general effects of base temperature are already included in the seasonal load cycle and only deviations from the norm will result in load changes [17,18].

Windspeed

A factor that can contribute significantly to weather dependent load is the wind. Wind effects are especially prevalent during winter and are as a direct consequence of the cooling power of the wind.

The cooling effect of the wind is dependent on both the windspeed and the dry bulb temperature, and has greatest significance at very low values of temperature when the infiltration of cold air into buildings usually results in a high degree of heat losses.

Research has shown [7], that the heat loss from a building is proportional to the product of the square root of the windspeed and the temperature deviation from the comfort level of approximately 18°C. This effect is relatively small in post winter seasons and for simplicity, are usually only included in winter models [7].

Some researchers prefer to use the windchill factor as a representation of the wind in their models, as the windchill factor is often strongly correllated with winter load [19,20]. Others contend that the windchill factor is only a measure of the discomfort level of the wind and temperature and as such, is not a true index for gauging the resulting load response [7,8]. High windchills however, do have the psycological effect of causing people to turn up their thermostats.

Humidity

A weather variable that greatly influences air conditioning and other related cooling loads in summer, is the level of humidity in the atmosphere. The effects of high humidity are generally only noticeable when the temperature is quite high, usually above room temperature.

The effects of humidity can be accounted for by modelling it as a function of either the relative humidity, the temperature humidity index or the dew point temperature, though use of the temperature humidity index is most common [7,8,19,20].

The temperature humidity index is a measure of the discomfort level or equivalent heat stress in summer and depends on both the temperature and relative humidity. It is the variable that normally shows greatest correllation with summer load and only influences the load above a predetermined cutoff temperature.

Illumination

The effects of daylight illumination on load are relatively small compared to those of the other meteorological factors previously discussed, and as such are often omitted from most load models

Low daytime illumination can cause an increase in daytime lighting load and advance the effects of nightfall, thereby altering the evening load pattern [7,8].

The level of luminous radiation recieved at ground level is the term used to model this effect, a factor that is influenced by such weather conditions as cloud cover, surface albido, dust, fog, haze etc..

### 2.2.3: Residual load

The residual load that occurs in load modelling usually accounts for a small percentage of total load and generally results from irregularities in the behaviour of the consuming public. Abnormal consumer demands, though quite frequent in occurence, are very difficult to model and predict and are not accounted for, in most load models.

The common facets of unpredictable load behaviour ranges from public response to major television events, to strikes, storms, disasters, time changes etc., and while these events are usually known well in advance they are not easily predictable and usually different with each occurence [21].

## 2.3: Short-term load forecasting models

Prior to the development of efficient computers, short term load forecasting was based primarily on simple relationships, heuristically derived from experience with the system load. The use of statistical techniques were limited, since early computers were unable to handle the large databases necessary to develop accurate relationships on which modern predictors are based.

The extensive computerisation of the power industry over the past forty years orso, has led to many new techniques developed. Most of these techniques are analytic in nature and based on statistical or signal processing techniques, with the exception of the most recent approach which is based on the principle of artifical intelligience.

Reviewers of short-term load forecasting methods have found that the modelling techniques foremost in use in todays industry, can be categorized as one of the following [4,21,22,23]

(1) Multiple linear regression

(2) General exponential smoothing

(3) Stochastic time series

(4) State space

(5) Expert systems approach.

These techniques are classified on the basis of the underlying mathematical principle employed in load modelling, and each possess distinct advantages and disadvantages over the others. In load forecasting, these approaches may be used individually or combined to increase accuracy. In the following subsections each of these methods are reviewed in turn.

### 2.3.1: Multiple linear regression

The multiple linear regression technique is the oldest of the existing load forecasting methods and was originally applied to short term forecasting by H.A.Dryar in 1944 [7]. In this method, the load is expressed as a function of explanatory weather and non-weather variables that influence the load. The influential variables are identified on the basis of correlation analysis with load, and their significance is determined through statistical tests such as the F and T tests [19,24].

Mathematically the load model is written as

$$y(t) = a_0 + a_1x_1(t) + \dots\dots\dots\dots + a_nx_n(t) + r(t) \qquad (2.1)$$

where $y(t)$ is the load at time t.

$x_1(t),...,x_n(t)$ are explanatory variables

$r(t)$ is the residual load at time t, and

$a_0$, $a_1,...,a_1$ are the regression coefficients relating the load $y(t)$ to the explanatory variables.

The regression coefficients are found using the least squares or least absolute value estimation technique. Past observations of load and explanatory variables are used to set up a system of overdetermined equations, whose minimised solution gives the value of the regression coefficients and completes the mathematical model for future predictions.

The multiple linear regression technique has found greatest application as an off-line forecasting method and is generally unsuitable for on-line forecasting as it requires many external variables that are difficult to introduce into an on-line algorithm.

These models are relatively simple to apply but require extensive initial analysis to identify the regressors and their place in the model. Also because the relationship between the load and weather variables is time specific, this model requires the continuous re-estimation of its coefficients to perform accurately.

### 2.2.2: General exponential smoothing

The theory of general exponential smoothing was originally developed by R.G.Brown (1969), but was first applied to short-term load forecasting by Christiaanse (1970) [25,26].

In this technique the load is modelled using a time dependent fitting function that satisfies the relationship

$$f(t) = Lf(t-1) \qquad (2.2)$$

where $f(t)$ is the fitting function at time t, and

$L$ is a constant matrix called the transition matrix.

Mathematically the model is expressed as

$$y(t) = \beta(t)^T f(t) + r(t) \qquad (2.3)$$

where $y(t)$ = load at time t

$\beta(t)$ = coefficient vector at time t, and

$r(t)$ = residual load or noise at time t.

The coefficient vector is estimated from a data window of previous observations, using the weighted or discounted least squares error minimisation technique.

The estimated coefficient vector is thus obtained by minimising the cost function [25,26]

$$J = \sum_{j=0}^{N-1} w^j [y(N-j) - f^T(-j)\beta]^2 \qquad (2.4)$$

where $0 \leq w \leq 1$ is called the weighting factor, and

$\alpha = 1-w$ is called the smoothing constant

The coefficient vector that minimises the cost function J can be written as

$$\hat{\beta}(N) = F^{-1}(N)h(N) \tag{2.5}$$

where

$$F(N) = \sum_{j=0}^{N-1} w^j f(-j)f^T(-j) \tag{2.6}$$

and

$$h(N) = \sum_{j=0}^{N-1} w^j f(-j)y(N-j) \tag{2.7}$$

The forecast at a lead time l, is then given by

$$\hat{y}(N+l) = f^T(l)\hat{\beta}(N) \tag{2.8}$$

and the coefficients and the forecasts can be updated using

$$\hat{\beta}(N+1) = L^T\hat{\beta}(N) + F^{-1}f(0)[y(N+1)-\hat{y}(N)] \tag{2.9}$$

and

$$\hat{y}(N+1+l) = f^T(l)\hat{\beta}(N+1) \tag{2.10}$$

This method can be used for both on and off-line forecasting, though its recursive nature and generally poor long range accuracy makes it much more suitable for on-line forecasting.

The low accuracy encountered for longer lead times stems from the fact that this technique cannot make use of weather information and as such cannot account for weather related load changes. Simplicity, recursiveness, and economy of use however, makes this method a very attractive forecasting tool.

### 2.2.3: Stochastic time series

The theory of stochastic time series methods was introduced by Weiner(1949) and later applied to load forecasting by Whittle(1968) [27,28]. Further work by Box and Jenkins (1970) established the time series approach as one of the more accurate forecasting methods available, and today it still stands as one of the more popular on-line forecasting techniques [29].

In this method the load is modelled as the output of a linear filter driven by white noise. Depending on the characteristics of the linear filter, different load models can be classified.

The autoregressive and the moving average processes are the two simplest form of stochastic time series and though neither of these processes is usually individually capable of accurately modelling the load, they form the basis for development of more complex processes.

In the autoregressive process the current value of load is expressed linearly in terms of previous values and a random noise. The order of this process depends upon the oldest previous value at which the load is regressed. The moving average process on the other hand expresses the load linearly in terms of current and previous values of a white noise series and again the order of the series depends upon the oldest previous value [19,29].

The autoregressive and the moving average processes are usually combined to give the popular ARMA or auto regressive moving average process which has found widespread use in the power industry. In the ARMA process the load at any instant t, is expressed as a linear combination of its past values and a white noise series and the order of this process is specified by the orders of the AR and MA series included in its composition [19,29].

Time series defined as AR, MA or ARMA are referred to as stationary processes in that their means and covariances are stationary with respect to time. If the process being modelled is nonstationary however, it is first transformed to a stationary series before being modeled as AR, MA or ARMA.

Stationarity of a nonstationary process is accomplished by the method of differencing and the order of a differenced process refers to the amount of times the process has been differenced before achieving stationarity. Differenced processes modelled as AR, MA or ARMA are now called integrated processes and are relabelled ARI, IMA and ARIMA [19,29].

The autoregressive integrated moving average or ARIMA process, like the ARMA process is a very popular load modelling technique that produces very accurate on line forecasts. For longer lead times ahead however a seasonal or periodic component must be included into these processes. This results in what is known as a seasonal process and the abbreviations SARMA and SARIMA are now used [30].

The lack of weather input into time series models usually limits their forecast lead time but by expressing these processes in transfer function form however, it becomes possible to input some weather information. This is usually limited to the single most influential variable, that is temperature, which generally accounts for most of weather induced load [19,29].

The popularity of the stochastic time series approach to on-line forecasting stems mainly from the level of accuracy available and their ease of on-line implementation. Identification of the time series models is a major disadvantage however, as the identification process requires extensive analysis of raw load data through the use of range-mean ,correlation and autocorrelation analyses.

### 2.2.4 State space methods

The application of state space formulation to short-term load forecasting was initiated by Toyoda et al.(1970), when they investigated the applicability of state estimation to time series forecasting [31]. Sharma and Mahalanabis (1972), later showed that this technique could be used to improve the forecasts of a general exponential smoothing algorithm [32].

This is a general forecasting approach that can include any of the previously discussed methods, though in short range forecasting, time series methods are most common.

In the state space method, the load is modeled as a function of state variables using state space formulation. Here two sets of discrete equations, namely the state and measurement equations, are used to identify the process [18,31,32].

This technique depends largely upon the initial model adopted and usually employs an optimal filter to generate its forecasts. The Kalman filter is most often used as its recursive nature makes for an ideal on-line predictor [31,32].

The state space approach can also be applied to off-line forecasting by incorporating a weather input in the state equation. This is usually accomplished through the use of transfer function models [19,29].

The identification process is the main stumbling block of this method, as an accurate load model is always required beforehand. Secondly since this method almost invariably makes use of an optimal filter, the covariances of the noise processes have to be identified, a task that is generally not easily accomplished. However, this disadvantage is usually compensated by an increase in accuracy over the original model.

#### 2.2.5: Expert systems approach

The forecasting technique that has been receiving the most attention in the past decade is the expert systems approach. This technique can trace its roots to the method of pattern recognition developed by Mathewson and Nicholson in 1968 [13].

Developments in the field of artificial intelligence have prompted modern forecasters to develop computer programs that can act as experts, with regards to load behaviour. The works of Jabbour et al., Rahman, Bhatnagar and Baba are among the earliest published on this subject [33,34,35].

Forecast models in this technique, are built using mathematical relationships extracted through statistical analysis of previous data and from the knowledge of experts in the field. A computer algorithm is developed that acts as a rule based expert i.e the program will respond in an IF-THEN manner according to rules drawn up by the programmers [19,20].

Forecasts are created by presenting forecast variables to the program through an interface and prompting for the desired lead time prediction. This approach to short term forecasting is becoming increasingly popular as researchers have been able to prove that its forecast accuracy is comparable to those obtained using the best of conventional techniques. Its major disadvantage lies in the fact that it requires the use of costly computer centres to store, analyse and predict data.

## 2.4: Design and performance criteria

The review of short range forecasting methods has indicated, that dependent on the forecasting technique employed, many different load models can be developed to predict the same load. For these models to be considered good or efficient however, their formulation must feature certain basic qualities and their performance must be within tolerable limits.

The literature indicates that some of the qualities preferred in a load forecasting algorithm include adaptiveness, recursiveness, economy, robustness and accuracy [21]. Under the following headings below, these qualities are briefly reviewed.

### Adaptiveness

The parameters of a short term load forecasting model are usually estimated from a fixed window of data and are only accurate for a specified period of time ahead. As the forecast period elapses and new measurement becomes available, the algorithm should be able to automatically update its data window and recompute its estimates.

### Recursiveness

As new data such as weather and load measurements become available the algorithm should be able to correct its forecasts and predict for the next step.

### Computational economy

The pursuit of accuracy can lead to very complicated models that require the use of excessive computing facilities. A forecasting algorithm however, should try to be computationally efficient with regards to execution time and core utilisation.

### Robustness

An algorithm should be robust to mis-specification and erronous data i.e reasonable forecasts should be produced even if the model is predicting for conditions for which it is not specified, or even if its database is contaminated with bad or anomalous data.

Accuracy

The performance of a short term load forecasting algorithm depends largely upon the forecasting lead time as well as upon such factors as load behaviour and model type.

For models with a 24 hour prediction period errors in the range of 2-3% are considered normal, whereas for models with a lead time of one hour the same error is considered large [20,23]. Models with longer lead times than 24 hours show reduced accuracy and for a lead time of one week, accuracies within 10% are to be expected.

## CHAPTER III

## PARAMETER ESTIMATION: THE STATIC CASE

In this chapter, the parameter estimation techniques applicable to off-line forecasting are considered. These techniques are sometimes referred to as static estimation techniques as they are required to produce independent estimates from fixed or static windows of previous data.

The chapter begins by stating the estimation problem and briefly reviewing its methods of solution. Next a more indepth look at the methods applicable to short term load forecasting is presented. Here the history and theoretical development of the least squares and the least absolute value methods of estimation are reviewed.

Included in the review of least absolute value techniques is the derivation of a new non-iterative method of LAV estimation that will feature in this thesis. Also included in this section is a proposed improvement to this new algorithm that would enable it to perform more robustly as an off-line load forecasting algorithm.

### 3.1: The static estimation problem

In short term load forecasting the load $z(t)$ at any instant in time $t$, can be modelled as

$$z(t) = h(t)\theta + r(t) \tag{3.1}$$

where $h(t)$ = (1xn) row vector of fitting functions

$\theta$ = (nx1) column vector of coefficients, and

$r(t)$ = noise or residual load at time t.

The vector $\theta$ of coefficients to be estimated for the forecast lead time ahead, is considered fixed for the previous m load observations, referred to as the data window. $\theta$ can then be estimated by solving the series of m equations in n unknowns, representing the measurements in the data window.

At any time t therefore, we have

$$\begin{bmatrix} z(t-1) \\ z(t-2) \\ \vdots \\ z(t-n) \end{bmatrix} = \begin{bmatrix} f(t-1) \\ f(t-2) \\ \vdots \\ f(t-n) \end{bmatrix} \begin{bmatrix} \theta \end{bmatrix} + \begin{bmatrix} r(t-1) \\ r(t-2) \\ \vdots \\ r(t-n) \end{bmatrix} \quad (3.2)$$

This system of equations can be expressed in compact vector form as

$$Z = H\theta + r \quad (3.3)$$

where $Z$ = (mx1) column vector of load measurements $(z_1,\ldots,z_m)$

$H$ = (mxn) matrix of fitting functions relating $Z$ to $\theta$

$\theta$ = (nx1) vector of coefficients to be estimated, and

$r$ = (mx1) vector of residuals $(r_1,r_2,\ldots,r_m)$

In general, the number of measurements n, exceeds the number of parameters m to be estimated, that is m > n, and the system of equations is overdetermined with no unique solution. A "good" solution can be obtained however, if it minimises the residual vector r, measured in some sense.

This can be accomplished in more than one ways, depending upon the cost function of the residuals to be minimised. In short-term load forecasting, the least sum of the squares of the residuals and the least sum of the absolute value of the residuals are the functions most commonly employed, through the least squares function is by far, the more common of the two.

A third cost function called the Chebyshev function is often used by statisticians but to date has not found any application in short-range load forecasting.

The least absolute value or LAV cost function is a minimum when the sum of the absolute values of the residuals is a minimum. This cost function can be expressed mathematically as [36,37]

$$J_1(\theta) = \sum_{i=1}^{m} |z_i - H_i\theta| \tag{3.4}$$

$$= \sum_{i=1}^{m} |r_i|$$

and the resulting estimate that makes this function a minimum is referred to as the $L_1$ or least absolute value estimate or solution.

On the other hand, the least squares, LS or $L_2$ estimate minimises the sum of the squares of the residuals via the least squares cost function given by [36,37,38]

$$J_2(\theta) = \{ \sum_{i=1}^{m} |z_i - H_i\theta|^2 \}^{1/2} \tag{3.5}$$

$$= \{ \sum_{i=1}^{m} (r_i)^2 \}^{1/2}$$

In general any order of estimate $L_P$ can be made by minimising the generalised cost function [38]

$$J_P(\theta) = \{ \sum_{i=1}^{m} |z_i - H_i\theta|^P \}^{1/P} \tag{3.6}$$

where P varies from 1 to infinity.

We note that when P = 1 the $L_1$ or least absolute value estimate results. Similarly when P = 2 the least squares or $L_2$ estimates results.* Again when P is infinity the $L_\infty$ or Chebyshev estimate is obtained. The Chebyshev estimate is also referred to as the minmax estimate as it minimises the largest absolute value of the residuals.

It should be noted that while it is possible to obtain estimates with the appelate P ranging from one to infinity, only the $L_1$, $L_2$ and minmax criteria have been extensively researched in the literature on parameter estimation.

In this thesis only the least squares and least absolute value estimation criteria are considered, and in the following sections the roles of these criteria in static parameter estimation are reviewed.

* The minimisation of the square of equation (3.5) is the equivalent to the minimisation of the same function.

## 3.2: A review of the LS and LAV methods of parameter estimation

The problem of finding the best estimate or fit to a data set is one that had plagued mathematicians for many centuries. The development of the least squares and least absolute value methods are direct results of early efforts to conquer this problem.

The least absolute value technique was first applied by Boswich in 1757 when he presented a procedure for finding the best measurement fit based on the sum of the absolute deviations [39]. In 1795 Gauss used the least squares estimation method to predict the motion of several heavenly bodies and later showed that this technique could be derived from his law of error distribution [40].

These two methods were the subject of numerous research papers by the advocates of each during the early part of the nineteenth century and it became generally accepted at that time, that the least squares method resulted in better estimates.

In 1939 Jefferys showed that the least squares method gave the best estimate only if the error distribution was Gaussian and concluded that for other unsymmetrical distributions with small measurement set the least absolute value method was the best choice [41].

Until the advent of modern computers, the least squares technique remained the more popular of the two techniques and researchers found that by using weighted matrices an even better method developed. This became known as the weighted least squares method and today, still remains as one of the more popular estimation methods.

Even with the availability of computers the least squares method continued to flourish, mainly because of the lack of efficient least absolute value algorithms and partly because the least squares method had become something of a fixture with mathematicians.

Early efforts to apply the LAV principle to computers relied on linear programming techniques that suffered the disadvantage of being iterative and requiring large computing effort in the form of time and storage. However with the development of more efficient and simple LAV algorithms, researchers have found it much simpler and sometimes even more accurate to apply the LAV method to a variety of parameter estimation problems [5,38,42].

The results of research to date indicate that LAV techniques give better approximations when the measurement set has an unknown distribution and also when the sample size is small. In addition, the nature of the LAV solution allows it to reject outlying data without any previous knowledge of its location [5,42]. The LS method on the other hand performs better when the error distribution is Gaussian, and the sample size is large.

### 3.2.1: Least squares estimation

Given the overdetermined system of equations

$$Z = H\theta + r$$

the least squares estimate is obtained by solving the normalised system of equations [40,43]

$$H^T Z = H^T H\theta \tag{3.7}$$

yielding

$$\theta_{LS} = (H^T H)^{-1} HZ \tag{3.8}$$

In many cases, weights are assigned to each measurement so that the measurements with the larger weights have a greater influence on the solution. In short-term load forecasting larger weights are usually assigned to the more recent observations as they are generally more indicative of immediate load trends.

The cost function for the weighted least squares estimation can be written as [44]

$$J_2(\theta) = (Z - H\theta)W^T(Z - H\theta) \tag{3.9}$$

where W is a diagonal (nxn) weighting matrix, i.e

W = diagonal ( $w_i$, i = 1, ....n)

Again from the overdetermined equations

$$Z = H\theta + r$$

we have, after multiplication by the weighting matrix W

$$WZ = WH\theta \tag{3.10}$$

Normalising this system of equation and solving gives

$$H^{T}WZ = H^{T}WH\theta \tag{3.11}$$

and

$$\theta_{WLS} = (H^{T}WH)^{-1}H^{T}WZ \tag{3.12}$$

The least squares and the weighted least squares methods are the most widely used parameter estimation techniques in short-term load forecasting. They find especial use in the multiple linear regression and the general exponential smoothing techniques and are usually employed at one stage or another in the other forecasting techniques.

Least squares algorithms are easy to implement on a digital computer and usually require a minimum of computing effort and time. They are also very robust and produce excellent estimates when the error distribution is Gaussian. The only recognizable disadvantage is that they do not inherently ignore outliers or so called bad data when estimating [5,42,43].

### 3.2.2: Least absolute value estimation

The use of the LAV criterion in parameter estimation and model fitting can be traced to the early works of Boswich (1757) and Laplace (1793) [39]. This criterion has been studied under a wide variety of labels such as the minimum or least sum of absolute error (MSAE, LSAE) criterion, the $L_1$ or Gershgorin norm, and the minimum or least absolute value/ deviations (LAV, LAE, MAV, MAD) [39].

Given the overdetermined system of equations of equation (3.1), the optimal LAV solution or estimate $\theta_{LAV}$ is reached only when the sum of the absolute value of the residuals is an absolute minimum.

Unlike the least squares method of estimation however, no mathematical procedure has been developed to give an optimal LAV estimate in a single computational sequence. Rather, true LAV solutions are only obtainable through the use of iterative linear programming techniques.

This iterative disadvantage has led to the development of many approximate LAV algorithms that greatly reduces the number of iterations required and in the case of a more recent study; to the development of a completely non-iterative algorithm [5,38].

In this thesis both the linear programming and the this newly developed non-iterative approach to LAV estimation are featured. The mechanics governing these methods, and their inherent advantages and disadvantages are discussed in turn, in the remainder of this chapter.

## Linear programming

The linear programming approach to LAV estimation is one that has played a central role in the development of LAV estimation theory. Linear programming techniques are iterative in nature and uses the methods of successive improvements to reach their solution.

Using this approach, the resulting LAV algorithm is usually formulated to solve a constrained or an unconstrained estimation problem. Research has shown that the unconstrained or unrestricted problem usually results in a non-unique solution, whereas the use of constraints severly limits the chances of a non-unique solution [39,45].

Inherent in any linear programming algorithm is a cost function to be minimised or maximised. If the problem formulation requires a cost function to be minimised, it is referred to as a primal formulation. On the other hand, a cost function to be maximised is labelled a dual formulation. Research has indicated that the primal formulation in linear programming is preferrable as it is generally much more efficient with regards to computing time and storage [39].

After the problem has been stated as a constrained primal or some similar formulation, an iterative technique is employed to converge to the best solution. The earliest and most widely known of these is the simplex method [39,46]. This method however, has been found to be computationally inefficient and subsequent studies have led to a revised simplex method that is computationally much more efficient, but still suffers from non-uniqueness of solution in many cases [46].

Further work on LAV linear programming has led to the development of many special purpose algorithms that are even more efficient and produce unique solution in most cases. These algorithms are generally derivatives of simplex techniques but are much more computationally efficient as they eliminate many or bypass many stages in the simplex procedure [38].

The formulation of the LAV estimation as a primal constrained linear programming problem is as follows [36,37,46].

Given the overdetermined system of equations as stated in equation (3.1). The LAV estimation problem is to determine $\theta_{LAV}$ such that the cost function given by equation (3.3) is minimised. This cost function can be rewritten as

$$J_1(\theta) = \sum_{i=1}^{m} \left| z_i - \sum_{j=1}^{n} H_{ij}\theta_j \right| \tag{3.13}$$

where $z_i$ is the $i^{th}$ element of Z

$\theta_j$ is the $j^{th}$ element of $\theta$

$H_{ij}$ is the element in the $i^{th}$ row and the $j^{th}$ column of the (mxn) matrix H, which defines the relationship between Z and $\theta$

and the $i^{th}$ residual is given by

$$r_i = z_i - \sum_{j=1}^{n} H_{ij}\theta_j \tag{3.14}$$

The linear programming formulation of the LAV estimation problem can now be stated as follows:

Minimise $\sum_{i=1}^{m}$ subject to the constraints

$$r_i + \sum_{j=1}^{n} H_{ij}\theta_j \geq z_i \qquad i=1,2,\ldots,m \tag{3.15}$$

$$r_i - \sum_{j=1}^{n} H_{ij}\theta_j \geq -z_i \qquad i=1,2,\ldots,m \tag{3.16}$$

It can be shown that the constraints restricts the value of $r_i$ to the absolute value of the $i^{th}$ residual. Minimising the sum of the $r_i$'s in the linear programming problem produces the LAV estimate [36,37].

The LAV based linear programming algorithm used in this thesis can be found in the IMSLSTAT Library and is based on the tecchnique developed by Barrodale and Roberts (1973) [38,47].

The main advantage of linear programming techniques is that they will present the optimal LAV estimate in almost all cases. Their disadvantages are that they require excessive computational effort due to their iterative nature, and that the solution obtained may not be unique.

New LAV technique

The iterative nature of linear programming methods has deterred many potential users of the least absolute value criterion, even though estimation via this criterion is advantageous in many cases. Even with highly efficient modified simplex algorithms, the computing effort required for linear programming solutions is still very much greater than that required for the equivalent least squares solution.

This has led researchers along the path of reducing the number of iterations in LAV estimation. Schlossmacher (1973), proposed a technique using weighted least squares estimation to converge to a final solution, but his method did not largely reduce the number of iterations over the then present techniques [48].

Sporito, Hand and McCormick (1977) later showed that by using the least squares estimate as an initial guess in the highly efficient algorithm of Barrodale and Roberts, many iterations in the linear programming procedure could be saved [49,50]. The use of the LS and WLS as initial estimates in linear programming techniques is now a well established procedure and features in the IMSLSTAT linear programming LAV package used in this thesis [47].

In 1987 Christensen and Soliman developed a new LAV technique that was non-iterative in nature and produces a unique solution if the matrix H is of full rank [5]. This novel approach manipulated a simple theoretical relationship between LS and LAV solutions and has produced estimates closely matching those resulting from conventional linear programming algorithms.

The method of Christensen and Solimon follows directly from the theorem governing LAV estimation, which reads as follows [5,42].

<u>Theorem</u>: If the column rank of the (mxn) matrix is k where ($k < n$), then the LAV estimate interpolates at least k of the m measurements.

From this theorem, Christensen and Soliman proposed, that by selecting the best n of these measurements with H being of full rank, a fully determined system of equations is obtained that could be solved to give the LAV estimate. The stages in this new algorithm are as follows:

(1) Given the overdetermined system of equations shown below

$$Z_m = H_m \theta + r_m \tag{3.16}$$

where $Z_m$ = (mx1) vector of measurements

$\theta$ = (nx1) coefficient vector to be estimated

$H_m$ = (mxn) matrix of rank n relating $Z_m$ to $\theta_m$, and

$r_m$ = (mx1) vector of residuals

find the least squares solution $\theta_{LS(m)}$, where

$$\theta_{LS(m)} = [H_m^T H_m]^{-1} H_m^T Z_m \qquad (3.17)$$

and the subscript LS(m) refers to a least squares estimate made from m measurement observations.

(2) Calculate the residuals resulting from the estimate $\theta_{LS(m)}$, i.e.

$$r_i = z_i - H_i \theta_m \quad (i=1,\ldots,m) \qquad (3.18)$$

where $H_i$ is the $i^{th}$ row of the (mxn) matrix H

$z_i$ is the $i^{th}$ measurement, and

$r_i$ refers to the $i^{th}$ residual

(3) Compute the standard deviation of the residuals as follows

$$\sigma = \left[ \frac{1}{m - n + 1} \sum_{i=1}^{m} (r_i - \bar{r})^2 \right]^{1/2} \qquad (3.19)$$

where $\bar{r}$ = mean value of residuals.

and reject those observations with residuals larger this value.

Note that points with residuals larger than the standard deviation are then considered as outliers and in keeping with the nature of LAV solutions are ignored.

(4) Recompute a LS estimate from the reduced measurement set. Thus assuming that p points are rejected, then the new least squares estimate is given by

$$\theta_{LS(m-p)} = [H^T_{m-p} H_{m-p}]^{-1} H^T_{m-p} Z_{m-p} \tag{3.20}$$

where the subscript m-p now indicates a reduced measurement set

(5) Rank the residuals beginning with the smallest and ending with the largest and select the first n measurements corresponding to the first n smallest residuals.

(6) A perfectly determined system of equations is now available that can be solved to give the least absolute value estimate for the overdetermined system of equations given by equation (3.16).

The LAV estimate is therefore found to be

$$\theta_{LAV} = [H_n]^{-1} Z_n \tag{3.21}$$

where the subscript k now refers to the final reduced set of n selected measurements.

(7) The residuals resulting from this LAV estimate can be computed as follows

$$r_i = z_i - H_i \theta_{LAV} \qquad (i=1,\ldots,m) \tag{3.22}$$

This algorithm, although only recently developed, has been applied to many power system estimation problems and found to be a comparable alternative to existing estimation methods [51,52,53].

The major advantages of this new estimation procedure are

(a) it is non-iterative and very efficient with regards to computing effort and time,

(b) it has inherent "bad data " rejection properties, and

(c) it produces a unique solution in all cases provided the final matrix $H_n$ selected is of full rank.

The major disadvantage of this algorithm is that, the matrix $H_n$ selected may not always be of full rank, in which case the solution produced will be erreneous. This has been a relatively non-existent situation in so far as the algorithm has been tested. However when the measurement equations are periodically repetitive, the chances of $H_n$ being of less than full rank are greatly increased.

This is a common situation in short term load forecasting where load behaviour is periodic and many load cycles are included in the measurement set. This difficulty can be high-lighted by considering a simple harmonic decomposition model given by

$$y(t) = a_0 + a_1 \sin(wt) + a_2 \cos(wt) \tag{3.23}$$

where $w = \pi/4$

The coefficients $[a_0, a_1, a_2]$ are to be determined from two (2) complete load cycles of four (4) observations each, that is, a total of eight (8) observations. Given that the observations are recorded at fixed discrete intervals apart beginning at t = 0, and can be denoted as $y_1, y_2, \ldots, y_8$, the overdetermined system of equations can be written as follows.

$$\begin{bmatrix} y_1 \\ y_2 \\ y \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} 1 & \sin(\pi/2) & \cos(\pi/4) \\ 1 & \sin(\pi) & \cos(\pi) \\ 1 & \sin(3\pi/2) & \cos(3\pi/2) \\ 1 & \sin(2\pi) & \cos(2\pi) \\ 1 & \sin(5\pi/2) & \cos(5\pi/2) \\ 1 & \sin(3\pi) & \cos(3\pi) \\ 1 & \sin(7\pi/2) & \cos(7\pi/2) \\ 1 & \sin(4\pi) & \cos(4\pi) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{3.24}$$

We note that rows of the H matrix (8x3) that occur one full cycle apart are essentially the same. For example, rows 1 & 5, 2 & 6, etc. are exactly the same.

Application of the proposed algorithm requires the best three (3) point to be selected. We note, that it is possible for the points (1,2,3), (2,3,4) etc. to be selected, in which $H_n$ will be of full rank and a good solution will result. Alternatively the points (1,5,6), (2,6,7), (3,7,8) etc. can be selected, in which case the rank of $H_n$ will no longer be n, and a valid solution is no longer possible.

This problem can be rectified by the implementation of a two stage ranking procedure in place of steps (3),(4) and (5) of the original algorithm, if it can be assumed (as is the general case in load forecasting) that

(1) the m observations are from L complete cycles of q observations per cycle, i.e. Lq = m, and

(2) the observations occur at fixed discrete intervals apart.

The modified LAV procedure that can be incorporated as an option for periodic observations in the original algorithm, is as follows:

(1) Given the overdetermined system of equations (3.16), find the least squares estimate (3.17)

$$\theta_{LS(m)} = [H_m^T H_m]^{-1} H_m^T Z_m$$

(2) Calculate the residuals corresponding to this estimate (3.18), and separate these residuals into q groups of L residuals, for each of the q discrete observations in a cycle.

(3) Rank each of the q groups of residuals, and select the measurements corresponding to the smallest residual of each group.

(4) Next, select the n smallest residuals from amongst the preselected best of each of the q groups, to give a perfectly determined system and solve to get the LAV estimate.

Referring to the example of equation (3.23), we note that there will now be four groups of residuals, each consisting of two residuals e.g. (1,5), (2,6), (3,7) and (4,8). The observations in each of these groups correspond to rows in the H matrix separated by a complete cycle, i.e. to the rows that are mathematically identical.

Since only one residual will be selected from each group, this new procedure will eliminate the chances of the final matrix $H_n$ selected being of less than full rank.

The effectiveness of this modified LAV procedure is further demonstrated with the help of the following numerical example.

Example

Consider a process to be modelled by equation (3.23) where we are required to estimate the coefficients $[a_0, a_1, a_2]$ using the LAV criterion. Given eight previous process measurements corresponding to $[y_1, y_2, \ldots, y_8]$ recorded at fixed discrete intervals of one quarter of a period apart, we can rewrite equation (3.24) to read

$$\begin{bmatrix} 106 \\ 99 \\ 94 \\ 101 \\ 105 \\ 99 \\ 95 \\ 104 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{3.25}$$

We will now find the LAV estimate using (a) linear programming, (b) the original LAV procedure and (c) the modified LAV procedure.

<u>Linear programming</u>

The LAV estimate determined using linear programming is given by

$$\theta_{LP} = [100 \quad 6 \quad -1]$$

with

Sum of absolute errors = 5

<u>Original technique</u>

Using this technique we first find the least squares estimate to be

$$\theta_{LS(8)} = [100.375 \quad 5.500 \quad -1.750]$$

with the residuals

$$r = \begin{bmatrix} -0.125 \\ -0.375 \\ 0.875 \\ 1.125 \\ 0.875 \\ -0.375 \\ -0.125 \\ -1.875 \end{bmatrix} \tag{3.26}$$

and a standard deviation of 1.0508.

We note that the fourth and eight residuals are larger in absolute value than the standard deviation, and therefore the corresponding measurements are rejected.

We recompute the least squares estimate and residuals to be

$\theta_{LS(6)}$ = [98.917   4.417   0.083], and

$$r = \begin{bmatrix} 2.667 \\ 0.000 \\ -0.500 \\ 2.167 \\ 1.667 \\ 0.000 \\ 0.500 \\ 5.167 \end{bmatrix} \tag{3.27}$$

We then select the measurements corresponding to the three(3) smallest absolute residuals. Points 2, 6 and 3 therefore, are selected to give a perfectly determined system of equations whose least squares solution is the LAV estimate. The reduced equations corresponding to these residuals can now be written as

$$\begin{bmatrix} 99 \\ 99 \\ 94 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{3.28}$$

We can immediately observe that these equations are ill conditioned and does not possess a unique solution. A possible solution to equation (3.28) is

$$\theta_{LAV} = [47 \quad -47 \quad 52]$$

an estimate that is far removed from the linear programming estimate given previously. We note that this solution is only possible due to the fact that the measurements corresponding to rows 1 and 2 in equation (3.28) are the same. In general however, a solution to this problem is highly unlikely.

<u>Modified technique</u>

Having found the initial least squares estimate and the corresponding residuals (3.26), we can now seperate the residuals into four groups of two residuals each, as shown below

| Group 1 | Group 2 | Group 3 | Group 4 |
|---|---|---|---|
| -0.125 | -0.375 | 0.875 | 1.125 |
| 0.875 | -0.375 | -0.125 | -1.875 |

and select the three smallest of the best four residuals. The residuals selected correspond to the measurements 1,7 and 2 which results in the perfectly determined equation below.

$$\begin{bmatrix} 99 \\ 95 \\ 106 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \tag{3.29}$$

which gives

$$\theta_{LAV} = [100.5 \quad 5.5 \quad -1.5]$$

with

Sum of absolute errors = 5

which is a true LAV solution to the problem set given.

# CHAPTER IV

## PARAMETER ESTIMATION: THE DYNAMIC CASE

The material presented in this chapter deals with the application of dynamic estimation theory to on-line short-term load forecasting. The dynamic parameter estimation problem is first introduced followed by the theory and development of the Kalman and a newly developed least absolute value based filter.

Next, the formulation of a recursive general exponential smoothing algorithm that will also feature in on line simulations, is presented. Finally the chapter concludes by discussing the procedure for the application of the Kalman, the new least absolute value based filter and the general exponential smoothing algorithm to on-line forecasting.

### 4.1: Stochastic estimation.

Stochastic estimation is the process of assessing unknown system states from noise corrupted observations of functions of the said state. In this process, the statistical properties of the noises are assumed known and the estimate is determined so that it minimises a specified cost function of the estimation error.

The literature indicates, that for a process where the error distribution is Gaussian, the least squares cost criterion results in optimal estimation. This is the case of the popular Kalman filter that has found application throughout modern control theory [54].

In keeping with the advantages of the least absolute value criterion however, researchers have theorised that least absolute value based stochastic estimation would inherit these advantages in cases of non-Gaussian error distribution [55,56].

The inherent iterative nature of least absolute value solutions however, has prevented the realisation of such an estimator, until more recently, when the development of a new non-iterative LAV estimation technique , has resulted in a practical formulation for a recursive LAV based stochastic estimator [5].

In the following sections, the dynamic stochastic estimation problem is presented followed by the theoretical development of the Kalman ,a weighted least absolute value filter and a general exponential smoothing algorithm.

## 4.2: Dynamic estimation problem

A discrete process whose states are to be estimated can be represented by the state equations [26,57]

$$x(k+1) = \phi(k)x(k) + w(k) \tag{4.1}$$

and the corresponding measurement equation

$$z(k) = H(k)x(k) + v(k) \tag{4.2}$$

where $x(k)$ = (nx1) process state vector at time $t_k$

$\phi(k)$ = (nxn) state transition matrix relating x(k) to x(k+1) in the absence of a forcing function.

w(k) = (nx1) noise vector with known covariance

z(k) = (mx1) vector measurement at time $t_k$

H(k) = (mxn) matrix giving the ideal relationship between measurement and state vector at time $t_k$

v(k) = (mx1) measurement error with known covariance

The noise processes w(k) and v(k) are assumed to be uncorrelated white noise processes, with the following characteristics:

$$E\,[w(k)w^T(k)] = \begin{cases} Q(k) & , i=k \\ 0 & , i \neq k \end{cases} \tag{4.3}$$

$$E\,[v(k)v^T(k)] = \begin{cases} R(k) & , i=k \\ 0 & , i \neq k \end{cases} \tag{4.4}$$

and

$$E\,[w(k)v^T(k)] = 0 \quad \text{for all i and k} \tag{4.5}$$

Given the estimated state vector at the $(k-1)^{th}$ instant for a known process, the value of the state at the $k^{th}$ instant can be predicted from the state equation as follows:

$$\hat{x}(k/k-1) = \phi(k-1)\hat{x}(k-1) \tag{4.6}$$

where $\hat{x}(k/k-1)$ refers to the value of the state vector at the $k^{th}$ instant predicted at the $k-1^{th}$ discrete instant, and the symbol " ^ " refers to estimated values.

Associated with this estimate will be an estimation error given by

$$e(k/k-1) = x(k) - \hat{x}(k/k-1) \tag{4.7}$$

where x(k) is the actual value of the state vector at the $k^{th}$ instant.

Substituting for x(k) and $\hat{x}(k/k-1)$ from equations (4.1) and (4.6) respectively, gives

$$\begin{aligned} e(k/k-1) &= \phi(k-1)x(k-1) + w(k-1) - \phi(k-1)\hat{x}(k-1) \\ &= \phi(k-1)e(k-1) + w(k-1) \end{aligned} \tag{4.8}$$

and thus the estimation error can also be predicted from the value given at the previous discrete instant. The error covariance corresponding to this predicted error can also be found from

$$P(k/k-1) = E[e(k/k-1)e^T(k/k-1)] \tag{4.9}$$

which can be simplified to read [58,59]

$$P(k/k-1) = \phi(k-1)P(k-1)\phi^T(k-1) + Q(k-1) \tag{4.10}$$

With the availability of the measurement observation z(k), it is now desirable to determine a suitable gain factor K(k), so that the predicted state estimate can be corrected by the measurement error and be optimal in a choosen manner. The corrected or smoothed estimate can now be represented as [57,58,59]

$$x(k) = \hat{x}(k/k-1) + K(k)[z(k) - H(k)\hat{x}(k/k-1)] \quad (4.11)$$

where $x(k)$ is the smoothed estimate at the $k^{th}$ interval and $K(k)$ is the gain or correction factor at the $k^{th}$ instant.

Associated with the smoothed estimate for the state vector will be a new value of estimation error given by

$$e(k) = x(k) - \hat{x}(k) \quad (4.12)$$

$$= x(k) - \{ \hat{x}(k/k-1) + K(k)[z(k) - H(k)x(k/k-1)] \}$$

$$= [x(k) - \hat{x}(k/k-1)] - K(k)[H(k)x(k) - H(k)\hat{x}(k/k-1) + v(k)] \quad (4.13)$$

and corresponding to this new value of estimation error will be a new or smoothed value of error covariance given by

$$P(k) = E\,[e(k)e^T(k)] \quad (4.14)$$

Substituting for $e(k)$ and simplifying gives

$$P(k) = [I - K(k)H(k)]P(k/k-1)[I - K(k)H(k)]^T + K(k)R(k)K^T(k) \quad (4.15)$$

where I is the (nxn) identity matrix.

It can be seen, that given an initial estimate for the state vector $\hat{x}(k-1)$, with the gain factor $K(k)$ known for all k, it is possible to recursively predict and correct the state vector and error covariance for the all k.

Equations (4.6), (4.10), (4.11) and (4.15) together constitute a recursive chain that allows for sequential data processing, and together with the gain vector K(k), forms the basis of the discrete filter.

Calculation of the gain is usually done so as to satisfy a choosen estimation error cost function, and when this function is the based on the stochastic least squares minimisation criteria the resulting filter is the discrete Kalman filter, named after its developer R.E.Kalman [54].

The gain vector K(k) can also be derived from the least absolute value criterion, as is the case of the more recently developed weighted least absolute value filter [60]. In the following sections, both the Kalman and the new filter gains are derived and their completete formulations presented.

### 4.2.1: Kalman filtering

The stochastic least squares cost function to be minimised can be written as [58]

$$J_2 = E\,[\sum_{k=1}^{m} e^T(k)e(k)] \tag{4.16}$$

where e is the error at the $k^{th}$ instant and m is the total number of observations. If the m measurements are processed sequentially as in the case of the discrete filter, the cost function will be a minimum when each individual cost element is an absolute minimum

Consider the error at the $k^{th}$ instant to be

$$e(k) = z(k) - H(k)\hat{x}(k) \qquad (4.17)$$

$$= H(k)x(k) - H(k)\hat{x}(k)$$

where $\hat{x}(k)$ is the estimated state vector at that instant and $x(k)$ is the actual value of the state vector at that instant.

The least squares cost at the $k^{th}$ instant is then given by

$$J_2(k) = E\{ e(k)e^T(k) \} \qquad (4.18)$$

$$= E\{(H(k)[x(k) - \hat{x}(k)])^T(H(k)[x(k) - \hat{x}(k)])\}$$

$$= E\{ tr(H(k)[x(k)-\hat{x}(k)])(H(k)[x(k) - \hat{x}(k)])^T$$

$$= E\{tr(H(k)[x(k) - \hat{x}(k)][x(k) - \hat{x}(k)]^T H^T(k))\}$$

$$= E\{tr(H(k)P(k)H^T(k))\} \qquad (4.19)$$

and we can see that by minimising the trace of the error covariance, the cost function is minimised. From equation (4.15) we have

$$P(k) = [I - K(k)H(k)]P(k/k-1)[I - K(k)H(k)]^T$$

$$+ K(k)R(k)K^T(k)$$

$$= P(k/k-1) - K(k)H(k)P(k/k-1) - P(k/k-1)H^T(k)K^T(k)$$

$$+ K(k)[H(k)P(k/k-1)H^T(k) + R(k)]^{-1}K^T(k) \qquad (4.20)$$

The expression for tr{P(k)} will be a minimum when its derivative with respect to K(k) is zero, thus given [58,59] that

$$d/dA[\mathrm{trace}(AB)] = B^T \qquad (4.21)$$

and

$$d/dA[\mathrm{trace}(ACA^T)] = 2AC \qquad (4.22)$$

and noting that P(k) is symmetric, we get after differentiation

$$dP(k)/dK(k) = 2K(k)[H(k)P(k/k-1)H^T(k) + R(k)] - 2[H(k)P(k/k-1)]^T \qquad (4.23)$$

Equating to zero gives

$$K(k) = P(k/k-1)H(k)[H(k)P(k/k-1)H^T(k) + R(k)]^{-1} \qquad (4.24)$$

Equation (4.24) can now be implemented along with equations (4.6), (4.9), (4.10) and (4.15) to complete the discrete Kalman filter. The complete filter formulation is presented below.

<u>Discrete Kalman filter</u>

Given initial values for P(k-1) and x(k-1), equations (4.6) and (4.9) can be used to predict the values of P(k) and x(k) as follows

(1) $P(k/k-1) = \phi(k-1)P(k-1)\phi^T(k-1) + Q(k-1)$

(2) $x(k/k-1) = \phi(k-1)x(k-1)$

Having predicted P(k/k-1) and x(k/k-1) with Q(k) and R(k) known, the Kalman gain vector at the $k^{th}$ instant can now be calculated from equation (4.24) as

(3) $K(k) = P(k/k-1)H^T(k)[H(k)P(k/k-1)H^T(k) + R(k)]^{-1}$

and the predicted estimates $P(k/k-1)$ and $x(k/k-1)$ corrected by equations (4.15) and (4.11) as follows

(4) $P(k) = [I - K(k)H(k)]P(k/k-1)[I - K(k)H(k)]^T + K(k)R(k)K^T(k)$

(5) $x(k) = x(k/k-1) + K(k)[z(k) - H(k)x(k/k-1)]$

The updated estimates $P(k)$ and $x(k)$ are now available and can be used as initial values for the next discrete step thereby creating a recursive chain for least squares based stochastic estimation.

### 4.2.2: Least absolute value filtering

The weighted least absolute value cost function to be minimised can be written as [57,58,61]

$$J_1 = \sum_{i=1}^{m} |N(x - \bar{x}) - w_i^{-1}(z_i - H_i x)| \qquad (4.25)$$

where $x$ = (nx1) actual state vector

$\bar{x}$ = (nx1) state vector to be estimated

$w_i$ = weighting element of $i^{th}$ measurement

$z_i$ = $i^{th}$ measurement observation

$H_i$ = (mxn) measurement vector at the $i^{th}$ instant, and

$$N = e^T P_x^{-1}$$

where $P_x$ = (nxn) error covariance matrix, and

$e$ = (mx1) column(1, 1, 1, . . . . . , 1) vector

The minimum value of this cost function is given by

$$J_1 = \sum_{i=1}^{m} | \varepsilon_i | = 0 \tag{4.26}$$

where

$$\varepsilon_i = N(x - \bar{x}) - r_i^{-1}(z_i - H_i x) = 0 \tag{4.27}$$

Rewriting equation (4.27) in compact vector form for all i, gives

$$E = M(x - \bar{x}) - W^{-1}(Z - Hx) = 0 \tag{4.28}$$

where $E$ = column $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$

$W^{-1}$ = column $(w_1^{-1}, w_2^{-1}, \ldots, w_m^{-1})$

$Z$ = column $(z_1, z_2, \ldots, z_m)$

$H$ = matrix of row vectors $H_i$, ( i=1, . . . . m), and

$M = Le^T P_x^{-1}$, where

$L$ = (mx1) column (1, 1, . . . 1) vector

Equation (4.28) can be rearranged to read

$$(M + W^{-1}H)x = M\bar{x} + W^{-1}Z \tag{4.29}$$

Adding and subtracting $W^{-1}H\bar{x}$ to thr right hand side of this equation yields

$$(M + W^{-1}H)x = (M + W^{-1}H)\bar{x} + W^{-1}(z - Hx) \tag{4.30}$$

This is an overdetermined system of equations that can be solved using the new LAV procedure of section 3.2.2.

If equation (4.30) is rewritten as

$$Ax = A\bar{x} + W^{-1}(z - Hx) \tag{4.31}$$

where

$$A = (M + W^{-1}H) \tag{4.32}$$

then

$$x = \bar{x} + (A^TA)^{-1}A^TW^{-1}(z - Hx) \tag{4.33}$$

is the least squares solution and the least absolute value estimate is given by

$$x = \bar{x} + \hat{A}^{-1}\hat{W}^{-1}(\hat{z} - \hat{H}x) \tag{4.34}$$

where $\hat{A}$ = (nxn) matrix of vectors corresponding to the m smallest residuals

$\hat{W}^{-1}$ = (nxn) diagonal of weighting elements corresponding to the measurements selected

$\hat{H}$ = (nxn) matrix of n selected measurement row vectors, and

$\hat{z}$ = (nx1) column vector of selected measurements

Equation (4.34) can be rewritten as

$$x = \bar{x} + K(z - H\bar{x}) \tag{4.35}$$

where

$$K = \hat{A}^{-1}W^{-1} \tag{4.36}$$
$$= (\hat{W}\hat{A})^{-1}$$

The gain vector calculated here, is for the static case using the m previous observations. For the dynamic (sequential) estimation however, the gain vector at the $k^{th}$ discrete instant becomes [58,61]

$$K(k) = [W(k)A(k)]^{-1} \tag{4.37}$$
$$= [W(k)M(k) + H(k)]^{-1} \tag{4.38}$$
$$= [W(k)Le^{T}P_{x}^{-1}(k) + H(k)]^{-1} \tag{4.39}$$

and substituting for the Kalman gain equation in the discrete Kalman filter gives the new weighted least absolute value filter. This filter is essentially the same as the Kalman filter except that the gain vector is calculated on the basis of the least absolute value criterion rather than the least squares criterion.

### 4.1.3: Adaptive general exponential smoothing

A discrete load process can be modelled by the equation

$$z(k) = f^T(k)\beta(k) + \varepsilon(k) \qquad (4.40)$$

where $\beta(k)$ is the vector of coefficients to be estimated, for the instant k

$\varepsilon(k)$ is residual load at the $k^{th}$ discrete instant

$z(k)$ is the load measurement at the $k^{th}$ interval, and

$f(k)$ is a vector fitting function that is related to successive values by the equation

$$f(k) = Lf(k-1) \qquad (4.41)$$

where L is called the transition matrix and is constant and specific to the process at hand

It is desired to estimate the coefficient vector $\beta(k)$, that minimises a weighted (discounted) least squares cost function given by [26]

$$J = \sum_{j=0}^{T} w^j[z(t-j) - f^T(-j)\beta(T)]^2 \qquad (4.42)$$

where w is a weighting factor choosen so that the weighted squared residuals are totally discounted in T previous observations.

The cost function given is a minimum when its derivative with respect to the coefficient vector is zero. Differentiating and equating to zero, therefore yields

$$\beta(T) = [\sum_{j=0}^{T} w^j f(-j) f^T(-j)]^{-1} [\sum_{j=0}^{T} w^j z(T-j) f(-j)] \tag{4.43}$$

which can be written in compact form as

$$\beta(T) = F^{-1}(T) g(T) \tag{4.44}$$

where

$$F(T) = \sum_{j=0}^{T} w^j f(-j) f^T(-j) \tag{4.45}$$

is called the coefficient matrix of weighted fitting functions, and

$$g(T) = \sum_{j=0}^{T} w^j z(T-j) f(-j) \tag{4.46}$$

is called the data vector

The data vector can be rearranged to read [25]

$$g(T) = z(T) f(0) + w L^{-1} g(T-1) \tag{4.47}$$

and multiplying throughout by $F^{-1}(T)$ gives

$$\beta(T) = z(T) F^{-1}(T) f(0) + w F^{-1}(T) L^{-1} F(T) \beta(T-1) \tag{4.48}$$

Given that F(T) will have reached a steady state value in T previous observations, we can represent it by F for purposes of simplification. The matrix F can also be rearranged to read [25]

$$F = F(T-1) + f(0)f^T(0) \quad (4.49)$$

Premultiplying equation (4.45) by $L^{-1}$ and post multiplying by the identity matrix $L^{-T}L^T$ gives

$$L^{-1}FL^{-T}L^T = \sum_{j=0}^{T} w^j [L^{-1}f(-j)][L^{-1}f(-j)]^T L^T \quad (4.50)$$

$$= (1/w)[F - f(0)f^T(0)]\ L^T \quad (4.51)$$

since

$$L^{-1}f(-j) = f(-j-1) \quad (4.52)$$

substituting for $WL^{-1}F$ in equation (4.48) gives

$$\begin{aligned}\beta(T) &= z(T)F^{-1}f(0) + F^{-1}[F - f(0)f^T(0)]L^T\beta(T-1) \\ &= z(T)F^{-1}f(0) + L^T\beta(T-1) - F^{-1}f(0)f^T(0)L^T\beta(T-1) \\ &= L^T\beta(T-1) + F^{-1}f(0)[z(T) - f^T(0)L^T\beta(T-1)] \quad (4.53)\end{aligned}$$

now since $f^T(0)L^T\beta(T-1)$ is an apriori estimate for the load at time T-1, we can rewrite equation (4.53) as

$$\beta(T) = L^T\beta(T-1) + F^{-1}f(0)[z(T) - \hat{z}(T)] \quad (4.54)$$

where $\hat{z}(T) = f^T(0)L^T\beta(T-1)$

Note that by making the following equivalents

$$\phi(k) = L^T$$

$$x(k) = \beta(k)$$

$$H(k) = f^T(k), \text{ and}$$

$$K(k) = F^{-1}f(0)$$

we obtain the general exponential smoothing equivalent of the discrete Kalman filter. We note also that the gain vector K is constant for all k, and is independent of the process noises.

## 4.3: Dynamic on-line estimation and forecasting.

An inherent characteristic of both the Kalman and the new WLAV filter, is that the gain vector and the error covariance will settle down to stationary values after a sufficiently large amount of measurement data is processed. At this stage it is no longer necessary to recursively re-estimate the gain and covariances, and stages 1,3 and 4 can be dropped from the discrete Kalman And WLAV filter formulations

The simplicity of the foregoing forecasting procedure is limited only by the large initial data analysis that is required to accurately determine the values of steady state gain for the stochastic filters.

This difficulty is compounded by the fact that both algorithms assume that the covariances of the noise processes are known for all k.

This is generally not the case with load forecasting data however, and Q(k) and R(k) must be accurately determined so that the estimation process is optimal.

Estimates of Q(k) and R(k) can be made from the recursive estimation equations of Sage and Husa [62]. These are estimation equations that can be implemented with the filter in question to give suboptimal estimates for Q(k) and R(k), so that K(k) can determined for all k. These equations are as follows

$$(6) \quad R(k) = (1/k)[(k-1)R(k-1) + r^2(k) - H(k)P(k/k-1)H^T(k)] \qquad (4.60)$$

$$(7) \quad Q(k) = (1/k)[(k-1)Q(k-1) + r^2(k)K(k)K^T(k) + P(k/k) - \phi(k)P(k-1/k-1)\phi^T] \qquad (4.61)$$

Equations (4.60) and (4.61) can be included as stages (6) and (7) in the respective filters. These equations possess the steady state characteristics of the stochastic filters and as such the values of Q(k) and R(k) will attain steady state along with K(k) and P(k) [62,63].

A simple on-line forecasting and smoothing algorithm is now available that requires a minimum of computing effort. We note, that these simplified filtering algorithms are now identical to the general exponential smoothing algorithm, except for differences in steady state gain values.

In general, if we represent the steady state gain vector by $K_s$, the on-line forecasting procedure can be described as follows.

### 4.2.1: On-line forecasting procedure

Given an initial estimate x(k-1) with $K_s$ known, the values of x(k/k-1) and z(k/k-1) can be predicted from

(1) $$x(k/k-1) = \phi(k)x(k) \tag{4.55}$$

and

$$z(k/k-1) = H(k)x(k/k-1) \tag{4.56}$$

Upon the availability of the measurement z(k) the estimated state vector can be smoothed as follows

(2) $$x(k) = x(k/k-1) + K_s[z(k) - z(k/k-1)] \tag{4.57}$$

where $r(k) = z(k) - z(k/k-1)$ (4.58)

is the prediction error for the $k^{th}$ discrete interval.

CHAPTER V

# LOAD MODELLING

In short-term load forecasting, the future load on a power system is predicted by extrapolating a pre-determined relationship between the load and its influential variables, namely time and/or weather. Determination of this relationship is a two stage process that requires (a) identifying the relationship between the load and related variables, and (b) quantifying this relationship through the use of a suitable parameter estimation technique.

In order to study the effects of parameter estimation techniques on short-term load forecasting accuracy therefore, it is first necessary to identify and develop suitable load models that will allow for the application of these estimation methods.

In chapter III, static parameter estimation techniques applicable to off-line forecasting were reviewed and in chapter IV dynamic estimation algorithms applicable to on-line forecasting were reviewed. In this chapter, load models are developed that will allow for the study and comparison of the effectiveness of these parameter estimation methods as they apply to short-term load forecasting. Here three off-line load models and a single on-line model are identified and developed for forecast simulations. These models will be used in both summer and winter forecasting modes and as such, where applicable, winter and a summer load formulations are included. These models and the mechanics governing their application are discussed, in turn, in the following two sections.

## 5.1: Off-line load models

In this section three off-line load models are identified for use with the static estimation techniques of chapter III. These models will be referred to as models A,B and C respectively and are identified on the basis of the author's familiarity with their underlying modelling technique.

Models A and B are based on the principles of multiple linear regression and general exponential smoothing respectively, whilst model C is a hybrid load model that embodies both these principles. These models are all developed for the twenty-four off-line forecasting problem i.e. they will be used to predict hourly loads up to twenty-four hours ahead.

In the following subsections each of these models are discussed in greater detail.

### 5.1.2: Model A

This is a multiple linear regression model that expresses the load at any discrete instant t, as a function of a base load and a weather dependent component. The base load is assumed constant for each discrete interval, as is the relationship between the load and its weather dependent variables.

This model will be used for both winter and summer load forecast simulations, and since the relationship between load and weather differ significantly over these two seasons, a different load formulation will be required in each case. This will result in two load models, namely a winter and a summer model.

These models are based on the assumptiom that a common daily base load cycle is experienced by days occuring during the week, i.e. days Monday to Friday, and that a constant but different base load cycle is experienced by weekend days, namely Saturday and Sunday. As such, to continously predict the load over a complete week, two models are required, i.e. one for predicting weekday loads and one for predicting weekend loads.

However, since the primary objective of this thesis is to study and compare the influences of various parameter estimation techniques on predictive accuracy, and not on load forecasting as such, only the weekday load model will be developed for forecast simulation.

Correlation analysis of load and temperature deviations from the norm, has indicated that the load to be modelled is dependent on both the immediate and previous values of temperature deviation. This correlation however, is strongest for immediate values of temperature deviation and dies out in approximately seventy-two hours.

The windchill and wind cooling factors also displayed similar relationships in winter, as did the temperature humidity factor in summer. The wind cooling factor however, was selected in favour of the windchill factor, as it generally resulted in smaller prediction errors during forecast trials.

Based on these analyses, initial winter and summer models were formulated and tested in off-line simulations, and after extensive re-formulation and re-testing using the methods of trial and error, the following two load model formulations were selected.

Winter model

Mathematically, the load at any discrete instant t, where t varies from one to twenty-four, can be expressed as

$$y(t) = a_0(t) + a_1(t)T(t) + a_2(t)T^2(t) + a_3(t)T^3(t) + a_4(t)T(t-1) + a_5(t)T(t-2) + a_6(t)T(t-3) + a_7(t)W(t) + a_8(t)W(t-1) + a_9(t)W(t-2) \quad (5.1)$$

where $y(t)$ = load at time t

$T(t)$ = temperature deviation at time t

$W(t)$ = wind cooling factor at time t

$a_0(t)$ = base load at time t, and

$a_1(t), a_2(t), \ldots, a_9(t)$ are the regression coefficients to be estimated at time t.

The temperature deviation at the instant t, is calculated as the difference between the dry bulb temperature at the time t, and the average dry bulb temperature of the previous twenty (four weeks) weekday temperature measurements, corresponding to the same discrete instant, i.e.

$$T(t) = T_d(t) - T_a(t) \quad (5.2)$$

where $T_d(t)$ = dry bulb temperature at time t, in °C, and

$$T_a(t) = [T_d(t-24) + T_d(t-48) + \quad + T_d(t-480)]/20 \quad (5.3)$$

is the average dry bulb temperature at time t.

It should be noted here, that equations (5.2) and (5.3) refer to a data base comprising only of weekday temperature recordings.

The wind cooling factor is calculated from

$$W(t) = [18 - T_d(t)][V(t)]^{1/2} \tag{5.4}$$

where $V(t)$ = wind speed in km/h at time t.

<u>Summer model</u>

The summmer equivalent of the load model given by equation (5.1) can be written as

$$\begin{aligned} y(t) = \; & a_0(t) + a_1(t)T(t) + a_2(t)T^2(t) + a_3(t)T^3(t) \\ & + a_4(t)T(t-1) + a_5(t)T(t-2) + a_6(t)T(t-3) + a_7(t)H(t) \\ & + a_8(t)H(t-1) + a_9(t)H(t-2) \end{aligned} \tag{5.5}$$

where again

$y(t)$ = load at time t

$T(t)$ = temperature deviation at time t

$a_0(t)$ = base load at time t

$a_1(t), a_2(t), \ldots, a_9(t)$ are the regression coefficients to be estimated at time t, and the temperature deviation is calculated as for the winter model.

In this model however, the wind cooling factor is replaced by the temperature humidity factor H(t), which is given by

$$H(t) = 0.55\ T_d(t) + 0.2\ T_p(t) + 5.05 \qquad (5.6)$$

where $T_p(t)$ = dew point temperature at time t, in °C.

It should be noted, that if the dry bulb temperature is less than twenty-five degrees centigrade, the temperature humidity factor is made equal to zero. This stems from the realisation that the effects of humidity are negligible, when the temperature is below room level.

In both the winter and summer models, the base load and regression coefficients are assumed fixed for each of the twenty-four discrete instants in a day, and as such, twenty-four seperate coefficient estimates are required to predict the next day hourly load profile. The coefficients corresponding at any given discrete interval therefore, are estimated using the previous four weeks of weekday data corresponding to the said discrete instant.

Given therefore, that the load models of equation (5.1) and (5.5) can be rewritten as

$$y(t) = f^T(t)\beta(t) \qquad (5.7)$$

where f(t) is a fitting function given by

$$f(t) = \begin{bmatrix} 1 \\ T(t) \\ T^2(t) \\ T^3(t) \\ T(t-1) \\ T(t-2) \\ T(t-3) \\ W(t) \\ W(t-1) \\ W(t-2) \end{bmatrix} \tag{5.8}$$

in winter, and

$$f(t) = \begin{bmatrix} 1 \\ T(t) \\ T^2(t) \\ T^3(t) \\ T(t-1) \\ T(t-2) \\ T(t-3) \\ H(t) \\ H(t-1) \\ H(t-2) \end{bmatrix} \tag{5.9}$$

in summer, and where $\beta(t)$ is a coefficient vector given by

$$\beta(t) = \begin{bmatrix} a_0(t) \\ a_1(t) \\ \vdots \\ a_8(t) \\ a_9(t) \end{bmatrix} \tag{5.10}$$

the overdetermined system of equations corresponding to the estimate at the instant t, will read

$$\begin{bmatrix} y(t-24) \\ y(t-48) \\ \vdots \\ y(t-960) \end{bmatrix} = \begin{bmatrix} f(t-24) \\ f(t-48) \\ \vdots \\ f(t-960) \end{bmatrix} \begin{bmatrix} \beta(t) \end{bmatrix} \tag{5.11}$$

and can be solved using an appropiate estimation technique. The estimated coefficient vector can now be substituted into equation (5.1) or (5.5) to give the load prediction for the time t.

### 5.1.2: Model B

This is a harmonic decomposition model that expresses the load at any time t, as a function of a constant base load and a Fourier harmonic series. This modelling approach is normally used in conjunction with the general exponential smoothing technique discussed in chapter II, as it satisfies the fitting function relationship given by equation (2.2).

Examination of previous load data revealed the presence of a weekly load cycle that is characterised by distinct daily periodicities. This behavioural pattern is confirmed by the autocorrellation plot shown in figure I.

In this model however, the weekly cycle is accounted for, by the use of a daily load model, whose coefficients are estimated seven times weekly. Also since this is a load shape model that does not take weather into consideration, a single load model will suffice for both winter and summer simulations.

The load at any time t therefore, can be written as

$$y(t) = a_0 + \sum_{i=1}^{m}[a_i \sin(i\omega t) + b_i \cos(i\omega t)] \tag{5.12}$$

where $y(t)$ = load at time t

$m = 9$

$w = 2\pi/24$

$a_0$ = constant base load for each day of the week, and

$a_1, b_1, \ldots, a_9, b_9$ are the coefficients corresponding to the harmonics in the load composition.

It should be noted that it is possible to decompose the load into a maximun of eleven harmonics, but from trial simulations it was found that the use of more than the first nine harmonics, did not result in any significant improvement in accuracy, and as such the tenth and the eleventh harmonics were dropped from the load model.

To predict the hourly load profile for any day of the week, an overdetermined system of equations is once again set up using data from the previous four weeks corresponding to the day in question.

Given again that equation (5.10) can be written in the form of equation (5.7), i.e.

$$y(t) = f^T(t)\beta \tag{5.13}$$

where now

$$f(t) = \begin{bmatrix} 1 \\ \sin(\omega t) \\ \cos(\omega t) \\ \vdots \\ \sin(9\omega t) \\ \cos(9\omega t) \end{bmatrix} \tag{5.14}$$

and

$$\beta = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_8 \\ a_9 \end{bmatrix} \tag{5.15}$$

The overdetermined system of equations can now be written as

$$\begin{bmatrix} y(t-168) \\ \vdots \\ y(t-192) \\ y(t-336) \\ \vdots \\ y(t-672) \end{bmatrix} = \begin{bmatrix} f^T(t-168) \\ \vdots \\ f^T(t-192) \\ f^T(t-336) \\ \vdots \\ f^T(t-672) \end{bmatrix} \begin{bmatrix} \beta \end{bmatrix} \qquad (5.16)$$

and $\beta$ estimated and substituted into equation (5.12) to give the forecast for the next twenty-four hours.

### 5.1.3: Model C

This is a hybrid load model that expresses the load as the sum of a time-varying base load and a weather dependent component. This model was developed with the aim of eliminating the disadvantages of the previous two models by combining their modelling approaches.

Model A possesses the advantage of being weather responsive, but suffer the disadvantages of requiring (a) twenty-four seperate coefficient estimates in order to predictthe next day load, and (b) the use of weekday and weekend models both with winter and summer formulations.

Model B on the other hand, only requires the use of a single model formulation and hence the estimation of a single coefficient vector in order to predict the next day load. However it suffers the major disadvantage of being weather insensitive.

It is evident therefore, that by combining these two load models, a computationally efficient and weather sensitive model will result. This new model will also eliminate the use of seperate weekday and weekend models, as is the case with model A. Also by limiting the weather input to temperature only, a single load model could be used for both winter and summer load forecast simulations. Its main disadvantage however, is that it assumes that the relationship between load and weather is constant for all time of day.

Mathematically therefore, we can express the load at any discrete instant as

$$y(t) = a_0 + \sum_{i=1}^{9}[a_i \sin(i\omega t) + b_i \cos(i\omega t)] + c_0T(t)$$

$$+ c_1T(t-1) + c_2T(t-2) + c_3(t-3) \qquad (5.17)$$

where T(t) is the temperature deviation at time t, and is given by

$$T(t) = T_d(t) - T_c(t) \qquad (5.18)$$

where $T_c(t)$ is the average dry bulb temperature for the discrete instant t, calculated from the previous twenty-eight daily temperature measurements corresponding to the said discrete instant, i.e.

$$T_c(t) = [T_d(t-24) + \ldots\ldots + T_d(t-672)] \qquad (5.19)$$

Alternatively equation (5.16) can be rewritten in the form of equation (5.13), i.e.

$$y(t) = f^T(t)\beta$$

where

$$f(t) = \begin{bmatrix} 1 \\ \cdot \\ \sin(\omega t) \\ \vdots \\ \cos(9\omega t) \\ \cdot \\ T(t) \\ \vdots \\ T(t-3) \end{bmatrix} \qquad (5.20)$$

and

$$\beta = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ b_9 \\ c_0 \\ \vdots \\ c_3 \end{bmatrix} \qquad (5.21)$$

and the vector of coefficients $\beta$ can be estimated as for model B, i.e. from the system of equations given by

$$\begin{bmatrix} y(t-168) \\ \vdots \\ y(t-192) \\ y(t-336) \\ \vdots \\ y(t-672) \end{bmatrix} = \begin{bmatrix} f^T(t-168) \\ \vdots \\ f^T(t-192) \\ f^T(t-336) \\ \vdots \\ f^T(t-672) \end{bmatrix} \begin{bmatrix} \beta \end{bmatrix} \qquad (5.22)$$

The next day forecast can then be made, by substituting for $\beta$ and the predicted values of temperature deviation into equation (5.17).

## 5.2: On-line model

In this section of the chapter, a single load model is identified and developed for on-line load simulations. This model will be referred to as model D, and will be expressed in state space formulation so that it can be directly implemented with the dynamic estimation algorithms of chapter IV.

Model D is a harmonic decomposition model that will be used to predict loads one discrete interval ahead, and is choosen on the basis of its simplicity and ease of application. It should be noted, that while this model will not result in ultimate predictive accuracy, the development of more complex and accurate on-line models is beyond the scope of this thesis.

In the following subsection, the basic principles governing the theory and application of this model, are presented.

### 5.2.1: Model D

Examination of one year of previous load data, has revealed the presence of three basic periodicities in the load structure to be modelled. There is a seasonal cycle with two periods per year, upon which is superimposed a weekly cycle made up of seven daily periodicities.

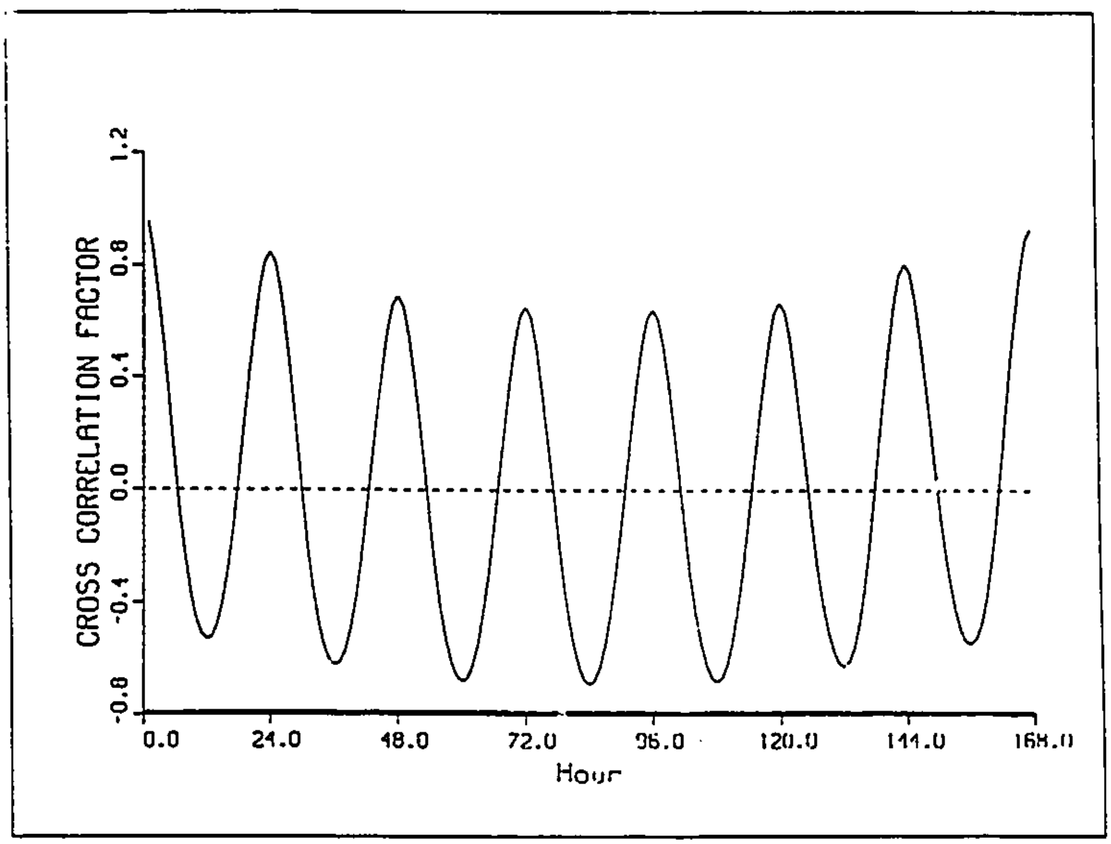
An autocorrelation plot of this data with the seasonal cycle removed, confirms the dominance of weekly cycle over the daily cycle (figure I). This is revealed by the higher correlation factor experienced by load measurements one week (168 hours) apart. A weekly load model is therefore selected for on-line simulation.

This choice of load model will avoid the discontinuities that would otherwise arise with the use of seperate weekday and weekend models, and will result in better accuracy over a daily model. Also because, the estimation process is adaptive and seasonal changes occur relatively slowly, this model will easily accomadate seasonal load changes.

The model choosen to represent the load is a Fourier harmonic series with a period of one week. This model can be expressed mathematically as

FIGURE 5.1

AUTOCORRELATION PLOT OF HOURLY LOAD DATA FOR 1989



$$z(t) = \sum_{i=0}^{m}[a_i \sin(\omega_i t) + b_i \cos(\omega_i t)] + \varepsilon(t) \tag{5.23}$$

$$= b_0 + \sum_{i=1}^{m}[a_i \sin(\omega_i t) + b_i \cos(\omega_i t)] + \varepsilon(t) \tag{5.24}$$

where $z(t)$ = load measurement at time t,

$\varepsilon(t)$ = residual load at time t,

$b_0$ = a constant, and

$a_1, b_1, \ldots, a_m, b_m$ are the coefficients quantifying the influence of their respective harmonics in the load model.

$\omega_i$ = $2\pi i/168$ is the frequency of the $i^{th}$ harmonic.

It should be noted that i is a positive integer less than the Nyquist limit (84), and represents the dominant harmonics in the load composition.

The load model given by equation (5.24) can be rewritten as

$$z(t) = f^T(t)\beta(t) + \varepsilon(t) \tag{5.25}$$

where

$$f(t) = \begin{bmatrix} 1 \\ \sin(\omega_1 t) \\ \cos(\omega_1 t) \\ \vdots \\ \sin(\omega_m t) \\ \cos(\omega_m t) \end{bmatrix} \tag{5.26}$$

and

$$\beta(t) = \begin{bmatrix} b_o \\ a_1 \\ b_1 \\ \vdots \\ a_m \\ b_m \end{bmatrix} \tag{5.27}$$

Since load observations are only available at intervals one hour apart, the load model can be considered to be a discrete process with a sampling time of one hour. We can therefore rewrite equation (5.25) to read

$$z(k) = f^T(k)\beta(k) + \varepsilon(k) \tag{5.28}$$

To accomodate fluctuations in load behaviour from week to week, we can assume that the parameter vector $\beta(k)$ undergoes small changes to account for these variations. Thus we can write

$$\beta(k+1) = \beta(k) + w(k) \tag{5.29}$$

where $w(k)$ = (nx1) noise vector representing the changes in $\beta(k)$

Equations (5.29) and (5.28) together, now constitute a discrete state space model for the load process to be modelled. Here the measurement vector f(k) is a function of k however, and will have to calculated for each discrete instant.

Given that the relationship between fitting functions f(t), seperated by one discrete interval k, is

$$f(k+1) = Lf(k) \tag{5.30}$$

where L is called the transition matrix and is given by

$$L = \left[\begin{array}{ccc|cc} 1 & & & & \\ 0 & \cos(\omega_1) & \sin(\omega_1) & & \\ 0 & -\sin(\omega_1) & \cos(\omega_1) & & \\ \hline 0 & & & \cos(\omega_m) & \sin(\omega_m) \\ 0 & & & -\sin(\omega_m) & \cos(\omega_m) \end{array}\right] \tag{5.31}$$

It can be proven, that if the state and measurement are correspondingly altered to read as follows [25,32]

$$\beta(k+1) = L^T\beta(k) + w(k) \tag{5.32}$$

and

$$z(k) = f^T(0)\beta(k) + \varepsilon(k) \tag{5.33}$$

an equivalent state space model with stationary state transition matrix and measurement vector is obtained.

Equations (5.32) and (5.33) are now in the form of equations (4.1) and (4.2) and can be readily implemented in any of the dynamic estimation algorithms of chapter IV.

From the foregoiong analogy with chapter IV therefore, we get the following equivalents.

$x(k) = \beta(k)$

$\phi(k) = L^T(k)$

$H(k) = f^T(0)$

$v(k) = \varepsilon(k)$

$w(k) = w(k)$, and

$z(k) = z(k)$

It should be noted here, that L(k) and hence $\phi(k)$ are constant for all k. Also since

$$f^T(0) = [1\ 0\ 1\ \ldots\ldots.0\ 1] \tag{5.34}$$

is constant for all k, then H(k) is also fixed for all k.

A power spectrum analysis was performed to identify the dominant harmonics to be used in the load model. Here the seasonal moving average was removed from the hourly load data, so that its periodic behaviour would not influence the relative strenghts of the weekly harmonics under scrutiny.

The power spectrum plot is shown in figure 2. From it, we can see that only twelve harmonics are essential for an accurate representation of the weekly load. The order of these harmonics have been identified as

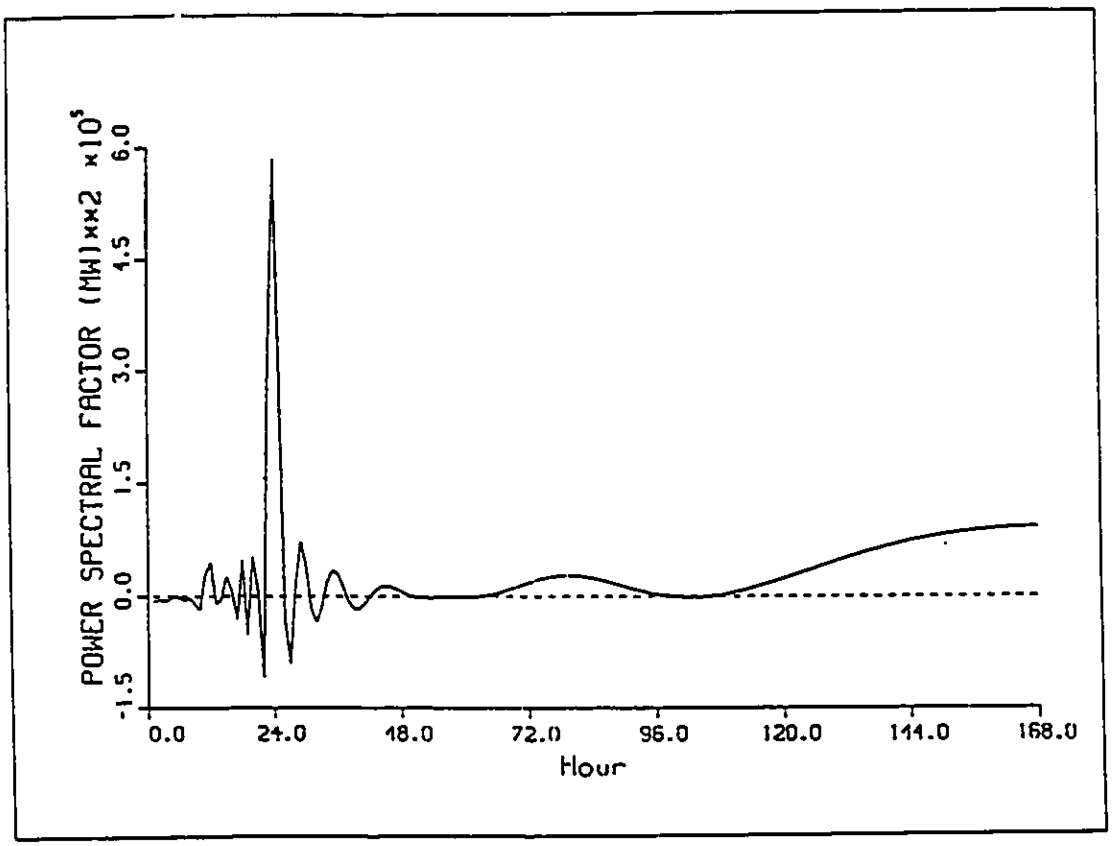
[ 1, 2, 4, 5, 6, 7, 8, 9, 11, 14, 21, and 28]

respectively.

It should be noted, that the addition of other than the dominant harmonics will result in additional noise in the estimation process.

FIGURE 5.2

POWER SPECTRUM OF HARMONICS IN LOAD COMPOSITION



## CHAPTER VI

## LOAD FORECAST SIMULATION

In this chapter, the results of off and on-line load forecast simulations are presented. These forecasts have been simulated using previous load and weather data recorded for the city of Edmonton, Alberta.

Hourly load data for the year 1989, was provided by the Edmonton Power Company, while weather data recorded at the Edmonton munincipal airport, was provided by the Atmospheric Environment Service of Alberta.

The results presented in this chapter has been separated into two sections, namely off and on-line simulations. In the first section, the results of off-line simulations made using the static parameter estimation methods of chapter III are presented, while the second section contains the results of on-line forecasts made using the dynamic estimation algorithms of chapter IV.

In each section, the results presented are followed by indepth discussions reviewing the advantages and disadvantages of each of the parameter estimation algorithm used in that section.

### 6.1: Off-line simulations

In this section, the three off-line load models developed in section 5.1 of chapter V, have been used to predict the next day hourly load profile for selected periods of both the winter and summer of 1989.

For each load model, forecasts are made using each of the three static parameter estimation methods , i.e the least squares , the new LAV and the linear programming techniques, and in each case, these estimation methods are applied to both contaminated and uncontaminated data bases.

In the following subsections, the results obtained via each load model, are presented.

### 6.1.1: Model A

This load model has been used to predict the twenty-four hour load profiles for one day in winter and one day in summmer. Simulations are restricted to a single day in each case, primarily because of the large number of estimates that are required to completely forecast the next day load.

The forecast simulations presented here however, are generally indicative of the responses to be expected during these seasons, and as such will allow for a generalised comparison of the responses of the given parameter estimation methods as they apply to this load model.

It should be noted here, that the over-determined system of equations to be solved for each discrete instant is non-periodic, and as such the new LAV technique will be applied in its unmodified state. Furthermore, since this technique makes use of two least squares estimate in its computations, it will also be used to provide both these least squares estimates.

In the remainder of this subsection, predictions made for the winter and summer day choosen for both un-contaminated and contaminated data sources are presented.

Here the results of winter simulations are presented in tables 6.1 and 6.2, followed by the summer forecast results in tables 6.3 and 6.4. These results are also presented graphically in figures 6.1 ,6.2, 6.3 and 6.4.

For purposes of simplicity of presentation, the following abbreviations will be used hereafter in this section:

$h$ = hour of day

$z$ = actual recorded load

$z_{LS}$ = Load forecast made from least squares estimates

$z_{LS2}$ = Load forecast made from the least squares estimate corresponding to the second estimation stage in the new LAV algorithm.

$z_{LAV}$ = Load forecasted using the new LAV algorithm, and

$z_{LP}$ = Linear programming estimated load

The percentage errors corresponding to the forecasted loads, are given by

$$\varepsilon_{LS} = [(z - z_{LS})/z]\ 100$$

$$\varepsilon_{LS2} = [(z - z_{LS2})/z]\ 100$$

$$\varepsilon_{LAV} = [(z - z_{LAV})/z]\ 100, \text{ and}$$

$$\varepsilon_{LP} = [(z - z_{LP})/z]\ 100$$

TABLE 6.1

MODEL A: WINTER DAY FORECAST

| h | $z$ | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 564 | 550 | 2.45 | 546 | 3.08 | 548 | 2.74 | 550 | 2.33 |
| 2 | 544 | 539 | 0.74 | 534 | 1.79 | 530 | 2.52 | 534 | 1.76 |
| 3 | 541 | 525 | 2.78 | 525 | 2.89 | 524 | 2.96 | 525 | 2.95 |
| 4 | 533 | 517 | 2.93 | 515 | 3.19 | 525 | 1.40 | 516 | 3.16 |
| 5 | 535 | 509 | 4.79 | 504 | 5.62 | 512 | 4.21 | 511 | 4.35 |
| 6 | 558 | 530 | 4.86 | 528 | 5.24 | 533 | 4.46 | 528 | 5.37 |
| 7 | 622 | 605 | 2.69 | 603 | 3.04 | 560 | 9.81 | 599 | 3.56 |
| 8 | 729 | 725 | 0.43 | 726 | 0.29 | 736 | -1.02 | 731 | -0.29 |
| 9 | 768 | 766 | 0.13 | 752 | 2.04 | 758 | 1.28 | 758 | 1.20 |
| 10 | 800 | 793 | 0.76 | 796 | 0.39 | 784 | 1.93 | 790 | 1.19 |
| 11 | 809 | 813 | -0.51 | 814 | -0.69 | 812 | -0.41 | 815 | -0.77 |
| 12 | 818 | 827 | -1.18 | 827 | -1.20 | 831 | -1.69 | 827 | -1.14 |
| 13 | 801 | 807 | -0.81 | 817 | -2.01 | 817 | -2.02 | 818 | -2.21 |
| 14 | 799 | 806 | -0.95 | 816 | -2.14 | 816 | -2.23 | 816 | -2.21 |
| 15 | 796 | 804 | -1.09 | 809 | -1.70 | 843 | -6.00 | 826 | -3.88 |
| 16 | 797 | 797 | -0.11 | 806 | -1.14 | 823 | -3.34 | 798 | -0.14 |
| 17 | 803 | 818 | -1.88 | 814 | -1.40 | 835 | -4.06 | 823 | -2.54 |
| 18 | 822 | 872 | -6.10 | 877 | -6.81 | 981 | -19.39 | 860 | -4.70 |
| 19 | 833 | 836 | -0.39 | 841 | -0.96 | 914 | -9.82 | 826 | 0.83 |
| 20 | 809 | 794 | 1.73 | 808 | 0.08 | 858 | -6.13 | 811 | -0.32 |
| 21 | 784 | 760 | 3.06 | 773 | 1.33 | 780 | 0.38 | 770 | 1.66 |
| 22 | 746 | 722 | 3.15 | 725 | 2.68 | 725 | 2.69 | 718 | 3.71 |
| 23 | 682 | 668 | 1.93 | 664 | 2.62 | 661 | 3.03 | 656 | 3.73 |
| 24 | 613 | 599 | 2.22 | 597 | 2.58 | 646 | -5.44 | 602 | 1.69 |

TABLE 6.2

MODEL A: WINTER DAY FORECAST WITH BAD DATA

| h | $z$ | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 564 | 476 | 15.48 | 550 | 2.41 | 551 | 2.30 | 550 | 2.33 |
| 2 | 544 | 473 | 12.91 | 539 | 0.85 | 609 | -11.99 | 524 | 3.52 |
| 3 | 541 | 471 | 12.76 | 526 | 2.75 | 523 | 3.21 | 525 | 2.95 |
| 4 | 533 | 465 | 12.61 | 518 | 2.74 | 517 | 2.88 | 516 | 3.16 |
| 5 | 535 | 502 | 6.10 | 509 | 4.77 | 503 | 5.95 | 511 | 4.34 |
| 6 | 558 | 526 | 5.65 | 530 | 4.85 | 528 | 5.34 | 528 | 5.37 |
| 7 | 622 | 611 | 1.71 | 604 | 2.76 | 594 | 4.38 | 599 | 3.63 |
| 8 | 729 | 714 | 1.98 | 726 | 0.38 | 732 | -0.53 | 720 | 1.11 |
| 9 | 768 | 630 | 17.89 | 767 | 0.02 | 761 | 0.89 | 751 | 2.09 |
| 10 | 800 | 1036 | -29.57 | 787 | 1.55 | 777 | 2.75 | 808 | -1.12 |
| 11 | 809 | 1066 | -31.85 | 806 | 0.26 | 816 | -0.91 | 821 | -1.53 |
| 12 | 818 | 892 | -9.13 | 829 | -1.36 | 825 | -0.93 | 832 | -1.79 |
| 13 | 801 | 832 | -3.94 | 807 | -0.78 | 809 | -1.02 | 820 | -2.42 |
| 14 | 799 | 941 | -17.85 | 807 | -1.03 | 852 | -6.74 | 824 | -3.15 |
| 15 | 796 | 365 | 54.12 | 751 | 5.63 | 1251 | -57.28 | 398 | 49.95 |
| 16 | 797 | 645 | 18.97 | 795 | 0.15 | 666 | 16.40 | 528 | 33.68 |
| 17 | 803 | 824 | -2.70 | 810 | -0.92 | 810 | -0.89 | 823 | -2.54 |
| 18 | 822 | 1058 | -28.81 | 856 | -4.25 | 851 | -3.60 | 860 | -4.70 |
| 19 | 833 | 1120 | -34.56 | 800 | 3.95 | 796 | 4.35 | 826 | 0.83 |
| 20 | 809 | 937 | -15.89 | 797 | 1.36 | 837 | -3.49 | 811 | -0.32 |
| 21 | 784 | 663 | 15.39 | 765 | 2.37 | 645 | 17.66 | 770 | 1.66 |
| 22 | 746 | 696 | 6.58 | 721 | 3.33 | 657 | 11.90 | 718 | 3.71 |
| 23 | 682 | 666 | 2.26 | 670 | 1.73 | 680 | 0.15 | 656 | 3.75 |
| 24 | 613 | 605 | 1.18 | 602 | 1.74 | 650 | -6.19 | 594 | 2.96 |

TABLE 6.3

MODEL A: SUMMER DAY FORECAST

| h | $z$ | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 537 | 522 | 2.61 | 503 | 6.19 | 498 | 7.12 | 499 | 7.05 |
| 2 | 499 | 491 | 1.47 | 468 | 6.09 | 484 | 2.94 | 475 | 4.64 |
| 3 | 488 | 506 | -3.75 | 532 | -9.05 | 454 | 6.87 | 526 | -7.93 |
| 4 | 487 | 500 | -2.83 | 500 | -2.71 | 502 | -3.23 | 488 | -0.28 |
| 5 | 490 | 493 | -0.64 | 484 | 1.22 | 478 | 2.43 | 487 | 0.53 |
| 6 | 493 | 487 | 1.11 | 487 | 1.22 | 483 | 1.99 | 487 | 1.12 |
| 7 | 529 | 531 | -0.53 | 530 | -0.33 | 482 | 8.74 | 533 | -0.79 |
| 8 | 627 | 649 | -3.51 | 638 | -1.91 | 657 | -4.79 | 642 | -2.49 |
| 9 | 719 | 711 | 1.01 | 713 | 0.76 | 712 | 0.88 | 716 | 0.30 |
| 10 | 760 | 771 | -1.58 | 787 | -3.58 | 790 | -4.05 | 783 | -3.08 |
| 11 | 803 | 804 | -0.18 | 799 | 0.49 | 783 | 2.39 | 792 | 1.28 |
| 12 | 840 | 829 | 1.21 | 818 | 2.52 | 820 | 2.34 | 816 | 2.74 |
| 13 | 845 | 842 | 0.29 | 839 | 0.64 | 841 | 0.46 | 837 | 0.88 |
| 14 | 847 | 847 | -0.02 | 847 | -0.11 | 847 | -0.02 | 846 | 0.01 |
| 15 | 849 | 850 | -0.16 | 849 | -0.09 | 833 | 1.78 | 843 | 0.69 |
| 16 | 844 | 854 | -1.21 | 856 | -1.52 | 858 | -1.69 | 858 | -1.75 |
| 17 | 832 | 839 | -0.90 | 834 | -0.31 | 845 | -1.66 | 841 | -1.15 |
| 18 | 804 | 794 | 1.14 | 795 | 1.02 | 842 | -4.82 | 798 | 0.71 |
| 19 | 749 | 755 | -0.88 | 744 | 0.66 | 701 | 6.34 | 753 | -0.55 |
| 20 | 716 | 721 | -0.82 | 711 | 0.62 | 708 | 1.02 | 713 | 0.36 |
| 21 | 697 | 690 | 0.94 | 692 | 0.64 | 692 | 0.61 | 694 | 0.39 |
| 22 | 670 | 665 | 0.65 | 661 | 1.20 | 669 | 0.11 | 671 | -0.23 |
| 23 | 660 | 626 | 5.07 | 614 | 6.93 | 636 | 3.58 | 622 | 5.69 |
| 24 | 600 | 582 | 2.90 | 581 | 3.03 | 591 | 1.44 | 589 | 1.83 |

TABLE 6.4

MODEL A: SUMMER DAY FORECAST WITH BAD DATA

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 537 | 788 | -46.92 | 575 | -7.09 | 544 | -1.36 | 590 | -10.04 |
| 2 | 499 | 687 | -37.79 | 481 | 3.46 | 613 | -22.90 | 488 | 2.04 |
| 3 | 488 | 547 | -12.22 | 491 | -0.80 | 432 | 11.38 | 507 | -4.02 |
| 4 | 487 | 510 | -4.77 | 497 | -2.18 | 1160 | -138.25 | 483 | 0.72 |
| 5 | 490 | 490 | -0.06 | 489 | 0.05 | 489 | 0.13 | 489 | 0.02 |
| 6 | 493 | 479 | 2.70 | 487 | 1.11 | 484 | 1.63 | 484 | 1.63 |
| 7 | 529 | 400 | 24.36 | 547 | -3.40 | 469 | 11.29 | 510 | 3.43 |
| 8 | 627 | 588 | 6.20 | 653 | -4.30 | 683 | -8.99 | 641 | -2.37 |
| 9 | 719 | 803 | -11.70 | 715 | 0.48 | 770 | -7.14 | 712 | 0.95 |
| 10 | 760 | 1097 | -44.35 | 800 | -5.32 | 785 | -3.41 | 809 | -6.52 |
| 11 | 803 | 867 | -8.00 | 801 | 0.15 | 798 | 0.55 | 792 | 1.28 |
| 12 | 840 | 933 | -11.15 | 830 | 1.19 | 792 | 5.71 | 816 | 2.74 |
| 13 | 845 | 810 | 4.06 | 842 | 0.26 | 850 | -0.60 | 836 | 1.00 |
| 14 | 847 | 789 | 6.81 | 847 | -0.11 | 846 | 0.03 | 846 | 0.01 |
| 15 | 849 | 804 | 5.27 | 852 | -0.41 | 854 | -0.61 | 843 | 0.69 |
| 16 | 844 | 438 | 48.10 | 860 | -1.95 | 860 | -1.93 | 849 | -0.61 |
| 17 | 832 | 826 | 0.68 | 838 | -0.77 | 870 | -4.65 | 841 | -1.15 |
| 18 | 804 | 782 | 2.68 | 797 | 0.84 | 783 | 2.52 | 798 | 0.71 |
| 19 | 749 | 755 | -0.91 | 760 | -1.53 | 593 | 20.72 | 753 | -0.55 |
| 20 | 716 | 590 | 17.50 | 735 | -2.68 | 745 | -4.13 | 709 | 0.88 |
| 21 | 697 | 646 | 7.26 | 681 | 2.20 | 691 | 0.83 | 691 | 0.83 |
| 22 | 670 | 701 | -4.70 | 660 | 1.39 | 631 | 5.79 | 657 | 1.80 |
| 23 | 660 | 647 | 1.90 | 631 | 4.29 | 626 | 5.05 | 622 | 5.67 |
| 24 | 600 | 564 | 5.84 | 584 | 2.55 | 597 | 0.40 | 586 | 2.28 |

FIGURE 6.1

MODEL A: WINTER DAY FORECAST CURVES

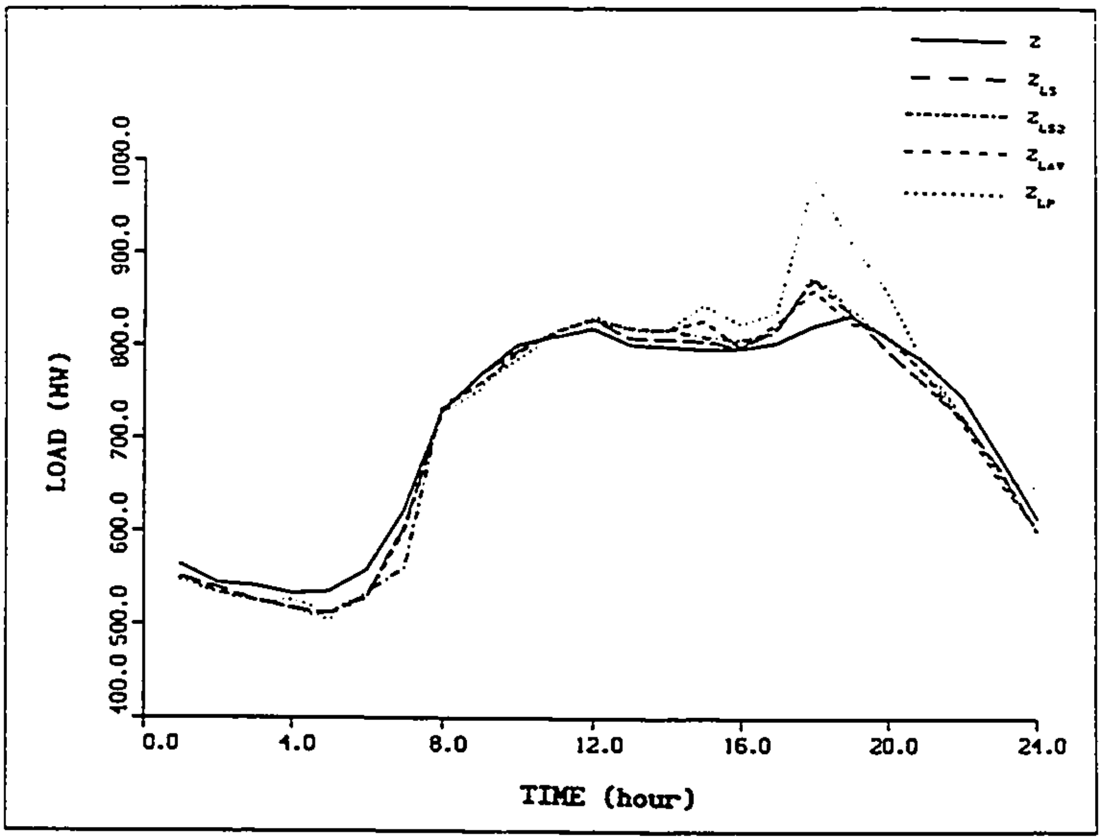


FIGURE 6.2

MODEL A: WINTER DAY BAD DATA FORECAST CURVES

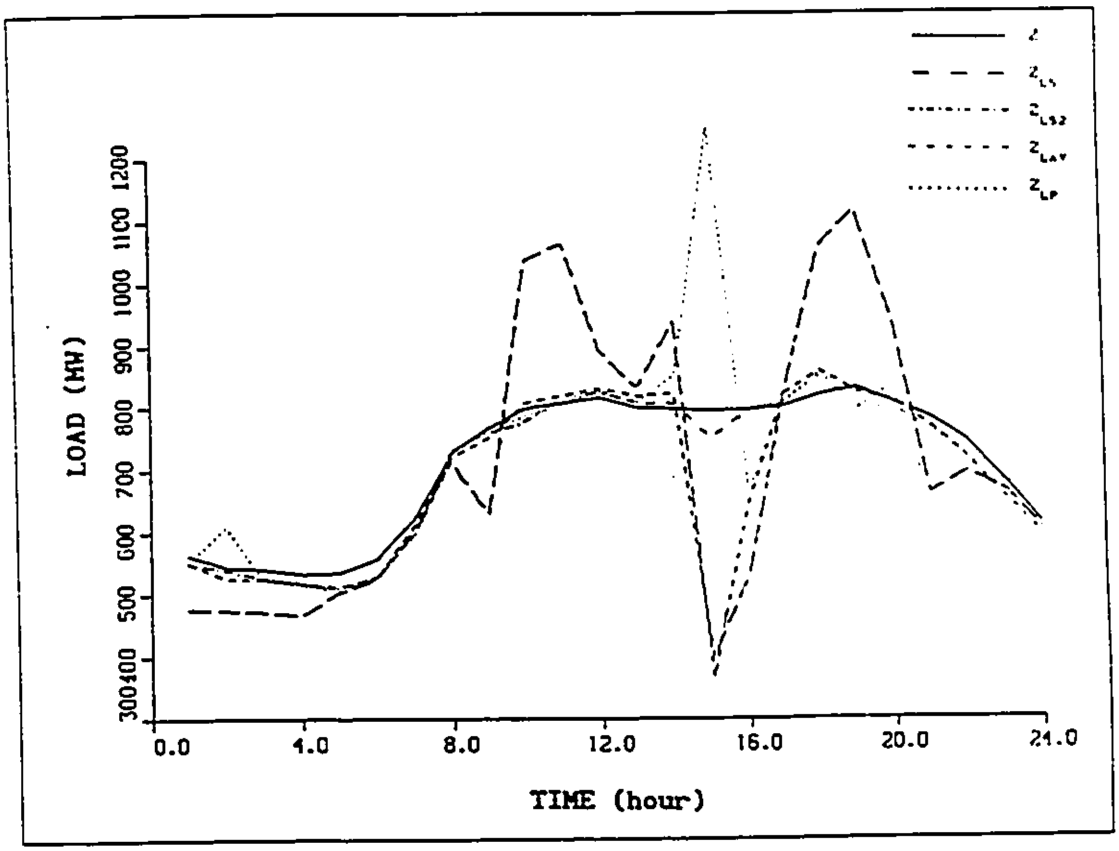


FIGURE 6.3

MODEL A: SUMMER DAY FORECAST CURVES

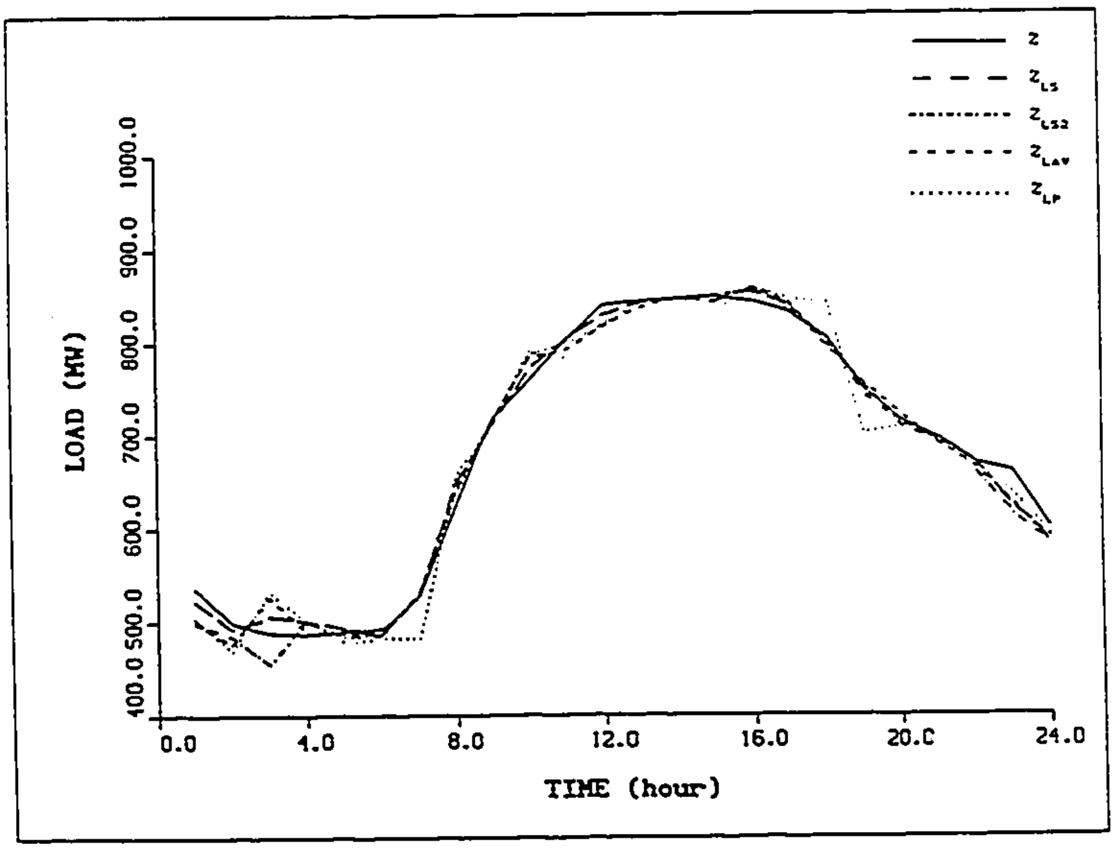
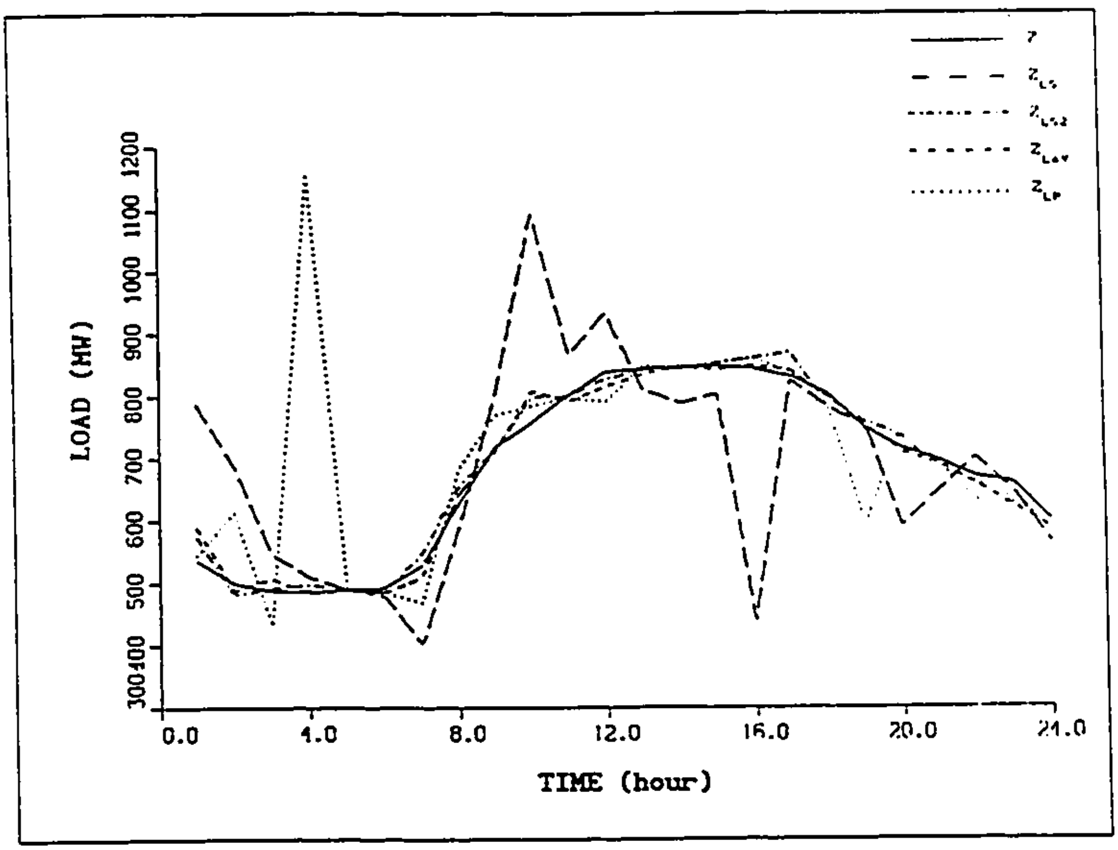


FIGURE 6.4

MODEL A: SUMMER DAY BAD DATA FORECAST CURVES



### 6.1.2: Discussion of Model A results

The multiple linear regression model described as model A, was used to predict the hourly load profile for a single days in winter and summer, corresponding to both contaminated and uncontaminated data bases. The contaminated data bases were created by replacing twenty percent of the original load measurements with gross errors.

Forecasting using this load model was restricted to a single day in either season, primarily because of the excessive volume of computations that are associated with a single twenty-four hour load prediction. For the same reason, the parameters estimated by each of the four methods, are not included in the results presented. It should be noted however, that since a single day's forecast requires twenty-four separate estimates for each estimation method, the results presented will allow for fairly generalised comparison of these techniques.

Examination of the results presented in subsection 6.1 for the uncontaminated data cases, reveals that the predictions made via the four estimates are essentially the same, except for a single hour's prediction where the linear programming estimate resulted in an exaggeration of the winter mid-afternoon peak load.

With twenty percent gross measurement error contamination of the data bases however, the forecasts produced via the various estimation methods are significantly dissimilar. Here we note, that the conventional least squares forecast is severly distorted by very large prediction errors, whilst the predictions made via the reduced data set least squares estimate is relatively unchanged.

The linear programming and the new LAV algorithms on the other hand, produced forecasts that were significantly better than those of obtained from the conventional least squares estimates, but still suffered from poor predictions in a few cases. These large errors can be attributed to the poor conditioning of the data set after the inclusion of the gross error measurements, and can only be remedied by a reformulation of the load model structure.

6.1.3: Model B

Forecasting using this load model involves solving for an overdetermined system comprising of periodically repetitive equations, and as explained in chapter V, this can inhibit the performance of the new LAV algorithm.

In order to combat this short coming however, a modification to the original algorithm was suggested, that resulted in what was referred to as the modified LAV algorithm. This modified LAV procedure will be used hereafter with both models B and C, in place of the original algorithm.

This technique was used to produce three estimates for each of the data sets involved. These are the estimates corresponding to the first, third and final stages in this algorithm, and together with estimates produced by a LAV based linear programming algorithm ,were used to predict the next day hourly load profile for two weeks in both winter and summer of 1989.

The mean absolute error obtained for each hour of day, for these two weeks of winter and summer forecast simulations, are listed in tables 6.5 and 6.6 and presented graphically in figures 6.5 and 6.6 respectively.

Next, the data bases corresponding to an average winter and an average summer day were selected and contaminated with twenty-five percent gross measurement errors. These contaminated data bases were then used to predict the twenty-four hour load for the days selected, and the responses compared to the original predictions.

The responses obtained for both contaminated and uncontaminated data are presented in tables 6.7, 6.8, 6.9 and 6.10 These responses are further illustrated graphically in figures 6.7, 6.8, 6.9, and 6.10. In each case, four sets of parameters were estimated. These are presented in tables 6.11, 6.12, 6.13 and 6.14.

In this and the next load model, the following symbols will be used to represent all estimated parameters.

$\theta_{LS}$ = least squares parameters

$\theta_{LS2}$ = least squares parameters corresponding to the forecasted load $z_{LS2}$

$\theta_{LAV}$ = new LAV parameters

$\theta_{LP}$ = linear programming parameters,

and i refers to $i^{th}$ coefficient in the parameter(coefficient) vector.

TABLE 6.5

MODEL B: HOURLY MEAN ABSOLUTE WINTER FORECAST ERRORS

| h | $\varepsilon_{LS}$ | $\varepsilon_{LS2}$ | $\varepsilon_{LAV}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|
| 1 | 7.393 | 7.930 | 7.325 | 7.857 |
| 2 | 8.183 | 8.780 | 8.266 | 8.458 |
| 3 | 8.535 | 8.799 | 7.885 | 8.853 |
| 4 | 8.305 | 9.052 | 8.066 | 8.915 |
| 5 | 8.790 | 9.358 | 8.207 | 8.556 |
| 6 | 8.376 | 9.181 | 7.635 | 8.993 |
| 7 | 6.586 | 7.279 | 6.978 | 6.701 |
| 8 | 4.952 | 5.863 | 6.738 | 5.575 |
| 9 | 4.485 | 3.893 | 6.770 | 4.519 |
| 10 | 5.036 | 4.596 | 6.182 | 4.487 |
| 11 | 5.379 | 5.026 | 6.258 | 5.321 |
| 12 | 5.333 | 5.091 | 5.920 | 5.110 |
| 13 | 5.700 | 5.241 | 5.676 | 4.983 |
| 14 | 5.569 | 5.159 | 5.863 | 5.305 |
| 15 | 5.375 | 5.218 | 6.112 | 4.927 |
| 16 | 4.983 | 4.665 | 5.777 | 4.272 |
| 17 | 4.500 | 4.549 | 6.205 | 4.713 |
| 18 | 5.054 | 5.512 | 7.081 | 5.031 |
| 19 | 4.314 | 4.256 | 5.228 | 4.237 |
| 20 | 4.302 | 4.034 | 4.627 | 4.470 |
| 21 | 4.312 | 4.199 | 4.769 | 4.904 |
| 22 | 4.604 | 4.181 | 4.955 | 4.959 |
| 23 | 5.423 | 5.124 | 5.464 | 5.475 |
| 24 | 6.327 | 5.828 | 6.445 | 6.287 |

TABLE 6.6

MODEL B: HOURLY MEAN ABSOLUTE SUMMER FORECAST ERRORS

| h | $\varepsilon_{LS}$ | $\varepsilon_{LS2}$ | $\varepsilon_{LAV}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|
| 1 | 5.063 | 5.692 | 6.767 | 5.175 |
| 2 | 5.485 | 6.033 | 6.380 | 5.168 |
| 3 | 4.950 | 5.215 | 6.570 | 5.641 |
| 4 | 5.067 | 5.307 | 6.834 | 4.974 |
| 5 | 4.862 | 4.981 | 6.862 | 4.602 |
| 6 | 5.549 | 5.906 | 8.625 | 5.293 |
| 7 | 5.113 | 5.448 | 7.501 | 5.103 |
| 8 | 4.145 | 4.702 | 6.278 | 4.709 |
| 9 | 3.771 | 4.429 | 5.714 | 3.747 |
| 10 | 3.819 | 4.004 | 5.274 | 3.899 |
| 11 | 3.897 | 4.223 | 5.286 | 4.783 |
| 12 | 4.100 | 4.567 | 5.546 | 4.504 |
| 13 | 4.068 | 4.284 | 5.600 | 4.219 |
| 14 | 4.271 | 4.309 | 5.743 | 4.536 |
| 15 | 4.374 | 4.520 | 5.565 | 4.350 |
| 16 | 4.591 | 4.624 | 5.695 | 4.931 |
| 17 | 4.564 | 4.648 | 5.517 | 4.664 |
| 18 | 4.651 | 4.628 | 5.491 | 4.585 |
| 19 | 4.713 | 4.667 | 5.575 | 5.066 |
| 20 | 5.185 | 5.229 | 6.115 | 5.316 |
| 21 | 5.391 | 5.594 | 6.234 | 5.699 |
| 22 | 5.007 | 5.050 | 6.406 | 4.929 |
| 23 | 6.021 | 5.940 | 7.580 | 5.678 |
| 24 | 5.756 | 6.031 | 6.596 | 6.435 |

FIGURE 6.5

MODEL B: MEAN ABSOLUTE WINTER FORECAST ERRORS

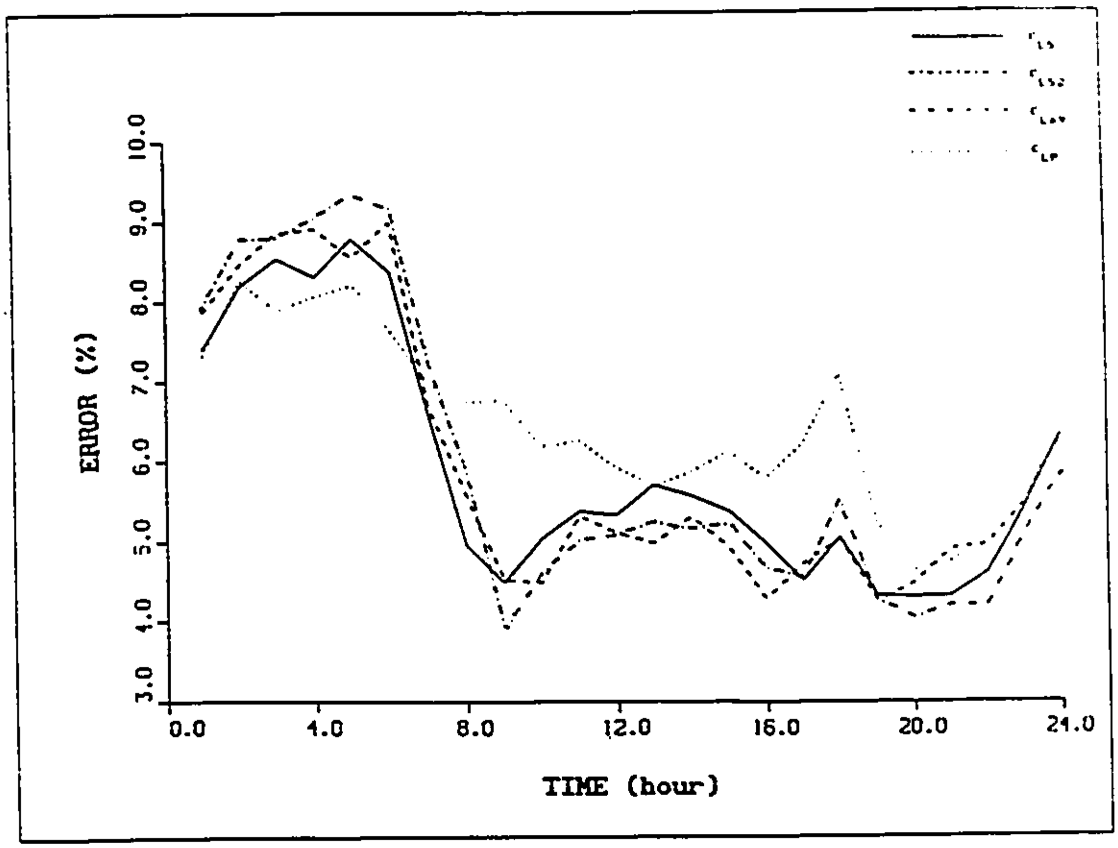


FIGURE 6.6

MODEL B: MEAN ABSOLUTE SUMMER FORECAST ERRORS

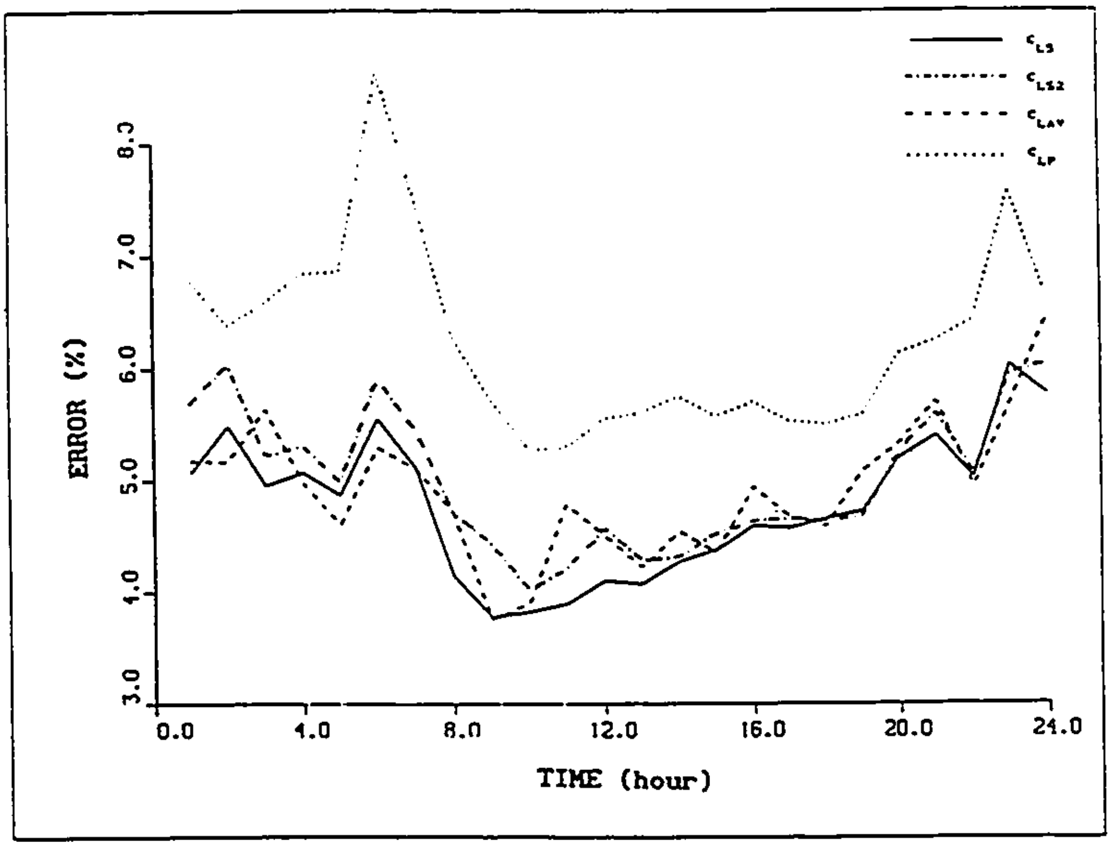


TABLE 6.7

MODEL B: WINTER DAY FORECAST

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 557 | 571 | -2.63 | 544 | 2.26 | 538 | 3.35 | 565 | -1.51 |
| 2 | 533 | 548 | -2.99 | 522 | 2.06 | 510 | 4.19 | 545 | -2.39 |
| 3 | 518 | 536 | -3.49 | 507 | 2.07 | 501 | 3.20 | 502 | 3.01 |
| 4 | 516 | 527 | -2.13 | 499 | 3.19 | 492 | 4.60 | 491 | 4.78 |
| 5 | 514 | 529 | -3.00 | 501 | 2.42 | 494 | 3.75 | 514 | -0.12 |
| 6 | 532 | 544 | -2.37 | 516 | 2.96 | 514 | 3.35 | 532 | -0.09 |
| 7 | 608 | 608 | -0.07 | 584 | 3.84 | 580 | 4.57 | 595 | 2.01 |
| 8 | 719 | 714 | 0.66 | 694 | 3.41 | 693 | 3.48 | 706 | 1.70 |
| 9 | 763 | 786 | -3.04 | 768 | -0.68 | 768 | -0.67 | 768 | -0.77 |
| 10 | 785 | 808 | -3.05 | 786 | -0.16 | 779 | 0.73 | 779 | 0.75 |
| 11 | 794 | 821 | -3.41 | 793 | 0.08 | 788 | 0.75 | 793 | 0.07 |
| 12 | 792 | 828 | -4.60 | 802 | -1.38 | 796 | -0.60 | 798 | -0.81 |
| 13 | 785 | 821 | -4.70 | 794 | -1.25 | 789 | -0.57 | 796 | -1.43 |
| 14 | 783 | 814 | -4.05 | 788 | -0.75 | 785 | -0.32 | 804 | -2.78 |
| 15 | 780 | 814 | -4.40 | 792 | -1.56 | 782 | -0.32 | 794 | -1.83 |
| 16 | 784 | 811 | -3.51 | 785 | -0.24 | 783 | 0.02 | 785 | -0.25 |
| 17 | 786 | 825 | -5.06 | 811 | -3.21 | 820 | -4.36 | 821 | -4.50 |
| 18 | 795 | 856 | -7.70 | 853 | -7.32 | 856 | -7.75 | 849 | -6.89 |
| 19 | 820 | 852 | -3.91 | 833 | -1.66 | 831 | -1.36 | 833 | -1.68 |
| 20 | 798 | 821 | -2.98 | 792 | 0.67 | 786 | 1.44 | 802 | -0.61 |
| 21 | 778 | 797 | -2.48 | 772 | 0.75 | 769 | 1.13 | 772 | 0.76 |
| 22 | 728 | 754 | -3.60 | 727 | 0.02 | 735 | -0.99 | 733 | -0.79 |
| 23 | 661 | 688 | -4.20 | 659 | 0.16 | 669 | -1.21 | 669 | -1.34 |
| 24 | 585 | 622 | -6.41 | 593 | -1.48 | 597 | -2.20 | 597 | -2.08 |

TABLE 6.8

MODEL B: WINTER DAY FORECAST WITH BAD DATA

| h | $z$ | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 557 | 457 | 17.83 | 551 | 0.95 | 536 | 3.75 | 537 | 3.52 |
| 2 | 533 | 452 | 15.03 | 583 | -9.45 | 537 | -0.91 | 507 | 4.86 |
| 3 | 518 | 433 | 16.34 | 526 | -1.66 | 531 | -2.55 | 499 | 3.56 |
| 4 | 516 | 415 | 19.49 | 432 | 16.18 | 487 | 5.44 | 493 | 4.38 |
| 5 | 514 | 431 | 16.13 | 506 | 1.43 | 494 | 3.73 | 494 | 3.85 |
| 6 | 532 | 447 | 15.94 | 576 | -8.40 | 530 | 0.28 | 513 | 3.49 |
| 7 | 608 | 485 | 20.21 | 605 | 0.49 | 598 | 1.54 | 580 | 4.55 |
| 8 | 719 | 564 | 21.44 | 705 | 1.88 | 705 | 1.88 | 693 | 3.51 |
| 9 | 763 | 625 | 18.07 | 780 | -2.33 | 768 | -0.78 | 768 | -0.67 |
| 10 | 785 | 642 | 18.20 | 821 | -4.62 | 797 | -1.55 | 777 | 0.98 |
| 11 | 794 | 641 | 19.18 | 831 | -4.76 | 812 | -2.30 | 786 | 0.97 |
| 12 | 792 | 646 | 18.38 | 805 | -1.70 | 798 | -0.79 | 798 | -0.77 |
| 13 | 785 | 654 | 16.62 | 820 | -4.50 | 789 | -0.54 | 788 | -0.41 |
| 14 | 783 | 647 | 17.37 | 833 | -6.39 | 799 | -2.10 | 782 | 0.03 |
| 15 | 780 | 635 | 18.51 | 815 | -4.50 | 808 | -3.67 | 787 | -0.93 |
| 16 | 784 | 636 | 18.75 | 801 | -2.22 | 798 | -1.88 | 785 | -0.24 |
| 17 | 786 | 653 | 16.81 | 814 | -3.64 | 804 | -2.29 | 797 | -1.47 |
| 18 | 795 | 673 | 15.34 | 867 | -9.18 | 845 | -6.35 | 825 | -3.85 |
| 19 | 820 | 663 | 19.14 | 857 | -4.58 | 841 | -2.61 | 829 | -1.18 |
| 20 | 798 | 643 | 19.33 | 801 | -0.48 | 798 | -0.01 | 802 | -0.62 |
| 21 | 778 | 634 | 18.41 | 797 | -2.55 | 774 | 0.42 | 772 | 0.75 |
| 22 | 728 | 601 | 17.35 | 766 | -5.32 | 732 | -0.64 | 730 | -0.38 |
| 23 | 661 | 545 | 17.51 | 696 | -5.30 | 673 | -1.89 | 668 | -1.18 |
| 24 | 585 | 489 | 16.26 | 608 | -4.02 | 602 | -2.93 | 598 | -2.32 |

TABLE 6.9

MODEL B: SUMMER DAY FORECAST

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 536 | 573 | -6.95 | 575 | -7.30 | 585 | -9.22 | 564 | -5.23 |
| 2 | 510 | 541 | -6.11 | 544 | -6.75 | 546 | -7.10 | 544 | -6.75 |
| 3 | 485 | 513 | -5.78 | 510 | -5.36 | 510 | -5.34 | 515 | -6.25 |
| 4 | 476 | 503 | -5.80 | 494 | -3.88 | 491 | -3.23 | 492 | -3.55 |
| 5 | 476 | 506 | -6.35 | 499 | -4.86 | 495 | -4.16 | 506 | -6.39 |
| 6 | 491 | 505 | -2.90 | 497 | -1.35 | 499 | -1.72 | 502 | -2.40 |
| 7 | 536 | 553 | -3.25 | 548 | -2.27 | 552 | -3.00 | 543 | -1.41 |
| 8 | 644 | 649 | -0.92 | 646 | -0.44 | 653 | -1.41 | 652 | -1.27 |
| 9 | 732 | 741 | -1.24 | 732 | -0.10 | 729 | 0.29 | 737 | -0.77 |
| 10 | 791 | 814 | -2.92 | 810 | -2.48 | 802 | -1.40 | 807 | -2.04 |
| 11 | 821 | 853 | -3.98 | 859 | -4.72 | 863 | -5.13 | 864 | -5.35 |
| 12 | 838 | 867 | -3.55 | 876 | -4.57 | 876 | -4.54 | 881 | -5.16 |
| 13 | 848 | 880 | -3.82 | 887 | -4.65 | 881 | -3.91 | 882 | -4.12 |
| 14 | 862 | 883 | -2.50 | 884 | -2.61 | 887 | -2.91 | 882 | -2.32 |
| 15 | 855 | 883 | -3.32 | 882 | -3.18 | 883 | -3.39 | 883 | -3.34 |
| 16 | 856 | 881 | -2.99 | 882 | -3.15 | 882 | -3.11 | 883 | -3.19 |
| 17 | 843 | 860 | -2.03 | 862 | -2.34 | 863 | -2.40 | 861 | -2.21 |
| 18 | 819 | 822 | -0.44 | 824 | -0.73 | 822 | -0.49 | 824 | -0.63 |
| 19 | 788 | 783 | 0.54 | 782 | 0.68 | 782 | 0.64 | 782 | 0.76 |
| 20 | 755 | 754 | 0.11 | 753 | 0.13 | 753 | 0.17 | 753 | 0.18 |
| 21 | 725 | 723 | 0.17 | 725 | -0.06 | 725 | -0.09 | 726 | -0.16 |
| 22 | 705 | 690 | 2.05 | 687 | 2.48 | 689 | 2.14 | 688 | 2.32 |
| 23 | 670 | 663 | 0.90 | 664 | 0.87 | 664 | 0.77 | 666 | 0.53 |
| 24 | 610 | 622 | -1.99 | 626 | -2.71 | 634 | -4.01 | 622 | -2.03 |

TABLE 6.10

MODEL B: SUMMER DAY FORECAST WITH BAD DATA

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 536 | 470 | 12.21 | 552 | -3.14 | 568 | -6.03 | 564 | -5.27 |
| 2 | 510 | 432 | 15.24 | 477 | 6.43 | 518 | -1.64 | 524 | -2.86 |
| 3 | 485 | 405 | 16.33 | 448 | 7.57 | 505 | -4.31 | 504 | -3.97 |
| 4 | 476 | 414 | 12.90 | 456 | 4.19 | 489 | -2.92 | 502 | -5.59 |
| 5 | 476 | 421 | 11.47 | 469 | 1.28 | 495 | -4.14 | 501 | -5.28 |
| 6 | 491 | 404 | 17.57 | 445 | 9.35 | 490 | 0.19 | 491 | -0.10 |
| 7 | 536 | 436 | 18.54 | 514 | 4.03 | 535 | 0.14 | 543 | -1.37 |
| 8 | 644 | 526 | 18.21 | 666 | -3.52 | 661 | -2.66 | 652 | -1.38 |
| 9 | 732 | 598 | 18.27 | 742 | -1.40 | 738 | -0.85 | 730 | 0.21 |
| 10 | 791 | 631 | 20.22 | 804 | -1.75 | 793 | -0.37 | 793 | -0.34 |
| 11 | 821 | 656 | 20.04 | 855 | -4.18 | 857 | -4.41 | 851 | -3.65 |
| 12 | 838 | 689 | 17.72 | 859 | -2.53 | 875 | -4.51 | 869 | -3.79 |
| 13 | 848 | 705 | 16.84 | 887 | -4.68 | 886 | -4.60 | 881 | -4.01 |
| 14 | 862 | 686 | 20.37 | 887 | -2.95 | 880 | -2.11 | 882 | -2.36 |
| 15 | 855 | 679 | 20.52 | 872 | -2.09 | 871 | -1.94 | 871 | -1.97 |
| 16 | 856 | 700 | 18.13 | 887 | -3.66 | 886 | -3.57 | 879 | -2.69 |
| 17 | 843 | 690 | 18.06 | 861 | -2.14 | 863 | -2.38 | 862 | -2.32 |
| 18 | 819 | 639 | 21.86 | 822 | -0.46 | 823 | -0.53 | 817 | 0.22 |
| 19 | 788 | 608 | 22.79 | 786 | 0.22 | 784 | 0.44 | 781 | 0.82 |
| 20 | 755 | 604 | 19.87 | 747 | 1.03 | 747 | 1.05 | 754 | 0.11 |
| 21 | 725 | 582 | 19.64 | 731 | -0.92 | 731 | -0.87 | 725 | -0.13 |
| 22 | 705 | 543 | 22.90 | 689 | 2.23 | 681 | 3.28 | 684 | 2.87 |
| 23 | 670 | 520 | 22.32 | 656 | 2.07 | 642 | 4.17 | 641 | 4.27 |
| 24 | 610 | 501 | 17.79 | 635 | -4.22 | 632 | -3.63 | 607 | 0.45 |

FIGURE 6.7

MODEL B: WINTER DAY FORECAST

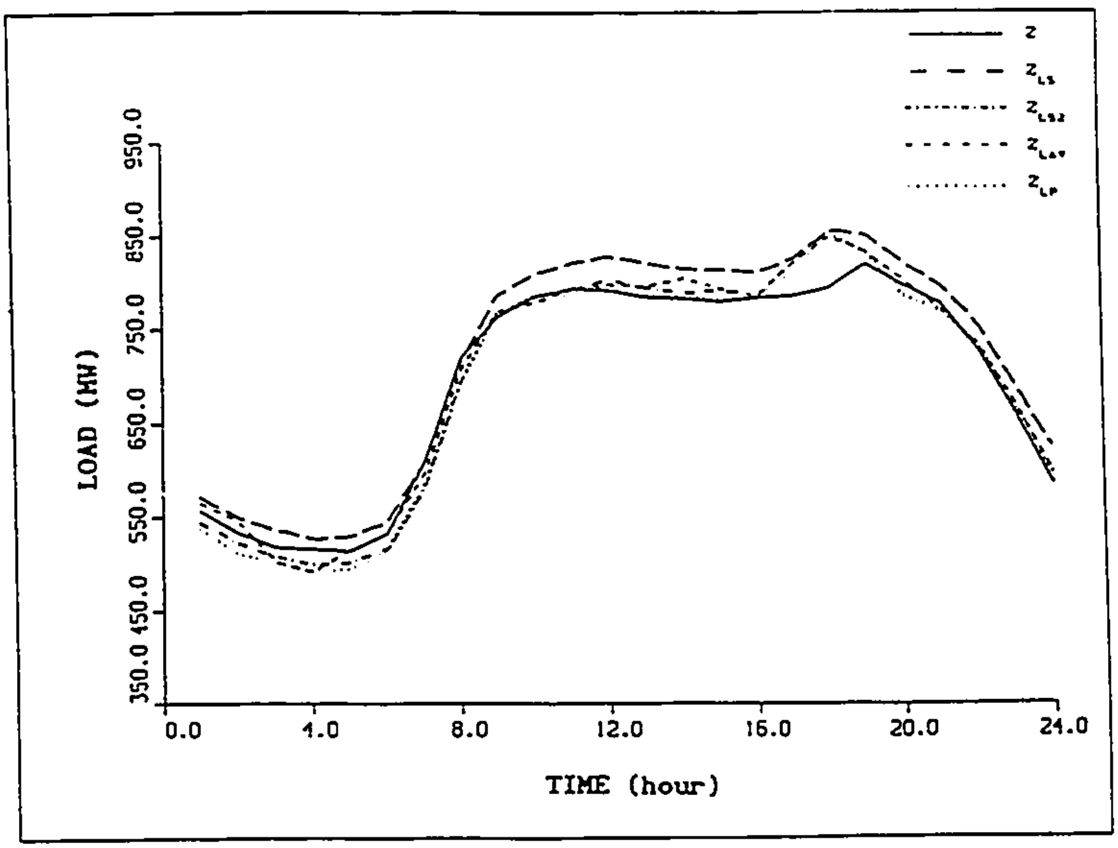


FIGURE 6.8

MODEL B: WINTER DAY FORECAST WITH BAD DATA

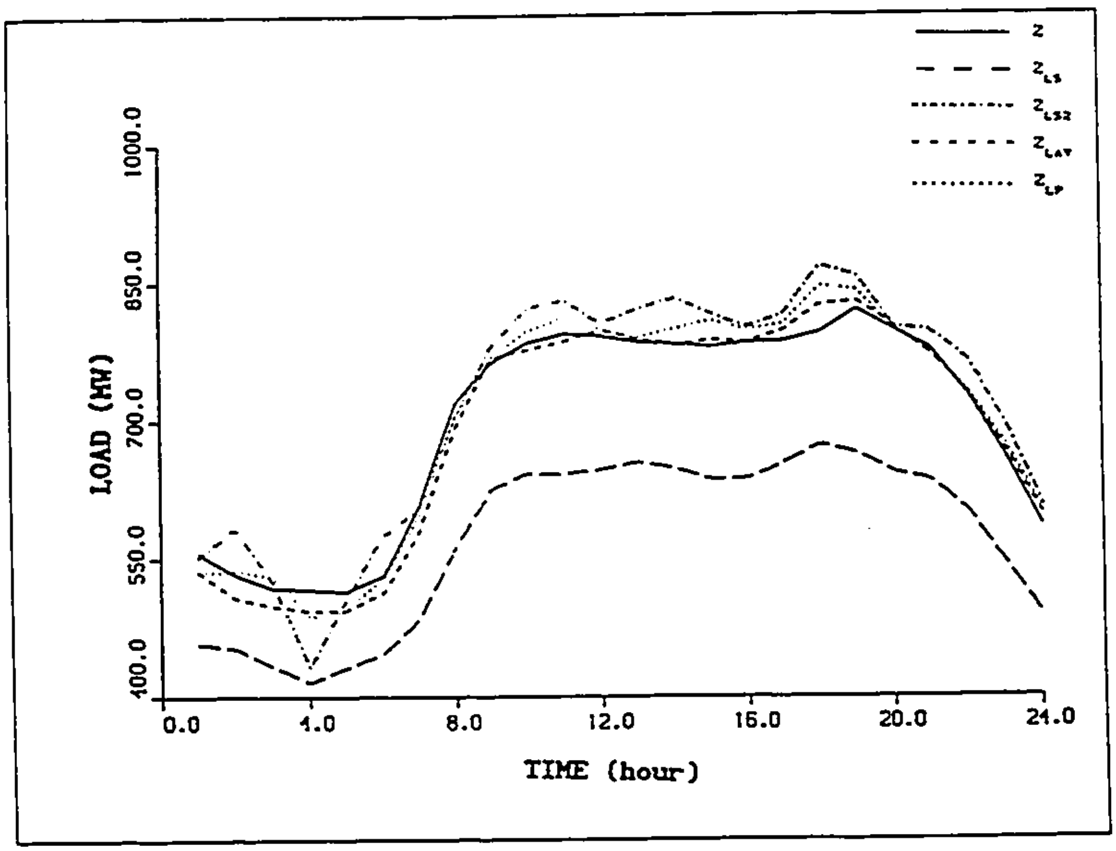


FIGURE 6.9

MODEL B:  SUMMER DAY FORECAST

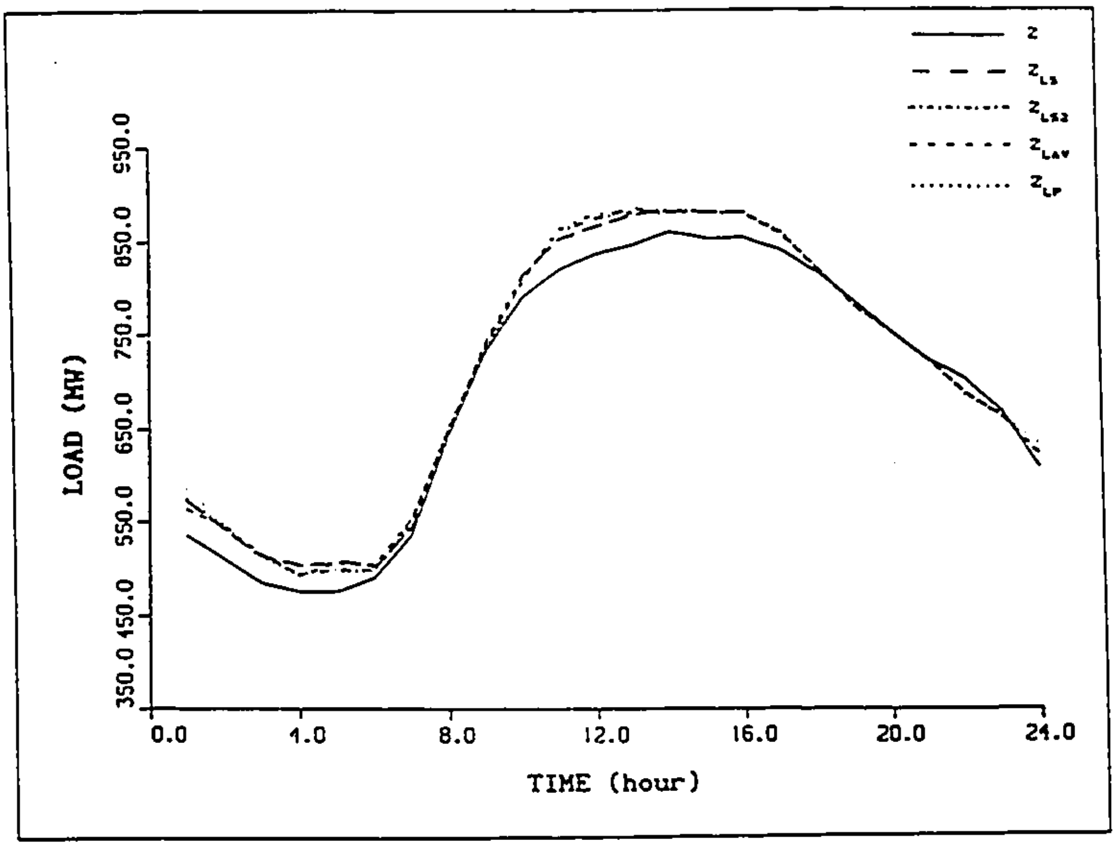


FIGURE 6.10

MODEL B: SUMMER DAY FORECAST WITH BAD DATA

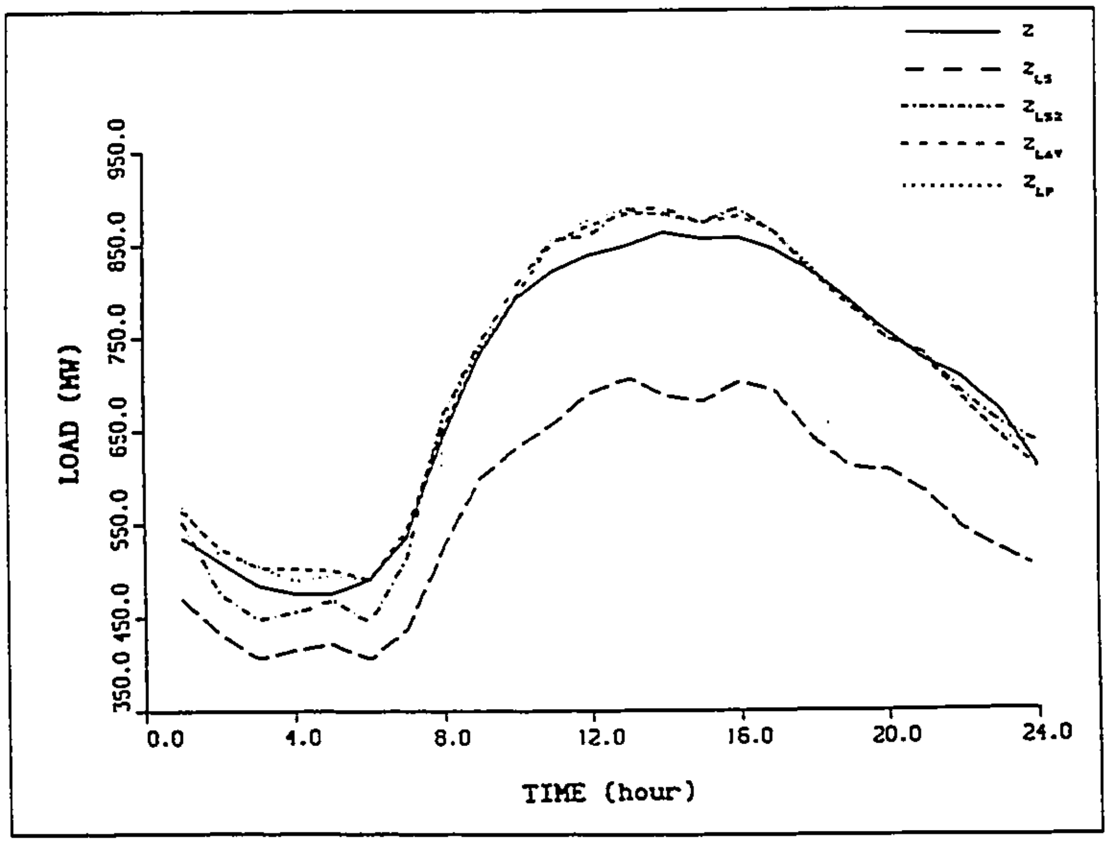


TABLE 6.11

MODEL B: WINTER DAY ESTIMATED PARAMETERS

| $i$ | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 721.059 | 696.937 | 694.355 | 702.326 |
| $a_1$ | -122.459 | -125.288 | -128.086 | -123.493 |
| $b_1$ | -100.004 | -102.281 | -100.660 | -98.743 |
| $a_2$ | -58.054 | -58.908 | -61.446 | -57.423 |
| $b_2$ | 8.354 | 4.487 | 3.879 | 4.292 |
| $a_3$ | 16.818 | 20.103 | 18.890 | 16.971 |
| $b_3$ | -4.271 | -2.031 | 1.196 | 1.843 |
| $a_4$ | 0.609 | 0.666 | -2.307 | 5.539 |
| $b_4$ | -10.319 | -9.235 | -7.417 | -6.738 |
| $a_5$ | -10.690 | -13.503 | -16.669 | -10.088 |
| $b_5$ | -0.148 | -0.957 | -2.503 | -1.165 |
| $a_6$ | 1.002 | 2.225 | 2.850 | 5.168 |
| $b_6$ | 0.040 | -1.655 | -1.946 | -4.223 |
| $a_7$ | 3.508 | 5.764 | 5.002 | 5.790 |
| $b_7$ | -3.890 | -3.841 | -3.177 | -5.604 |
| $a_8$ | -1.244 | -1.603 | -2.471 | -5.270 |
| $b_8$ | 1.842 | 3.451 | 3.380 | -2.041 |
| $a_9$ | -0.947 | -1.846 | -0.060 | -0.091 |
| $b_9$ | 1.023 | 0.106 | 0.948 | -0.463 |

TABLE 6.12

MODEL B: WINTER DAY ESTIMATED BAD DATA PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 571.751 | 716.974 | 702.768 | 692.444 |
| $a_1$ | -91.073 | -128.147 | -123.419 | -125.466 |
| $b_1$ | -75.307 | -104.353 | -102.658 | -99.836 |
| $a_2$ | -43.872 | -65.270 | -59.664 | -63.509 |
| $b_2$ | 6.270 | 11.929 | 5.996 | 6.685 |
| $a_3$ | 11.684 | 14.966 | 18.989 | 14.750 |
| $b_3$ | -3.180 | 3.397 | 0.776 | -1.538 |
| $a_4$ | 1.403 | 7.393 | 2.367 | 0.853 |
| $b_4$ | -8.940 | -6.726 | -13.205 | -11.437 |
| $a_5$ | -7.504 | -2.982 | -9.653 | -12.436 |
| $b_5$ | -2.079 | -3.658 | -1.531 | 1.115 |
| $a_6$ | 5.549 | 0.573 | -6.956 | -0.168 |
| $b_6$ | -4.442 | -25.005 | -6.079 | 3.003 |
| $a_7$ | 0.224 | -3.520 | 1.965 | 2.502 |
| $b_7$ | -3.534 | -9.846 | -3.715 | -4.972 |
| $a_8$ | -2.057 | -9.593 | -4.379 | -1.566 |
| $b_8$ | 0.843 | 2.681 | 3.561 | 2.527 |
| $a_9$ | -1.463 | -4.350 | -2.328 | 0.431 |
| $b_9$ | 1.738 | 9.989 | 4.484 | 0.610 |

TABLE 6.13

MODEL B: SUMMER DAY ESTIMATED PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 711.332 | 710.863 | 711.562 | 711.203 |
| $a_1$ | -135.731 | -139.150 | -139.019 | -137.737 |
| $b_1$ | -136.177 | -137.233 | -135.158 | -138.310 |
| $a_2$ | -17.096 | -16.456 | -16.393 | -17.293 |
| $b_2$ | 37.497 | 41.335 | 42.288 | 39.942 |
| $a_3$ | 14.487 | 15.734 | 15.018 | 15.147 |
| $b_3$ | 4.824 | 4.390 | 6.953 | 2.967 |
| $a_4$ | -10.177 | -9.368 | -7.614 | -11.269 |
| $b_4$ | -7.061 | -4.712 | -3.361 | -5.779 |
| $a_5$ | -2.698 | -1.703 | -0.179 | -2.053 |
| $b_5$ | 3.486 | 1.548 | 1.033 | 0.749 |
| $a_6$ | 2.327 | 2.232 | 1.084 | 1.198 |
| $b_6$ | 2.406 | 3.193 | 3.961 | 3.381 |
| $a_7$ | -0.360 | -0.279 | 1.946 | -0.380 |
| $b_7$ | -0.551 | -0.157 | 0.367 | -2.709 |
| $a_8$ | -1.776 | -2.838 | -3.749 | -6.135 |
| $b_8$ | -2.108 | -2.228 | -2.390 | -1.438 |
| $a_9$ | -1.910 | -3.012 | -1.067 | -2.785 |
| $b_9$ | 0.904 | 1.742 | 2.708 | 2.972 |

TABLE 6.14

MODEL B: SUMMER DAY ESTIMATED BAD DATA PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 572.195 | 708.385 | 713.691 | 709.069 |
| $a_1$ | -112.699 | -170.051 | -159.152 | -149.926 |
| $b_1$ | -98.876 | -144.070 | -136.860 | -128.646 |
| $a_2$ | -18.042 | -39.274 | -24.875 | -22.417 |
| $b_2$ | 27.393 | 44.044 | 40.0[illegible]3 | 35.219 |
| $a_3$ | 11.566 | 14.080 | 18.326 | 14.551 |
| $b_3$ | 4.781 | 14.438 | 8.767 | 5.591 |
| $a_4$ | -8.653 | -12.977 | -13.876 | -12.833 |
| $b_4$ | -3.954 | -4.497 | -5.020 | -3.357 |
| $a_5$ | -4.537 | -11.870 | -9.353 | -7.450 |
| $b_5$ | 1.078 | 7.703 | 7.419 | 4.670 |
| $a_6$ | 14.038 | 10.313 | 10.097 | 16.145 |
| $b_6$ | 11.516 | 12.566 | 7.710 | 15.031 |
| $a_7$ | -2.817 | 4.415 | 2.111 | 0.581 |
| $b_7$ | 2.135 | 6.170 | 4.061 | -0.061 |
| $a_8$ | -3.052 | -4.058 | -8.281 | -4.576 |
| $b_8$ | 2.481 | 2.195 | 3.463 | 4.278 |
| $a_9$ | -1.153 | -3.686 | -3.205 | -0.724 |
| $b_9$ | 0.201 | 8.328 | 3.554 | 3.504 |

6.1.4: Discussion of Model B results

The harmonic decomposition model labelled as model B, was used to forecast the next day load hourly load profile for two weeks periods of both winter and summer load. The mean absolute values of the daily errors are presented in tables 6.5 and 6.6, and shown in figures 6.5 and 6.6. These results were obtained from uncontaminated data sources and were simulated for the purposes of generalised comparison.

Examination of the daily error profiles indicates that in general, the forecast obtained via the different parameter estimation methods are essentially the same, with the exception of the linear programming technique case where significantly larger errors were recorded than for the other methods.

These errors are especially larger in winter and can be attributed to the fact that the load model used here is weather insensitive, and as such, this algorithm treated the measurement observations corresponding to relative extreme weather conditions as bad data points. Further since such extreme weather conditions are less frequent in summer, the errors obtained there are correspondingly smaller.

From the results of the days selected for bad data predictions, it is easily seen that with twenty-five percent gross error measurements in the databases, only the conventional least squares predictions are significantly affected. Here the reduced least squares and the two least absolute value predictions are relatively unchanged.

Further scrutiny of these results however, reveals that the forecasts made via the new LAV technique are the least affected by the bad data contamination.

### 6.1.5: Model C

This load model, like previous model has been used to predict the next day responses for two weeks each, of winter and summer load. Here, the new LAV algorithm is once again used in place of the original algorithm as this load model formulation is essentially similar to its predecessor.

The mean absolute error profiles obtained from winter and summer simulations, are presented in tables 6.15 and 6.16 respectively. These results are further illustrated graphically in figures 6.11 and 6.12 respectively.

This load model was also used to predict the twenty-four hour load for a selected winter and a selected summer day. Here also, forecasts were made with these data bases contaminated with twenty-five percent gross measurement errors.

The results of these simulations are listed in tables 6.17, 6.18, 6.19 and 6.20, and presented graphically in figures 6.13, 6.14, 6.15 and 6.16. The parameters estimated in each case, are presented in tables 6.21, 6.22, 6.23 and 6.24 respectively.

TABLE 6.15

MODEL C: HOURLY MEAN ABSOLUTE WINTER FORECAST ERRORS

| h | $\varepsilon_{LS}$ | $\varepsilon_{LS2}$ | $\varepsilon_{LAV}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|
| 1 | 4.122 | 3.816 | 4.714 | 4.439 |
| 2 | 4.182 | 3.571 | 4.503 | 4.054 |
| 3 | 4.362 | 3.711 | 4.139 | 4.218 |
| 4 | 4.608 | 3.669 | 5.239 | 4.604 |
| 5 | 4.926 | 3.796 | 4.738 | 4.893 |
| 6 | 5.526 | 4.093 | 5.582 | 5.295 |
| 7 | 4.677 | 3.565 | 4.766 | 3.973 |
| 8 | 3.405 | 3.426 | 4.515 | 3.320 |
| 9 | 2.704 | 2.753 | 3.397 | 2.848 |
| 10 | 2.556 | 2.519 | 3.761 | 2.768 |
| 11 | 2.699 | 2.763 | 4.370 | 3.118 |
| 12 | 2.592 | 2.426 | 3.981 | 2.868 |
| 13 | 2.780 | 2.870 | 3.249 | 2.932 |
| 14 | 2.513 | 2.589 | 3.373 | 2.766 |
| 15 | 2.328 | 2.867 | 3.031 | 2.559 |
| 16 | 2.389 | 2.899 | 2.723 | 2.747 |
| 17 | 3.252 | 4.680 | 3.733 | 3.594 |
| 18 | 3.957 | 5.350 | 4.433 | 4.083 |
| 19 | 2.460 | 2.983 | 2.293 | 2.305 |
| 20 | 2.316 | 2.364 | 2.254 | 2.134 |
| 21 | 2.305 | 2.683 | 2.245 | 2.282 |
| 22 | 2.451 | 2.879 | 2.507 | 2.405 |
| 23 | 2.821 | 3.290 | 4.084 | 2.956 |
| 24 | 3.399 | 3.884 | 4.681 | 3.786 |

TABLE 6.16

MODEL C: HOURLY MEAN ABSOLUTE SUMMER FORECAST ERRORS

| h | $\varepsilon_{LS}$ | $\varepsilon_{LS2}$ | $\varepsilon_{LAV}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|
| 1 | 1.612 | 1.804 | 1.776 | 1.784 |
| 2 | 2.390 | 2.293 | 2.425 | 2.321 |
| 3 | 1.961 | 1.772 | 2.221 | 1.729 |
| 4 | 1.937 | 2.104 | 2.177 | 2.017 |
| 5 | 2.114 | 2.301 | 2.576 | 2.152 |
| 6 | 2.843 | 3.542 | 3.834 | 3.142 |
| 7 | 2.438 | 2.931 | 3.002 | 2.692 |
| 8 | 2.093 | 2.297 | 2.706 | 2.368 |
| 9 | 1.646 | 1.770 | 1.778 | 1.713 |
| 10 | 1.781 | 1.795 | 1.642 | 1.853 |
| 11 | 1.839 | 1.540 | 1.946 | 1.669 |
| 12 | 2.130 | 1.779 | 1.706 | 1.928 |
| 13 | 2.038 | 1.891 | 1.898 | 1.839 |
| 14 | 1.890 | 1.829 | 1.834 | 1.841 |
| 15 | 2.053 | 1.923 | 1.938 | 1.758 |
| 16 | 2.370 | 2.293 | 2.616 | 2.134 |
| 17 | 2.645 | 2.428 | 2.454 | 2.711 |
| 18 | 2.474 | 2.201 | 2.682 | 2.186 |
| 19 | 2.442 | 2.147 | 2.424 | 2.198 |
| 20 | 2.474 | 2.613 | 3.078 | 2.825 |
| 21 | 2.645 | 2.716 | 2.853 | 2.781 |
| 22 | 2.760 | 2.749 | 3.113 | 2.735 |
| 23 | 3.490 | 3.466 | 3.556 | 3.532 |
| 24 | 2.531 | 2.567 | 2.691 | 2.585 |

FIGURE 6.11

MODEL C: MEAN ABSOLUTE WINTER FORECAST ERRORS

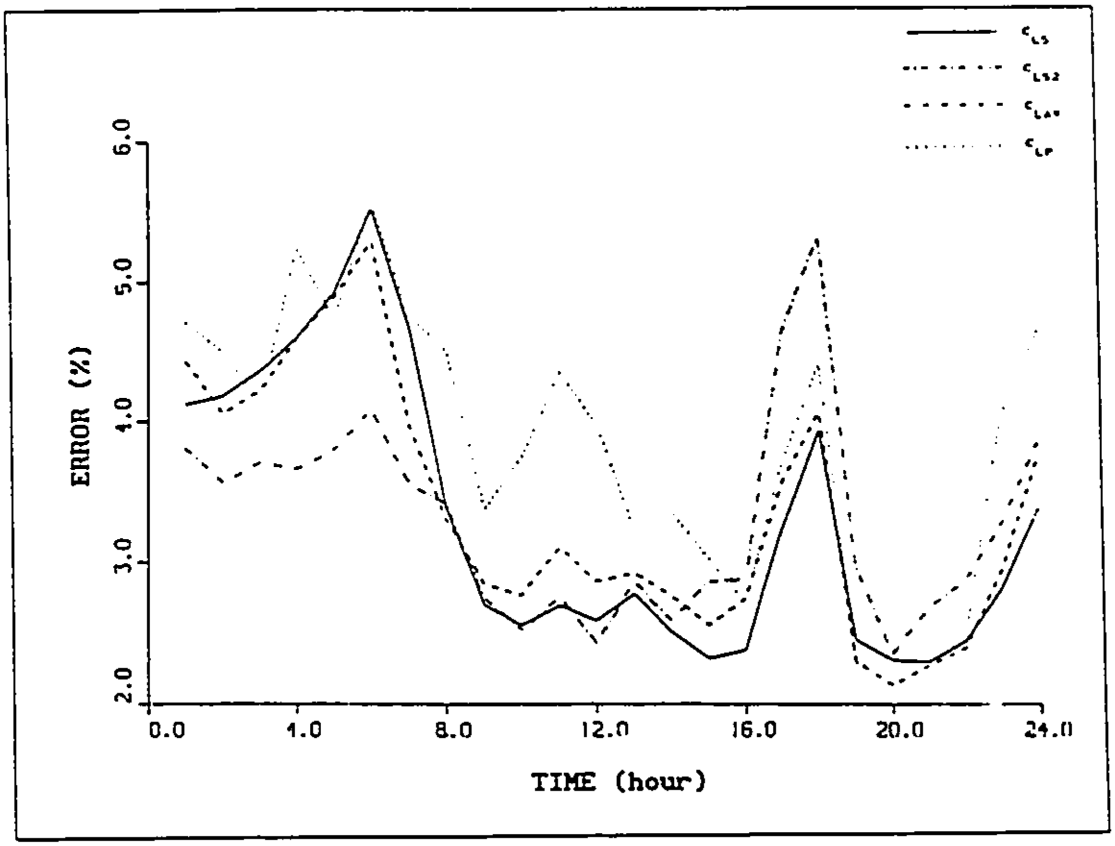


FIGURE 6.12

MODEL C: MEAN ABSOLUTE SUMMER FORECAST ERRORS

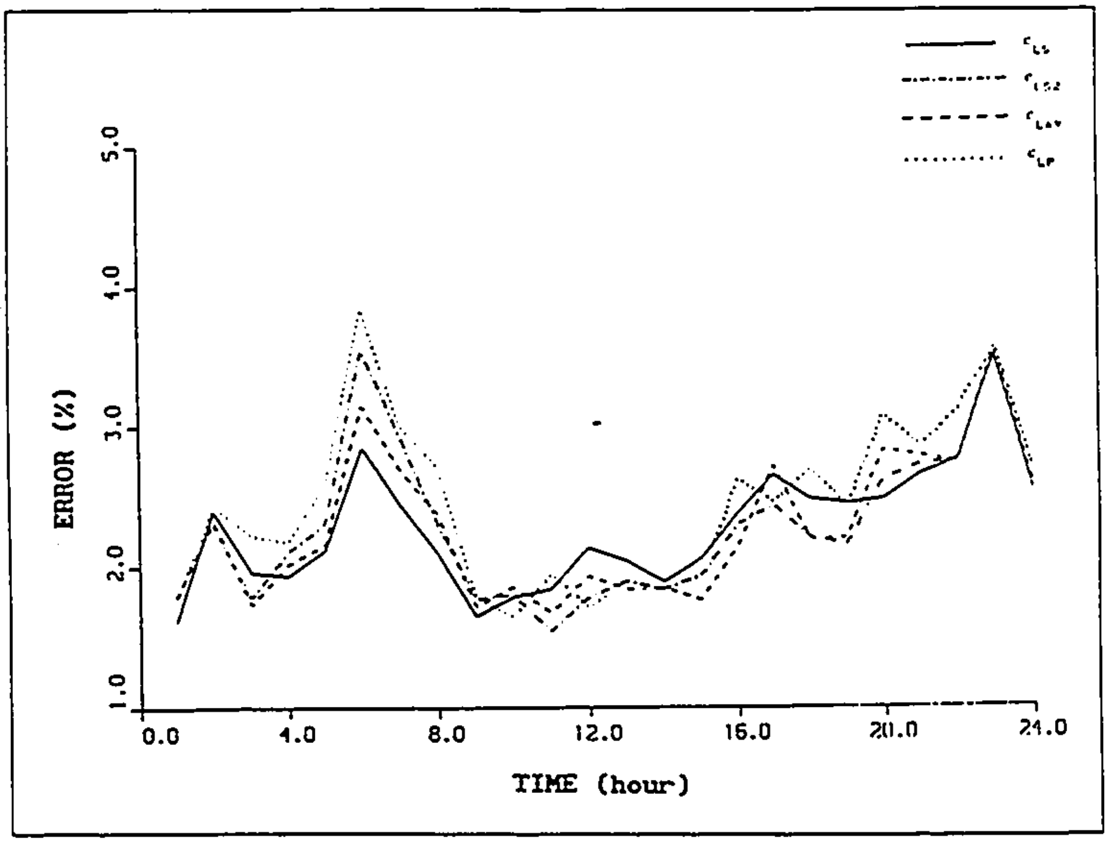


TABLE 6.17

MODEL C: WINTER DAY FORECAST

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 557 | 566 | -1.76 | 565 | -1.57 | 570 | -2.34 | 563 | -1.22 |
| 2 | 533 | 543 | -2.01 | 541 | -1.54 | 540 | -1.50 | 543 | -2.00 |
| 3 | 518 | 528 | -2.00 | 523 | -1.10 | 511 | 1.29 | 518 | -0.03 |
| 4 | 516 | 512 | 0.62 | 509 | 1.34 | 498 | 3.44 | 502 | 2.52 |
| 5 | 514 | 513 | 0.06 | 510 | 0.62 | 512 | 0.21 | 513 | 0.01 |
| 6 | 532 | 532 | -0.18 | 528 | 0.67 | 530 | 0.28 | 534 | -0.52 |
| 7 | 608 | 597 | 1.72 | 594 | 2.15 | 595 | 1.99 | 591 | 2.76 |
| 8 | 719 | 709 | 1.36 | 711 | 1.01 | 696 | 3.08 | 705 | 1.90 |
| 9 | 763 | 767 | -0.59 | 764 | -0.22 | 760 | 0.28 | 762 | 0.02 |
| 10 | 785 | 789 | -0.53 | 793 | -1.06 | 791 | -0.87 | 790 | -0.74 |
| 11 | 794 | 803 | -1.14 | 808 | -1.85 | 814 | -2.62 | 808 | -1.79 |
| 12 | 792 | 813 | -2.75 | 818 | -3.33 | 829 | -4.75 | 817 | -3.22 |
| 13 | 785 | 817 | -4.16 | 825 | -5.17 | 827 | -5.37 | 825 | -5.11 |
| 14 | 783 | 809 | -3.40 | 814 | -4.06 | 816 | -4.25 | 813 | -3.93 |
| 15 | 780 | 806 | -3.35 | 812 | -4.18 | 816 | -4.65 | 807 | -3.58 |
| 16 | 784 | 801 | -2.24 | 810 | -3.34 | 814 | -3.92 | 799 | -1.92 |
| 17 | 786 | 815 | -3.72 | 828 | -5.44 | 827 | -5.26 | 818 | -4.11 |
| 18 | 795 | 839 | -5.63 | 849 | -6.82 | 849 | -6.83 | 846 | -6.42 |
| 19 | 820 | 834 | -1.76 | 831 | -1.40 | 831 | -1.46 | 832 | -1.48 |
| 20 | 798 | 806 | -1.08 | 805 | -0.92 | 811 | -1.63 | 807 | -1.16 |
| 21 | 778 | 786 | -1.04 | 792 | -1.87 | 797 | -2.52 | 795 | -2.20 |
| 22 | 728 | 741 | -1.87 | 741 | -1.83 | 744 | -2.25 | 743 | -2.20 |
| 23 | 661 | 678 | -2.68 | 678 | -2.61 | 682 | -3.30 | 683 | -3.41 |
| 24 | 585 | 611 | -4.49 | 613 | -4.85 | 619 | -5.91 | 611 | -4.47 |

TABLE 6.18

MODEL C: WINTER DAY FORECAST WITH BAD DATA

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 557 | 498 | 10.52 | 612 | -9.91 | 608 | -9.32 | 570 | -2.36 |
| 2 | 533 | 469 | 11.95 | 481 | 9.62 | 545 | -2.44 | 527 | 1.10 |
| 3 | 518 | 393 | 23.98 | 381 | 26.40 | 492 | 4.97 | 498 | 3.76 |
| 4 | 516 | 405 | 21.44 | 483 | 6.26 | 513 | 0.50 | 492 | 4.56 |
| 5 | 514 | 352 | 31.48 | 489 | 4.81 | 482 | 6.03 | 505 | 1.68 |
| 6 | 532 | 567 | -6.66 | 476 | 10.46 | 575 | -8.25 | 536 | -0.89 |
| 7 | 608 | 475 | 21.76 | 671 | -10.49 | 576 | 5.22 | 609 | -0.25 |
| 8 | 719 | 545 | 24.13 | 656 | 8.65 | 721 | -0.31 | 683 | 4.98 |
| 9 | 763 | 580 | 23.86 | 744 | 2.45 | 758 | 0.58 | 757 | 0.70 |
| 10 | 785 | 526 | 32.91 | 731 | 6.75 | 739 | 5.80 | 779 | 0.68 |
| 11 | 794 | 629 | 20.71 | 747 | 5.86 | 807 | -1.71 | 804 | -1.37 |
| 12 | 792 | 664 | 16.13 | 867 | -9.52 | 827 | -4.52 | 813 | -2.73 |
| 13 | 785 | 675 | 13.98 | 855 | -8.98 | 838 | -6.85 | 807 | -2.87 |
| 14 | 783 | 640 | 18.14 | 812 | -3.74 | 833 | -6.51 | 820 | -4.78 |
| 15 | 780 | 689 | 11.63 | 814 | -4.45 | 834 | -6.97 | 806 | -3.41 |
| 16 | 784 | 634 | 19.03 | 839 | -7.03 | 774 | 1.20 | 793 | -1.24 |
| 17 | 786 | 625 | 20.47 | 777 | 1.08 | 805 | -2.47 | 796 | -1.36 |
| 18 | 795 | 622 | 21.73 | 775 | 2.42 | 836 | -5.18 | 817 | -2.80 |
| 19 | 820 | 649 | 20.77 | 809 | 1.29 | 840 | -2.52 | 824 | -0.56 |
| 20 | 798 | 613 | 23.16 | 826 | -3.62 | 790 | 0.88 | 809 | -1.43 |
| 21 | 778 | 626 | 19.41 | 777 | 0.04 | 778 | -0.05 | 785 | -0.98 |
| 22 | 728 | 557 | 23.46 | 734 | -0.85 | 723 | 0.64 | 742 | -2.01 |
| 23 | 661 | 548 | 16.96 | 653 | 1.10 | 670 | -1.42 | 670 | -1.41 |
| 24 | 585 | 446 | 23.74 | 613 | -4.95 | 593 | -1.46 | 610 | -4.32 |

TABLE 6.19

MODEL C: SUMMER DAY FORECAST

| h | z | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 536 | 522 | 2.49 | 523 | 2.34 | 529 | 1.28 | 524 | 2.18 |
| 2 | 510 | 485 | 4.72 | 487 | 4.45 | 484 | 4.98 | 485 | 4.79 |
| 3 | 485 | 469 | 3.14 | 471 | 2.87 | 463 | 4.49 | 470 | 2.94 |
| 4 | 476 | 462 | 2.91 | 463 | 2.56 | 459 | 3.47 | 463 | 2.60 |
| 5 | 476 | 465 | 2.30 | 471 | 0.92 | 474 | 0.40 | 466 | 1.93 |
| 6 | 491 | 470 | 4.13 | 468 | 4.50 | 463 | 5.58 | 472 | 3.81 |
| 7 | 536 | 528 | 1.32 | 520 | 2.92 | 506 | 5.50 | 525 | 1.95 |
| 8 | 644 | 634 | 1.42 | 630 | 2.09 | 627 | 2.51 | 629 | 2.25 |
| 9 | 732 | 726 | 0.69 | 715 | 2.20 | 716 | 2.12 | 718 | 1.78 |
| 10 | 791 | 790 | 0.02 | 778 | 1.52 | 774 | 2.05 | 782 | 1.05 |
| 11 | 821 | 824 | -0.46 | 821 | -0.07 | 829 | -0.99 | 826 | -0.71 |
| 12 | 838 | 834 | 0.39 | 830 | 0.89 | 836 | 0.12 | 835 | 0.25 |
| 13 | 848 | 838 | 1.16 | 835 | 1.50 | 833 | 1.72 | 837 | 1.26 |
| 14 | 862 | 840 | 2.44 | 836 | 2.99 | 828 | 3.85 | 837 | 2.80 |
| 15 | 855 | 843 | 1.39 | 844 | 1.22 | 852 | 0.28 | 848 | 0.78 |
| 16 | 856 | 836 | 2.22 | 839 | 1.91 | 837 | 2.13 | 839 | 1.89 |
| 17 | 843 | 827 | 1.83 | 821 | 2.50 | 831 | 1.37 | 832 | 1.28 |
| 18 | 819 | 805 | 1.60 | 804 | 1.81 | 804 | 1.80 | 798 | 2.44 |
| 19 | 788 | 753 | 4.33 | 755 | 4.17 | 752 | 4.53 | 759 | 3.61 |
| 20 | 755 | 736 | 2.42 | 729 | 3.42 | 721 | 4.40 | 747 | 1.03 |
| 21 | 725 | 724 | 0.13 | 725 | -0.01 | 723 | 0.14 | 731 | -0.91 |
| 22 | 705 | 689 | 2.24 | 694 | 1.56 | 688 | 2.31 | 691 | 1.88 |
| 23 | 670 | 667 | 0.33 | 665 | 0.71 | 664 | 0.88 | 669 | 0.01 |
| 24 | 610 | 628 | -3.01 | 628 | -3.11 | 634 | -3.94 | 630 | -3.30 |

TABLE 6.20

MODEL C: SUMMER DAY FORECAST WITH BAD DATA

| h | $z$ | $z_{LS}$ | $\varepsilon_{LS}$ | $z_{LS2}$ | $\varepsilon_{LS2}$ | $z_{LAV}$ | $\varepsilon_{LAV}$ | $z_{LP}$ | $\varepsilon_{LP}$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 536 | 416 | 22.33 | 479 | 10.62 | 502 | 6.30 | 529 | 1.29 |
| 2 | 510 | 366 | 28.15 | 396 | 22.22 | 459 | 9.99 | 477 | 6.44 |
| 3 | 485 | 344 | 28.90 | 401 | 17.15 | 451 | 6.99 | 467 | 3.68 |
| 4 | 476 | 348 | 26.80 | 365 | 23.13 | 431 | 9.27 | 461 | 3.04 |
| 5 | 476 | 348 | 26.80 | 332 | 30.24 | 422 | 11.27 | 460 | 3.18 |
| 6 | 491 | 357 | 27.28 | 365 | 25.63 | 422 | 13.94 | 465 | 5.17 |
| 7 | 536 | 416 | 22.21 | 496 | 7.33 | 504 | 5.80 | 522 | 2.57 |
| 8 | 644 | 526 | 18.18 | 655 | -1.79 | 636 | 1.15 | 631 | 1.88 |
| 9 | 732 | 638 | 12.72 | 755 | -3.25 | 724 | 1.06 | 726 | 0.79 |
| 10 | 791 | 679 | 14.07 | 813 | -2.89 | 795 | -0.60 | 788 | 0.26 |
| 11 | 821 | 646 | 21.29 | 819 | 0.17 | 823 | -0.26 | 822 | -0.13 |
| 12 | 838 | 665 | 20.57 | 847 | -1.11 | 827 | 1.20 | 841 | -0.37 |
| 13 | 848 | 687 | 18.87 | 831 | 1.98 | 829 | 2.23 | 849 | -0.19 |
| 14 | 862 | 698 | 19.01 | 863 | -0.16 | 854 | 0.91 | 838 | 2.72 |
| 15 | 855 | 664 | 22.31 | 823 | 3.68 | 824 | 3.56 | 835 | 2.31 |
| 16 | 856 | 684 | 20.02 | 899 | -5.04 | 854 | 0.22 | 857 | -0.20 |
| 17 | 843 | 628 | 25.46 | 749 | 11.10 | 790 | 6.24 | 837 | 0.66 |
| 18 | 819 | 558 | 31.86 | 795 | 2.92 | 789 | 3.63 | 804 | 1.71 |
| 19 | 788 | 546 | 30.71 | 702 | 10.82 | 717 | 8.91 | 744 | 5.50 |
| 20 | 755 | 578 | 23.33 | 691 | 8.36 | 709 | 5.97 | 728 | 3.49 |
| 21 | 725 | 557 | 23.04 | 697 | 3.80 | 708 | 2.22 | 729 | -0.61 |
| 22 | 705 | 518 | 26.43 | 698 | 0.95 | 682 | 3.23 | 692 | 1.73 |
| 23 | 670 | 496 | 25.86 | 631 | 5.77 | 641 | 4.32 | 653 | 2.52 |
| 24 | 610 | 473 | 22.29 | 619 | -1.58 | 617 | -1.23 | 626 | -2.78 |

FIGURE 6.13

MODEL C: WINTER DAY FORECAST

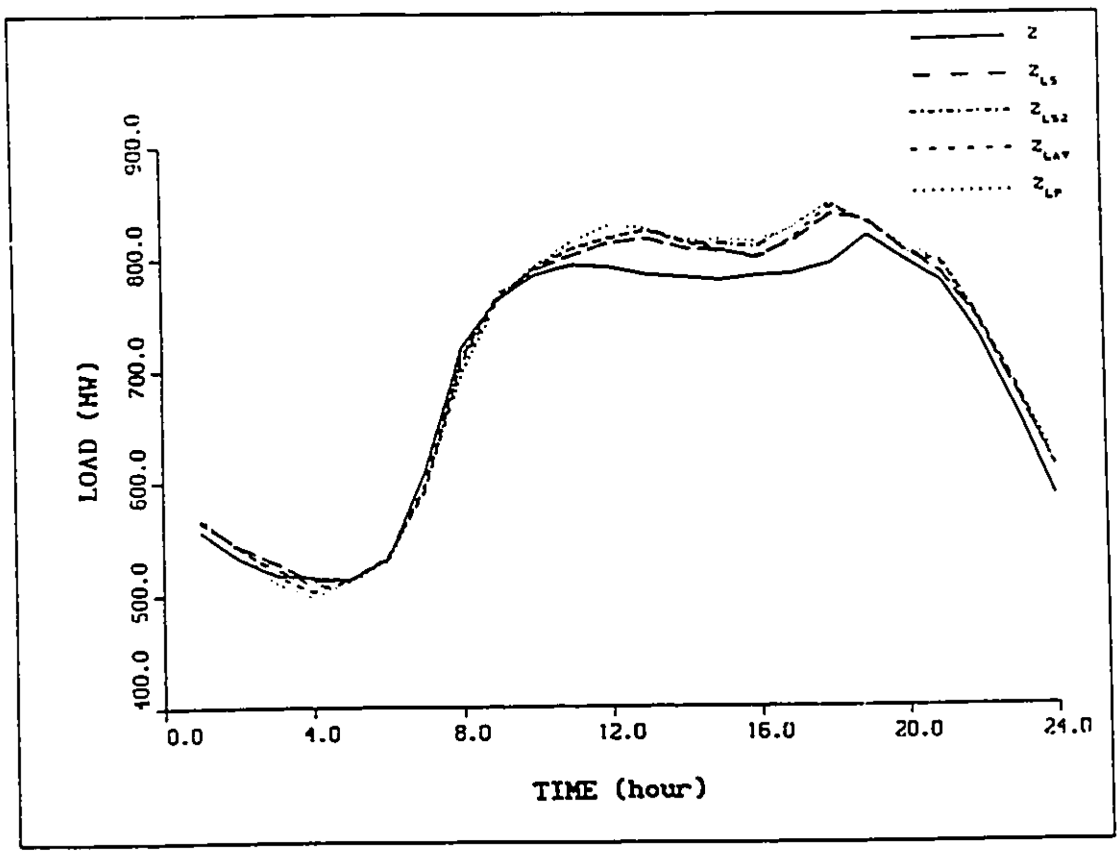


FIGURE 6.14

MODEL C: WINTER DAY FORECAST WITH BAD DATA

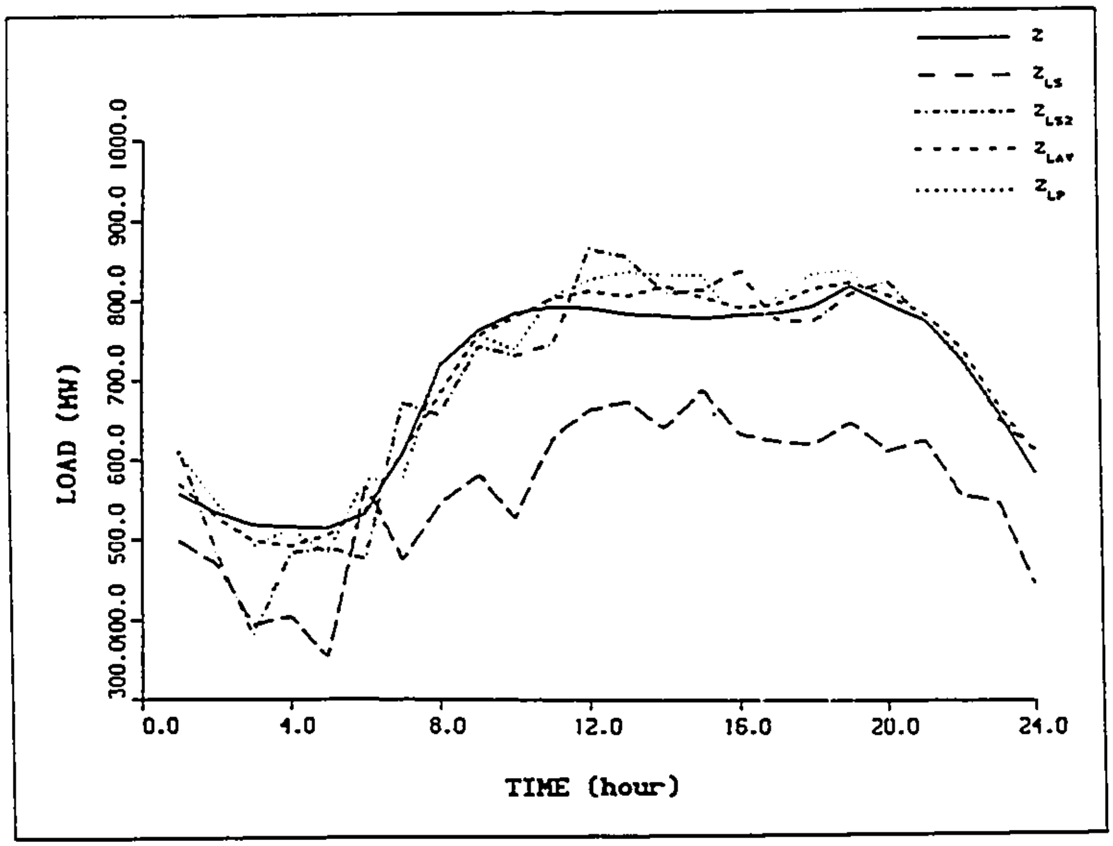


FIGURE 6.15

MODEL C: SUMMER DAY FORECAST

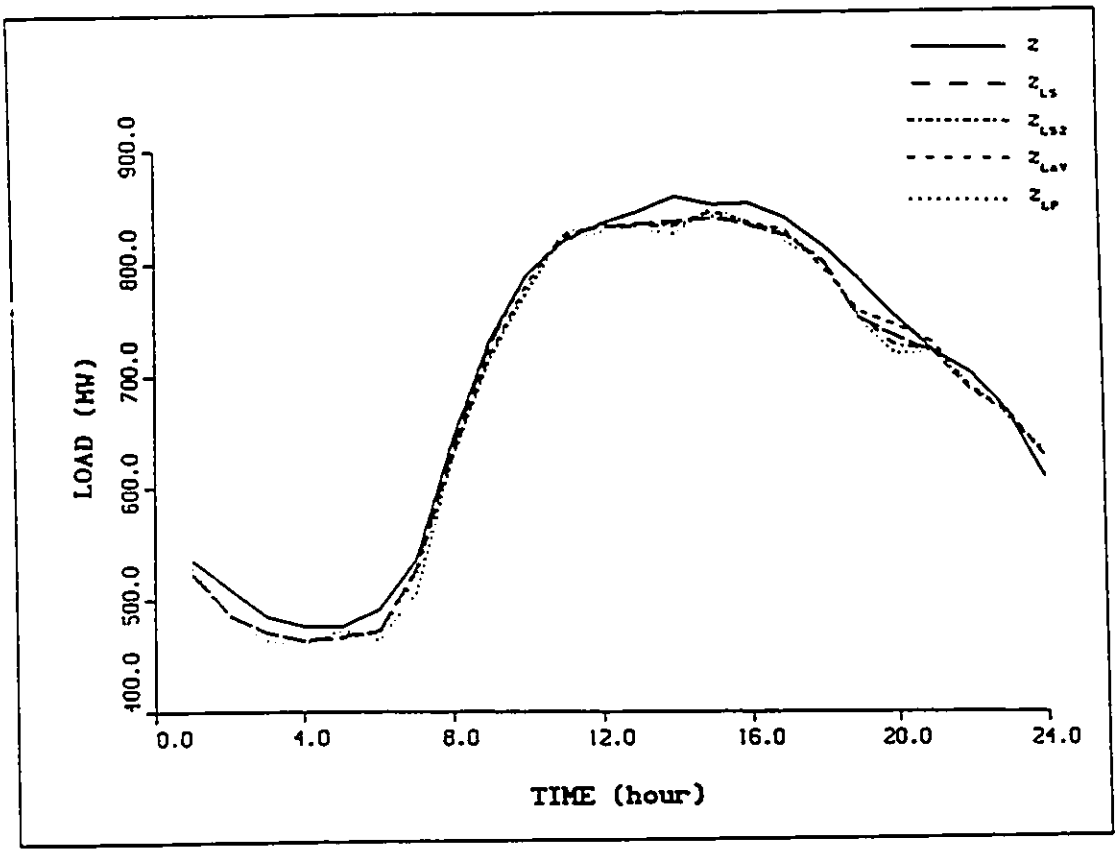


FIGURE 6.16

MODEL C: SUMMER DAY FORECAST WITH BAD DATA

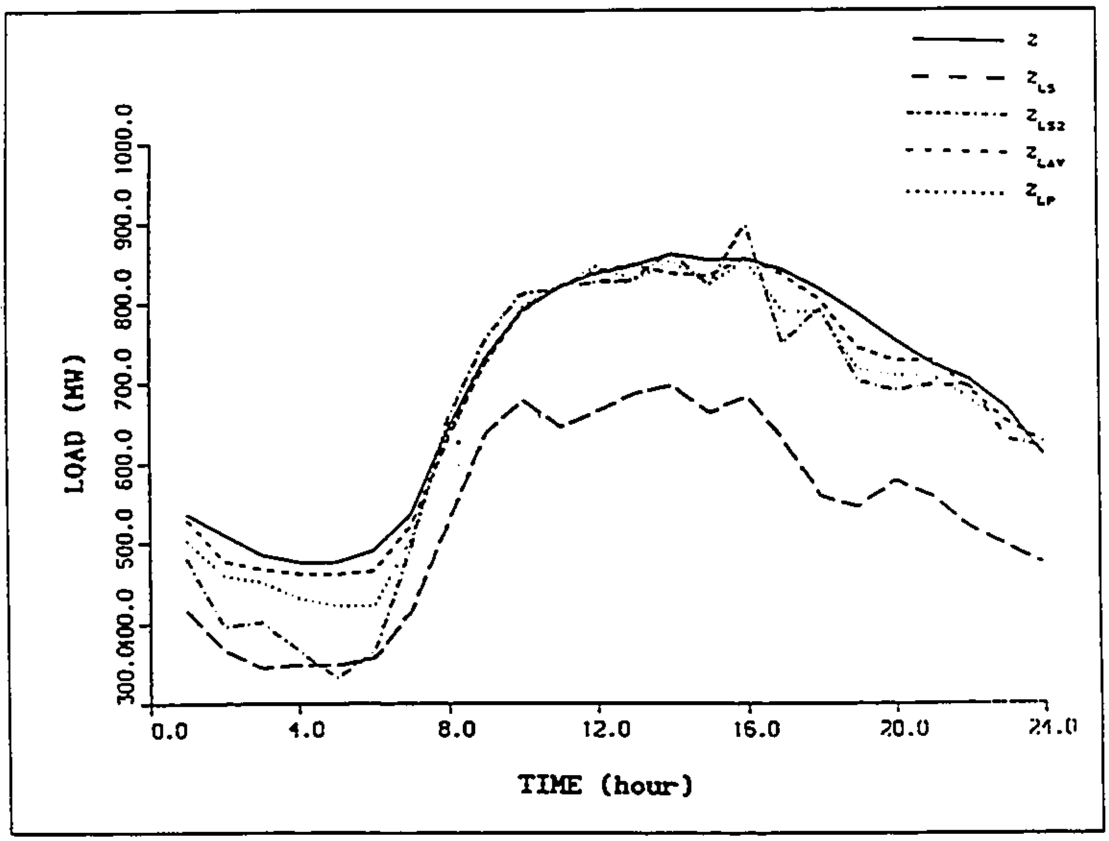


TABLE 6.21

MODEL C: WINTER DAY ESTIMATED PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 732.320 | 733.564 | 734.153 | 732.241 |
| $a_1$ | -128.496 | -133.281 | -136.332 | -132.281 |
| $b_1$ | -101.385 | -104.646 | -105.621 | -102.825 |
| $a_2$ | -60.213 | -60.210 | -60.590 | -62.094 |
| $b_2$ | 9.394 | 10.956 | 12.323 | 11.425 |
| $a_3$ | 15.972 | 16.154 | 15.326 | 14.665 |
| $b_3$ | -4.292 | -4.233 | -3.066 | -5.525 |
| $a_4$ | 0.325 | -1.830 | -2.422 | -0.681 |
| $b_4$ | -9.919 | -8.340 | -7.648 | -8.144 |
| $a_5$ | -10.875 | -9.573 | -8.112 | -10.139 |
| $b_5$ | 0.145 | -0.766 | -0.263 | -0.076 |
| $a_6$ | 1.272 | 3.017 | 3.143 | 3.399 |
| $b_6$ | 0.650 | 1.108 | 1.061 | -1.197 |
| $a_7$ | 3.201 | 4.009 | 4.195 | 3.410 |
| $b_7$ | -3.577 | -2.112 | -2.425 | -3.219 |
| $a_8$ | -1.549 | -1.746 | -3.630 | -3.200 |
| $b_8$ | 2.081 | 2.582 | 1.945 | 1.831 |
| $a_9$ | -1.162 | -2.567 | -3.066 | -3.698 |
| $b_9$ | 0.933 | 0.411 | 1.768 | 0.929 |
| $c_0$ | 0.339 | 0.296 | 1.488 | 1.228 |
| $c_1$ | 0.300 | 1.020 | -0.116 | -0.928 |
| $c_2$ | -0.253 | 1.037 | -3.187 | 0.570 |
| $c_3$ | -4.282 | -6.179 | -1.933 | -4.702 |

TABLE 6.22

MODEL C: WINTER DAY ESTIMATED BAD DATA PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 574.776 | 719.677 | 724.514 | 722.469 |
| $a_1$ | -105.450 | -152.350 | -132.543 | -130.313 |
| $b_1$ | -77.498 | -117.811 | -104.905 | -103.859 |
| $a_2$ | -47.717 | -80.376 | -58.454 | -65.139 |
| $b_2$ | 7.326 | 14.188 | 15.412 | 11.030 |
| $a_3$ | 13.739 | 9.006 | 15.080 | 12.312 |
| $b_3$ | -9.055 | -3.458 | -6.508 | -4.402 |
| $a_4$ | 3.771 | -2.266 | 11.134 | 0.136 |
| $b_4$ | -8.960 | -11.850 | -2.625 | -11.810 |
| $a_5$ | -5.997 | -16.387 | -11.893 | -8.441 |
| $b_5$ | -1.635 | 15.088 | 3.161 | 5.432 |
| $a_6$ | -0.708 | 8.400 | 4.992 | 0.596 |
| $b_6$ | -2.290 | 27.765 | 0.051 | -0.814 |
| $a_7$ | 1.272 | 16.978 | 12.704 | 5.830 |
| $b_7$ | -5.380 | -2.306 | -5.052 | -0.338 |
| $a_8$ | -1.047 | 7.535 | 0.191 | -3.575 |
| $b_8$ | 1.027 | -11.471 | 0.888 | -0.475 |
| $a_9$ | -8.976 | 1.524 | -1.508 | 2.707 |
| $b_9$ | 3.737 | 2 089 | -2.433 | 4.144 |
| $c_0$ | 43.599 | 21.826 | 15.570 | 5.011 |
| $c_1$ | -43.339 | 4.191 | -19.060 | -0.004 |
| $c_2$ | 2.447 | -46.005 | 9.898 | -10.533 |
| $c_3$ | -6.129 | 14.762 | -9.682 | 2.049 |

TABLE 6.23

MODEL C: SUMMER DAY ESTIMATED PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 704.018 | 702.204 | 702.319 | 704.277 |
| $a_1$ | -148.402 | -149.561 | -150.567 | -151.327 |
| $b_1$ | -124.472 | -122.839 | -121.227 | -123.194 |
| $a_2$ | -27.770 | -24.544 | -23.635 | -27.162 |
| $b_2$ | 36.053 | 36.351 | 37.985 | 36.016 |
| $a_3$ | 15.677 | 14.216 | 14.517 | 13.846 |
| $b_3$ | 4.792 | 4.729 | 4.181 | 3.267 |
| $a_4$ | -12.297 | -13.484 | -14.534 | -12.319 |
| $b_4$ | -6.388 | -7.367 | -4.917 | -7.062 |
| $a_5$ | -6.433 | -6.035 | -5.335 | -3.830 |
| $b_5$ | 2.126 | 2.087 | 2.510 | 2.241 |
| $a_6$ | 2.025 | 1.731 | 3.279 | 0.945 |
| $b_6$ | 2.816 | 2.976 | 4.391 | 4.711 |
| $a_7$ | 0.714 | 2.948 | 4.786 | 0.064 |
| $b_7$ | 2.326 | 0.374 | -1.139 | 1.988 |
| $a_8$ | -3.752 | -2.826 | -4.125 | -3.787 |
| $b_8$ | 1.204 | 1.215 | 3.169 | 0.972 |
| $a_9$ | -2.980 | -3.439 | -3.258 | -1.889 |
| $b_9$ | 0.648 | 3.097 | 3.619 | 1.055 |
| $c_0$ | 1.341 | 3.075 | 4.634 | 3.307 |
| $c_1$ | 1.158 | 0.424 | -5.720 | -1.632 |
| $c_2$ | 1.536 | 0.605 | 4.214 | 2.820 |
| $c_3$ | 3.938 | 4.135 | 5.240 | 3.430 |

TABLE 6.24

MODEL C: SUMMER DAY ESTIMATED BAD DATA PARAMETERS

| i | $\theta_{LS}$ | $\theta_{LS2}$ | $\theta_{LAV}$ | $\theta_{LP}$ |
|---|---|---|---|---|
| $a_0$ | 558.914 | 684.409 | 693.179 | 701.044 |
| $a_1$ | -101.276 | -166.700 | -152.083 | -149.051 |
| $b_1$ | -94.261 | -137.861 | -125.914 | -124.977 |
| $a_2$ | -28.675 | -50.697 | -30.073 | -27.064 |
| $b_2$ | 35.640 | 54.951 | 44.281 | 36.736 |
| $a_3$ | 10.443 | 13.495 | 17.734 | 13.774 |
| $b_3$ | -6.371 | 16.937 | 4.781 | 4.444 |
| $a_4$ | -1.083 | -6.766 | -11.489 | -14.031 |
| $b_4$ | -8.389 | -15.991 | -11.484 | -6.122 |
| $a_5$ | 0.571 | -12.965 | -7.554 | -5.720 |
| $b_5$ | 5.576 | -3.232 | -0.014 | -1.692 |
| $a_6$ | 15.563 | -4.204 | 1.870 | 6.820 |
| $b_6$ | 9.760 | 10.732 | 5.295 | 6.662 |
| $a_7$ | -8.620 | 5.319 | 3.984 | 0.129 |
| $b_7$ | -0.374 | 3.428 | 4.666 | 3.321 |
| $a_8$ | -1.767 | 4.554 | -3.261 | -0.118 |
| $b_8$ | 0.305 | 4.411 | -0.400 | 2.270 |
| $a_9$ | -0.935 | -3.665 | -4.228 | -2.333 |
| $b_9$ | -1.259 | 8.125 | 5.323 | 3.450 |
| $c_0$ | -22.440 | -18.931 | -9.402 | -0.434 |
| $c_1$ | 8.451 | 28.036 | 14.627 | 1.565 |
| $c_2$ | 16.757 | -19.603 | -5.203 | -0.241 |
| $c_3$ | 3.135 | 19.581 | 9.008 | 5.811 |

## 6.1.6: Discussion of Model C results

The observations made from the results of model B, are further confirmed by the results of simulations made with load model C. This is a weather sensitive hybrid of the two previous two load models, that was applied to the same data bases as model B.

The results of the two weeks each of winter and summer simulations, once again indicate that with uncontaminated data sources, the predictions made via the different estimation techniques are basically similar. In this load model however, the linear programming technique did not result in larger winter prediction errors, which only serves to validate the explanation offered previously.

With the contaminated data sources, we once again note the unchanged performance of the two least absolute value forecasts from the uncontaminated predictions. The reduced least squares method also performed within tolerable limits, but its predictions are somewhat more erratic than the least absolute value predictions.

## 6.2: Discussion of off-line simulation results

In subsections 6.1.1, 6.1.3, and 6.1.5 of this chapter, the results of off-line forecasts simulations for each of the three off-line load models developed in section 5.1 of chapter V are presented. These results were simulated using the static parameter estimation methods of chapter III, and will form the basis for generalised conclusions reflecting the effectiveness of the said algorithms as off-line forecasting tools.

It should be noted, that in short-term load forecasting, the use of static parameter estimation techniques is generally limited to multiple linear regression and general exponential smoothing models, even though they are also applicable to certain time series methods. This is a relatively rare situation however, and since the development of time series models is beyond the scope of this thesis, the off-line load models developed have been restricted to the multiple linear regression and the general exponential smoothing cases.

Each of the three off-line load models developed, have been used to predict the next day load corresponding to four separate parameter estimates. Of these, two are based on the least squares minimisation criterion, while the remaining two are results of the least absolute value minimisation criterion.

The first least squares estimate is obtained from a conventional least squares algorithm, while the second estimate is made using the same technique, but from a reduced data set where measurement outliers with least squares residuals larger than the standard deviation have been removed. As explained earlier in this chapter, both these estimates are outputs of the new LAV algorithm.

The least absolute value estimates on the other hand were produced by the new LAV algorithm and a linear programming algorithm respectively. It should be noted here, that a modified version of this algorithm was required for use with models B and C. The modified procedure is essentially the same as for the original algorithm, except that the ranking procedure has been modified to accomodate periodic data.

From the results obtained in general, it is evident that least absolute value estimates is to be preferred in cases where the data source is likely to be contaminated with errors. For uncontaminated data sources however, the performance of the conventional least squares method is the best choice, since these estimates require the minimum of computing effort and the use of least absolute value techniques will offer no gain in accuracy.

A comparison of the techniques used in these off-line simulations reveals that the new LAV technique will result in predictive accuracies matching that of the of the linear programming technique in all cases. As such, in cases where a least absolute estimation method is required, one will be well advised to use this technique as it is computationally much more efficient than its linear programming counterpart and will generally result in equivalent or better predictive accuracies.

## 6.3: On-line simulations

In this section, the on-line load model developed in section 5.2, is used to simulate the forecast responses of the three dynamic forecasting algorithms developed in Chapter IV. Here, the steady state gain factors are first determined off-line, and then used to predict the "one step ahead" load for each of three data sets.

The first two data sets, each consists of eight weeks (1344 hours) of winter and summer observations respectively, and forecasts made here will allow for a generalised comparison of the effectiveness of the three dynamic estimation algorithms, as on-line load forecasting methods.

The third data set consists of two days of hourly load data, of which two measurement observations have been replaced by gross error values. One gross error point is made excessively larger than its actual value, while the other is made significantly lower than its corresponding true value.

The forecasting algorithms are allowed to recursively track the load data, with and without the injected error points, and the resulting output recorded in each case.The results obtained for these and the forementioned simulations,are presented in the following subsections.

### 6.3.1: Steady state gain vectors

Twelve weeks of hourly load observations have been used to determine the steady gain vectors for both the Kalman and the weighted least absolute value filter. Of these, the first two weeks were used to make least squares approximations for the initial state vector and its corresponding error covariance matrix.

The general procedure used in determining the steady state gain values were as follows.

Given that the two weeks (336 hours) of hourly data can be written in vector form as

$$Z = Ax + B \tag{6.1}$$

where $Z$ = column $[z_1, z_2, \ldots\ldots\ldots z_{336}]$ of previous load observations

$x$ = state vector to be determined

$B$ = column $[b_1, b_2, \ldots\ldots\ldots b_{336}]$ of residuals corresponding to the estimate x, and

$A$ = $e^T H$ is matrix of output row vectors with

$e^T$ = (336x1) column $[1, \ldots\ldots\ldots\ldots 1]$ vector

We first find the least squares estimate for the state vector x as follows

$$x = (A^T A)^{-1} A^T Z \tag{6.2}$$

and then approximate its corresponding error covariance [63] to be

$$P = (H^T H)^{-1} R \tag{6.3}$$

where R is the initial value of measurement noise covariance.

Since the values of the noise covariances are unknown, we begin the estimation process with initial guesses of both these values. In this study, the following initial values for the noise covariances were used:

$R = 100,$

and $Q =$ (nxn) diagonal matrix, with

$Q_{ii} = 5$

With initial values of x, P, R and Q now available, the filters were allowed to recursively track the remaining eight weeks of load data until steady state was achieved.

In figures 6.17 and 6.18, the estimated values of the Kalman and the WLAV filter gains are plotted as a function of discrete time. It should be noted that in these figures, the value of the first element of the two gain vectors are actually used to represent the respective gain vectors.

The steady state gain values recorded for the Kalman and the WLAV filter as well as the calculated value of the general exponential smoothing vector are presented in table 6.25.

FIGURE 6.17

KALMAN FILTER GAIN RESPONSE

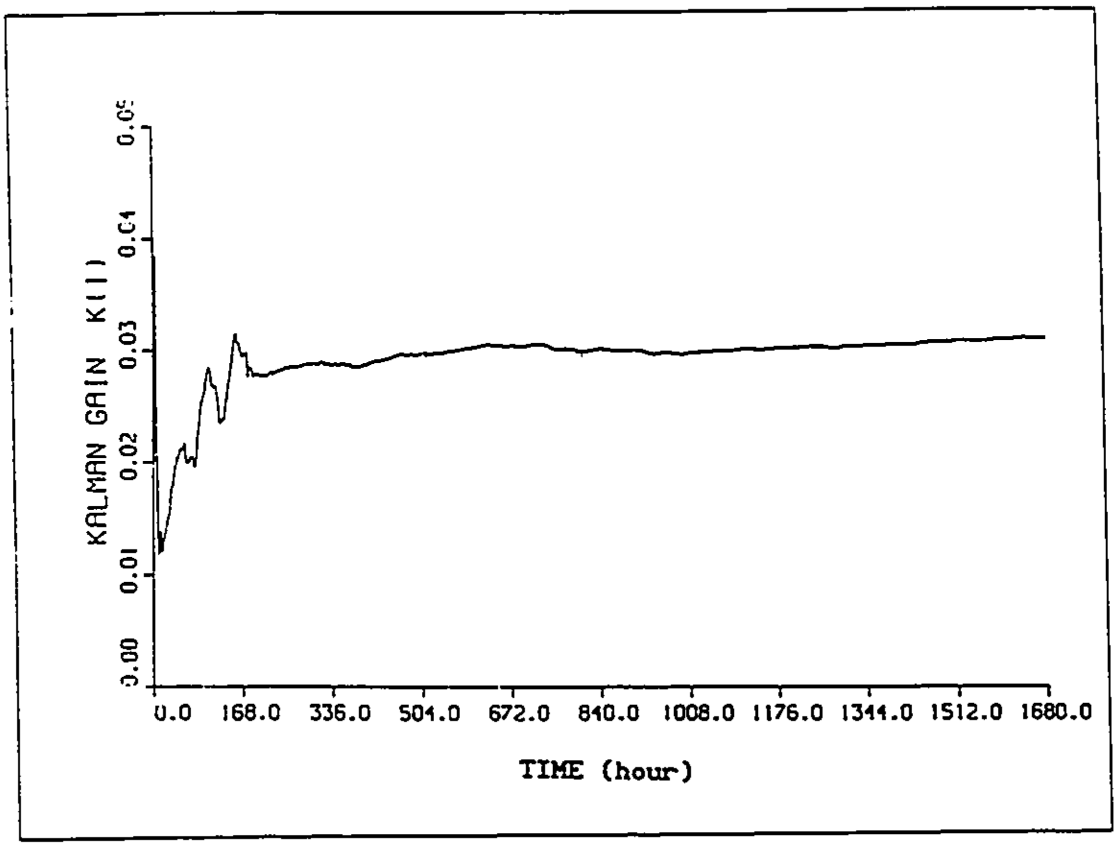


FIGURE 6.18

WLAV FILTER GAIN RESPONSE

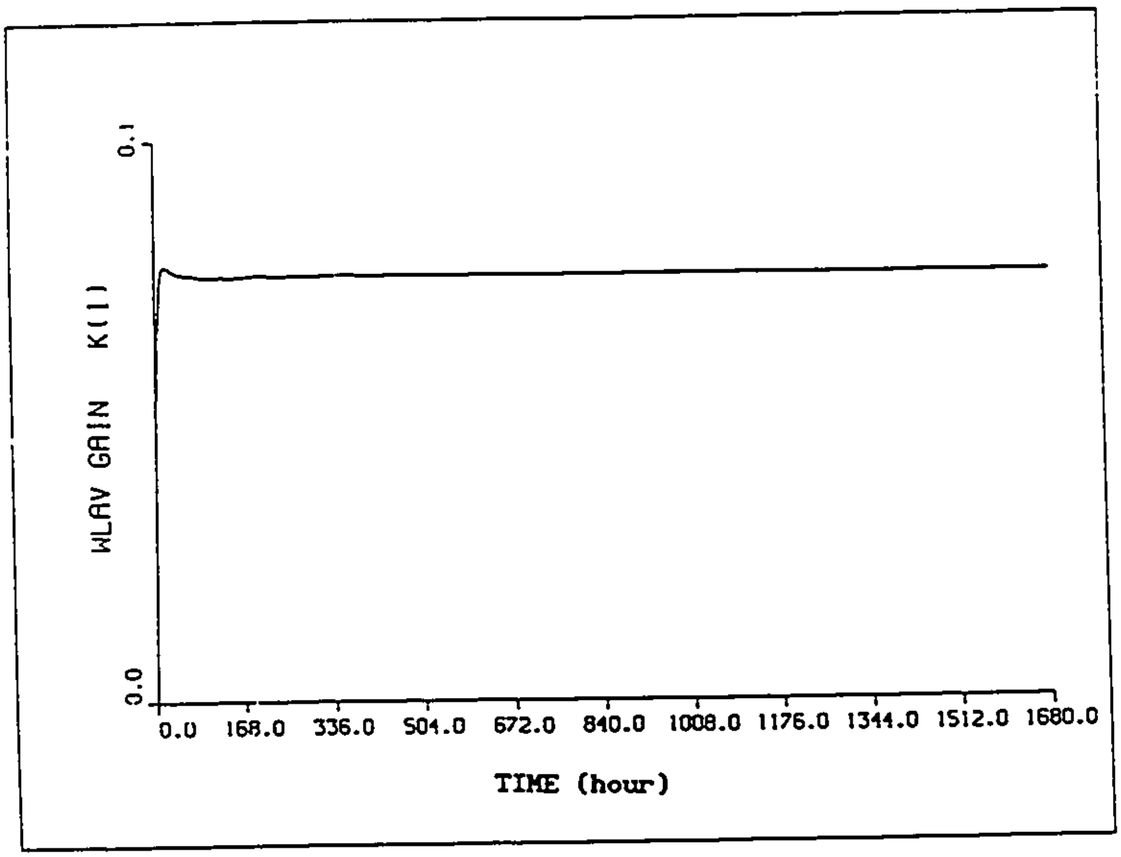


TABLE 6.25

SREADY STATE GAIN VECTORS

| Kalman gain $K_K$ | WLAV gain $K_W$ | GES gain $H$ |
|---|---|---|
| 0.03114 | 0.07598 | 0.00486 |
| 0.01187 | 0.00009 | 0.00055 |
| 0.04691 | 0.07596 | 0.00970 |
| 0.02145 | 0.00002 | 0.00148 |
| 0.04221 | 0.07595 | 0.00954 |
| -0.00270 | 0.00076 | -0.00052 |
| 0.04740 | 0.07597 | 0.00964 |
| 0.01020 | 0.00055 | 0.00043 |
| 0.04700 | 0.07603 | 0.00971 |
| 0.01785 | 0.00048 | 0.00099 |
| 0.04433 | 0.07605 | 0.00968 |
| 0.02400 | 0.00042 | 0.00151 |
| 0.04140 | 0.07607 | 0.00961 |
| 0.03100 | 0.00037 | 0.00214 |
| 0.03553 | 0.07608 | 0.00947 |
| 0.03928 | 0.00029 | 0.00321 |
| 0.02375 | 0.07611 | 0.00909 |
| 0.03497 | 0.00038 | 0.00255 |
| 0.02753 | 0.07611 | 0.00919 |
| 0.03199 | 0.00050 | 0.00216 |
| 0.02946 | 0.07614 | 0.00926 |
| 0.02287 | 0.00088 | 0.00129 |
| 0.03636 | 0.07617 | 0.00940 |
| 0.02275 | 0.00102 | 0.00116 |
| 0.03675 | 0.07623 | 0.00941 |

6.3.2: Simulation Results

Using the steady state values previously determined, the three forecasting algorithms were applied to eight weeks of both winter and summer data. Initial values for the state vectors in each case, was determined by "running" the algorithms on two previous weeks of corresponding data using least squares initial values.

The average error statistics obtained for both winter and summer simulations, as well as the overall average values are presented below in Table 6.26.

**Table 6.26**

**AVERAGE ERROR STATISTICS**

| | Winter | | Summer | | Overall | |
|---|---|---|---|---|---|---|
| | % MAE | % RMSE | % MAE | % RMSE | % MAE | % RMSE |
| Kalman | 2.615 | 3.296 | 2.263 | 2.971 | 2.439 | 3.134 |
| WLAV | 2.794 | 3.516 | 2.718 | 3.410 | 2.756 | 3.463 |
| GES | 4.301 | 5.358 | 3.667 | 4.544 | 3.984 | 4.951 |

MAE = mean absolute error RMSE = root mean square error

Average daily error profiles for winter and summer forecast simulations are shown in Figures 6.19 & 6.20 respectively, while in Figures 6.21, 6.22 and 6.23, the responses of each of the three algorithms on the contaminated data base, are presented separately.

FIGURE 6.19

MODEL D: WINTER DAILY MEAN ABSOLUTE ERROR PROFILES

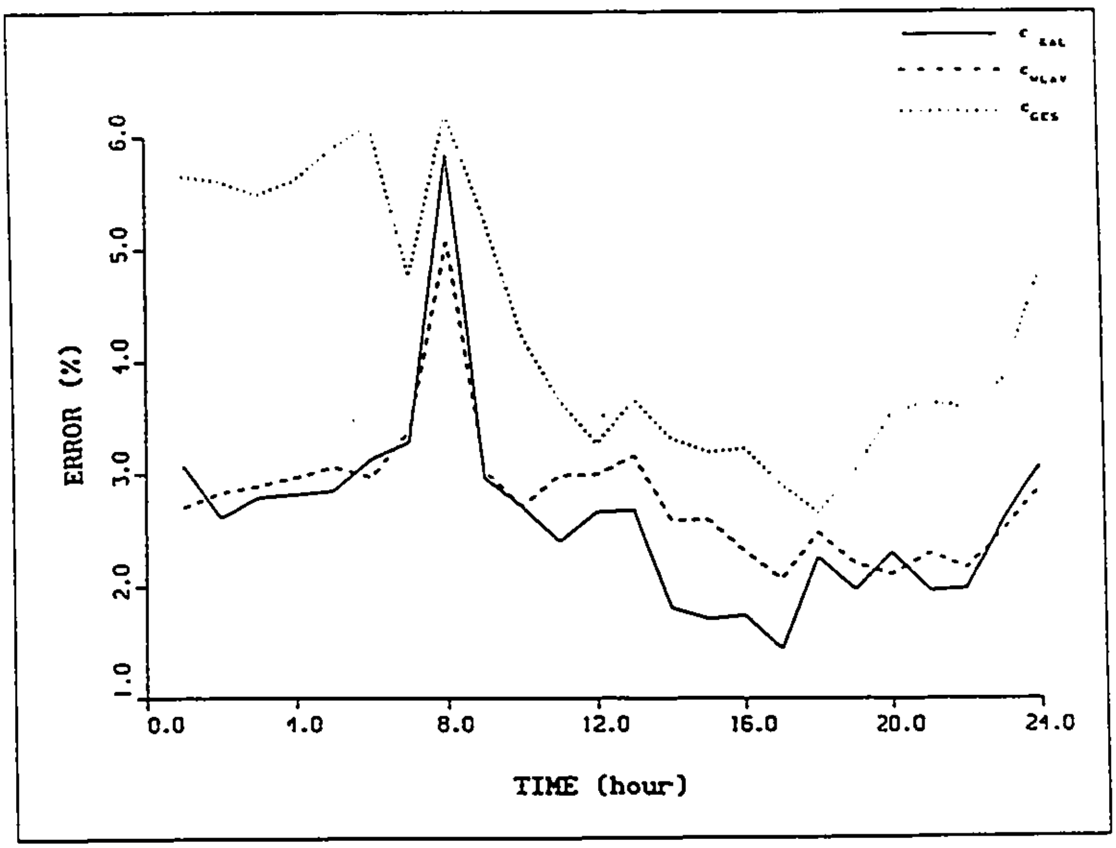


FIGURE 6.20

MODEL D: SUMMER DAILY MEAN ABSOLUTE ERROR PROFILES

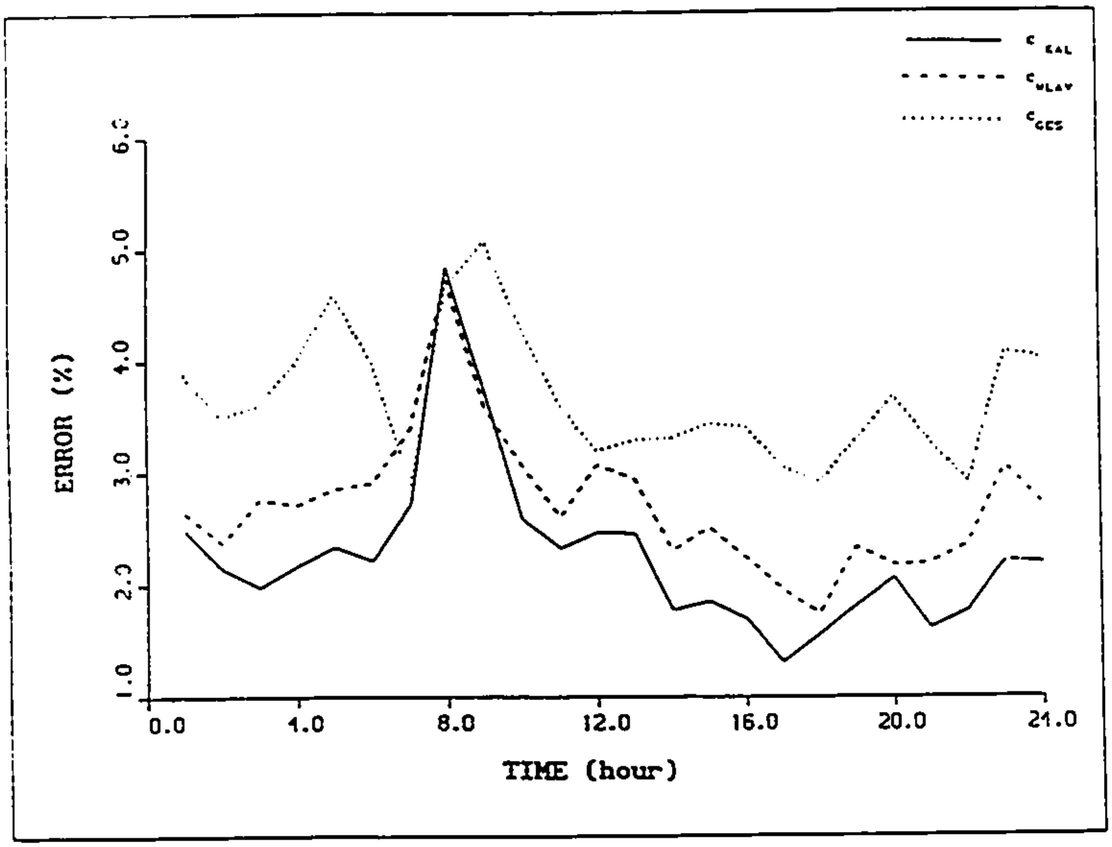


FIGURE 6.21

KALMAN FILTER BAD DATA RESPONSE

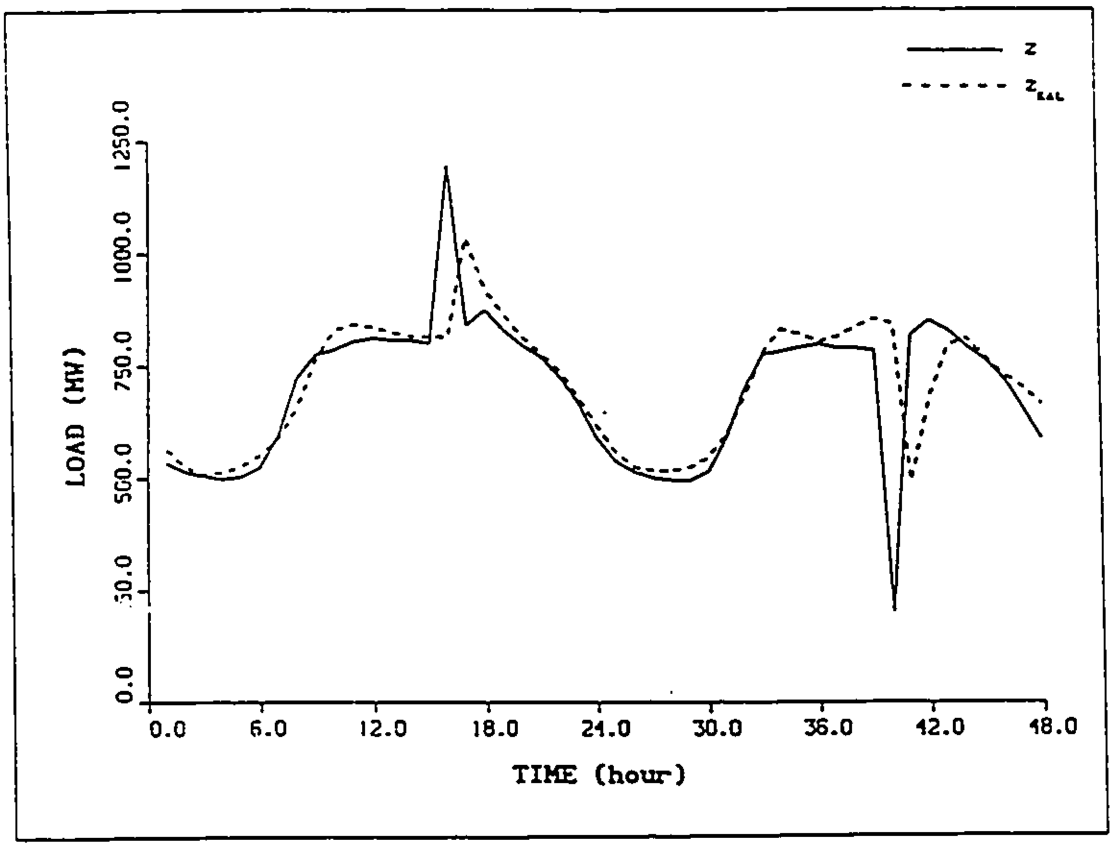


FIGURE 6.22

WLAV FILTER BAD DATA RESPONSE

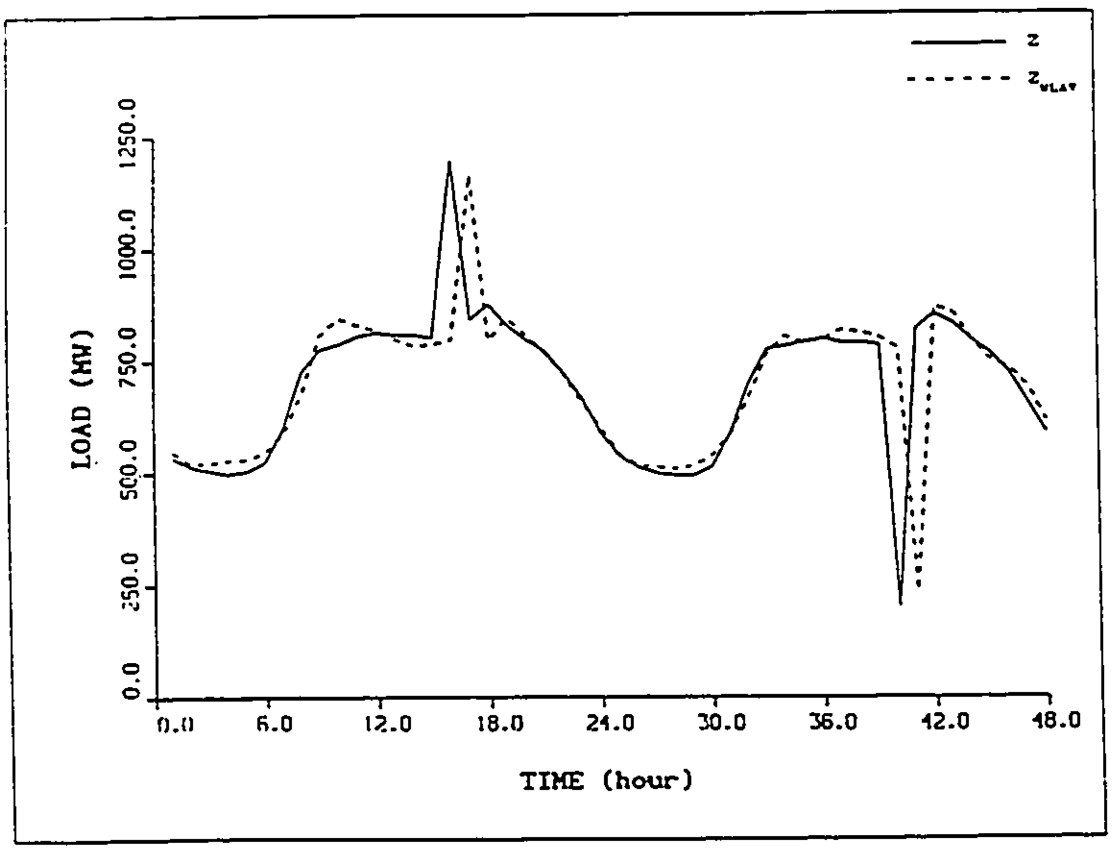
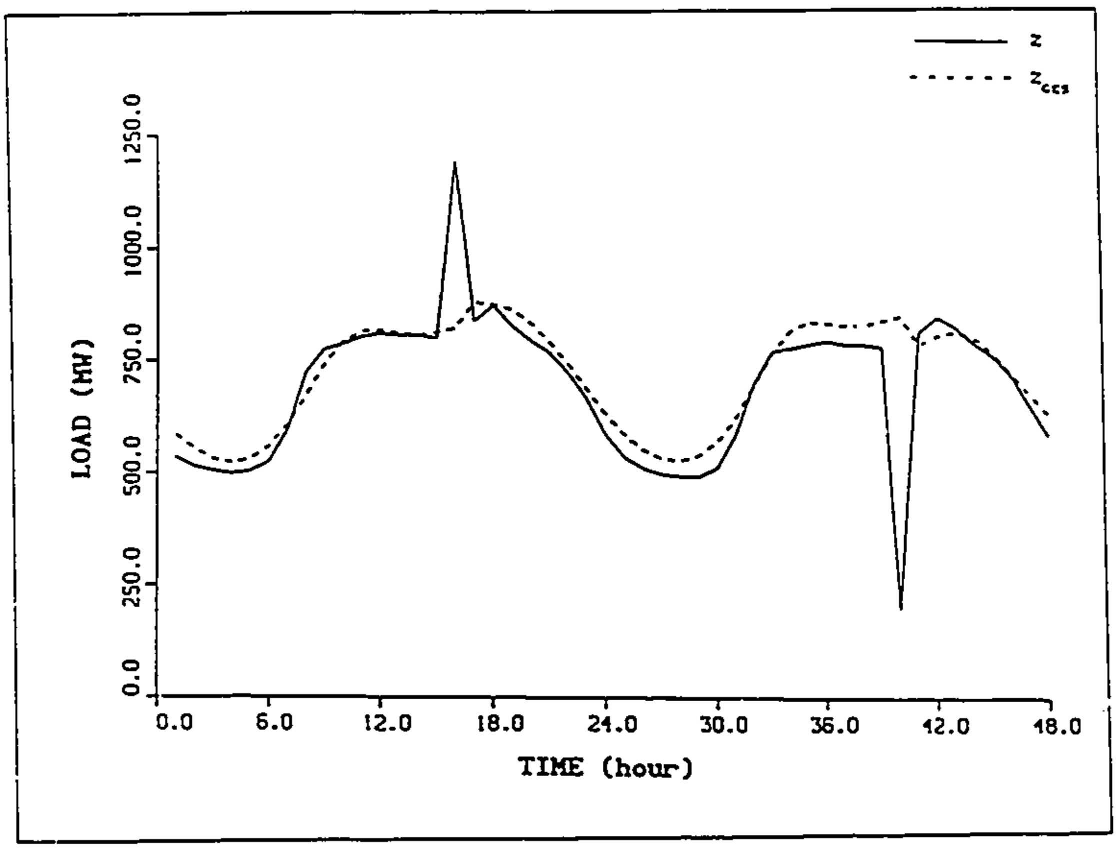


FIGURE 6.23

GES ALGORITHM BAD DATA RESPONSE



### 6.3.3: Discussion

A very important factor in any load forecasting algorithm, is the choice of load model, as it ultimately determines the degree of accuracy achieveable by the algorithm. For the purposes of this study however, a simple load model was identified and developed for on-line implementation. As such, the level of overall accuracy obtained may be significantly less than that to be expected from an actual forecasting algorithm, employed by an electric utility.

The fitting function chosen to model the load was Fourier harmonic series made up of a constant and twelve harmonics identified on the basis of a power spectral analysis (Fig.5.2) of all possible harmonics. Since the Fourier series, in this case, could consist of a maximum of eighty-four harmonics, thereby resulting in an impractical fitting function, a small percentage of modelling accuracy was compromised by restricting the harmonics to those with high power spectral factors.

An examination of the results presented in Table 6.26, shows a significant increase in accuracy from winter to summer for each of the three algorithms. This can be attributed to the fact that the load model used was weather insensitive, and could not respond to the sudden temperature changes that occured with greater frequency during winter.

Also, from Figures 6.19 and 6.20, it can be seen that relatively large errors occur at the eight hour in both, the winter and summer daily average error profiles. This error, much larger than for any other time of day, is due to the sudden large load demand that occurs as a result of combined businesses and industries beginning their daily activities.

While it is apparent that this error can be greatly reduced by adding a pre-determined fixed amount to the forecasted load at this time of day, it should be noted that this step would have interfered with the comparison of the responses of the algorithms in their intrinsic state and as such, this mechanism was not incorporated into the forecasting procedure.

Although, it is now obvious that the load model used in this study has limited potential, it should be recognised that it does provide a fair basis for the comparison of the simulated responses of the different on-line forecasting algorithms. Also the development of an accurate on-line load model in state space formulation is extremely complex, and beyond the scope of this thesis.

The efficiency of any forecasting algorithm is measured by the computing effort it requires to perform the desired forecast function. Since it is obvious that all three estimation techniques make use of the same forecasting algorithm then, the computing effort required in each case must be the same.

The difference in effort stems however from the initial analysis required to determine the respective steady state gain vectors. For the general exponential smoothing case, a minimum of effort was needed since this gain vector is time independent and depends only on the weighting factor and fitting function.

It should be noted here, that a weighting factor of 0.995 was used in this study as it exponentially discounts previous observations in approximately three weeks, and from experimentation was found give better accuracy than other tested values.

The Kalman and weighted least absolute value based algorithms on the other hand, both required more extensive initial analysis, as it was necessary to "run" each respective filter on previous load data before the steady state gain vectors could be determined. From figures 6.17 and 6.18 however, we can see that the Kalman filter required approximately eight weeks of initial data before it achieved steady state, whereas the weighted least absolute value filter only required two weeks of the same data.

It should be noted, that even though the WLAV filter requires a fraction of the initial data needed by the Kalman filter, it does require a larger initial effort per measurement, due to an increased matrix inverse requirement in its recursive structure. This increased effort per measurement is marginal however, and is amply compensated by the large saving in initial computing effort.

The extra initial effort required by the two filtering algorithms is compensated by the fact that both algorithms yielded significantly superior forecasts over those obtained from the less complex general exponential smoothing algorithm. This is to be expected however as this algorithm does not make allowances for the estimation and measurement noise process.

The results listed under Table 6.26, show that the Kalman based forecasting algorithm produces better winter and summer forecasts than both the WLAV and GES based algorithms even when measured on both the mean absolute error and root mean squared error criteria. On the other hand the results shown in Figures 6.15, 6.16 and 6.17 clearly indicate, that the GES based algorithm which produces the least accurate forecasts on regular data, has better on-line bad data rejection properties than either of the other two algorithms.

It is quite easy to spot the correlation between the bad data responses and the gain vector values listed in Table 6.25. The WLAV filtering algorithm with its large value of gain results in the largest next step errors while the Kalman based algorithm with a slightly smaller value of gain produces correspondingly smaller errors. The GES based algorithms however, with its very low value of gain results in significantly smaller errors almost indistinguishable from its regular performance.

It is now quite obvious that when applied to on-line functions the weighted least absolute value based algorithm cannot inherit the bad data rejection properties expected of an LAV estimator. This is not totally surprising, since the forecasting algorithms only make use of steady state gain vectors that does not take the error into account.

One should not summarily discuss the WLAVF as being inadequate however as it does possess a quick error response due to its high gain value and with this load model it is on the average only marginally less efficient than the Kalman filter. Also, other ongoing research [61] using the WLAV filter indicates that it does indeed possess some bad data rejection property and its performance in general is comparable to that of the Kalman filter.

All in all, one can surmise that since forecast error reduction is the ultimate aim of any forecasting algorithm, an ideal algorithm would be one that possess the properties of the Kalman filter when processing regular data, and the low gain value property of the general exponential smoothing algorithm on contaminated data points.

This idea is not far fetched and can be incorporated on the Kalman or WLAV filter, if a data point corresponding to a predetermined fixed or greater value of absolute error is deemed a "bad data" point and used to trigger a clause in the algorithm that corresponding alters the value of gain at the next step.

# CHAPTER 7

## CONCLUDING REMARKS

In this thesis, the use of the least squares and least absolute value methods of parameter estimation in short-term load forecasting was researched. First, the subject of short-term load forecasting was reviewed in chapter II, with especial emphasis on load modelling techniques and the role of parameter estimation methods therein.

From this review, it was found that parameter estimation in short-term load forecasting, could be described as either static or dynamic depending upon the modelling strategy adopted.

In chapter III, the static parameter estimation problem was introduced and its least squares and least absolute methods of solution reviewed. Here, the theory of conventional least squares, linear programming and a newly developed method of least absolute value estimation were presented.

The dynamic estimation problem was dealt with in chapter IV. Here the application of the Kalman and a recently developed weighted least absolute value filter to on-line forecasting were introduced, and a generalised dynamic forecasting algorithm developed for on-line simulations, later on in the thesis.

In chapter V, off and on-line load models were identified and developed to facilitate comparison of the estimation techniques reviewed in the two previous chapters, and in chapter VI, the results of these forecast simulations were presented.

From the results of off-line simulations using the static parameter estimation techniques reviewed in chapter III, it can be concluded that if the data source is free of errors, the use of either the least squares or the least absolute value minimisation criterion will result in the same degree of predictive accuracy.

On the other hand, it can also be concluded, that if the data source is contaminated with gross error points, then the use of the least absolute value criterion, will result in greater predictive accuracy.

From these results, it was also observed that both the linear programming and the new least absolute value algorithm resulted in similar predictions, with contaminated and uncontaminated data. As such it can be concluded that since the new LAV algorithm is computationally much more efficient and a precise least absolute value estimate is of lesser importance, then this new technique will offer the best choice in all cases.

From the results of on-line simulations presented in section 6.2, it can be seen, that the Kalman filtering algorithm resulted in greater predictive accuracy with uncontaminated data in both cases, even though the performance of the weighted least absolute filter was not far removed, and especially close in winter.

It should be noted here, that in the estimation of the steady state gain vectors for both filters, the equations of Sage and Husa were used to approximate the noise covariances at each step [equations 4.60 and 4.61]. These equations however, were derived for use on the Kalman filter and based on the least squares minimisation criterion, and as such, may have resulted in a WLAV filter steady state gain vector somewhat different from its actual value.

Since a change in steady state gain value could result in a significant change in the level of predictive accuracy, no firm comparison can be made between the Kalman and the new filter, other than the latter requires the processing of a smaller number of initial measurements in order to achieve steady state.

The results presented in section 6.2 also indicate, that for sequential on-line forecasting, neither filter will inherently be able to identify and reject bad data points. This stands to reason however, since these bad data points were not reflected in values of noise covariances estimated during steady state analysis.

## 7.1: Recommendations

A desirable quality in any load forecasting algorithm, is the ability of the algorithm to forecast within expected limits, even when the forecasting data base contains a small percentage of measurement errors. In short-term forecasting, this quality of a predictor is usually ensured by prefiltering the data base of suspected error points.

In prefiltering, errors in the data base are identified and replaced by corresponding "good" measurements further down the data source. These so called bad data points, are usually measurements recorded for unusual daily events such as holidays, major television events etc., and in rare cases can be introduced through operator mistakes at the man-machine interface.

Since this prefiltering procedure is an integral part of the new LAV algorithm, it is recommended that for off-line load forecasting, the new LAV algorithm be used in place of the conventional least squares estimation methods. However, if the possibility exists that the H matrix selected may sometimes be close to ill conditioned, as was the case with the multiple linear regression model of section 5.1.1; one should refrain from selecting just the best number of measurements corresponding to number of unknowns and use the reduced least squares estimate instead.

An interesting observation made during off-line simulations, was that the reduced least squares estimates resulting from the rejection of outliers with residuals larger than the standard deviation, were in most cases, very close to the estimates produced by the new LAV and the linear programming algorithms.

This seems to suggest, that there is a possible relationship between the measurements rejected and the optimal least absolute value estimate, that could ultimately result in a much more robust and efficient LAV algorithm. As such a possible topic for future research would be to investigate the relationship between the rejection of outlying measurements and the resulting estimates with the aim of further simplying the new LAV procedure.

In on-line forecasting, errors in the data base will be reflected in the values of estimated noise covariances and the resulting steady gain vector, and as can be seen from figures 6.13 and 6.14, these have less influence on the WLAV filter gain than they do on the Kalman filter gain. However, there is a need for more indepth comparative studies of these two filters using a wider variety of on-line load models, before any firm conclusions can be made.

An interesting prospect for future research would be to formulate the off-line load model of subsection 5.1.3 (Model C), as a state space model, and apply it via these filters to on-line forecasting. The rationale for this, stems from the degree of accuracy that was gained by model D, when allowances were made for the noise processes in the cases of the Kalman and the WLAV filter.

This can be seem more clearly by considering the improvement in predictive accuracy from models B to C, and then noting that model D which is very similar to model B, showed a significant improvement in accuracy after it was used in conjunction with an optimal filter.

Another possibility for futute studies, would be to investigate the possibility of including a time propagating standard deviation equation in either the Kalman or WLAV filter, so that it can be used as error detector when forecasting sequentially as in figures 6.15, 6.16 and 6.17. Here, values of prediction error larger than the standard deviation would be considered to be the result of a bad data point, and would be limited in its influence on the estimate for the next step.

## REFERENCES


[1] IEEE Committee Report, "Load forecast bibliography, Phase I," IEEE transaction on power apparatus and systems, vol. PAS-99, no.1, 1980, pp. 53-58

[2] IEEE Committee Report, "Load forecast bibliography, Phase II," IEEE transaction on power apparatus and systems, vol. PAS-100, no.7, 1981, pp. 3217-3220

[3] A.H.El Abiad, "Load forecasting and fast growth problems," in Power system analysis and planning, A.H.El Abiad editor, Hemisphere publishing corporation, 1983, 195-200

[4] M.S.Sachdev and R.A.Billington, "A review of the techniques for forecasting power and energy demands," CEA Transactions, Engineering and operations division, vol.75, no.12, Dec.1987, 76-SP-141

[5] G.S.Christensen, S.A.Soliman and A.Rouhi, "A new technique for curve fitting based on minimum absolute deviations," Computational statistics and data analysis, vol.6, no.4, 1988, pp.341-351

[6] D.W.Bunn and E.D.Farmer, "Economic and operational context of electric load prediction," Comparative Models for Electrical load forecasting. D.W.Bunn and E.D.Farmer editors, New York, J.Wiley and sons, 1985, pp. 3-9

[7] H.A.Dryar, "The effect of weather on system load," AIEE Transactions, vol.63, 1944, pp. 1006-1013

[8] M.Davies, "The relationship between weather and electricity demand," IEE monograph 3145, 1958, pp. 27-37

[9] E.D.Farmer, "Online load prediction," Optimal Power System Operation Conference. UMIST, 1969

[10] E.D.Farmer and M.J.Potton, "Development of online load prediction techniques with results from the south-west region of the CEGB," Proceedings of the IEE, 115, no.10, 1968, pp. 1549-1558

[11] E.D.Farmer, "A method of prediction of nonstationary processes and its application to the problem of load estimation," Proceedings of the 2nd IFAC congress, Basle, 1963

[12] E.D.Farmer and M.J.Potton, "the prediction of load on a power system," Proceedings of the 3rd IFAC congress, London, 1966

[13] P.D.Mattewman and H.Nicholson, "Techniques for load prediction in the electricity supply industry," Proceedings of the IEE, 115, (10), 1968, pp. 1451-1457

[14] S.L.Corpening, N.D.Reppen and R.J.Ringlee, "Experience with weather sensitive load models for short and long term forecasting," IEEE winter power meeting, New York, 1973, pp. 1966-1972

[15] C.E.Ashbury, "Weather load model for electric demand and energy forecasting," IEEE. Transactions on power apparatus and systems, PAS-94, no.4, 1975, pp. 1111-1116

[16] G.T.Heinman, D.A.Nordman and E.C.Plant, The relationship between summer weather and summer loads," IEEE Transactions on power apparatus and systems., PAS-85, no.11, 1966, pp. 1144-1154

[17] A.M.Schneider, T.Takenawa and D.A.Schiffman, "24-hour Electric Utility Load Forecasting," Comparative models for electric load forecasting, D.W.Bunn and E.D.Farmer editors, New York, J.Wiley and sons, 1985, pp. 87-108

[18] F.D.Galiana, E.Handschin and A.R.Fiechter, "Identification of stochastic electric load models from physical data," IEEE Transactions on automatic control, vol. AC-19, no.6, Dec. 1976, pp. 887-893

[19] I.Moghram and S.Rahman, "Analysis and evaluation of five short-term load forecasting techniques," IEEE Transactions on power systems, vol.4, no.4 , Oct. 1989, pp. 1484-1490

[20] S.Rahman, "Formulation and analysis of a rule based short-term load forecasting algorithm," Proceedings of the IEEE, vol.78, no.5, May 1990, pp. 805-815

[21] D.W.Bunn and E.D.Farmer, "Review of short-term forecasting methods in the electric power industry,"in Comparative models for electric load forecasting, D.W.Bunn and E.D.Farmer editors, John Wiley and sons ltd., 1985, pp.13-30

[22] M.A.Abu-El-Magd and N.K.Sinha, "Short-term load demand modelling and forecasting: A Review," IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-12, no.3, May/June 1982, pp. 370-382

[23] G.Gross and F.D.Galiana, "Short-term load forecasting," Proceedings of the IEEE, vol.75, no.12, Dec. 1987 1558-1573

[24] B.Abraham and J.Ledolter, "Statistical methods for forecasting," J.Wiley and sons Inc., 1983

[25] R.G.Brown, "Smoothing, forecasting and prediction of discrete time series," Englewood Cliffs, NJ, Prentice-Hall 1965

[26] W.R.Christiaanse, "Short-term load forecasting using general exponential smoothing," IEEE Transaction on power apparatus and systems, vol. PAS-90, no.2, 1971, pp. 900-910

[27] N.Weiner, "Extrapolation, interpolation and smoothing of stationary time series," New York, Wiley 1949

[28] P.Whittle, "Prediction and regulation by linear least squares methods," London: English Univ., 1963

[29] G.E.P.Box and G.M.Jenkins, "Time seies analysis, forecasting and control," Holden-Day, 1970

[30] E.H.Barakat, M.A.Qayyum, M.N.Hamed and S.A.Al Rashed, "Short-term peak demand forecasting in fast developing utility with inherent dynamic characteristics," IEEE Transactions on power systems, vol.5, no.3, Aug.1990, pp. 813-824

[31] J.Toyoda, M.Chen and Y.Inonye, "An application of state estimation to short-term load forecasting, Part I and Part II," IEEE Transactions on power apparatus and systems, vol. PAS-89, no.7, 1970, pp. 1678-1688

[32] K.L.S.Sharma and A.K.Mahalanabis, "Recursive short-term load forecasting algorithm," Proceedings of the IEEE, vol.121, no.1, 1974, pp. 59-62

[33] K.Jabbour, J.F.V.Rivers, D.Lansberger and W.Meyer, "ALFA: Automated load forecasting assistant," IEEE Transactions on power systems, vol.3, no.3, 1988, pp. 1908-1914

[34] S.Rahman and R.Bhatnagar, "A expert system based algorithm for short term load forecasting," IEEE Transactions on power systems, vol.3, no.2, 1988, pp. 392-399

[35] S.Rahman and M.Baba, "Software design and evaluation of a micro-computer based automated load forecasting system," IEEE Transactions on power systems, vol.4. no.2, 1989, pp.782-787

[36] N.Liu, "Frequency estimation in power systems," M.Sc. Thesis, University of Alberta, 1989

[37] A.H.Rouhi, "A new LAV power system state estimator," M.Sc. Thesis, University of Alberta, 1988

[38] I.Barrodale and F.D.K.Roberts, "Algorithms for restricted least absolute value estimation," Communications in statistics-simulation and computation, B6(3). 1977, pp. 353-363

[39] J.E.Gentle, "Least absolute values estimation: An introduction," Communications in statistics-simulation and communications, B6(4), 1977, pp. 313-328

[40] J.M.Mendel, "Lessons in digital estimation theory," Prentice Hall, Englewood Cliffs, New Jersey, 1987

[41] H.Jeffereys, "Thoery of probability," Oxford university press, London, 1939

[42] S.A.Soliman and G.S.Christensen, "Parameter estimation in linear static systems based on LAV estimation," Journal of optimisation theory and applications, vol.61, no.2, May 1989, pp.276-292

[43] Eisenhart, "The meaning of "least" in least squares," Journal of the Washington Academy of Sciences, vol.54, 1964, pp.24-33

[44] S.C.Narula, "An algorithm for the minimum sum of weighted absolute errors regression," Communications in statistics - simulation and computation, B6(3), 1977 pp.341-352

[45] V.Chvtal, "Linear programming," W.H.Freeman and company, New York, 1983

[46] Y.Dodge, "An introduction to statistical data analysis $L_1$ - norm based," Statistical data analysis based on the $L_1$ - norm and related methods, Y.Dodge editor, Elsevier science publishing company, New York, 1987, pp. 1-21

[47] STAT/LIBRARY ,"Alternatives to least squares," IMSL. incorporation, 1990, pp.251-252

[48] E.J.Schlossmacher, "An iterative technioque for absolute deviations curve fitting," Journal of the American Statistical Association, vol.68, no.344, December 1973, pp. 857-859

[49] G.F.Mc.Cormick and V.A.Sposito, "Using the $L_2$ estimator in $L_1$-estimation," SIAM journal of numerical analysis, vol.13, no.3, June 1976, pp. 337-343

[50] G.F.Mc.Cormick, M.L.Hand and V.A.Sposito, "Using an approximate $L_1$ - estimator," Communications in statistics - simulation and computation, B6(3), 1977, pp. 263-268

[51] S.A.Soliman, S.E.A. Emam and G.S.Christensen, "A new algorithm for optimal parameter estimation of a synchronous machine from frequency test based on LAV approximations," Canadian journal of electrical and computer engineering, vol.14, no.3, 1989, pp. 98-102

[52] S.A.Soliman, G.S.Christensen and S.S.Fouda, "On the application of the least absolute value parameter estimation algorithm to distance relaying," Electric power systems research, 1990 pp. 23-35

[53] S.A.Soliman, G.S.Christensen, D.H.Kelly and N.Liu, "An algorithm for frequency relaying based on least absolute value approximations," Electric power systems research, 1990, pp.73-84

[54] R.E.Kalman, "A new approach to linear filtering and prediction problems," Journal of basic engineering, Transactions of the ASME, March 1960. pp. 35-45

[55] S.C.A.Thomopoulos, "Estimation with information theoretic criteria: an interpretation," Proceedings of the American control conference, 1987, pp.1963-1987

[56] B.A.Kosko, "Foundation of fuzzy estimation theory," Ph.D. thesis, University of California, Irvine, 1987

[57] A.E.Bryson Jr. and Y.C.Ho, "Applied optimal control," Hemispherepublishing corporation, New York, 1985

[58] F.L.Lewis, "Optimal estimation with an introduction to stochastic control theory," Wiley, New York, 1986

[59] R.G.Brown, "An introduction to random signal analysis and Kalman filtering," John Wiley and sons, 1983

[60] G.S.Christensen and S.A.Soliman, "Optimal filtering of discrete dynamic systems based on least absolute value estimations," Automatica, vol.26, no.2, 1990, pp. 389-395

[61] Discussions with K. El Naggar, Ph.D student, Department of electrical engineering, University of Alberta

[62] A.P.Sage and G.W.Husa, "Adaptive filtering with unknown priori statistics," Proceedings of the joint automotive control conference, 1969, pp. 760-769

[63] K.L.S.Sharma and A.K.Mahalanabis, "Harmonic analysis via Kalman filtering technique." Proceedings of the IEEE, 61, (3), 1973, pp. 391-392