An Investigation Into The Application Of Artificial Neural Networks And Cluster Analysis In Long-Term Load Forecasting

Prepared by: G Bougaardt, Master of Science student in the Department of Electrical and Electronic Engineering at the University of Cape Town

Prepared for: The Department of Electrical and Electronic Engineering at the University of Cape Town

1 January 2002

Thesis prepared in fulfillment of the requirements of the degree of Master of Science in Electrical and Electronic Engineering

ACKNOWLEDGEMENTS

No thesis can be completed without the help of many individuals. I would like to acknowledge the following people/organizations, with thanks:

My supervisor, Professor C. T. Gaunt, for his invaluable guidance, assistance in preparing this thesis and support throughout the project. To also gratefully acknowledge the ideas contributed by Professor C. T. Gaunt.

Thanks to Professor J. Greene, for his comments and suggestions to improve the thesis.

The author is also indebted to Eskom for the customer load data and the financial support for the research.

Thanks also goes to the South African Weather Bureau in Cape Town for the weather data that was made available.

And last but not least many thanks to my parents, brother, sister, and friends, whose support is always appreciated.

TERMS OF REFERENCE

Professor C. T. Gaunt commissioned this thesis at the University of Cape Town in the Electrical Engineering Department in February 2000.

Professor Gaunt's specific instructions were:

1. To investigate how to possibly produce an effective long-term load forecasting model by studying the load forecasting methodologies.

2. To investigate how to possibly produce an effective long-term load-forecasting model by segmenting customers according to the pattern of use of the customer.

ABSTRACT

This thesis investigates the problem of electric long-term load forecasting based on weather conditions (specifically temperature) and also investigates load forecasting by segmenting customers according to their pattern of use using clustering techniques in order to produce an effective long-term load forecast.

In an attempt to produce a possible effective long-term load forecast a methodology that uses linear regression analysis is compared to a methodology that implements artificial neural networks (ANNs). Long-term load forecasts require years of data however only a limited data set is available. The two methodologies have to be adjusted since the limited data set stretches across two seasons and long-term load forecasts are only performed for a season. For the limited data set, little or no load growth can be assumed contrary to long-term forecasts where long-term load growth is expected. The results show that for a small amount of data the regression model is more accurate than the ANN model however when a bigger data set (seasonal data set) is used the ANN model is slightly (0.1 %) more accurate than the regression model. The seasonal data set is the correct size data used since the long-term models use the same size data. Since the linear model is so successful, the polynomial non-linear regression analysis is tested and compared to the linear and ANN analysis. The results show that the polynomial and ANN analysis is more accurate than the linear regression model and therefore can be used to perform a more effective long-term load forecasts.

In an attempt to produce a possible effective long-term load forecast customers are segmented based on the load parameters of the customers using cluster analysis and compared to a load forecast that groups all the customers together as one aggregate load. A load forecast is performed separately for each cluster of customers. The load forecast of any customer is based on the load forecast of the cluster to which the customer is identified to. Only the available 12 customers are used in the cluster analysis and the load is forecasted for 2 customers. The results are mixed. In some cases the cluster analysis approach is effective and in other cases the cluster analysis approach is not effective.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Page

Figures

Tables

Chapter 1: Introduction

1 Content

This thesis investigates methods for producing an effective long-term load forecasting model using the available limited load data set for each customer, including the improvements to the forecast by characterising or classifying the limited number of customers into groups.

1.1 Background to the study

Power system expansion planning starts with a forecast of anticipated future load requirements. The accuracy of the load forecast is important to any electric utility. The utility needs to plan ahead so that it can supply the load that the system requires. Errors in this respect can lead to financial problems.

Years of load data are required to perform long-term load forecasts of about 10 to 20 years. The problem is that the author does not have years of load data available. The ideal case for the thesis is to have years of load data available for each customer. The author, Bougaardt G, only has a limited load data set for each customer since the electricity supplier only captured this data. This thesis discusses how the available limited load data set will be used and justified to investigate methods to produce an effective long-term load- forecasting model. The long-term models will be modified and designed for the limited load data set. Since the thesis investigates how to best utilise the available limited load data set to produce an effective long-term load forecasting model, the historical load data falls outside of the scope of the thesis and is not needed anymore since it is not limited load data. Historical load data can be many years of load data. This thesis does not discuss producing an effective long-term load-forecasting model when the load data is limited or small i.e. a long-term model that only uses a limited load data set. An effective long-term load forecast is worth investigating because there is a need for long-term load forecasts for developing countries. Presently in the literature there appears to be no long-term load forecasting models. Long-term models as reviewed in chapter 2 date back to before the 1980's. The thesis aims to improve on these models for an effective load forecast. Clayton et al (1973) and Davey et al (1973) focus on weather sensitive load forecasting and recognises the weather as an important variable.

In most cases, e.g. Clayton et al (1973) and Davey et al (1973), customers are all grouped together as one load for a load forecast. In some cases the customers are classified into classes such as a residential, commercial or industrial customer. This approach classifies customers according to the type of customer and does not recognise that each customer has its own characteristics. The load forecast result is the same whether all customers are grouped together as one load or classified according to the type of customer since the individual load forecast results of classes residential, commercial and industrial are added together to produce the aggregate load forecast. In this sense classification of customers is not used to improve the load forecast but merely as an alternative to observe load trends in the different customer groups i.e. residential, commercial and industrial. In other approaches, Sullivan (1977), suggests classifying customers according to the geographical area and tariff schedule. Sullivan (1977) then discusses that there can be a wide range of customers in a geographical area and tariff schedule classification and that this would not be ideal for load forecasting purposes. The problem is to define the characteristics of each customer. The customers with the same characteristics can then be grouped together. The problem is also how to classify the customers into groups. In order to produce an effective load forecast the author attempts to characterise the customers into groups based on each customer's characteristics. The author only has a limited number of customers since the electricity supplier only captured the data of a few customers. The ideal case for the thesis is to have many customers' load data since in any analysis the more data the better the conclusions that can be drawn. The data set for each customer is also limited as discussed before.

1.3 Objectives

1.3.1 The First Objective

The first objective is to possibly develop a more effective long-term load-forecasting model than the long-term models used in the literature using the available limited load data set. This can be achieved by investigating methods to produce an effective long-term load forecast by studying the methodology used in long-term load forecasts. If necessary, consider altering the methods used for short-term load forecasts or introducing techniques not presently used in long-term load forecasts. Since a limited load data set is used the experiments will be designed and modified for the limited load data set and the load forecasting results obtained from using the limited data set are used to justify if an effective long-term load-forecasting model can be developed.

The scientific approach is to compare different load forecasting methods and determine which method is most effective for long-term load forecasting based on the accuracy of the load forecast. The methods are compared in terms of the load forecast percentage error (%). The error load forecast (%) is determined by comparing the actual peak load (kW) in the future to the peak load forecast calculated using a load-forecasting model.

1.3.2 The Second Objective

The second objective is to test whether classifying or characterising customers into groups using the limited customer load data set improves the load forecast in the long-term load-forecasting model. Classifying or grouping customers could provide a better or more accurate load forecast than grouping together all the customers as one aggregate load. The load forecasting results obtained from using the limited load data set are used to justify if an effective long-term load-forecasting model can be developed.

The scientific approach is to compare the two approaches to load forecasting i.e. classifying customers into groups compared to not classifying the customers and determine if classifying customers into groups produces a more effective long-term load forecast based on the accuracy of the load forecast. The accuracy of the two different classification methods is measured in terms of the load forecasting percentage error (%).

If classifying customers into groups does produce a successful load forecast then this could lead to possibly developing a classification type load-forecasting model that can be used by other customers, whose load is to be forecasted. The customer, whose load is to be forecasted, would then be based on other customers' load forecasts (load data). These other customers' are in a group based on some form of classification. This method of load forecasting is different from the traditional methods of load forecasting (e.g. Clayton et al (1973)) where a load forecast of customers is based on the respective customers load data. This classification type load-forecasting model can then be used in any area where there are customers whose load is to be forecasted.

1.4 Scope of Limitations

1) The customer load data is limited to the Cape Town area. This is the only customer load data available. The load data for each customer is also limited.
2) Only a limited load data set is needed since the thesis investigates how to best utilize the available limited load data set to develop an effective long-term load-forecasting model.
3) Industrial and commercial customers are only considered in the thesis. Industrial and commercial customers provide a better variety of customers than residential customers do. Possible classifications or groups of residential customers are urban, rural, wealthy suburban, poor suburban. Industrial and commercial customers provide a wider variety. Eskom made twenty customers load data available over a period of one year, of which fourteen is complete.

Chapter 2: Literature Review

2 Content

This chapter reviews the literature on load forecasting. At present in the literature there appears to be no methods of long-term load forecasting in network system planning. In practice most of the network planner's use very simplistic approaches to long-term load forecasting that are reviewed here. The survey will therefore be supplemented by considering short-term load forecasting literature and methods for extrapolating data. A survey on conditioning customer load data such as classification methods is also presented.

2.1 The Approaches to Long-term Load Forecasting

This section review the methodologies and techniques used to perform long-term load forecasts and compare the different methodologies used in long-term load forecasts.

A regression analysis model relates summer weather to summer loads in the paper Heinemann et al (1966). The main principle of their model is that the system load is broken down into two components namely weather sensitive load and non-weather sensitive load and each forecasted separately. The model is described in the equation

$$\textbf{Daily Peak Load} = \textbf{Basic Load} + \textbf{Weather sensitive Load} \quad \ldots 2.1$$

$$\text{Or } \mathbf{DPL} = \mathbf{B} + \mathbf{CDF} * \mathbf{WV}$$

Where CDF = cooling demand factor

WV = weather variable

The cooling demand factor is a coefficient that relates to the summer air-conditioning saturation. Heinemann et al (1966) thus found it significant to separate weather sensitive load from the system load because the weather sensitive load grows at a different rate to that of the non-weather sensitive load. The weather variable is a complex composite of various weather variables, which includes both coincident and antecedent weather conditions. The basic load is the non-weather sensitive load. In order to develop the model described in equation 2.1 Heinemann et al (1966) uses multiple linear regression analysis. The paper uses 16 years of hourly-integrated loads and historical weather data in the regression analysis. The unknown constants in equation 2.1 are B and CDF. The daily peak loads are not homogeneous (i.e. there is load growth) and therefore the variables B and CDF will vary with time. The variables B and CDF are determined for each year in the study for the summer season. Once the model in equation 2.1 is known (i.e. B and CDF) it can be used for load forecasting for a summer season. The model can therefore study the separate growth trends of the basic load (i.e. non-weather sensitive load) and the weather sensitive load components. Of course a similar model can be developed for the winter season with appropriate changes in detail.

Stanton (1971) develops a weather load model that only includes temperature as a weather variable whereas Heinemann et al (1966) investigates various weather variables. Both papers investigate seasonal load forecasting. Stanton (1971) however focuses on separating the weather sensitive load from the non-weather sensitive load by using a weather load model. The weather load model is used to subtract the weather sensitive load from the system load to determine the non-weather sensitive load. Stanton (1971) uses the weather load model by projecting certain coefficients of the weather load model into the future to forecast load. The methodology used to develop the weather-load model is linear regression analysis. The weather-load model is however developed using a discontinuous linear regression model. This means that the weather-load model in Stanton (1971) is broken down into two main sections. The two sections are linear equations, which of course has slopes. The slopes are the characteristics of each weather-load model. The slopes of each weather-load model for each year are projected forward for load forecasting purposes.

Clayton et al (1973) is the first paper to discuss a program for long-term weather load modeling that uses about 20 years of historical weather data. The theme of this paper is also to develop a weather load model using regression analysis. Besides using several years of historical load and weather data to build the weather load model, an additional number of years of historical weather data is used as input to the weather load model to produce as output the peak load. Weather load models are developed for each year of the five years preceding the years for which the peak load forecast is to be carried out. Each year of the additional years of historical weather data used as input to the weather load model produces a possible range of peak loads. The mean peak load is calculated for each weather load model and used to forecast the load in the future. Davey et al (1973) is similar to Clayton et al (1973) in that the basic methodology is the same. The difference in Davey et al (1973) is that individual models are developed for each service area and peak load forecasts are carried out for each service area. The total peak load forecast is a summation of the individual service area's peak load forecast.

A paper by Corpening et al (1973) discussing long-term load forecasting is similar to Heinemann et al (1966). The model equation is the same, but Corpening et al (1973) calculates the weather variable WV in a different way. The weather function by Corpening et al (1973) includes 1) a non-linear function of weighted temperature, 2) non-linear dew point function, and 3) a combined cloud cover and wind speed function.

The difference between the load forecasting models described above is the way the weather-load model is developed. In some cases the weather-load model is broken down into weather-sensitive and non-weather sensitive load. The methods used to develop the weather-load models are regression analysis or variations of regression analysis. The methods or techniques used to perform load forecasting are fairly similar and rely on accurate weather-load models for accurate load forecasts. In order to produce a more effective long-term model is to possibly produce an accurate weather-load model since the whole long-term load forecasting methodology revolves around weather-load models. The next section reviews short-term load forecasting methods with the view to possibly amend the short-term methods and techniques for long-term load forecasting for a possibly more effective long-term load forecast.

2.2 The Approaches to Short-term Load Forecasting

This section review the methodologies and techniques used to perform short-term load forecasts and compares the different methodologies used in short-term load forecasts.

The short-term load-forecasting model of Gupta and Yamada (1972) uses a weather-load model to forecast the load for a 24-hour period. The weather-load model in Gupta and Yamada (1972) is not a static model as in the long-term load forecasting models. The weather-load model in Gupta and Yamada (1972) is an adaptive model. An adaptive model is one that updates its own parameters according to its past performance. The weather-load model is updated with the latest weather and current hourly load data. The weather variables are forecasted daily. The approach of Abou-Hussien et al (1981) is similar to that of Gupta and Yamada (1972) in that the model is also adaptive.

Several papers describe load forecasts being implemented using artificial neural networks (ANN) or variations of ANN technology. Most of the papers being reviewed are methodologies for short-term load forecasts i.e. one hour to a week ahead with the odd exception of monthly load forecasts. These papers are reviewed next.

A short-term feedforward backpropagation network is implemented in Chen et al (1992). This weather sensitive load model is implemented using load and temperature data as input. It requires forecasted temperature, indices listing the day of the week and the time of the day and the most recent load and temperature data as input. The output is the forecasted load. The main feature of Chen et al (1992) is that the backpropagation ANN network is not fully connected and hence the network is faster than the normal fully connected ANN network. This means that for example the neurons of layer one is not all connected to layer two. The fast algorithm of Chen et al (1992) is achieved by implementing a non-fully connected ANN network, but Dash et al (1994) proposes an improved Kalman filter based learning algorithm, which is faster than the backpropagation algorithm that Chen et al (1992) uses. Dash et al (1994) presents this new algorithm with the view to improve on the slow convergence rate of the backpropagation algorithm. Dash et al (1994) compares the backpropagation algorithm to that of the Kalman filter algorithm and shows that the number of iterations required for convergence for the Kalman filter is less than that of the backpropagation algorithm. At the same time the forecasting errors are compared which shows that the Kalman filter is superior. The input parameters are similar to that of Chen et al (1992).

A short-term load forecasting system called ANNSTLF or artificial neural network short-term load forecaster is implemented by Khotanzad et al (1995) in 20 electric utilities across the US using the backpropagation learning algorithm. Khotanzad et al (1995) forecasts hourly load differently from Chen et al (1992) and Dash et al (1994). The main feature of Khotanzad et al (1995) is that they recognize that there are different trends in a load-weather relationship and thus they break up the load-weather relationship into three modules. The three modules are 1) the weekly module, 2) the daily module and 3) the hourly module. Each module contains several ANN's and each module independently forecasts load from one hour to 24 hours. For the final hourly load forecast an adaptive combiner combines the hourly load forecasts from each module for that hour.

The artificial neural network short-term load forecaster or ANNSTLF, which is discussed in a paper by Khotanzad et al (1995) is also discussed in Khotanzad et al (1997). The latter paper gives more detail about the ANNSTLF program and the neural network includes relative humidity as a weather variable, which the previous paper did not discuss. A weather forecasting engine is discussed and is part of the ANNSTLF software package. The weather forecasting engine forecasts hourly temperature and relative humidity. The forecasters are separate and are ANN networks. The forecasted temperature and relative humidity are needed as input to the neural network in order to produce the hourly load forecast. The weather ANN networks is multilayered feedforward and is trained using the backpropagation learning algorithm.

The papers reviewed thus far discuss artificial neural networks. However another branch of neural networks is discussed in Dash et al (1996) and is called fuzzy neural networks. For an artificial neural network to be called a fuzzy neural network (FNN) the input and weights of the network should be fuzzified. This means that the input to the network, which is the load and weather data, is classified into small, medium and large. So each set of the input data i.e. load and weather is first classified according to small, medium and large and then passes through the rest of the neural network. The output of the network, which is the forecasted load, is then defuzzified according to the classification small, medium and large. Fuzzification and defuzzification really means classifying the load and weather variables into small, medium and large. Dash et al (1996) discuss fuzzy set theory/rules, which is how each class i.e. small, medium and large is defined so that the classes by definition do not overlap. The algorithm used for training the fuzzy neural network is the backpropagation algorithm. The neural network forecasts load from one hour to several hours and one of the input variables is forecasted temperature and forecasting the temperature is a separate model.

Bashir and El-Hawary (2000) discuss another branch of artificial neural networks called wavelet neural networks or WNN. Fuzzy neural networks classify the input variables into small, medium and large whereas wavelet neural networks transform the input variables by using wavelet neurons. The resulting weights in the network are called wavelet weights. The WNN is trained using a backpropagation algorithm and the main feature for transforming a neural network into a wavelet neural network is because of its fast convergence rate during training. The paper compares WNN to ANN by means of the average forecasting percentage error. WNN proves to be more accurate than ANN.

In conclusion the ANN technology used to perform short-term load forecasts could possibly prove to be a useful tool in long-term load forecasting if a method is found since they prove to produce very accurate short-term load forecasts. The next section reviews ways to generate load forecasts.

2.3 The Methods or Technology used to Generate Load Forecasts

The success of any load-forecasting model is the accuracy of the load forecast. The literature does not mention the level of accuracy required for a load-forecasting model to be regarded as a success. The following sections review the techniques used to generate load forecasts.

2.3.1 Extrapolation of Annual Peak Loads

These methods involve fitting a trend curve to past values of annual peak load and make the forecast by extrapolating this trend curve forward to the desired year of forecast. This is a regression analysis of one variable i.e. annual peak load. In this respect the effect of weather on the peak load is ignored on the basis that similar weather conditions can be expected at the time of the annual peak loads. The assumption is that the annual peak load always occurs in the same season.

2.3.2 Energy and Load Factor Method

Annual energy forecasts can be used to determine the annual peak load forecast using the annual load factor, Glover and Sarma (1994), since the

$$\text{Load Factor} = \frac{Annual\ Energy\ Used}{(Annual\ Peak\ Load)*(Time)} \quad \ldots 2.2$$

Where Annual Energy Used in (kWh)
Annual Peak Load in (kW)
Time in (hours) for a year

Sullivan (1977) suggests that annual energy forecasts are more reliable than peak load forecasts however forecasting annual load factors is very difficult since system load factors can vary erratically.

2.3.3 Regression Analysis

The most commonly used methods to perform load forecasts are multiple regression methods. There are various variations of regression methods. There is linear regression analysis, which assumes a linear relationship between the dependent variable and the independent variable. Some of the weather-load models as discussed in section 2.1 use linear regression analysis. The regression analysis methods use weather data as the independent variable and load data as the dependent variable. This approach is different to section 2.3.1, which only considers load data in the regression analysis. There are also non-linear estimation methods and they are exponential, logistic, Gompertz and polynomial models among others.

2.3.4 Separate Extrapolation of Weather Sensitive and Non-Weather Sensitive Load

This approach considers separately forecasting the weather sensitive and non-weather sensitive load of the annual peak load. The final annual peak load forecast is the summation of the weather sensitive and non-weather sensitive load. In order to perform such a load forecast a weather-load model is required. The reason for developing the two components of the annual peak load is because the two components grow at different rates, Heinemann et al (1966). The advantage of this approach is that the load forecaster can study the growth of the two components.

2.3.5 Artificial Neural Networks

The literature in the 1990s makes extensive use of artificial neural networks to perform load forecasts. There are also variations of ANNs and they are fuzzy neural networks and wavelet neural networks. The ANN technology is used to determine the relationship between the dependent variable (load) and the independent variables (e.g. weather and load) in the weather-load model as discussed in section 2.2. The use of ANN technology in short-term load forecasting provides very accurate load forecasts as discussed by many of the papers. The ANN technology could be of use in developing an effective long-term load-forecasting model since ANN technology proves to provide accurate short-term load forecasts.

2.4 Customer Load Data Conditioning

This section reviews the different ways to classify customers such as residential, commercial, industrial (type of customer) and variations of this.

2.4.1 Classification of Customers

Most of the load forecasting methodologies discussed does not attempt to classify customers in any way. In some cases customers are classified broadly as residential, commercial and industrial. This is a classification according to the type of customer as discussed in Sullivan (1977). Sullivan (1977) discusses that there are further subdivisions within each of the classes. Sullivan (1977) discusses that this is unfortunate for load forecasting purposes since these classifications overlap in the sense that customers in a given class do not have characteristics unique to that class. For example residential customers can further be subdivided into rural and urban. Sullivan (1977) suggests that classification by the geographic area, tariff schedule, or level of use can make more useful classifications for load forecasting purposes. There can still be a wide range of customers in a geographic area if classified by this method. There can also be a wide range of customers in the tariff schedule classification method. The level of use refers to the amount of energy (kWh) or load consumption (kW, kVAr, kVA) of the customer. The level of use classification method identifies the unique characteristics of a customer and could prove to be a possible way to classify customers.

2.4.2 Cluster Analysis

This is a special statistical technique for classifying objects. It classifies objects based on the objects' characteristics as discussed in Everitt (1974) and Romesburg (1984). Cluster analysis concentrates of different ways to group or cluster objects together and involves many clustering algorithms and techniques, Everitt (1974) and Romesburg (1984). The objects in this thesis would be the customers. The approach of cluster analysis could prove to be a useful tool in grouping the customers by considering the level of use of each customer. The level of use of each customer is unique to each customer and cluster analysis recognizes unique characteristics or attributes of the objects or customers.

2.4.3 Spatial or Small Area Load Forecasting

Spatial load forecasting can be described as a means of classifying customers since the methodology involves dividing the utility service area (geographical area) into a number of small areas and forecasting the future load in each. The small areas are cells of a uniform square grid. There can be various types of loads (i.e. residential, commercial and industrial) in each small area. Willis and Tram (1983) discuss small area load forecasting. There are two categories of small area load forecasting namely simulation and trending. Trending methods produce future load forecasts by extrapolating past trends. Simulation methods make use of much more involved procedures that require much more data. The paper discusses a new small area trending method that only works on a rectangular grid basis. The paper presents the new technique, which combines two other trending methods. The first concept upon which this new technique is based is called Vacant Area

Inference (VAI). VAI is a technique that trends the load of a small area (cell) with no load history (vacant area) by considering the total projected group load of the surrounding cells. The load forecast of the vacant area (cell with no load history) is dependent on the load forecast of the surrounding small areas (cells). The second concept upon which this new technique is based is called Cluster Template Matching, a trending method. Cluster template matching extrapolates the load not on an individual small area (cell) basis but on a set basis. A set can include 5 or 6 small areas (cells). The paper describes the new forecasting method that uses VAI's concept of subdivision, applied to the cluster template matching of extrapolation for improved or more accurate load forecasting. The paper presents a flowchart of the new forecasting procedure.

A paper by Willis et al (1983) discusses data collection for load forecasting purposes in particular small area load forecasting and other forecasting methods. A land-use method requires information on the type and amount of each type of customer within each small area. Types of customers are residential, commercial and industrial. It is described that the most popular method of collecting this data is by manual interpretation of aerial photographs. A few low altitude aerial photographs are collected that cover the study area. Grid lines are drawn onto the photographs. Each square is identified as a small area. Each small area is inspected to identify the amount and type of customer. The actual interpretation of the amounts and types of customers is based entirely on human judgment. This method of collecting customer data is labor intensive and prone to error as discussed in the paper however this method is widely used. Another source of land-use data is the utility customer database. Another way to acquire land-use data according to the paper is by using NASA's Landsat satellites. The satellites produce digital pictures.

Willis and Northcote-Green (1983) present a review of spatial load forecasting techniques. They discuss the different classes of spatial load forecasting. They discuss non-analytic and analytic methods and the degree of complexity of the methods. Non-analytic methods rely almost completely on user intuition whereas analytic methods perform analysis of past and present data that involves trending and multivariate analysis. Trending methods include regression-based methods and V.A.I that work with load data whereas multivariate analyses include methods that work with more than load data. Non-analytic methods are regarded as simple methods whereas analytic multivariate analysis is regarded as very complex methods and therefore also produces accurate load forecasts. The non-analytic methods produce poor or inaccurate load forecasts as presented in the paper.

2.5 Conclusions

There is a need for better long-term load forecasting methods since the methodologies discussed in the literature review are simplistic. There seems to be little progression in methodologies for long-term load forecasting since the methodologies constantly use regression analysis or variations of regression analysis. There is a progression of technologies in short-term load forecasting, which starts with the use of regression analysis and progresses to the use of ANN technology and variations of ANN technology.

In order to produce a more effective long-term load forecast, Sullivan (1977) suggests that a more useful classification system can be developed other than classifications by type i.e. residential, commercial or industrial. The level of use of a customer could prove to be useful in classifying customers since this is a unique characteristic of a customer. In spatial load forecasting classification is by means of a small area (square or rectangular cell). There can be a wide variety of customers (residential, commercial and industrial) in a small area as discussed in Willis et al (1983).

Chapter 3: Basic Theory

3 Content

This chapter discusses the basic theory that will be used throughout the thesis. The chapter concentrates on the basics of forecasting, regression analysis, weather-load models, artificial neural networks (ANNs) and cluster analysis because these are the techniques identified in chapter 2.

3.1 The Basics of Forecasting

The success of any load-forecasting model is determined by the accuracy of the model. The accuracy of the load forecast is represented by the absolute percentage error (APE in %) and is calculated as follows, Dash et al (1994) and Elkateb et al (1995).

$$\text{APE} = \frac{|Forecasted\ Peak\ Load - Actual\ Peak\ Load|}{Actual\ Peak\ Load} * 100\% \ldots 3.1$$

Where Peak Load is measured in kW

And in some cases the additional mean absolute percentage error (MAPE in %) is used to determine the overall performance of the load forecast and is calculated as follows, Dash et al (1994) and Elkateb et al (1995).

$$\text{MAPE} = \frac{\sum_{1}^{N}(APE)}{N} \ldots 3.2$$

There are basically two types of load forecasts i.e. demand and energy forecasts. Energy forecast is a forecast of kWh (kilowatt-hours) whereas the forecast of demand is the forecast of real power or kW (kilowatts). As discussed in chapter 2 demand forecasts can depend on energy forecasts if the method of "Energy and Load Factors" is used.

3.2 Regression analysis

The general purpose of multiple regression analysis is to analyze the relationship between several independent variables and the dependent variable. This relationship between the dependent and independent variables is in the form of an equation. The equation is a line, which best fits the data points. The line can be computed using the least squares approach. The least squares approach will compute a line so that the squared deviations of the observed points from the line are minimized. The line equation can take the form of a straight, polynomial and exponential line etc., Sullivan (1977).

The regression analysis process is a simple approach and the user chooses which line should fit the data points. If linear regression analysis is used the assumption is that the relationship between the variables is a linear one. The major conceptual limitation of all regression techniques is that one can only ascertain relationships but can never be sure.

3.3 Weather-Load Models

System loads respond to changes in weather conditions, Sullivan (1977). The peak load demands in summer corresponds to the hot summer days and the peak load demands in winter corresponds to the cold winter days. In the summer season air conditioning loads contributes to the system load and in the winter season electric heating loads contribute to the system load, Heinemann et al (1966).

There are many more weather variables used in long-term load forecasting than in short-term load forecasting. Long-term load forecasting models can concentrate on weather variables such as dry-bulb and wet-bulb temperatures, relative humidity, wind velocity, wind direction, dew point etc. Short-term load forecasting models include temperature data and might include humidity data. The temperature and the humidity data are recorded in °C and % respectively. A weather station supplies hourly and daily weather data to the electric utilities. The load data used in load forecasting models are supplied by the electric utilities. The load data are recorded in terms of half-hourly integrated kWh and kVArh. The half-hourly kWh can be converted to real power (kW).

The weather and load data (kW) are used to develop weather-load models by means of regression analysis and ANN technology. In regression analysis an equation is developed which relates the load as a function of the weather variables. ANN technology is a form of non-linear regression.

3.4 Artificial Neural Networks

A brief background of neural networks is presented and how this background corresponds to the model of neural networks as implemented in the literature. A basic neural network model is presented and the architecture of the model is discussed.

3.4.1 Introduction to Artificial Neural Networks

The human body or more specifically the central nervous system inspired artificial neural networks (ANNs). ANNs are simplified models of the central nervous system. Consider a neuron model from the human neural system shown in figure 3.1 from Patterson (1996).

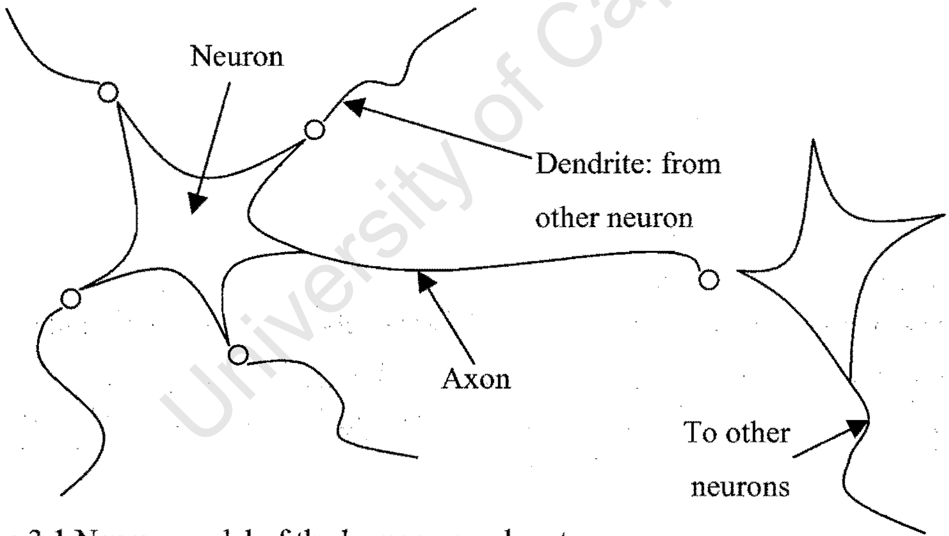


Figure 3.1 Neuron model of the human neural system

From the above model each neuron is connected to another neuron via the dendrite or the axon. A neuron takes input via the dendrites and presents an output to the next neuron via the axon. The neuron model above is used to develop the ANN model shown in figure 3.2 from Patterson (1996). The interconnections are the dendrites.

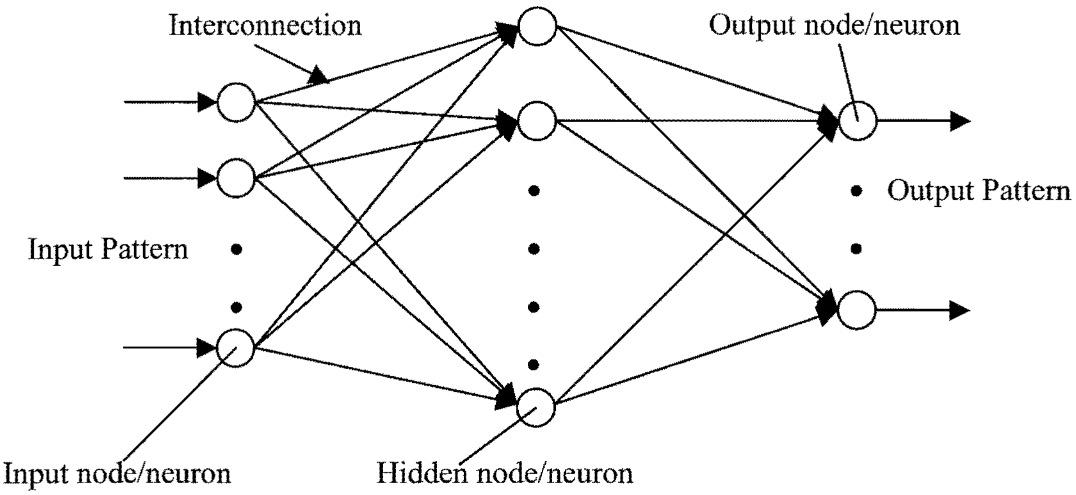
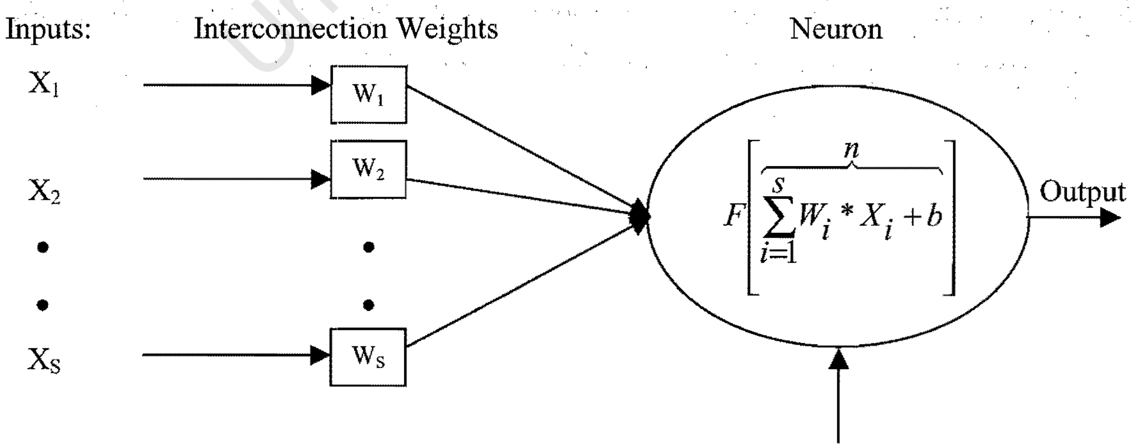


Figure 3.2 Artificial neural network model

ANNs are systems that comprise of a number of processing units or neurons that are linked via weighted interconnections. A processing unit is really a mathematical equation, which is also referred to as a transfer function. The transfer function is typically a non-linear, bounded differential function such as the hyperbolic tangent or sigmoid function. The transfer function is a two-dimensional function that takes an input and produces an output. The input to the transfer function is a summation of the inputs to the neuron multiplied by the weighted interconnections plus a bias. This is shown in figure 3.3 from Patterson (1996).



Where F is the transfer function and b is the bias.

Figure 3.3 Schematic diagram of a single neuron.

Output $= a = F\left[\overbrace{\sum_{i=1}^{S} W_i * X_i + b}^{n}\right]$. The above mentioned Tan-sigmoid and

log-sigmoid transfer functions are shown in figure 3.4 from Demuth and Beale (2000).

Tan-Sigmoid Transfer Function **Log-Sigmoid Transfer Function**

Where $n = \sum_{i=1}^{S} W_i * X_i + b$

Figure 3.4 The Tan-sigmoid and Log-sigmoid transfer function

3.4.2 Characteristics of Artificial Neural Networks

The ANN network in figure 3.2 is a feedforward network because signals propagate only in a forward direction, from the input nodes to the output nodes. The behavior of a neural network, how it maps input data to output data, is determined by the transfer functions of the neurons, how they are interconnected, the weights of those interconnections and the bias of each neuron, Demuth and Beale (2000). The architecture of a neural network is first established, Demuth and Beale (2000). This includes how many input neurons, hidden neurons and output neurons are used and also which transfer functions are used in the neurons. Then a mathematical algorithm is used to determine what the weights of the interconnections and bias should be to maximize the accuracy of the outputs produced. The mathematical algorithm is a learning algorithm that aims to learn the relationship between the input variables and the output variables by using previous examples. This is also referred to as the training process where the neural network is trained to determine the relationship between input and output variables by setting the weights and biases.

Once the relationship is established (i.e. the weights and biases are set), the neural network can be presented with new inputs and can generate predictions.

3.4.3 The Learning or Training process

As discussed before, to train a neural network, data needs to be gathered which includes the input and output data, Demuth and Beale (2000). The network structure or architecture needs to be selected and then the network is trained by using a learning algorithm. Two sets or samples of data are set aside for a neural network. The first set of data is used to train the network, the other set is used to test the network. The network is trained, not programmed, meaning that it starts out not knowing anything about the patterns presented to it. The network will determine the relationship between the variables by repeatedly showing the patterns to the network. It is because of this process that a network can be overtrained. An overtrained network means that the network will memorize the training patterns. This is undesirable because a network that has memorized the training patterns results in a network that generalizes poorly. Generalization means that when new, unique data patterns are presented to an overtrained neural network, the output will tend to resemble the nearest neighbor recalled from the set of memorized patterns.

On the other hand a neural network can also be undertrained. When an insufficient amount of data is presented to the neural network during the training process, the network will not be able to determine the relationship between the variables. An undertrained network is evident from a high error on the training data.

The test data mentioned above is a set of data used to test the performance of the network. So once a network is trained the test data is used to investigate whether the network may be overtrained, Demuth and Beale (2000). Once the performance of the network is satisfactory then it can be used for prediction.

3.4.3.1 Supervised and Unsupervised Learning

Learning methods for ANNs can be classified as one of two basic types: supervised or unsupervised, Demuth and Beale (2000). Learning methods are procedures for modifying the weights and biases of a network. In supervised learning a set of example patterns, which includes the input pattern and the corresponding output/target or desired pattern, is presented to the network during training. The target or desired pattern is the correct answer. During the learning process a comparison is made between the calculated output and the target, which is the correct output. The difference between the two, which is the error, is used to change the network's weights and biases. How the error is used to change the weights and biases of the network is dependent on the learning algorithm used. The resultant error is feedback to adjust the weights and biases until the error has reduced to an acceptable level for each training pair in the training set.

In unsupervised learning there are no target output available. The weights and biases of the network are modified only in response to the inputs.

3.5 Cluster Analysis

The objective of cluster analysis is explained which is basically to form clusters or groups. The process to achieve the clusters is discussed.

3.5.1 Introduction to Cluster Analysis

Cluster analysis is a branch of multivariate statistics that classifies objects/cases (e.g. individuals, species etc.) into classes/groups or clusters, Everitt (1974) and Romesburg (1984). The idea is to group objects together that are similar. The aim is to establish a set of clusters such that objects within a cluster are more similar to each other than they are to objects in other clusters. Similarity can be measured in a number of ways and it is because of this that classification of objects depends upon the particular method used. There are two groups of ways to actually measure similarity, this is a similarity measure and a distance measure. A similarity measure includes, among others, correlation and binary matching methods. A distance measure includes, among others, Euclidean distance, squared Euclidean distance, Chebychev and city block methods. Each object has attributes or characteristics and it is these attributes that are used to measure similarity between objects.

3.5.2 Choice of Clustering Algorithm

The next step in the cluster analysis procedure is to choose the clustering algorithm, Everitt (1974) and Romesburg (1984). A clustering algorithm is a rule, which governs which objects are grouped together. There are various groups of clustering techniques such as hierarchical techniques, partitioning techniques, density search techniques, clumping techniques and other techniques. Hierarchical techniques are further subdivided into agglomerative methods and divisive methods. Agglomerative methods are, among others, the nearest neighbor method, the furthest neighbor, group method and Ward's method as presented in Everitt (1974) and Romesburg (1984). Agglomerative implies that each object starts out as a cluster on its own and clusters are formed until one cluster is formed. The literature, Everitt (1974) and Romesburg (1984), does not suggest which method is suited for a particular problem and the literature, Everitt (1974) and Romesburg (1984), suggests that the choice of clustering algorithm requires an expert opinion. The consequence of a choice of clustering algorithm is that the cluster formation would be slightly different for each clustering algorithm.

3.5.3 The Output of the Cluster Analysis

The output of the cluster analysis is a dendogram or tree. It shows where objects combine to form a class/cluster and the measure of similarity when they combine. An example of a tree is shown in figure 3.5, Romesburg (1984). The horizontal axis represents the objects in the analysis and the vertical axis represents the Euclidean distance (or linkage distance) at which objects combine.

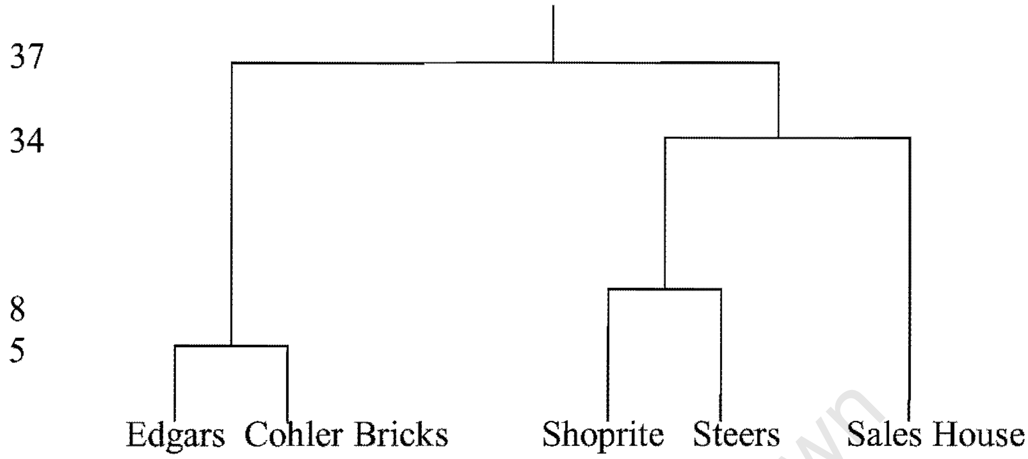


Figure 3.5 An example of a dendogram

The dendogram has to be interpreted. The question is, where should the tree be cut. Where the tree is cut determines the number of clusters in the classification. For example from the above tree, if the tree is cut just before the linkage distance of 34 then there will be 3 clusters i.e. cluster 1 = Edgars and Cohler Bricks, cluster 2 = Shoprite and Steers and cluster 3 = Sales House. If the tree is cut just after the linkage distance of 34 then there will be 2 clusters i.e. cluster 1 = Edgars and Cohler Bricks and cluster 2 = Shoprite, Steers and Sales House. This step is discussed in Romesburg (1984) and says that it is sometimes left to that of an expert opinion however the most popular method as discussed in Romesburg (1984) is to cut the tree where the cluster structure remains stable for a long distance. This can be computed by calculating the linkage distance between one cluster and the next cluster formation. This is computed for the whole tree. The largest of these distances would represent where the cluster structure remains stable for a long distance. This is where the tree is cut. For the above dendogram, the calculations would be as follows: linkage distance 8 – 5 = 3, linkage distance 34 – 8 = 26 and linkage distance 37 – 34 = 3. The largest of these distances is 26 and this is where the tree is going to be cut. Thus the clusters are cluster 1 = Edgars and Cohler Bricks, cluster 2 = Shoprite and Steers and cluster 3 = Sales House.

3.5.4 Flowchart of Cluster Analysis Process

This section summarizes the cluster analysis process in figure 3.6.

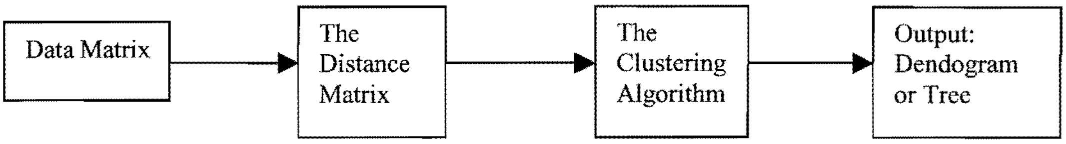


Figure 3.6 Flowchart of cluster analysis process

The data matrix is a matrix of the objects and the attributes of the objects. The distance matrix is calculated using the data matrix. The distance matrix is a matrix that determines the similarity between the objects based on the attributes of the objects. The clustering algorithm uses the distance matrix to determine the dendogram, which is the output of the cluster analysis.

Chapter 4: Theoretical Development

4 Content

This chapter develops the theory discussed in chapters 2 and 3. Ideas are presented for the purpose of developing an effective long-tem load-forecasting model using ANN technology and cluster analysis.

4.1 Development of Possible Effective Long-term Load Forecasting Model

This section builds on the literature discussed in chapter 2 to develop a possible effective long-term load-forecasting model. An interest in ANN technology is identified because of its accuracy in the short-term model. This section investigates and analyses why the short-term ANN model cannot be extended to perform long-term load forecasts and how possibly to develop a long-term model using ANN technology. A flowchart of a possible effective ANN model is presented and a flowchart of the model which inspired the possible effective model. The changes to the model are identified.

4.1.1 Investigating and Analyzing the Short-term ANN Model

It is interesting to note that there appears to be no papers that implement ANN technology for long-term load forecasting.

The question that arises is why does it appear that there are no papers that use ANN technology for long-term load forecasting since the technology produces very accurate short-term load forecasts as discussed in the papers that implement ANN technology. Consider the flowchart in figure 4.1of the short-term load forecasting using ANN technology.

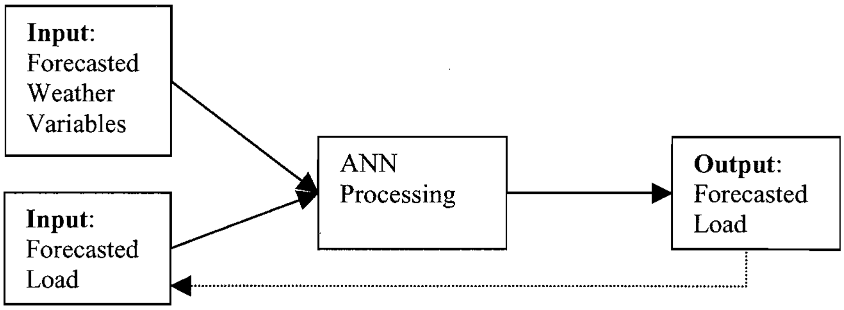


Figure 4.1 Flowchart of typical short-term ANN Model

The input to the ANN model is forecasted weather data and previously forecasted load data. The output of the ANN model is forecasted load and the accuracy of this load forecast depends on accurate forecasted weather data and on previously forecasted load. This is an iterative load forecasting process. Weather variables can only be forecasted accurately over a short-term period. Forecasting weather variables over a long-term period would produce inaccurate weather forecasts and hence inaccurate load forecasts. In an iterative process errors accumulate which makes the method unsuited to long-term forecasting. Analyzing the ANN papers shows that it is not the ANN technology that is the problem but rather the method of load forecasting as explained above. In conclusion the ANN technology could possibly prove to be a useful tool in long-term load forecasting if a method is found since it provides very accurate short-term load forecasts.

4.1.2 Long-term Load Forecasting Methodology using ANN Technology

This section investigates how to possibly apply the ANN tool to long-term load forecasting.

The long-term load forecasting methodology of Clayton et al (1973) and Davey et al (1973) could provide a way to implement the ANN tool for long-term load forecasting. The models of Clayton et al (1973) and Davey et al (1973) were reviewed in chapter 2. A flowchart of the model is presented in figure 4.2.

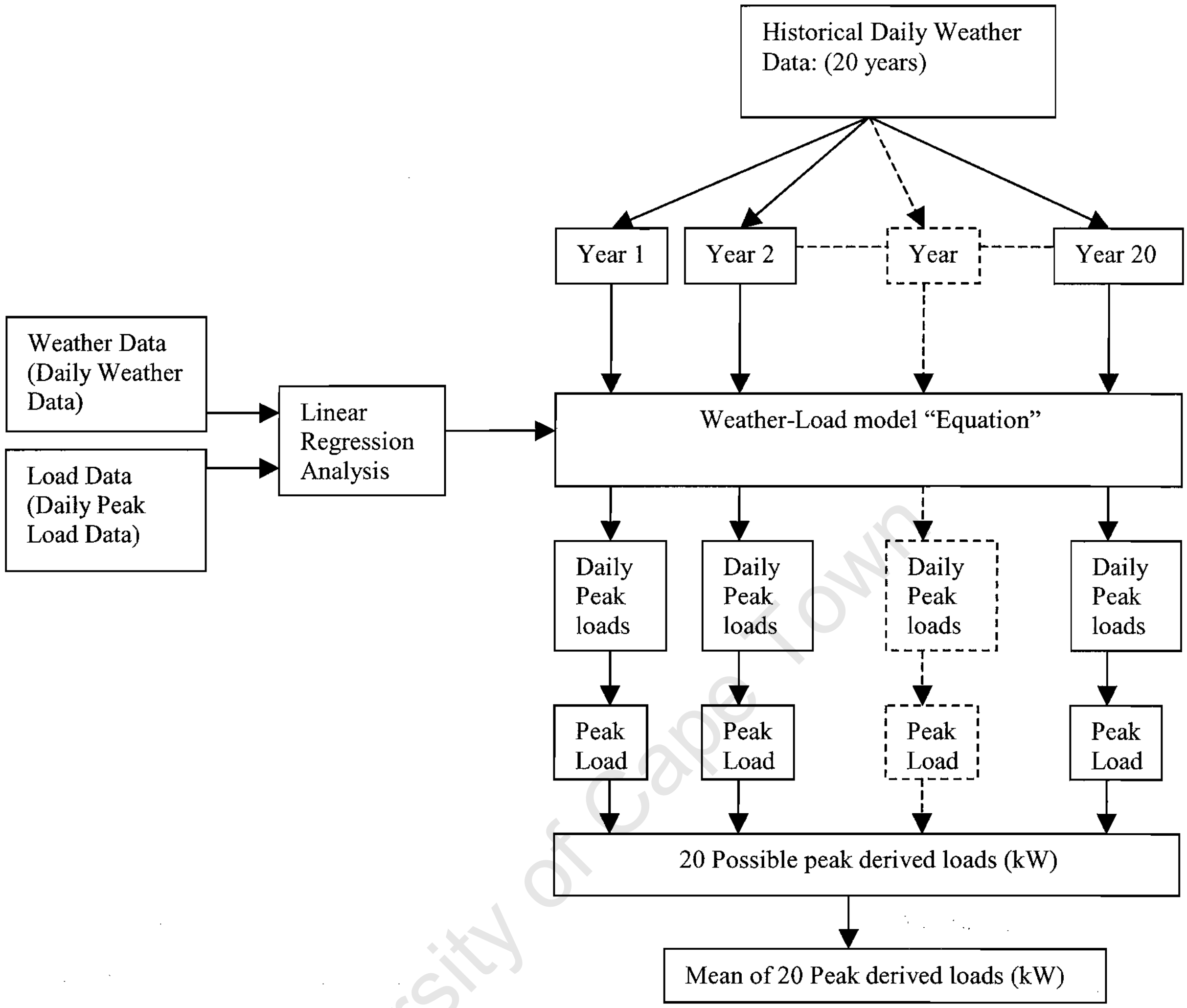


Figure 4.2 Flowchart of Long-term methodology

If the annual winter peak load is to be forecasted for the years 2000, 2001 and 2002 then the process in figure 4.2 is performed for typically 5 years preceding the forecasted period (i.e. 2000, 2001, 2002). The process in figure 4.2 is performed for each of the 5 years separately. The 5 years would be 1999, 1998, 1997, 1996 and 1995. The winter daily peak loads and the corresponding winter daily weather data for each of these years are used to develop the weather-load model for that year. The historical daily winter weather data used as input to each of the 5 weather-load models are typically 20 years of daily winter weather data from the period 1994 to 1975. This historical daily weather data produces daily peak loads as output from the weather-load model. Each year of historical weather data produces one peak-derived load. An example of 20 peak derived loads for the year 1995 is presented in figure 4.3.

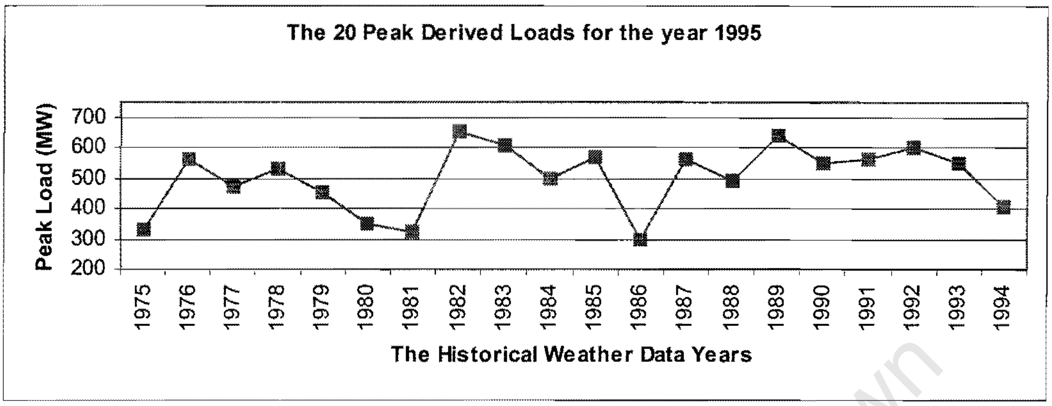


Figure 4.3 Example of 20 peak derived loads for the year 1995

The Mean Load (figure 4.3) for the year 1995 is 500 MW. This Mean Load is presented in figure 4.4 for the year 1995. The final load forecast process is shown in figure 4.4.

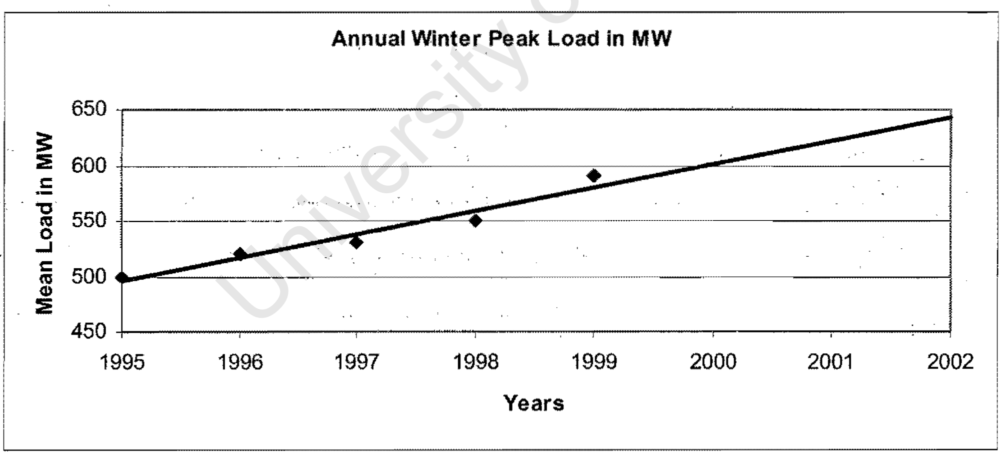


Figure 4.4 Annual Winter Peak Load Forecast for years 2000 to 2002

The Mean Loads as shown in figure 4.2 are calculated for each of the years 1995 to 1999. The Mean Loads for each of the years 1995 to 1999 are projected forward to forecast the annual winter peak load for the years 2000 to 2002 as shown in figure 4.4.

Clayton et al (1973) and Davey et al (1973) use historical weather data to produce long-term load forecasts and do not forecast weather data. This is different to the ANN short-term model that forecasts weather data. The methodology of Clayton et al (1973) and Davey et al (1973) can be used with the ANN technology to possibly produce long-term load forecasts. Clayton et al (1973) and Davey et al (1973) uses linear regression analysis to perform long-term load forecasts whereas the ANN model uses ANN technology. A flowchart of a possible ANN model is presented in figure 4.5.

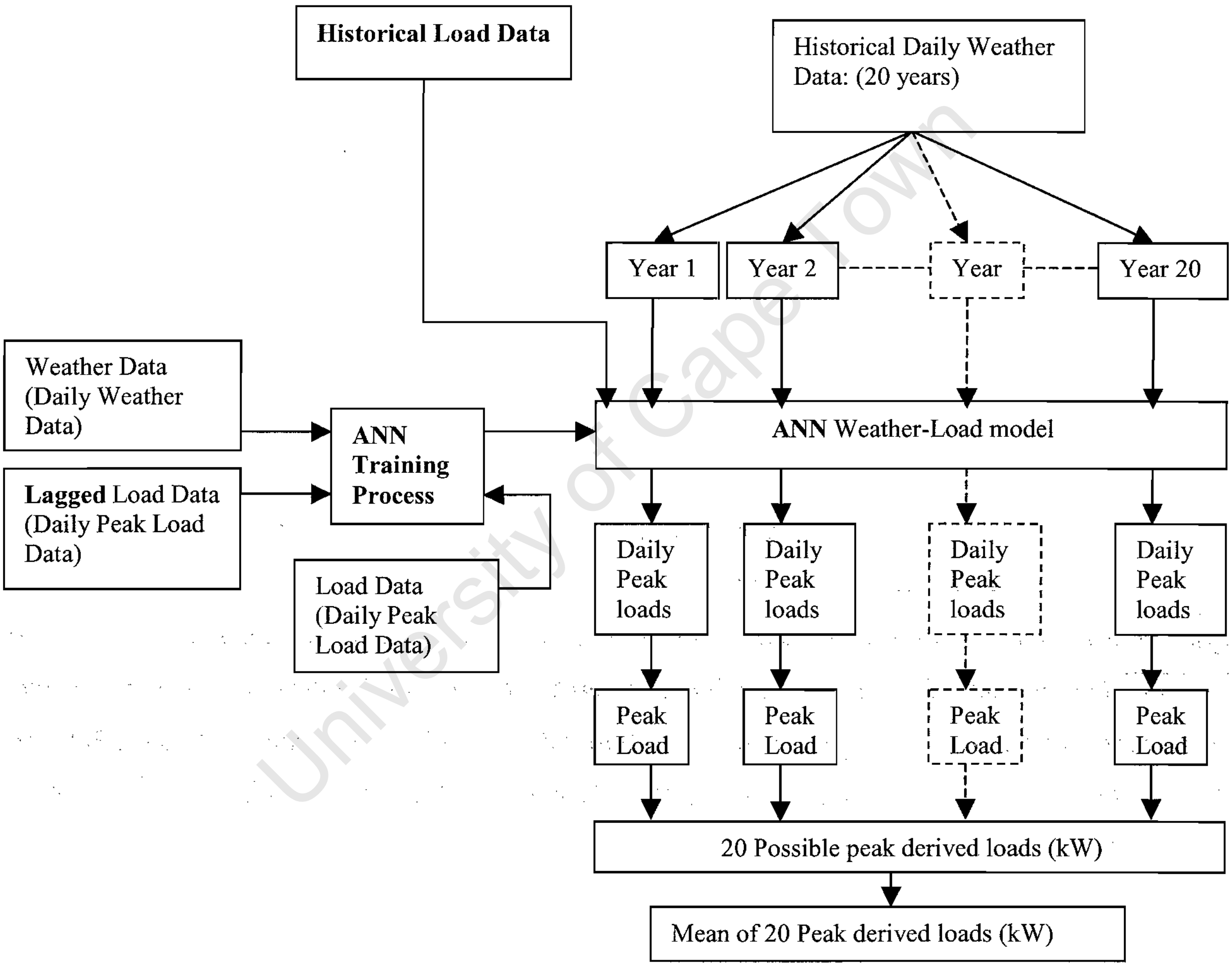


Figure 4.5 Flowchart of ANN Long-term model

The model of Clayton et al (1973) and Davey et al (1973) in figure 4.2 only makes use of historical weather data and therefore only recognises the load as a function of the weather. On the other hand as shown in figure 4.1 for the short-term ANN model the load is a function of the weather and previously forecasted load (lagged load). The methodology of the short-term ANN model (figure 4.1) is an iterative process and an iterative process for load forecasting is not suited for a long-term model. The long-term model of Clayton et al (1973) and Davey et al (1973) is a non-iterative model. The lagged load approach can be used differently as follows. The weather-load model can be developed based on weather data and lagged load data as shown in figure 4.5. The load is therefore a function of weather and lagged load. The historical load data as shown in figure 4.5 could then be used as input to the weather-load model together with the historical weather data to produce daily peak loads, the peak load, 20 peak loads and then eventually the mean of the 20 peak loads. It is still a non-iterative process.

4.2 The Use of the Limited Load Data Set for Long-term Load Forecasting

This section discusses how the use of the limited load data set will be used and justified for the load forecast.

Long-term load forecasts require years of load data to forecast load in the long-term as discussed in Clayton et al (1973) and Davey et al (1973). The author does not have years of data available. No other data is available. No other data is needed since the long-term models will be designed and modified for the limited load data set as discussed in chapter one. The author has 18 weeks of load data available. The first 10 weeks of load data will be used to forecast the load for the following 8 weeks. The only difference between a few weeks of data and years of data is the size or amount of data. Years of data produce several years of load forecasts whereas a few weeks of data produce several weeks of load forecasts. The difference is the forecast period. The basic load forecasting methodology that is used, i.e. Clayton et al (1973) and Davey et al (1973), is a non-iterative process and this remains the same whether years or weeks of data are used. For a long-term load forecast long-term load growth is taken into account whereas in a limited data set little or no load growth can be assumed.

4.3 Development of Possible Effective Long-term Model using Classification of Customers

This section builds on the literature discussed in chapter 2 to develop a possible effective long-term load-forecasting model by classification of customers. An interest in cluster analysis was first discussed in chapter 2 and it identified that each object is classified based on its unique characteristics. This interest in cluster analysis is further discussed here. The parameters that will be used for the customers are identified. The development of the model and load forecast process is discussed here. A flowchart of the cluster analysis process along with the load forecast process of a cluster is presented. A flowchart for identifying a customer (a customer whose load is to be forecasted) to a cluster is presented.

4.3.1 The Choice of the Load Parameters for the Cluster Analysis

In chapter 2 the use of parameters such as kW, kVAr and kVA was identified as possible characteristics of load customers. This section identifies other possible load parameters that could be used in the cluster analysis.

The customer load data that is made available is the kWh and the kVArh in half-hourly integrating periods for each customer. From the load data that is made available i.e. kWh and kVArh, the kVAh can be calculated using the formula, Glover and Sarma (1994):

$$\text{kVAh} = \sqrt{(kWh)^2 + (kVArh)^2} \quad \ldots\ 4.1$$

Where kVAh = Apparent Energy

kWh = Active Energy

kVArh = Reactive Energy

The Apparent, Active and Reactive Energy load data will be used to determine the load parameters. The initial load parameters that are considered for the cluster analysis using the load data kWh, kVArh and kVAh are shown in table 4.1. The relevance of each load parameter will be discussed.

Table 4.1: Load Parameters for the cluster analysis

Load Parameter	Units
Max Apparent Power	kVA
Max Active Power	kW
Max Reactive Power	kVAr
Ave Apparent Power	kVA
Ave Active Power	kW
Ave Reactive Power	kVAr
Min Apparent Power	kVA
Min Active Power	kW
Min Reactive Power	kVAr
Max Power Factor	---
Ave Power Factor	---
Min Power Factor	---
Load Factor	---

Through consultation with Gaunt [Ref. (Personal Communication)] it was suggested that the min[kVA, kW, kVAr] can be neglected. The reasoning behind this is that the minimum values of kVA, kW and kVAr has little significance for most networks when it comes to load forecasting and is only needed in special cases where the operator wants to know how low the load might get. The objective of load forecasting is to ensure that the network capacity and generation output will be sufficient and therefore the maximum and average values of kVA, kW and kVAr are models of how big the network capacity is. Since the min[kVA] and the min[kW] is neglected, the min[Power Factor] will also be neglected because to calculate the min[Power Factor] the min[kVA] and the min[kW] is required (i.e. min[Power Factor] = min[kW] / min[kVA]).

The initial argument about the power factor is that it should be neglected. There is a bound on the power factor i.e. values have to fall within the range 0 to 1. Practically the bound is for e.g. in the range 0.85 to 0.97. Most customers would probably fall within this range. What this implies is that there will be very little difference between the customers' power factor. So in a cluster analysis the customers will be so close to each other that it will be difficult to determine cluster groups. Consultation with Gaunt [Ref. (Personal Communication)] suggests that the range of power factors is probably correct but within this range there might be clusters or groups. There might be 2 main clusters within this range and they are customers that implement power factor correction and those customers that do not implement power factor correction. Customers with power factor correction would for example like to keep their power factor close to the 0.97 area, this is definitely a cluster possibility, and those without power factor correction might be in the 0.85 to 0.9 area. This is also a cluster possibility. The power factor load parameter will be considered for the cluster analysis. The load parameter is the Power Factor[ave, max]. This is the average and maximum power factor. After investigating the Power Factor[ave, max] of several customers there is a very small difference between the average and the maximum power factor. So for the cluster analysis only the average power factor is considered because in cluster analysis it is of little benefit to have 2 attributes of an object (2 load parameters of a customer) that are so close to each other.

The author initially considered neglecting the load factor because although the bound of the load factor is 0 to 1, practically the author thought that the range would be somewhere in the range of 0.5 to 0.9. What this means is that the customers' load factor values are so close to each other that it would make clustering very difficult but what Gaunt [Ref. (Personal Communication)] suggests is that the range of load factors is wider than the range of 0.5 to 0.9. Gaunt [Ref. (Personal Communication)] has measured load factor values as low as 0.2 for industrial customers and therefore the range is wider. Investigating the load factor values of several customers shows that the load factors vary from as low as 0.244 to as high as 0.731. The load parameter, load factor, is therefore considered in the cluster analysis. The load factor is dependent on the Maximum Active Power (kW) and the total energy (kWh) used as presented in equation 4.2, Glover and Sarma (1994).

$$\text{Load Factor} = \frac{Total\ Energy\ Used}{(Max\ Active\ Power)*(Time)} \quad \ldots\ 4.2$$

Where Total Energy Used in (kWh)

Max. Active Power in (kW)

Time in (hours)

The initial load parameters in table 4.1 are reduced to table 4.2.

Table 4.2: Final Load Parameters for the cluster analysis

Load Parameter	**Units**
Max Apparent Power	kVA
Max Active Power	kW
Max Reactive Power	kVAr
Ave Apparent Power	kVA
Ave Active Power	kW
Ave Reactive Power	kVAr
Ave Power Factor	---
Load Factor	---

4.3.2 The Cluster Analysis Process Flowchart

This section presents the flowchart for the cluster analysis process. The data matrix in the cluster analysis has the objects and the attributes of the objects, Romesburg (1984). In this case the objects are the customers and the attributes of the customers are the customers load parameters as shown in the flowchart in figure 4.6.

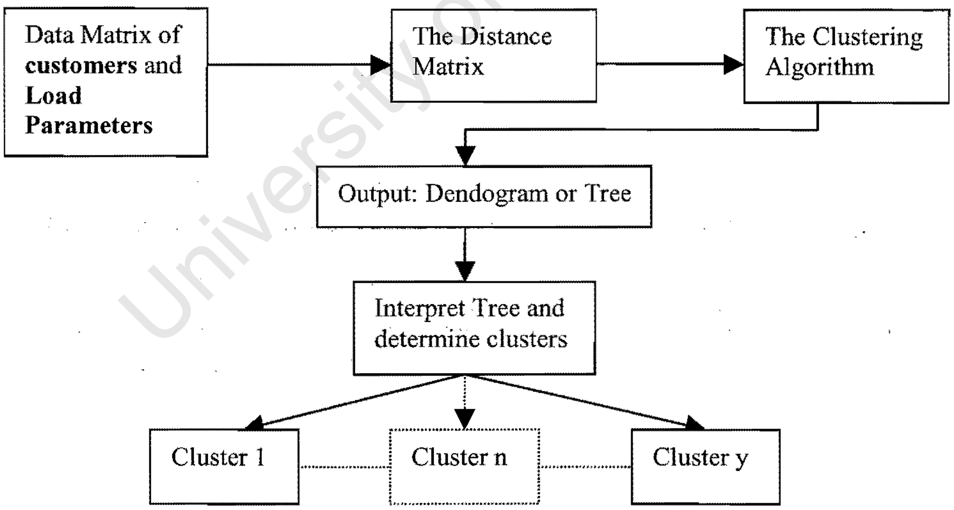


Figure 4.6 Flowchart of cluster analysis process

The tree, which is the output of the cluster analysis, has to be interpreted and can produce any number of clusters indicated by cluster y. Each cluster can consist of any number of customers.

4.3.3 The Load Forecast Process of a Cluster

There will be a load forecast associated with each cluster. The customers in the cluster will determine the load forecast of the cluster. The load forecast of any customer is dependent on the load forecasts of the customers in a cluster. A flowchart of the load forecast process for a cluster is presented in figure 4.7.

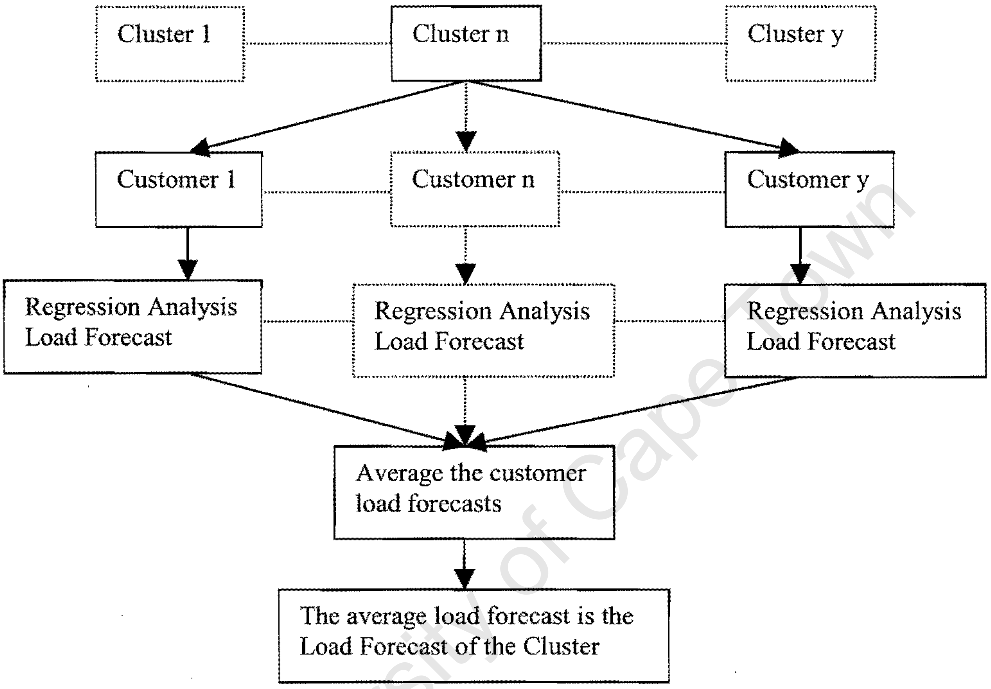


Figure 4.7 Load Forecast process of a cluster

As indicated in figure 4.7, the load forecast of a cluster is represented by the average load forecast of all the customers in the cluster. The reason why the average load forecast is calculated is because customers in a cluster, as calculated by cluster analysis, are expected to be similar and therefore the average would be a good representation. The aim of performing a cluster analysis is to cluster customers together that are similar in respect of their load parameters.

Linear regression analysis is used to forecast the future load using the load data that precedes the forecasting period. The linear regression analysis is the load model. Linear regression analysis is used because it is a simple approach to load forecasting. The method used to forecast the load is not important since the same load forecast approach is used throughout the analysis. The reason why non-linear regression analysis and ANN technology is not implemented is because it would be much more difficult to implement. If a polynomial regression analysis is used then various degree polynomials (i.e. 2nd, 3rd, 4th etc.) must be tested for each customer to determine which degree polynomial produces the most accurate load forecast. For example a 2nd degree polynomial could produce the most accurate load forecast for one customer whereas a 4th degree polynomial could produce the most accurate load forecast for another customer. If ANN technology is used then an ANN must be designed for each customer before a load forecast is produced.

The reason why load data is only used in the load forecast is so that the cluster model can be applied in any area (see section 4.3.4). Introducing weather-sensitive load forecasting will be very limiting to the cluster analysis load-forecasting model. Consider using customers that are in Cape Town to build the cluster model and then using weather-sensitive load forecasting as a means to perform load forecasts. Since the customers are in Cape Town, Cape Town's weather data must be used. One cannot take a customer from Durban, whose load is to be forecasted, and input this into the cluster model that used weather-sensitive load forecasting using Cape Town's weather and customers. The weather in Durban is different to the weather in Cape Town. This cluster analysis load-forecasting model would thus only be able to be used in Cape Town from the load forecasting point of view. This is limiting and therefore the thesis concentrates on load data in the load forecasting process.

4.3.3.1 Long-term Load Forecast for a Cluster

For the limited data set little or no load growth can be assumed and for a long-term load forecast long-term load growth is expected. Consider the long-term load forecast of a cluster: If the annual winter peak load is to be forecasted for one customer in the cluster for the years 2000, 2001 and 2002 then typically 5 years preceding the forecasted period (i.e. 2000, 2001, 2002) is considered for long-term load forecasting. The 5 years would be 1999, 1998, 1997, 1996 and 1995. The winter peak load of one customer is determined for each of the 5 years using the winter load data in each year. These winter peak loads for one customer are projected forward to forecast the future winter peak load for the years 2000, 2001 and 2002. This is shown in figure 4.8.

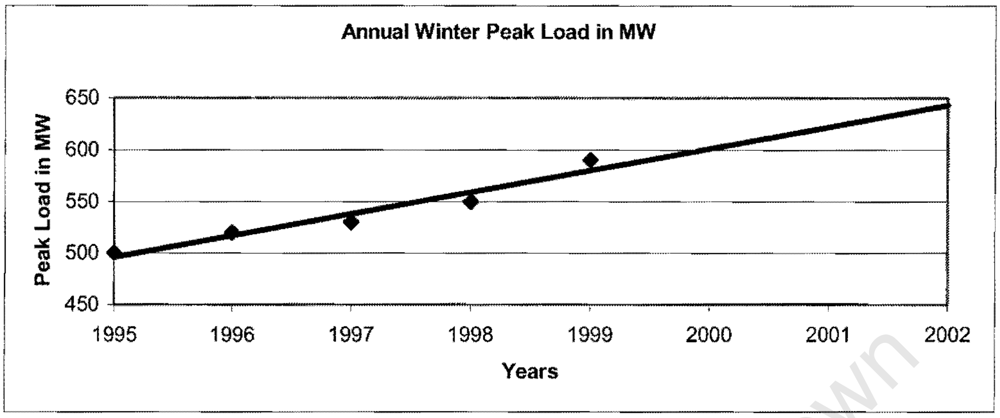


Figure 4.8 Annual Winter Peak Load Forecast for years 2000 to 2002

This process is performed for each customer in the cluster. If there are 30 customers in a cluster then there will be 30 winter peak loads forecasted for each of the years 2000, 2001 and 2002. The average load forecast is determined for each of the years 2000, 2001 and 2002. The winter peak load forecast for the year 2000 is the mean or average of the 30 winter peak loads. The same applies to the years 2001 and 2002.

If the load is to be forecasted for any customer then the customer must be identified to a cluster (see section 4.3.4) using the load parameters of that customer. The load parameters for the customer are calculated over the 5-year period i.e. 1999, 1998, 1997, 1996 and 1995. The load parameters of the customers in the cluster are also calculated over the 5-year period.

4.3.4 Process of Identification of a Customer to a Cluster

This section describes how the load is forecasted for a customer. The customer must first be identified to a cluster. The load parameters of a customer are the only things that are needed to identify a customer to a cluster and once the customer is identified to a cluster the load forecast for that customer is the average load forecast of the customers in the cluster. Consider the flowchart in figure 4.9 that presents the process of identification of a customer to a cluster.

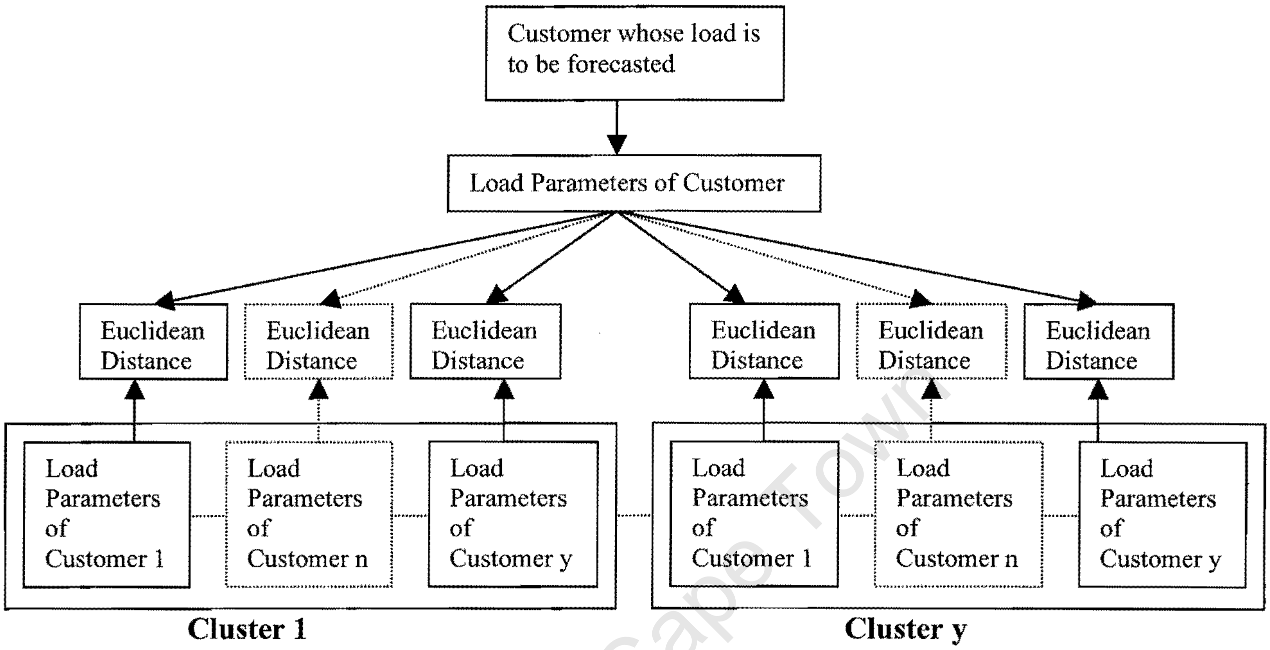


Figure 4.9 Flowchart of Customer Identification

The customer, whose load is to be forecasted, is identified to a cluster by using the customer's load parameters to calculate the Euclidean distance between the customer and each customer in each cluster. To identify a customer to a cluster the minimum Euclidean distance is determined. The customer in the cluster that helped make the minimum Euclidean distance is where the customer is identified to. The minimum Euclidean distance refers to the shortest distance between the customer, whose load is to be forecasted, and a customer in a cluster. This process is discussed in Romesburg (1984) and is described as a process to identify an object to a cluster.

4.4 The Use of the Limited Customer Load Data Set for Classification System

Only 14 customers are available and the load data is complete. These are industrial and commercial customers. Industrial and commercial customers provide a better variety of customers than residential customers do.

Ideally the more customers that are used in the cluster analysis, a statistical technique, the better the conclusions that can be drawn from the analysis in respect of the size and prominence of the clusters with respect to the other clusters. The more customers in a cluster the better the load forecast of that cluster since the load forecast of the cluster is represented by the average load forecast of all the customers in the cluster. The confidence in an average value is best represented by taking the average value of many values than by taking the average value of only a few values.

The load data for each of the 14 customers is limited. The author has 18 weeks of load data available for each customer. Long-term load forecasts require years of load data to forecast load in the long-term as discussed in section 4.3.3.1. Years of data produce several years of load forecasts whereas a few weeks of data produce several weeks of load forecasts. The difference is the forecast period. The basic load forecasting methodology as explained in section 4.3.3.1 is a non-iterative process and this remains the same whether years or weeks of data are used. For a long-term load forecast long-term load growth is expected whereas in a limited data set little or no load growth can be assumed.

Chapter 5: The Implementation of the Linear Regression Load Forecasting Model for the Limited Data Set

5 Content

This chapter discusses the linear regression long-term load-forecasting weather sensitive model of Clayton et al (1973) and Davey et al (1973) and the implementation of this model for a limited load data set.

5.1 The Linear Regression Long-term Load Forecasting Methodology

This section presents a flowchart of the methodology of Clayton et al (1973) and Davey et al (1973) and briefly discusses the process.

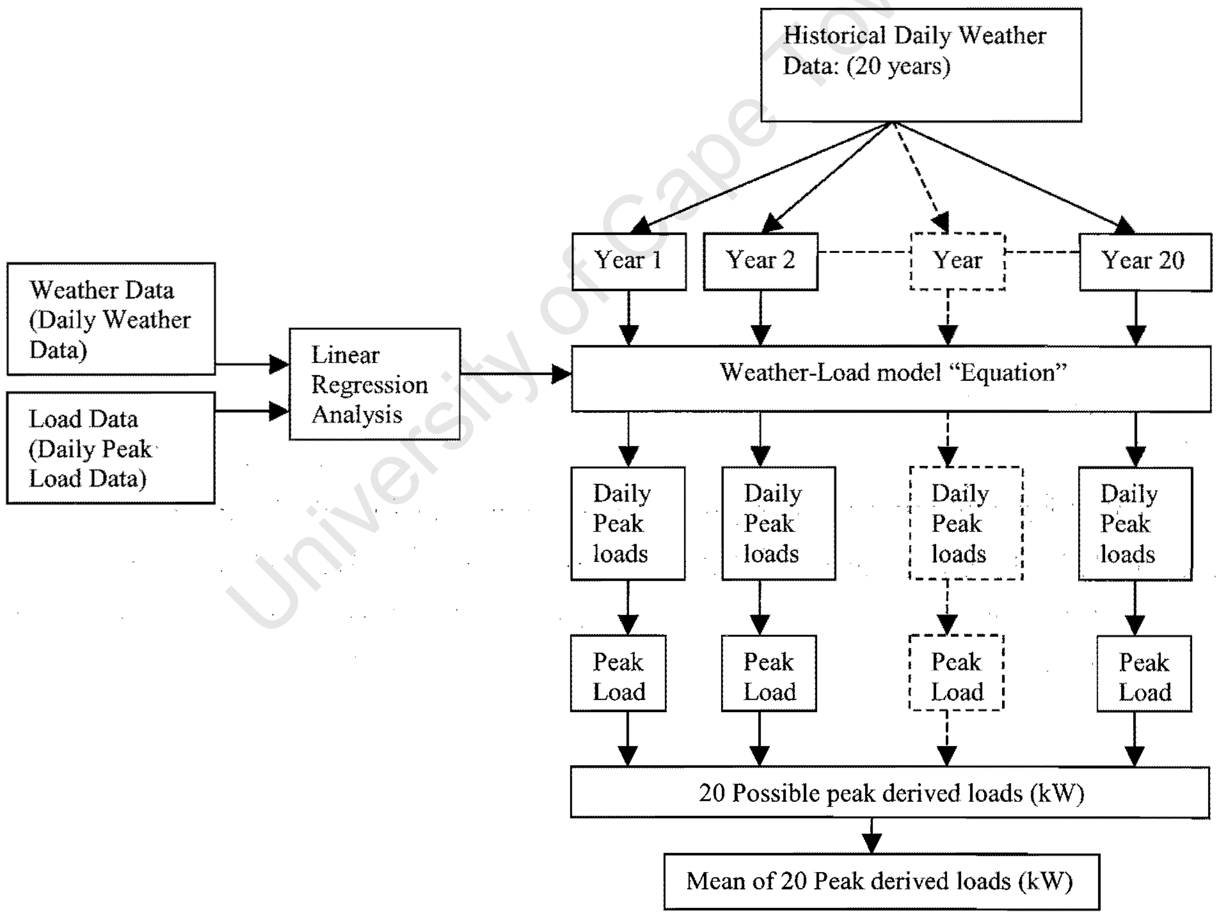


Figure 5.1 Flowchart of Long-term methodology

The model in figure 5.1 is used to forecast future annual peak load. The model is applied separately to the winter and summer season because the annual peak loads usually occurs within one of these seasons, Heinemann et al (1966). The weather data used in figure 5.1 is therefore either winter weather data (i.e. June, July and August months) or summer weather data (i.e. December, January and February). The model is applied separately to each season because the weather data within each season is homogenous.

5.2 The Development of the Linear Regression Model for a Limited Data Set

This section discusses a slightly different approach to the long-term model for the limited data set. A problem is identified to why the limited data set cannot be directly applied to the long-term model. A flowchart of the new process is presented with the changes.

For such a limited data set, little or no load growth can be assumed. The advantage of this is that a linear load growth relationship can be assumed and in this respect it would not be difficult for the linear regression analysis to determine the relationship. The limited data set is from the period 21 February to the 23 June 2000. The data from the period 21 February to 28 April 2000 is used to forecast the load for the period 1 May to 23 June 2000. The period 21 February to 23 June 2000 mainly stretches crosses two seasons i.e. autumn and winter. The weather conditions are not homogenous within the two seasons. The model in figure 5.1 does not perform load forecasts across different seasons. The model concentrates on one season. The model in figure 5.1 has to be adjusted for load forecasts across two different seasons. The new approach to figure 5.1is presented in figure 5.2.

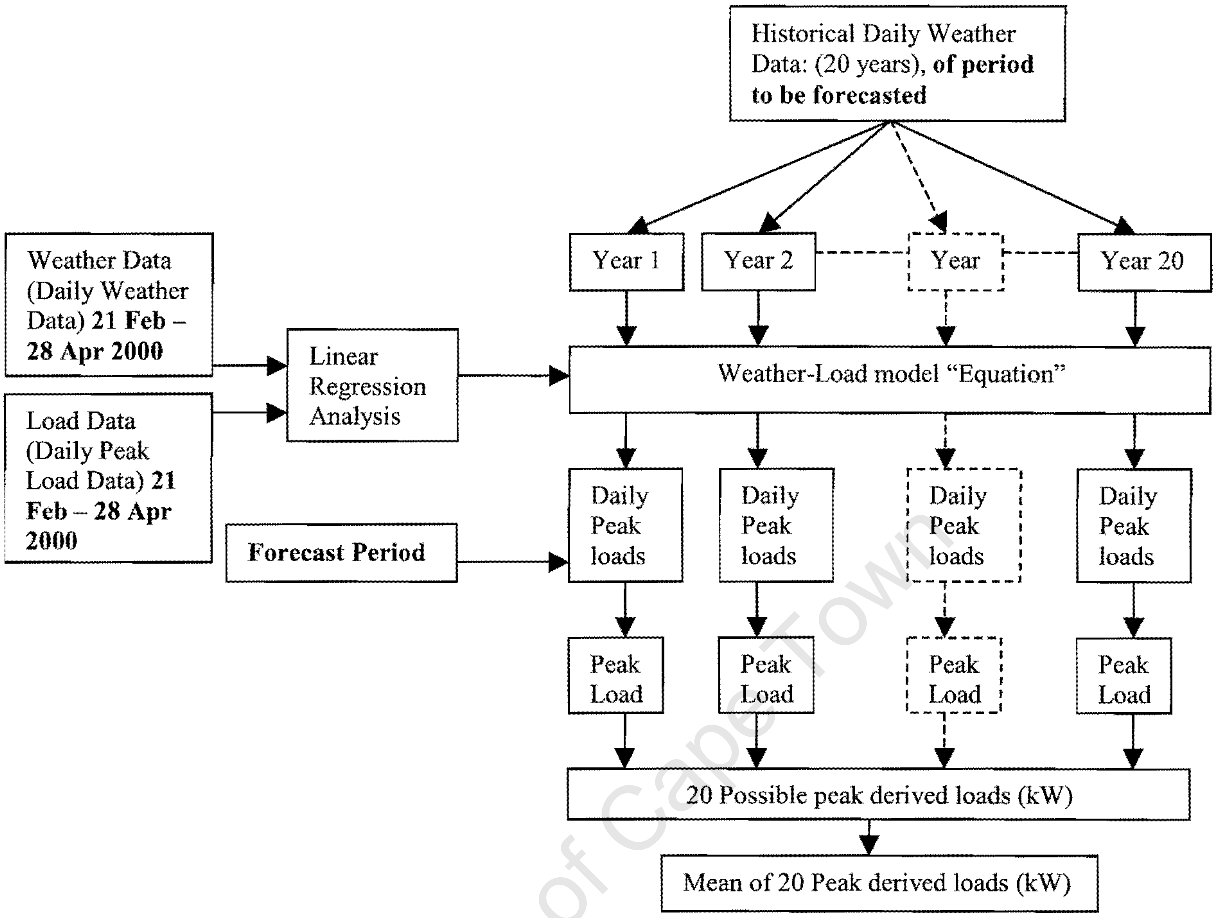


Figure 5.2 Flowchart of model for limited data set

The main change in the model in figure 5.2 is how the historical daily weather data is used. The historical daily weather data corresponds to the period (i.e. 1 May to 23 June 2000) for which the load is to be forecasted. This change was suggested by Gaunt [Ref. (Personal Communication)]. Only one weather-load model is developed as shown in figure 5.2. The weather-load model is developed based on the daily weather and corresponding daily peak load data preceding the forecasting period. This weather-load model is used for the period(s) for which the load is to be forecasted. This assumes that there is little load growth for a limited data set by using the weather-load model in the forecast period(s). The reason for using only one weather-load model is so that there is enough data to develop the relationship between the daily weather and daily peak load data. The "Mean of 20 Peak derived loads (kW)" is the peak load forecast for the period concerned. This is different to model in figure 5.1, which projects the means into the future for the annual peak load forecast (e.g. figure 4.3).

The period for which the peak load is to be forecasted is 1 May to 23 June 2000 and there will be one percentage error (%) associated with this peak load forecast. In order to observe and study the performance or accuracy of the load forecast over the whole forecasting period the forecasting period is broken up into several periods. The peak load forecast is therefore determined for each of the periods and a percentage error (%) is determined for each period. The forecasting period is an 8-week period. The 8-week period will be broken up into 4 fortnights (i.e. 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000). It could also have been broken up into weekly periods however the "Peak Load" would have been calculated from only "5 Daily Peak Loads" whereas for a fortnight the "Peak Load" is calculated from "10 Daily Peak Loads". Only weekdays are considered for forecasting in this thesis. Weekend days can be forecasted in the same manner but for simplicity and understanding weekdays were only considered. The load forecast for weekend days is a separate component of a load forecast because the weather load relationship is different to that of the weather load relationship of the weekdays. The effects of public holidays on the load forecast have not been considered in order to keep the forecast model simple and thus easy to understand.

5.3 The Load Forecast Results of the Linear Regression Model

This section presents the load-forecast results of the Linear Regression model. It also determines the accuracy of the Linear Regression model by comparing it to the actual or real load values.

The objective is to forecast the peak load for the fortnights 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000. This is based on load and weather data for the period 21 February 2000 – 28 April 2000. The weather data used is the daily temperature data (°C). To be specific, it is the dry-bulb temperature at 14h00pm (2pm) obtained from the South African Weather Bureau for the area of Cape Town. Eskom in Cape Town made the corresponding customer daily peak load data available. The daily peak load and corresponding daily weather data for the period 21 February 2000 – 28 April 2000 that is used to develop the weather-load model is presented in figure 5.3.

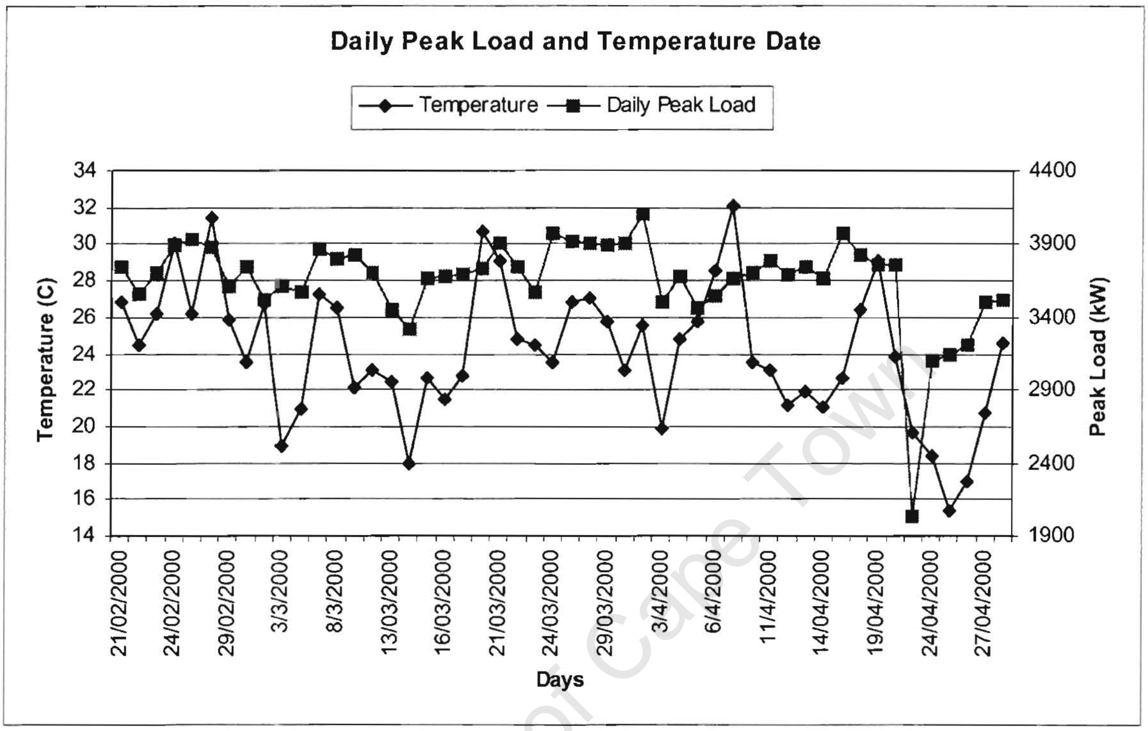


Figure 5.3 Graph of Load and Temperature data for 21 February 2000 – 28 April 2000

The 20 years of historical daily weather data dates from 1980 to 1999. This is daily temperature data at 14h00pm (2pm). Sample weather data of the year 1980 corresponding to the first fortnight (i.e. 1-12 May) is presented in figure 5.4. The corresponding calculated daily peak loads for 1-12 May 2000 are also presented in figure 5.4. The daily peak loads are calculated by using the 1980 weather data as input to the weather-load model.

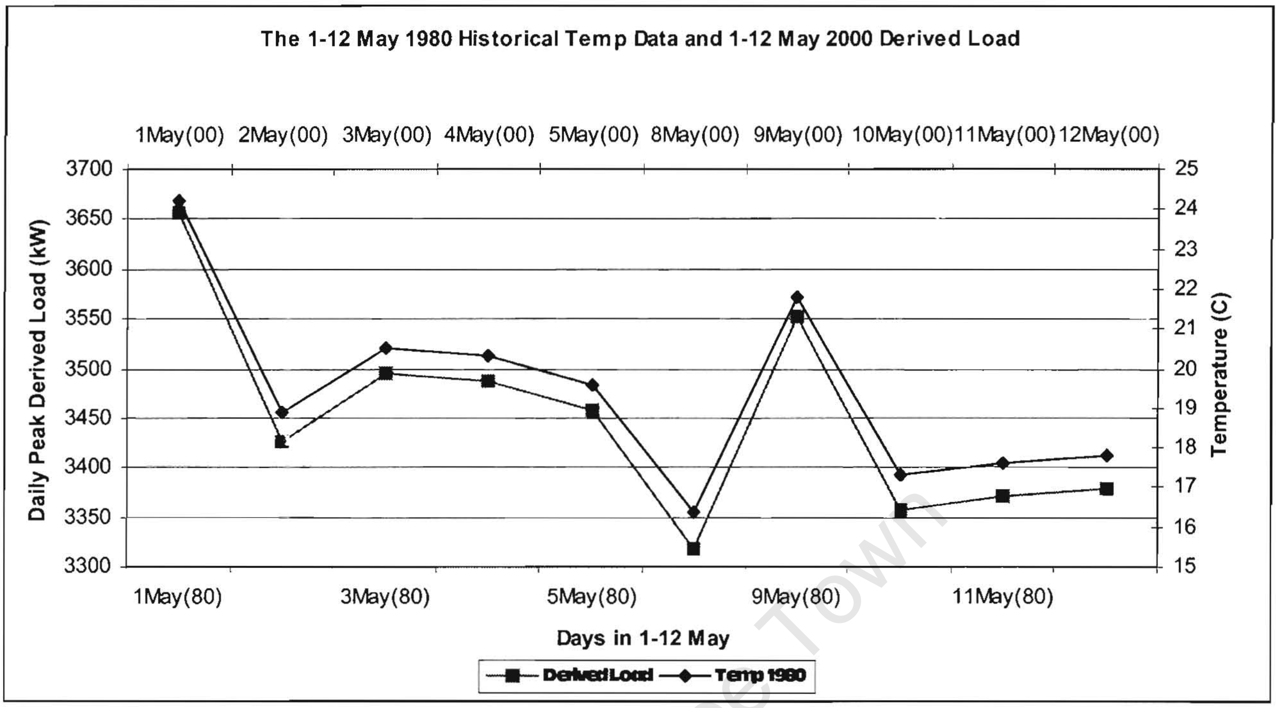


Figure 5.4 Graph of 1980 Historical Temperature Data and Daily Peak Load

For each of the 20 years of historical daily temperature data as input to the weather-load model, a peak load is determined. The peak load in figure 5.4 is 3656 kW. For the fortnight 1-12 May 2000 the 20 peak loads are presented in figure 5.5.

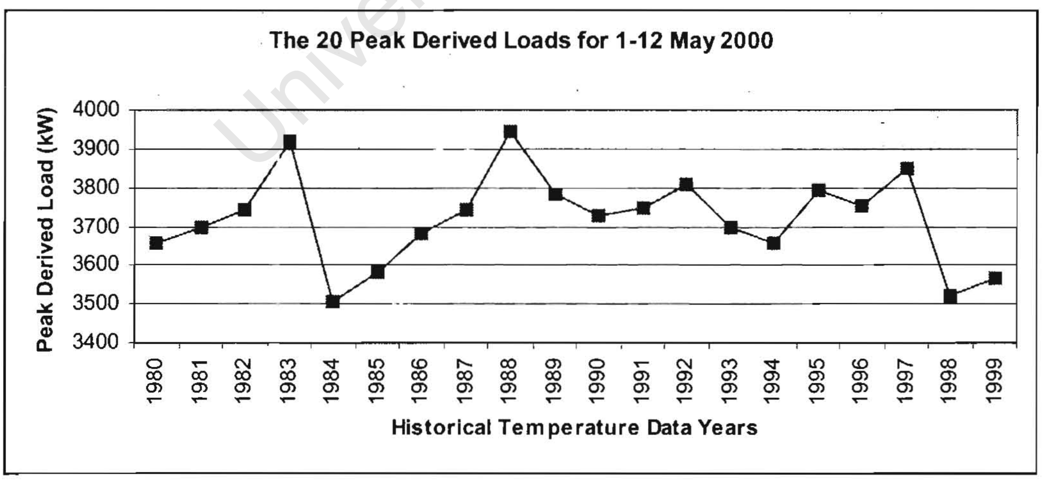


Figure 5.5 Graph of the 20 Peak Derived Loads for the fortnight 1-12 May 2000

This process is performed for each of the 4 fortnights. The average of the 20 peak derived loads is the peak load forecast for the fortnight. The peak load forecast (for 1-12 May 2000) is 3719 kW.

The forecasted peak loads for each of the 4 fortnights is presented in figure 5.6 and compared to the actual or real peak loads for the same fortnights.

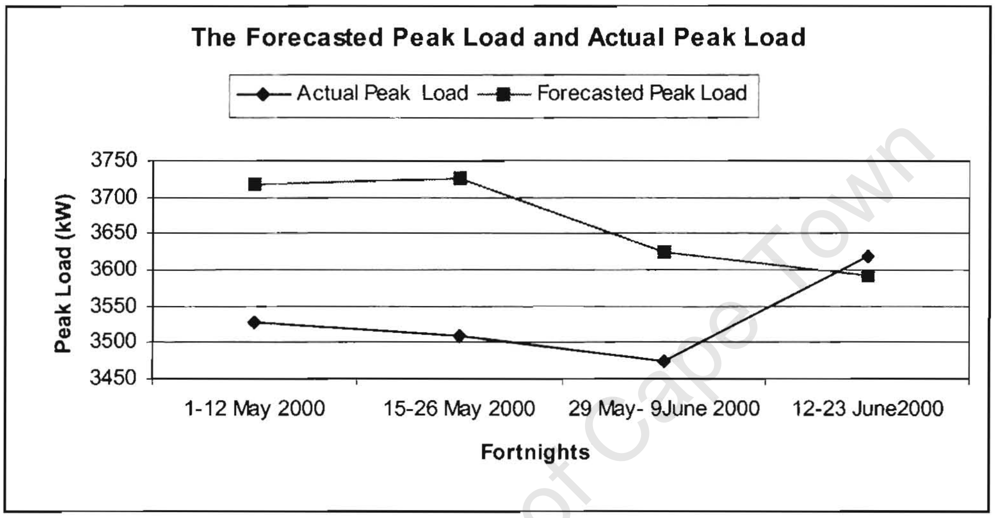


Figure 5.6 Graph of Forecasted Peak Load and Actual Peak Load

The error of the load forecast is expressed in terms of the absolute percentage error (%) and is calculated using the forecasted load (kW) and the actual load (kW) of figure 5.6 in equation 3.1. The error results are presented in table 5.1.

Table 5.1: The load forecasting results of the Linear Regression Model

Fortnight to be forecasted	**APE (%)**
1-12 May	5.41
15-26 May	6.16
29 May-9 June	4.36
12-23 June	0.74
MAPE (%)	**4.17**

The mean absolute percentage error (%) (MAPE) is calculated using the APE in equation 3.2 and the MAPE represents the overall performance of the load forecast.

5.4 Linear Regression Model implemented in Excel Spreadsheet

The file is on the disk and is called Regression_Based_Model.xls. The first sheet named "Data for Regression" presents the data that is used to perform the regression analysis. This data is presented in figure 5.3. The second sheet is named "Regression Analysis" and is the output from the regression analysis. The regression coefficients are important in this sheet. In this case it is the intercept and the DT14pm. The DT14pm is the dry-bulb temperature at 14h00pm (or 2pm). The third, fourth, fifth and sixth sheet in the spreadsheet are named 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000 respectively. The peak load is forecasted for each of these fortnights. The 20 peak derived loads for the fortnights 15-26 May 2000, 29 May-9June 2000 and 12-23 June 2000 are presented in figure 5.7.

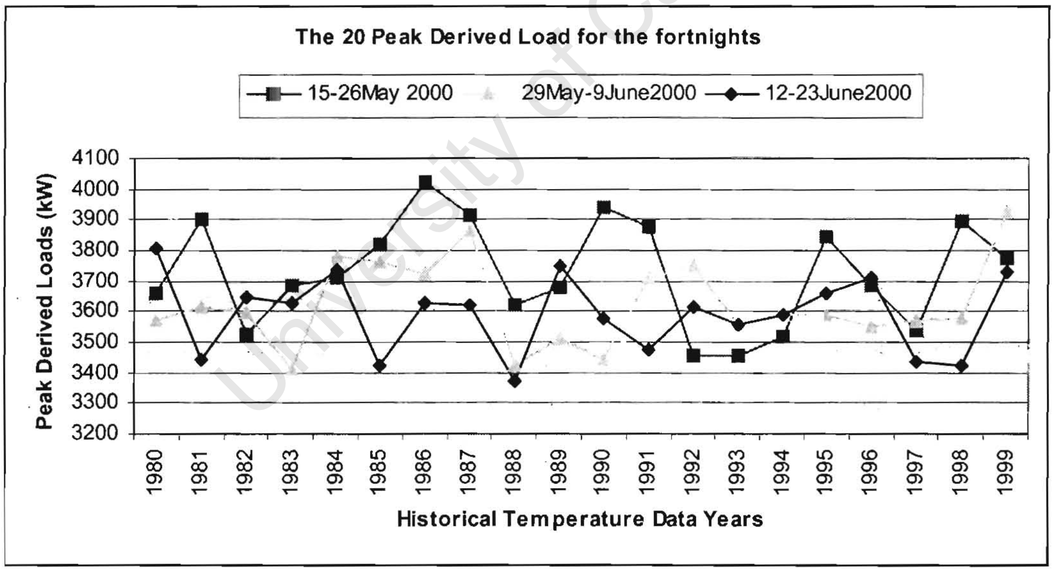


Figure 5.7 The 20 Peak Derived Loads for the three fortnights

The load-forecast error is also determined for each fortnight in the form of the Mean Absolute Percentage Error (MAPE) as presented in table 5.1.

Chapter 6: The Implementation of the ANN Load Forecasting Model for the Limited Data Set

6 Content

This chapter discusses how an ANN model is developed for the limited data set. The methodology of the ANN model is presented and designed here.

6.1 The Development of the ANN Model for a Limited Data Set

In section 4.1 the use of ANN technology is identified for a long-term load-forecasting model. This model requires years of data since it is a long-term model and the author only has a limited data set available. This section discusses the ANN model that is developed for the limited data set.

Figure 4.5 in chapter 4 presents the flowchart for the ANN model applied to long-term load forecasting. This model makes use of historical weather data and historical load data. Twenty years of historical weather data is available however historical load data is not available. Historical load data is not needed since it falls outside the scope and objectives of the thesis as discussed in chapter one. Historical load data can be many years of data and therefore is not a limited data set. The ANN model in figure 4.5 will therefore not consider the historical load data as input to the weather-load model. The ANN weather-load model will therefore only consider the load as a function of the weather.

The limited data set is from the period 21 February to the 23 June 2000. The data from the period 21 February to 28 April 2000 is used to forecast the load for the period 1 May to 23 June 2000. The period 21 February to 23 June 2000 mainly stretches crosses two seasons i.e. autumn and winter. The weather conditions are not homogenous within the two seasons. The model in figure 4.5 does not perform load forecasts across different seasons. The model concentrates on one season. The model in figure 4.5 has to be adjusted for load forecasts across two different seasons. The new approach to figure 4.5 is presented in figure 6.1.

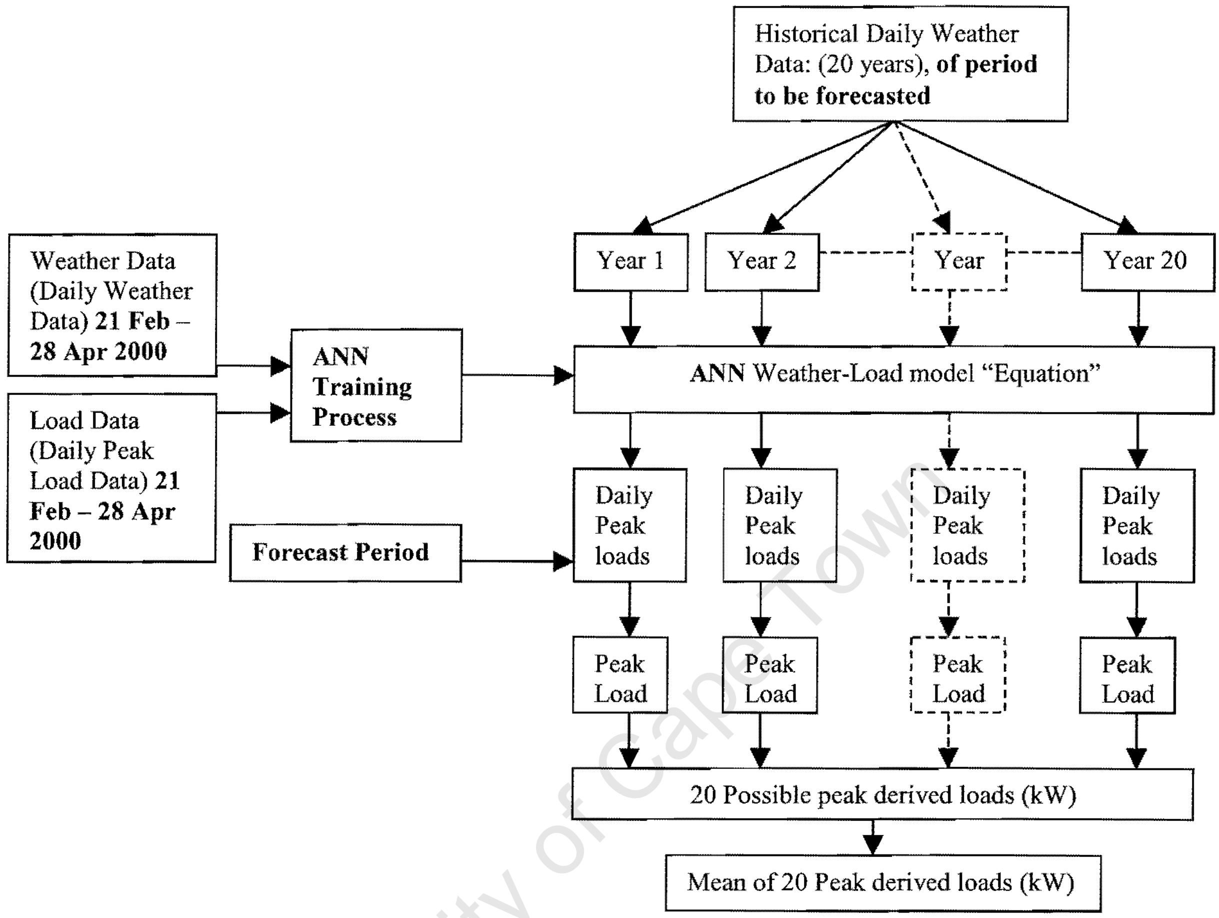


Figure 6.1 Flowchart of ANN model for limited data set

The main change in the model in figure 6.1 is how the historical daily weather data is used. The historical daily weather data corresponds to the period (i.e. 1 May to 23 June 2000) for which the load is to be forecasted. This change was suggested by Gaunt [Ref. (Personal Communication)]. Only one weather-load model is developed as shown in figure 6.1. The weather-load model is developed based on the daily weather and daily peak load data (21 February to 28 April 2000) preceding the forecasting period. This weather-load model is used for the period(s) for which the load is to be forecasted. The reason for using only one weather-load model is so that there is enough data to develop the relationship between the daily weather and daily peak load data. The "Mean of 20 Peak derived loads (kW)" is the peak load forecast for the period concerned.

The forecasting period is an 8-week period. The 8-week period will be broken up into 4 fortnights (i.e. 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000). The peak load will be forecasted for each of the fortnights.

6.2 The Design of the ANN Model

The methodology of the ANN model is presented in the previous section. The next step is to design the ANN model which involves the architecture of the model, the size of the model, the data used for training and testing, the preparation of the data (scaling), the choice of the transfer functions and the algorithm used in the training process.

6.2.1 The Data for the ANN Model

The ANN model has the architecture as shown in figure 6.2 (Chen et al (1992) and Dash et al (1994)) below with the daily temperature data as the input and the daily peak load as the output variable.

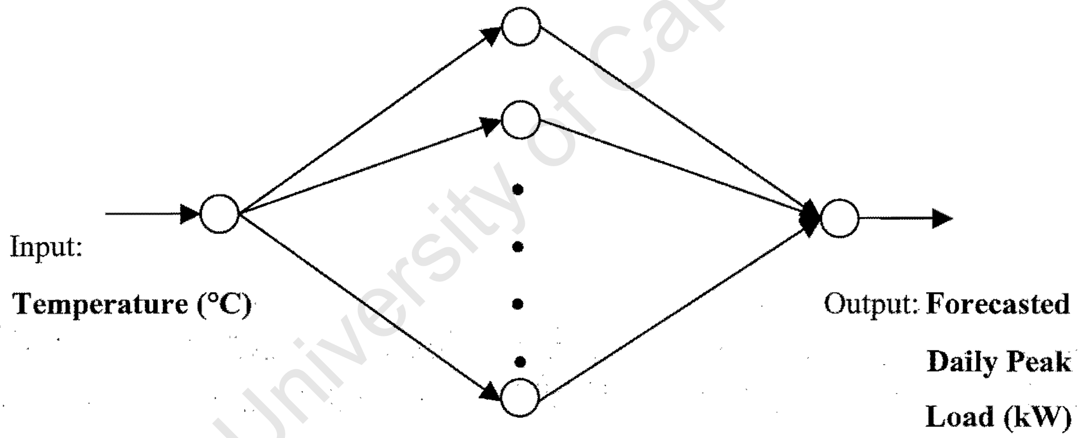


Figure 6.2 The ANN Model

Both the daily temperature data and corresponding daily peak load data for the period 21 February 2000 – 28 April 2000 is used to train the neural network. The neural network will find the relationship between the daily temperature data and the daily peak load data. Once the pattern or relationship is determined for the period 21 February 2000 – 28 April 2000, then the load will be forecasted for each fortnight in the period 1 May 2000 – 23 June 2000 using the historical daily temperature data. The daily temperature data and the corresponding daily peak load data for the period 21 February 2000 – 28 April 2000 is presented in figure 6.3.

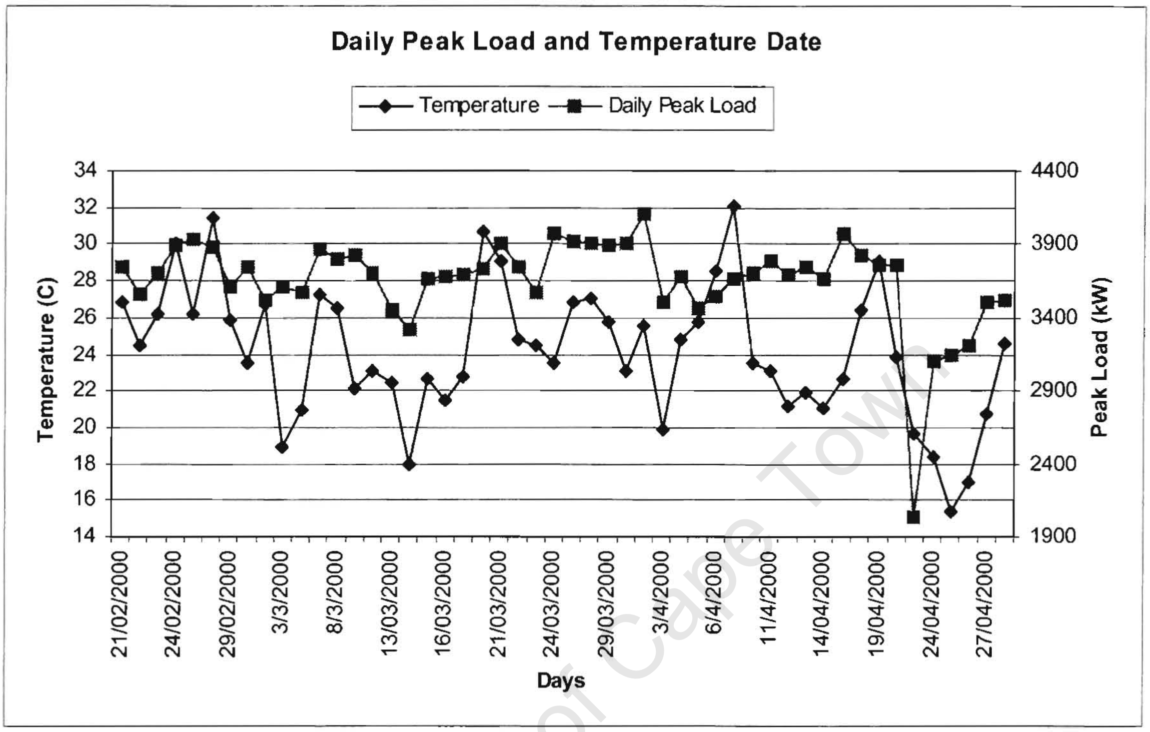


Figure 6.3 Daily Load and Temperature Data for 21 February 2000 – 28 April 2000

6.2.2 Scaling of the Input and Output Data

In the previous section the input and output variables are represented in their real units i.e. temperature in °C and load in kW. The input and output data is first scaled before the training process begins, as recommended by most authors (e.g. Dash et al (1994) and Riad et al (1994)). Scaling or normalizing the input and output data of the ANN is important so as to make the neural network training more efficient. The data is normalized so that each value falls within the range –1 to 1. Remember that each neuron has a transfer function such as the tan-sigmoid and log-sigmoid transfer function and these two transfer functions have output limits which are (-1 to 1) and (0 to 1) respectively. Scaling the data is done to prevent the neurons from being driven too far into saturation. Once saturation is reached, changes in the input data results in very little or even no change in the output. This severely limits the neural network to gain or understand the relationship or pattern between the input data and the output data as discussed in Demuth and Beale (2000). After the ANN is trained with the normalized data, the weights and biases of the ANN, which is set after training, is also normalized.

Once the ANN is trained with the normalized data, it is ready for prediction. This means that new data can be presented as input to the ANN for prediction. The new data has to be normalized and the resultant output is also normalized. Of course to interpret the output correctly in real units, the output has to be unnormalized as described in Demuth and Beale (2000). Matlab has preprocessing and postprocessing functions, which help to, normalize and unnormalize the data.

6.2.3 Choice of Training Algorithm for the Neural Network

There are many training algorithms available to train the ANN. In this case because of the limited data set being used for training the ANN the Bayesian regularization training algorithm will be used. This algorithm is implemented in The Matlab Neural Network Toolbox and it refers to this training algorithm as trainbr. This choice of training algorithm is further explained below.

One of the problems that occur during neural network training is called overfitting. Instead of overfitting the designer would like the ANN to generalize. Overfitting means that the error is very small on the training data but when new data is used as input to the ANN then the output error is large, Demuth and Beale (2000). This means that the ANN has memorized the training data. One way that overfitting can occur is when the incorrect size of the network is used for the training data. The idea is to have an ANN just large enough so as to provide an adequate fit and thus improve generalization. The problem is to find the correct size network. If the network is large the ANN will overfit the training data, but if the network is small the network will not be able to learn the relationship in the training data. Demuth and Beale (2000) suggest that the way to find the correct network size is to either start off with a small size network and gradually increase the size of the network until generalization is achieved or to start off with a large size network and gradually decrease the network size until generalization is achieved.

To optimize generalization with the limited data set available, the Bayesian regularization-training algorithm is chosen. A training set of data is deemed small when the number of parameters (i.e. weights and biases) in the network is close to the total number of training data points. Demuth and Beale (2000) state that if the number of parameters in the network is much smaller than the total number of points in the training set, then there is little or no chance of overfitting. The process of achieving generalization with a small training data set is difficult but the Bayesian regularization-training algorithm has the best chance of ensuring generalization. The advantage of this algorithm is that it shows the designer how many network parameters (i.e. weights and biases) are being effectively used by the network. What this means is that if the ANN is made large enough (i.e. large number of weights and biases) then training the network will give the designer an idea of how many network parameters are effectively being used by the network and the designer can reduce the network size. The designer would want to reduce the network size so as to reduce the training time for convergence. Also an important point concerning this algorithm is that making the network size larger, within reason, the network will still not overfit the training data. This eliminates the guesswork required for determining the optimum network size. "Within reason" means that if the designer sees that after training the network size is large enough then adding on a few neurons will not affect the network's response to overfitting the training data. Adding on say extra hidden layers of neurons, which is a large increase in network size, will affect the network's response.

One way to determine if the Bayesian algorithm has converged is to observe after training whether the network parameters (i.e. # Par) remain approximately constant out of the total number of network parameters over several training epochs. This of course assumes that the designer has trained the network for a sufficient number of iterations (epochs) to ensure convergence. The convergence of the network parameters is one measure of convergence; there are three more ways or indications to determine whether after training the network has converged. They are 1) sum of the squared errors (SSE), 2) sum of the squared weights (SSW) and 3) the training can stop with the message "Maximum MU reached", Demuth and Beale (2000). MU is the Marquardt adjustment parameter, which is associated with the Levenberg-Marquardt algorithm that is built into the Bayesian algorithm to update the weights and biases. The SSE and SSW have to remain approximately constant over several training epochs. "Maximum MU reached" is a good indication that the algorithm has truly converged. Another good indication of convergence is to have all of the following to converge: the number of network parameters (# Par), the SSE and the SSW.

6.2.4 Choice of Transfer Functions in the Neurons

The basic ANN architecture is the same as shown in figure 6.2 with one hidden layer and one output neuron. The input neurons do not actually have transfer functions within them; they are only input data. Only the hidden layer of neurons and the output neuron have transfer functions. Each neuron in the hidden layer has a tan-sigmoid transfer function. The reason for using this transfer function and not a linear transfer function is that the relationship that the ANN is trying to establish is assumed to be non-linear. If linear transfer functions are used in the hidden layer then it will be impossible for the ANN to find a non-linear relationship if there is one, Demuth and Beale (2000). The linear transfer functions can find non-linear relationships but not as well as non-linear transfer functions. Demuth and Beale (2000) present this theory in their manual called Neural Network Toolbox. The output neuron will have a linear transfer function so as to give the ANN the ability to take on any value in the output. The ANN output can take on any value because a non-linear transfer function such as the tan-sigmoid transfer function is bounded between –1 and +1 and the linear transfer function is not bounded.

6.2.5 The Architecture of the ANN Model

The architecture of the ANN model refers to the size of the ANN model that is used. The basic architecture is presented in figure 6.2. This section discusses how many hidden layers and neurons per layer are to be used.

To determine the architecture of the ANN model is a trial and error process. The designer, Bougaardt G, chose to start off with one hidden layer and 10 neurons in this hidden layer. The ANN model is therefore a 1-10-1 network. Both 1's in 1-10-1 refer to the temperature input and the load output respectively. The network is trained with the daily temperature and daily peak load data presented in figure 6.3 which is from the 21 February 2000 – 28 April 2000. The training results for the 1-10-1 network are presented in figure 6.4.

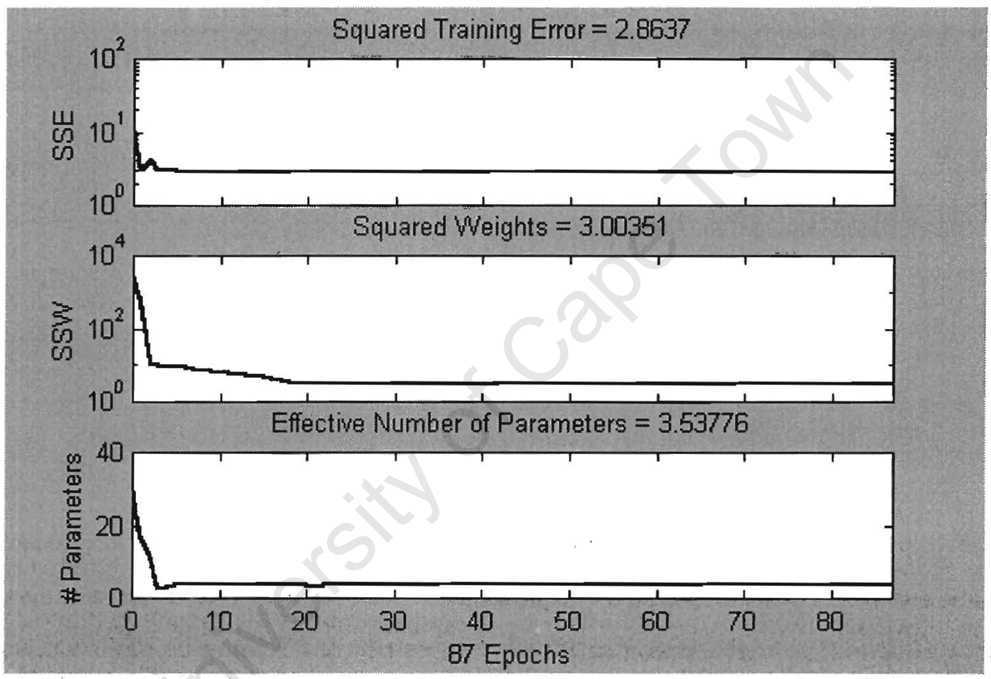


Figure 6.4 Training results for 1-10-1 network

The 1-10-1 network converged with the output "Maximum MU reached" but this does not mean that the network is large enough. To test whether the network is large enough, the neurons in the hidden layer is increased to 30 neurons and the training results will be analysed. The network is a 1-30-1 network. The training results are presented in figure 6.5.

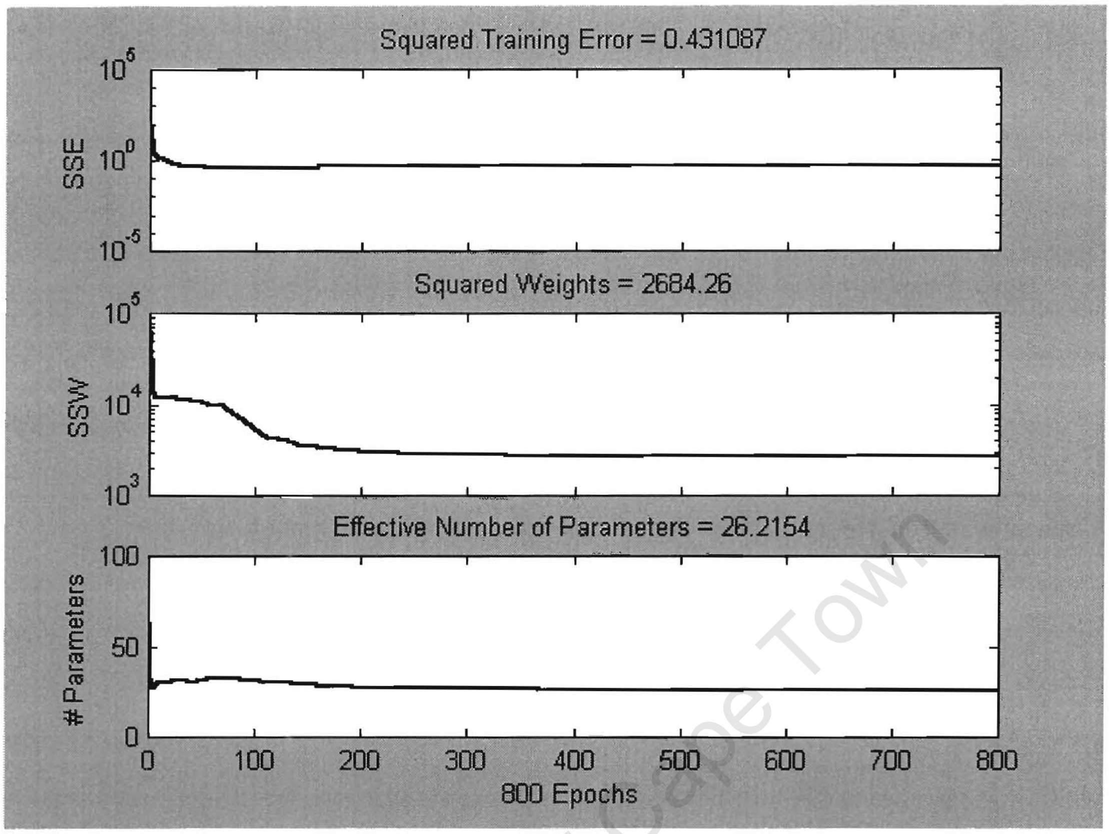


Figure 6.5 Training results for 1-30-1 network

The 1-30-1 network is trained for 800 epochs and did not converge with the output "Maximum MU reached". If a network does not converge with this output, then the number of parameters (i.e. # Par) will be looked at for convergence. It is clear that from figure 6.4 the # Par converged to 3.54 out of 31 parameters and from figure 6.5 the # Par converged to 26.22 out of 91. So clearly the network 1-10-1 was not big enough since there is a big difference in the number of parameters (# Par) between the 1-10-1 network and the 1-30-1 network. The number of parameters (# Par) must remain approximately constant when the network size is increased for the network to be declared large enough. This means that the network size must be increased until the number of parameters (# Par) remains approximately constant. The network is thus increased to a 1-40-1 network and the training results are displayed in figure 6.6.

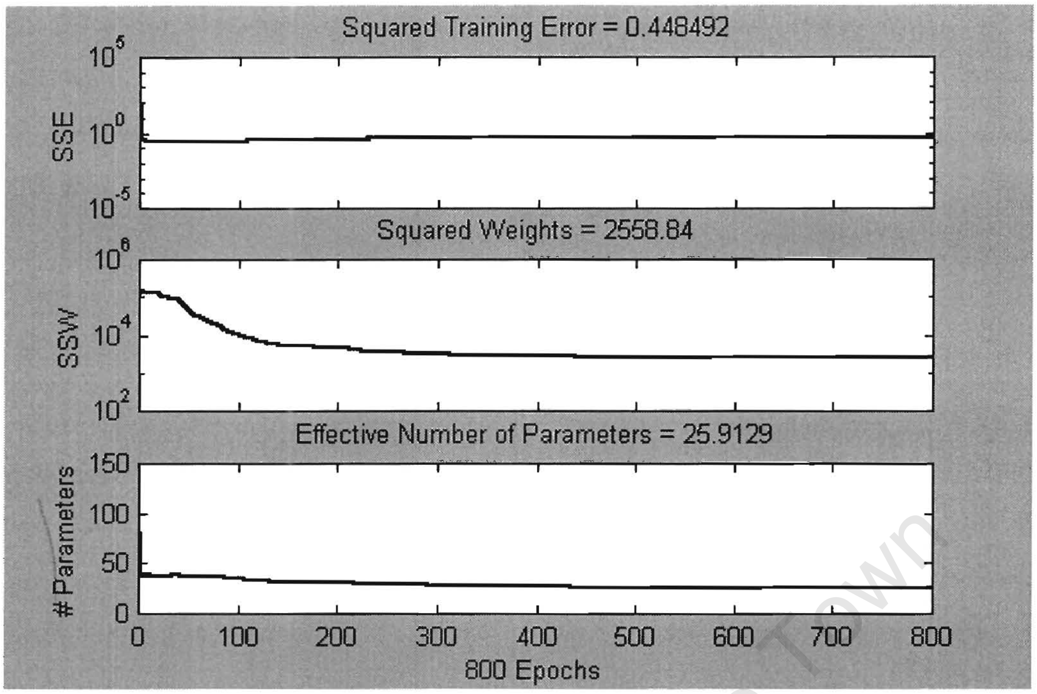


Figure 6.6 Training results for 1-40-1 network

The 1-40-1 network converged to # Par = 25.91. This is relatively constant compared to the # Par = 26.21 for the 1-30-1 network. This suggests that the 1-30-1 network is large enough but is the 1-30-1 network robust. Robust means that the output training results of the 1-30-1 network must be approximately the same over several training runs i.e. the convergence must be approximately the same by randomly changing the weights and biases of the network. This is testing the sensitivity of the convergence by randomly changing the weights and biases of the network. Remember the above ANN does not converge with the message "Maximum MU reached" and thus the output training results will be different on each "training run". So to make sure that the ANN output training results is approximately the same and converges, the network is trained several times and the output training results are studied. This is presented in table 6.1. The network did converge on each "training run".

Table 6.1: Several training runs for network 1-30-1

Number of training runs	#Par
1	2.62e+001/91
2	2.66e+001/91
3	2.65e+001/91
4	2.63e+001/91
5	2.66e+001/91
6	2.65e+001/91
7	2.65e+001/91

The 1-30-1 network is robust since the # Par is approximately the same on each training run. Since the convergence results for the 1-30-1 network are similar over 7 training runs, the load forecasting results over the 7 training runs will be similar though not identical. The reason why the training results are not the same is because the ANN uses an initialization function to initialize the weights and biases of the network so as to start the network training process, Demuth and Beale (2000). The initialization function uses a random number based system to assign values to the weights and biases to start the training process.

The 1-30-1 network will be used. The 1-40-1 network could also have been used. There is no concern for overtraining the network since the Bayesian regularization training algorithm, Demuth and Beale (2000), is used. The advantage of the Bayesian regularization is that overtraining is eliminated no matter how large the network becomes therefore the 1-30-1 or the 1-40-1 network can be used. The number of parameters used is approximately the same for the 1-30-1 and for the 1-40-1 network. This indicates that the network's response will not overfit the data as discussed in Demuth and Beale (2000).

The weights and biases of the 1-30-1 network are fixed after it has been trained. This section tests the accuracy of the 1-30-1 network by using the same daily temperature data (21 February – 28 April 2000) that was used to train the 1-30-1 network as input to the 1-30-1 network. The output, which is the daily peak loads, will be compared to the actual or real daily peak load data that was used to train the 1-30-1 network. Figure 6.7 presents the actual daily peak load and the forecasted daily peak load data for the period 21 February 2000 – 28 April 2000 i.e. 50 days.

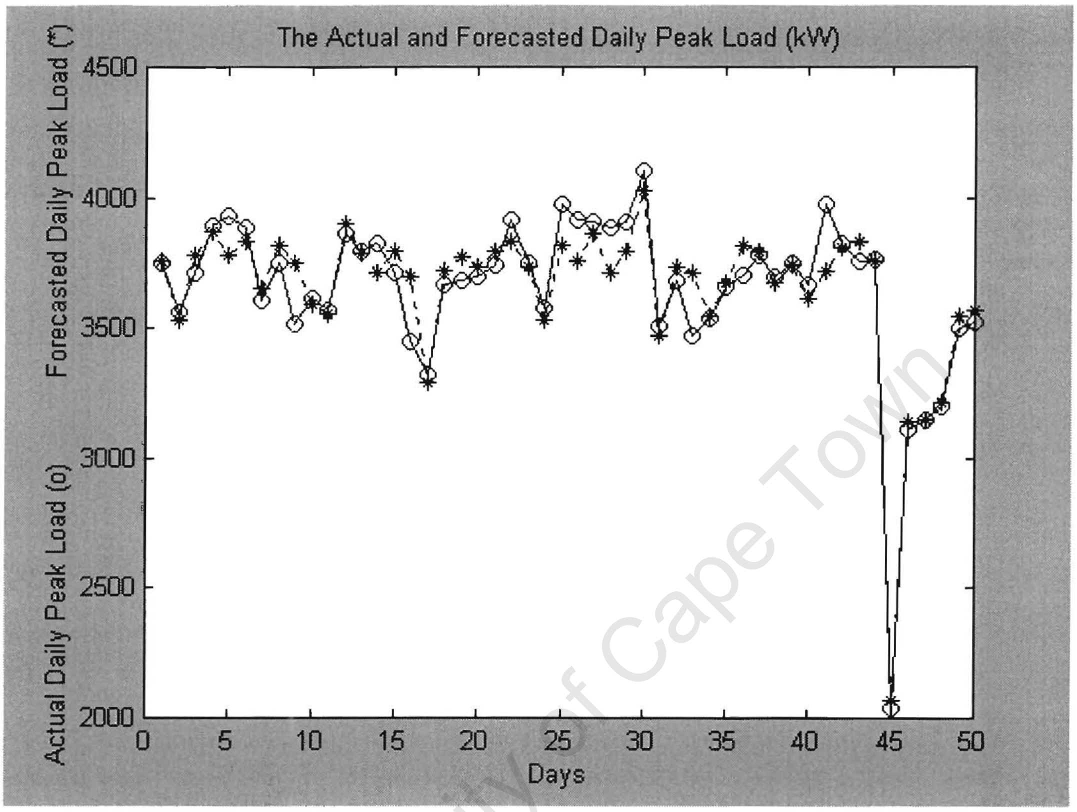


Figure 6.7 Actual and Forecasted Daily Peak Load for 1-30-1 network

The error between the actual and forecasted daily peak load is presented in figure 6.8. The absolute percentage error (APE) is calculated using equation 3.1.

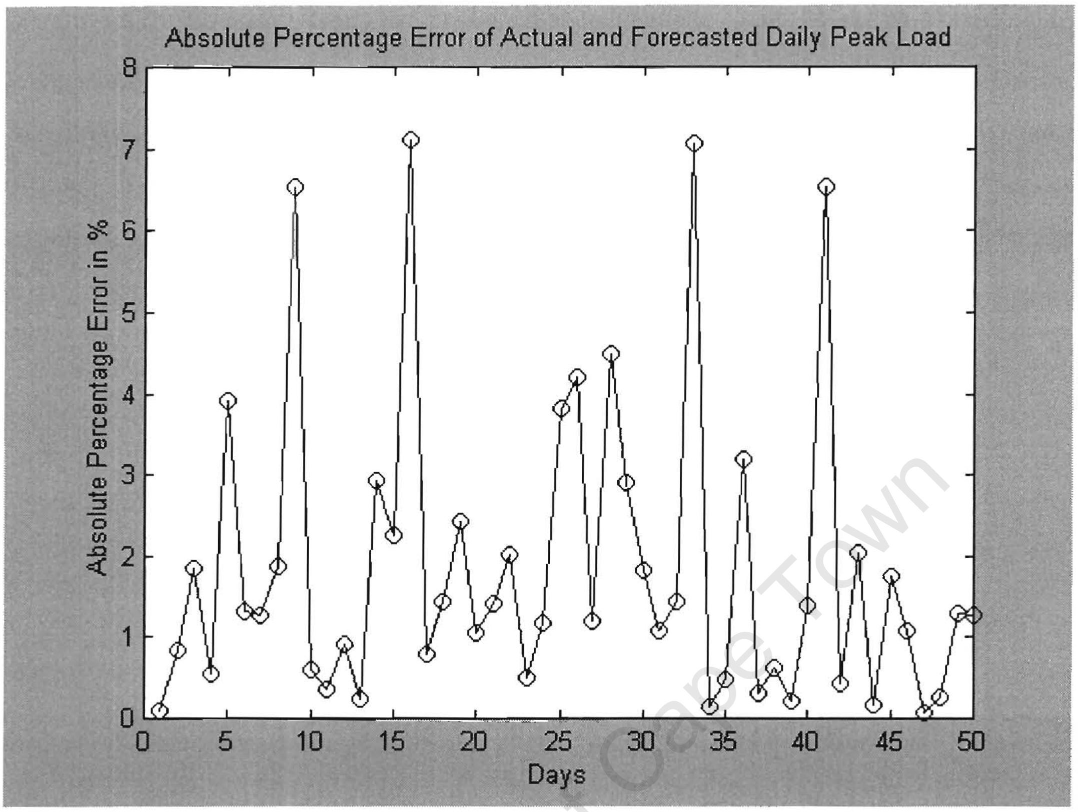


Figure 6.8 APE of Actual and Forecasted Daily Peak Load

The next section uses the 1-30-1 network to forecast the peak load for the 4 fortnights in the period 1 May 2000 – 23 June 2000. Historical temperature data is used as input to the 1-30-1 network.

6.3 The Load Forecast Results of the ANN Model

This section presents the load-forecast results of the ANN model. It also determines the accuracy of the ANN model by comparing it to the actual or real load values. The accuracy is measured in terms of the Mean Absolute Percentage Error in %.

The objective is to forecast the peak load for the fortnights 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000. The 20 years of historical daily weather data dates from 1980 to 1999. This is daily temperature data at 14h00pm (2pm). Sample weather data of the year 1980 corresponding to the first fortnight (i.e. 1-12 May) is presented in figure 6.9. The corresponding calculated daily peak loads for 1-12 May 2000 are also presented in figure 6.9. The daily peak loads are calculated by using the 1980 weather data as input to the weather-load model.

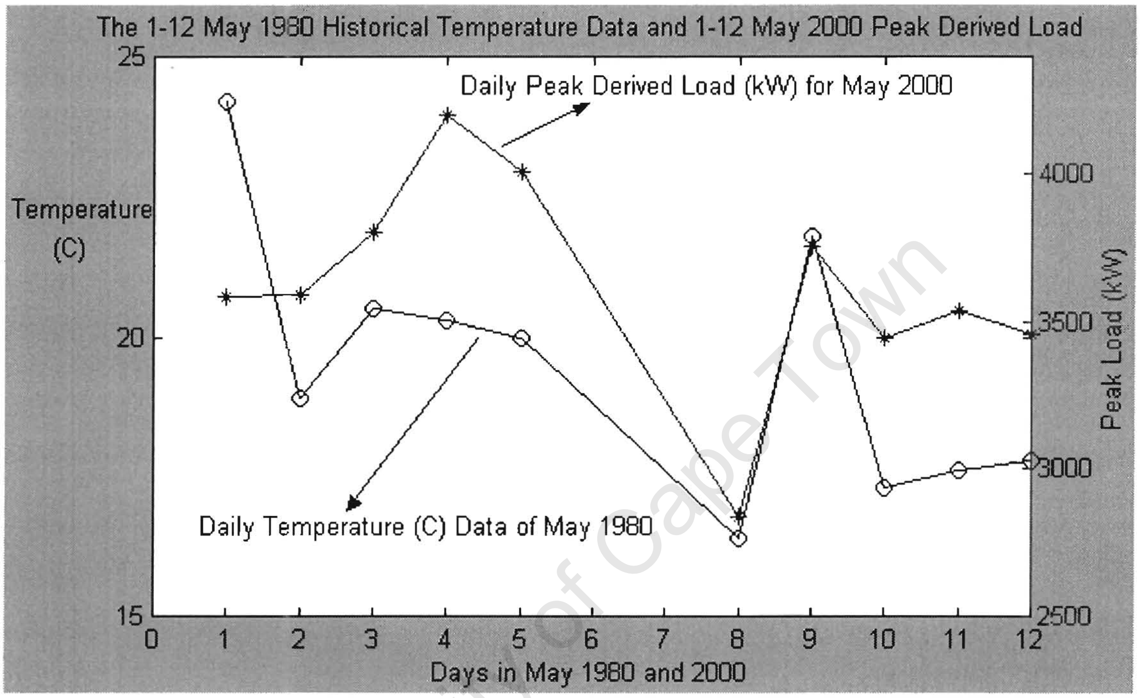


Figure 6.9 Graph of 1980 Historical Temperature Data and Daily Peak Derived Load

For each of the 20 years of historical daily temperature data as input to the weather-load model, a peak load is determined. The peak load in figure 6.9 is 4200 kW. For the fortnight 1-12 May 2000 the 20 peak loads are presented in figure 6.10.

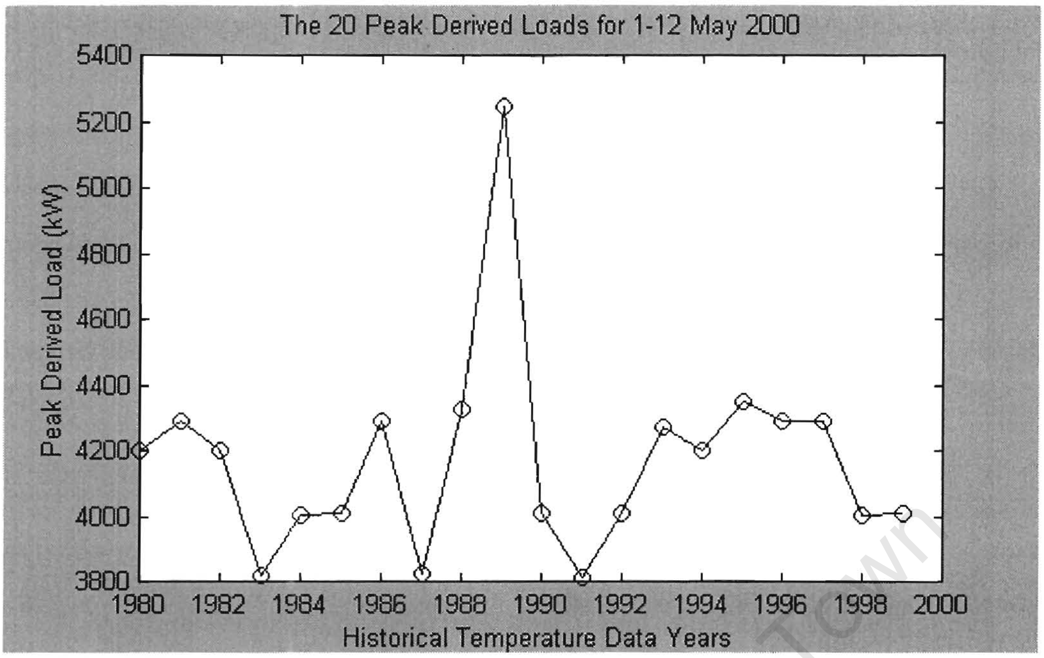


Figure 6.10 Graph of the 20 Peak Derived Loads for the fortnight 1-12 May 2000

This process is performed for each of the 4 fortnights. The average of the 20 peak derived loads is the peak load forecast for the fortnight. The peak load forecast (for 1-12 May 2000) is 4172 kW.

The forecasted peak loads for each of the 4 fortnights is presented in figure 6.11 and compared to the actual or real peak loads for the same fortnights.

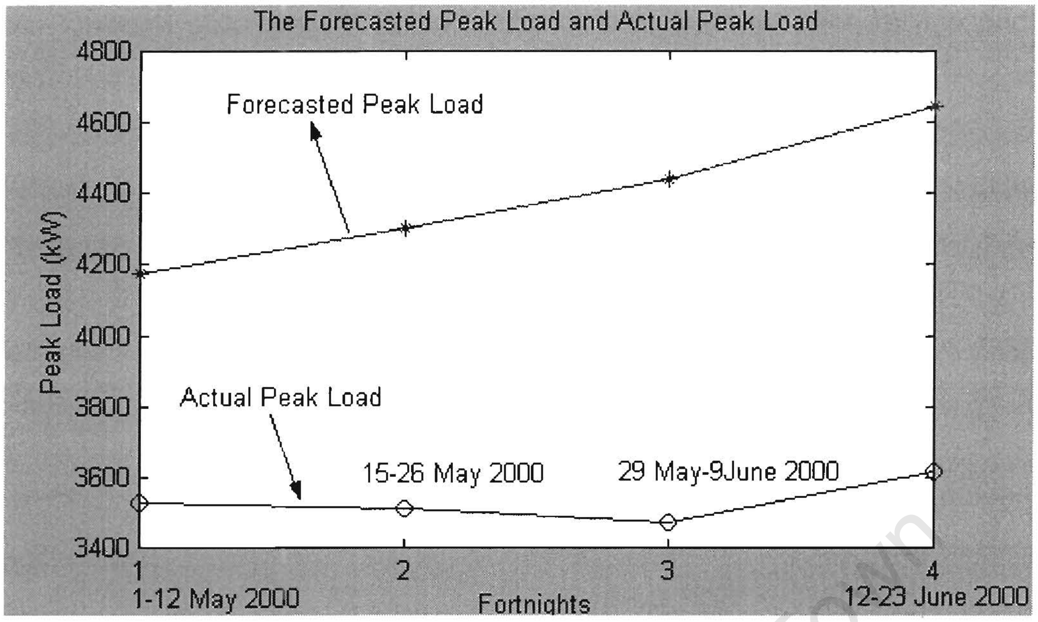


Figure 6.11 Graph of Forecasted Peak Load and Actual Peak Load

The error of the load forecast is expressed in terms of the absolute percentage error (%). The absolute percentage error is calculated using the forecasted load (kW) and the actual load (kW) of figure 6.11 in equation 3.1. The error results are presented in table 6.2.

Table 6.2: The load forecasting results of the ANN Model

Fortnight to be forecasted	**APE (%)**
1-12 May	18.26
15-26 May	22.59
29 May-9 June	27.83
12-23 June	28.31
MAPE (%)	**24.25**

The mean absolute percentage error (%) (MAPE) is calculated using the APE in equation 3.2 and the MAPE represents the overall performance of the load forecast.

The above load forecast results were only for one run of the ANN model. The ANN model is run 4 times to observe the load forecasting result. The results are presented in table 6.3.

Table 6.3: The load forecasting results for 4 runs of the ANN Model

Fortnight to be forecasted	APE (%)	APE (%)	APE (%)	APE (%)
1-12 May	18.29	18.01	18.05	18.28
15-26 May	22.19	21.39	20.82	22.17
29 May-9 June	27.16	26.03	25.25	27.13
12-23 June	27.47	26.08	25.00	27.43
MAPE (%)	**23.78**	**22.88**	**22.28**	**23.75**

The load forecasting results from table 6.3 is approximately constant.

The Matlab code that is used to implement the ANN Model is presented in the appendix. The Matlab code is also on a disk. The file is called ANNModel.m.

Chapter 7: Comparison of the Linear Regression Model and the ANN Model

7 Content

This chapter compares the linear regression model to that of the ANN model in terms of the accuracy of the load forecast in percentage error (%).

7.1 Comparison of the Linear Regression and the ANN Model

In this section the accuracy results of the ANN model in chapter 6 is compared to the accuracy results of the linear regression model in chapter 5.

Table 7.1 presents the load forecast accuracy results (in terms of APE and MAPE) of the Linear Regression Model and the ANN model for the 4 fortnights.

Table 7.1: The load forecasting results of the Linear Regression Model and ANN Model

Fortnight to be forecasted	**Linear Regression Model**: APE (%)	**ANN Model**: APE (%)	**ANN Model**: APE (%)	**ANN Model**: APE (%)	**ANN Model**: APE (%)
1-12 May	5.41	18.29	18.01	18.05	18.28
15-26 May	6.16	22.19	21.39	20.82	22.17
29 May-9 June	4.36	27.16	26.03	25.25	27.13
12-23 June	0.74	27.47	26.08	25.00	27.43
MAPE (%)	**4.17**	**23.78**	**22.88**	**22.28**	**23.75**

From table 7.1 it is clear that the linear regression model is more accurate than the ANN model. The overall performance of the linear regression model is a MAPE of 4.17 % whereas the overall accuracy of the ANN model ranges from 22.28 % to 23.78 %.

The MAPE from the ANN model is significantly larger than the MAPE from the linear regression model. When the peak load and temperature relationship was tested in the ANN model the maximum error in figure 6.7 was approximately 7 % as shown in figure 7.1.

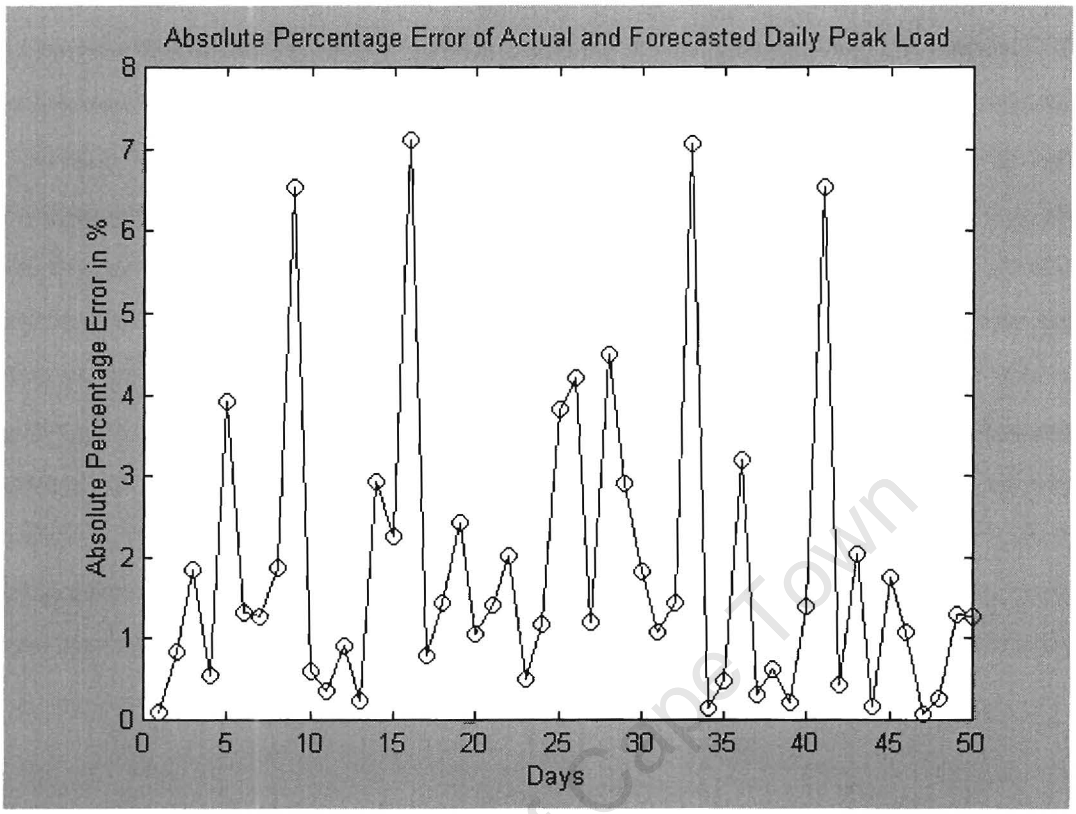


Figure 7.1 APE of Actual and Forecasted Daily Peak Load for 21 Feb. to 28 April 2000

The error on the training data set (i.e. approximate maximum of 7 %) is significantly smaller compared to the error in table 7.1 when new temperature data (historical temperature data) is presented to the ANN. A possible reason why the error is so large is because not enough data was used to train the ANN. In order to test this more data will be presented to train the ANN. This is discussed in the next section.

7.2 ANN Model with a Larger Training Data Set

The training data set that was used to train the ANN was the peak load and temperature data from the 21 February – 28 April 2000. The new training data set that will be used is from the 21 February – 26 May 2000. This new training data set includes the data from the 2 fortnights i.e. 1-12 May 2000 and 15-26 May 2000. The peak load will therefore only be forecasted for the other 2 fortnight i.e. 29 May – 9 June 2000 and 12-23 June 2000. The reason why the training data set is from the 21 February – 26 May 2000 is because this period contains 70 weekdays and this is predominately an autumn season period (i.e. March, April and May). There will therefore be 70 data points in the training data set. The reason why the autumn data set is so significant is because in the long-term model of Clayton et al (1973) and Davey et al (1973) seasonal data is used in the linear regression analysis to develop a weather-load model. If the future winter peak load is to be forecasted then the winter data set is considered to develop the weather-load models in Clayton et al (1973) and Davey et al (1973). The winter data set would consist of the daily peak load and weather data for the 3 winter months i.e. June, July and August. There are approximately 66 weekdays in June, July and August. The load data set would consist of approximately 66 daily peak loads. The linear regression analysis would use approximately 66 data points to develop a weather-load model. The ANN model above is going to use a training data set that has 70 data points. This is therefore a realistic data set size that will be used in the ANN model. The new training data set from the 21 February – 26 May 2000 is presented in figure 7.2.

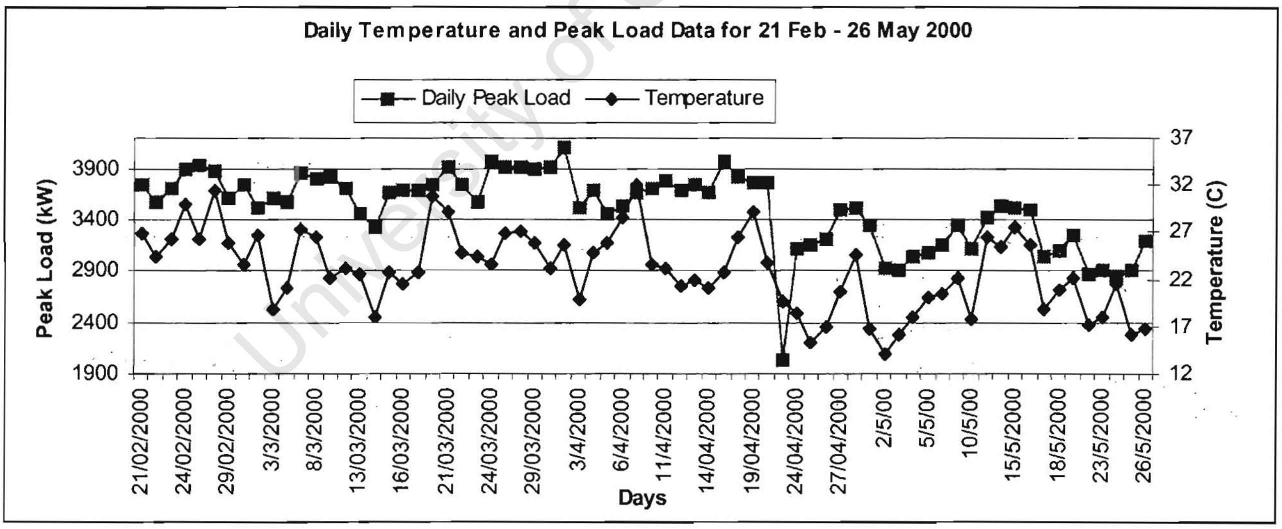


Figure 7.2 Daily Load and Temperature Data for 21 February 2000 – 26 May 2000

In figure 7.2 it can be seen that there is a very low daily peak load value (of 2034 kW) on the 21/4/2000 with respect to all the other 69 daily peak load values. This very low daily peak load value could have been caused by an abnormal condition and therefore will be removed (deleted) from the analysis. The corresponding temperature value (of 19.7 °C) on the 21/4/2000 will also be removed from the analysis. There will therefore only be 69 data points used in the analysis.

7.2.1 The Load Forecast Results of the ANN Model

The objective is to forecast the peak load for the fortnights 29May – 9 June 2000 and 12-23 June 2000.

The 1-30-1 and 1-40-1 network was tested and the 1-30-1 network will be used in the ANN model. The training results of the 1-30-1 network are presented in figure 7.3.

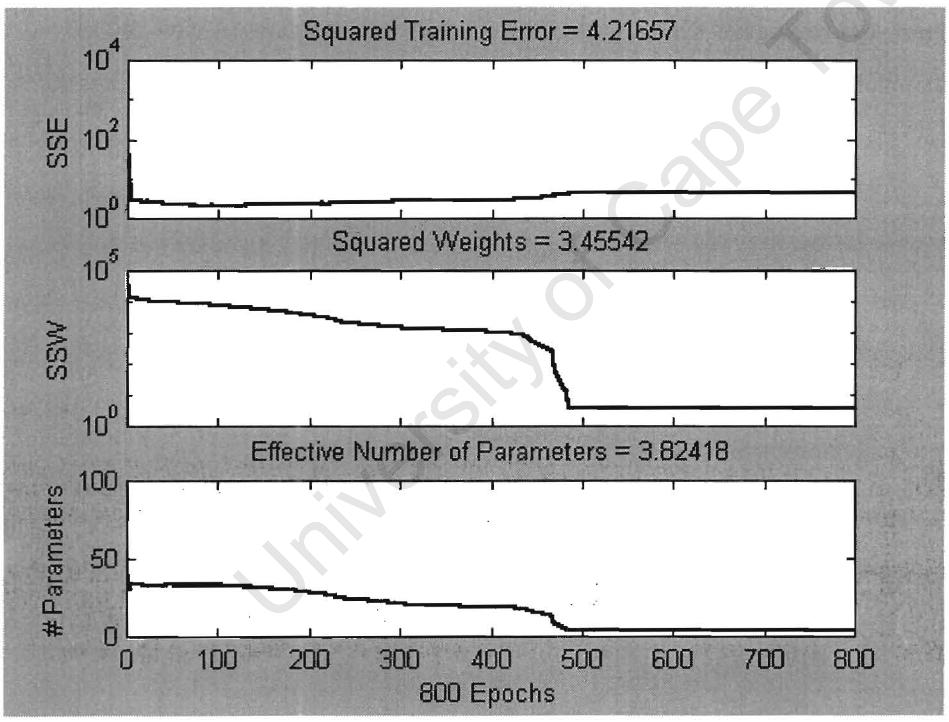


Figure 7.3 Training Results for the 1-30-1 network

The weights and biases of the 1-30-1 network are fixed after it has been trained. This section tests the accuracy of the 1-30-1 network by using the same daily temperature data (21 February – 26 May 2000) that was used to train the 1-30-1 network as input to the 1-30-1 network. The output, which is the daily peak loads, will be compared to the actual or real daily peak load data that was used to train the 1-30-1 network. Figure 7.4 presents the actual daily peak load and the forecasted daily peak load data for the period 21 February 2000 – 26 May 2000 i.e. 69 days.

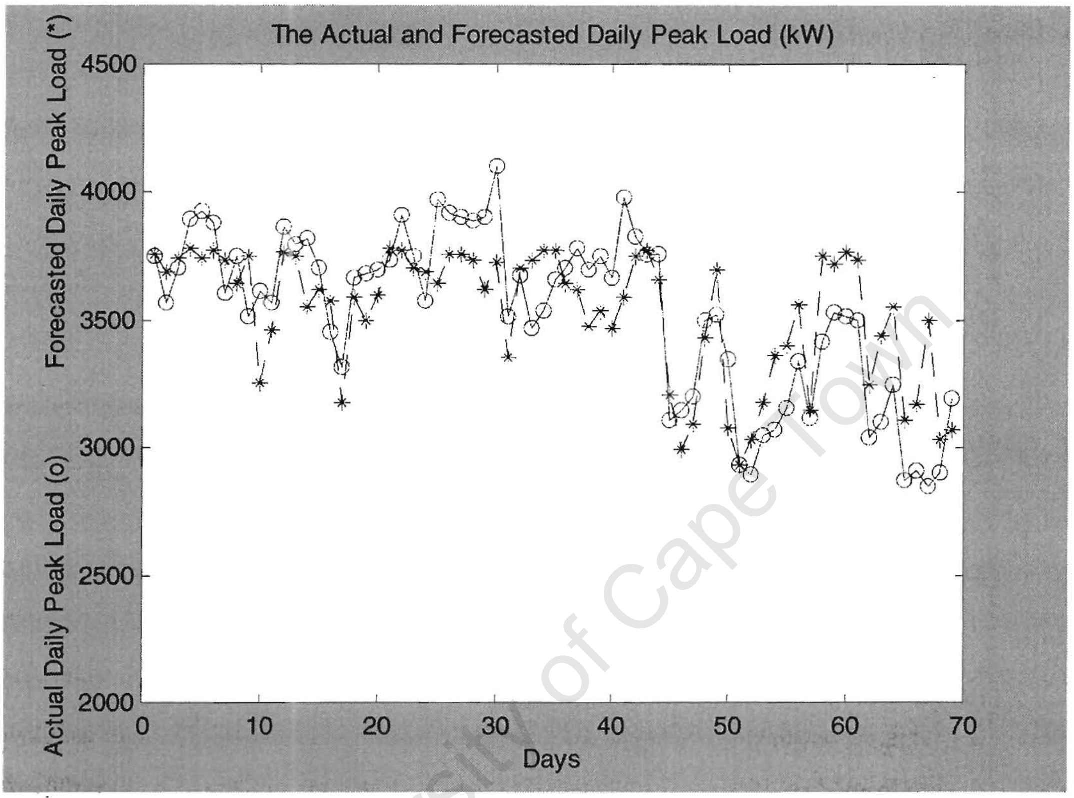


Figure 7.4 Actual and Forecasted Daily Peak Load for 1-30-1 network

The error between the actual and forecasted daily peak load is presented in figure 7.5. The absolute percentage error (APE) is calculated using equation 3.1.

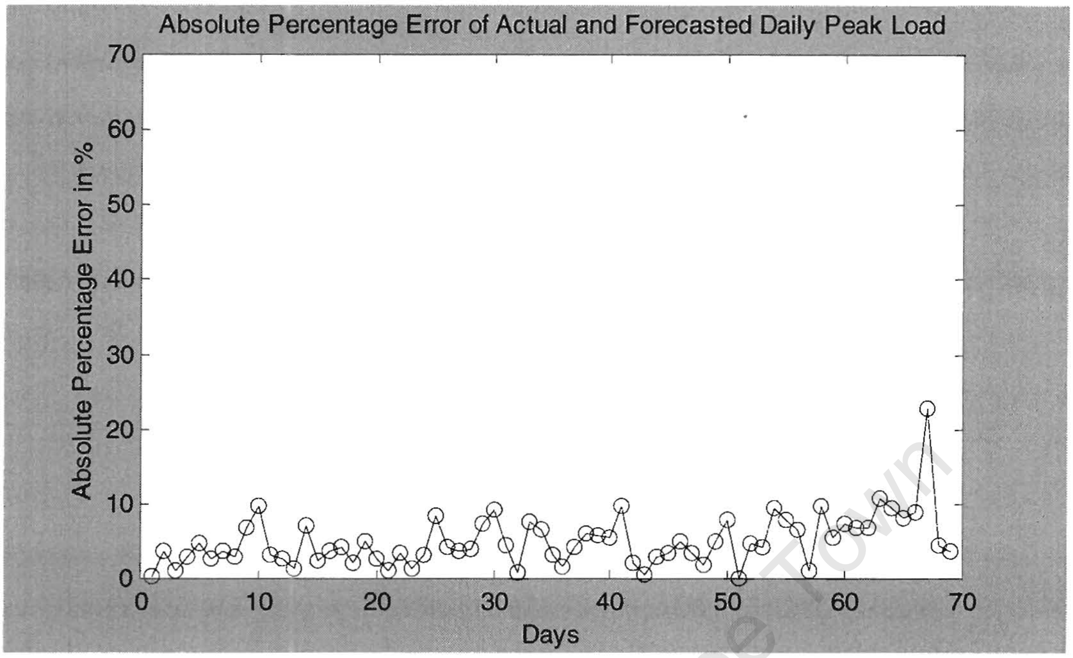


Figure 7.5 APE of Actual and Forecasted Daily Peak Load

This section uses the 1-30-1 network to forecast the peak load for the 2 fortnights i.e. 29 May – 9 June 2000 and 12-23 June 2000. Historical temperature data is used as input to the 1-30-1 network. The ANN model is run 4 times to observe the load forecasting result. The results are presented in table 7.2.

Table 7.2: The load forecasting results of the ANN Model

Fortnight to be forecasted	APE (%)	APE (%)	APE (%)	APE (%)
29 May-9 June	3.08	3.08	3.08	3.08
12-23 June	2.03	2.03	2.03	2.03
MAPE (%)	**2.56**	**2.56**	**2.56**	**2.56**

The errors in table 7.2 are significantly smaller than the errors in table 7.1, which ranges from 22.28 % to 23.78 %. It is clear from the load-forecast results that the ANN needed more data to train the network in order to determine the relationship/pattern between the load and weather data. The next section examines the load forecasting errors of the linear regression model using the same training data set.

7.3 Linear Regression Model with a Larger Data Set

The load and weather data from the 21 February – 26 May 2000 is used to develop the weather-load model. The objective is to forecast the peak load for the fortnights 29 May – 9 June 2000 and 12-23 June 2000. Table 7.3 presents the load-forecast results for the 2 fortnights.

Table 7.3: The load forecasting results of the Linear Regression Model

Fortnight to be forecasted	**APE (%)**
29 May-9 June	2.38
12-23 June	2.92
MAPE (%)	**2.65**

The program file for section 7.2 is ANNModelAECIPF2.m and the file for section 7.3 is AECIPF_Linear_&_Polynomial_Regression2.xls. Both files are on a disk.

The overall performance of the linear regression model as shown in table 7.3 is 2.65 % whereas the overall performance of the ANN model as shown in table 7.2 is approximately 2.56 %. The ANN model is slightly more accurate than the linear regression model when seasonal data (in this case autumn data set) is used. This is the same realistic data size that is used in the long-term model of Clayton et al (1973) and Davey et al (1973). Since linear regression analysis produces such an accurate load forecast could the traditional non-linear regression analysis approach produce an even more accurate load forecast? ANN technology is a form of non-linear regression analysis but the traditional non-linear regression here refers to for e.g. a polynomial regression analysis, exponential type regression etc. The next section discusses the non-linear regression.

7.4 Non-linear Regression Analysis with the Larger Data Set

A polynomial non-linear regression analysis will be performed. The reason why a polynomial is to be used is because it has a linear component in it and is possibly the closest to the linear regression analysis. The reason why exponential type regression is not performed is because exponential growth is not expected in a limited data set, little or no load growth is expected. What degree polynomial (i.e. 2nd, 3rd, 4th etc.) should be used? There is no justification to use only one-degree polynomial. The author will therefore evaluate various degree polynomials and determine if there is a limit on the degree polynomials for the data that is used in the analysis.

The load, customer AECIPF, and weather data from the 21 February – 26 May 2000 is used in the polynomial regression analysis to develop the weather-load model. The customer AECIPF load data is the same data that was used by the linear regression and ANN analysis. The equation for a 5th degree polynomial is:

$$y = \mathbf{a} + \mathbf{b}x + \mathbf{c}x^2 + \mathbf{d}x^3 + \mathbf{e}x^4 + \mathbf{f}x^5 \quad \ldots 7.1$$

Where y is the Daily Peak Load (kW),

x is the Daily Temperature (°C),

and a, b, c, d, e and f are the coefficients to be determined by the regression analysis

Table 7.4 presents the coefficients and the polynomial analysis is performed in Statistica.

Table 7.4: Coefficients of the 2nd, 3rd, 4th and 5th degree polynomials for AECIPF

Degree of Polynomial	Coefficient: a	Coefficient: b	Coefficient: c	Coefficient: d	Coefficient: e	Coefficient: f
2nd	77.78751	250.5076	**-4.23093**	---	---	---
3rd	2536.399	-83.9004	10.50961	**-0.211114**	---	---
4th	22486.72	-3752.35	257.9343	-7.47223	**0.078308**	---
5th	36395.17	-7006.20	556.8081	-20.9511	0.376957	**-0.002603**

The coefficients did not converge for degree polynomials (i.e. 6th, 7th, 8th etc.) higher than the 5th degree polynomial. This is the limit on the degree polynomial on this data set. Investigating the coefficient of the highest power for each of the 2nd, 3rd, 4th and 5th degree polynomial shows that this coefficient decreases in value (magnitude) from the 2nd degree polynomial to the 5th degree polynomial. The coefficient of the highest power of the 5th degree polynomial is very close to zero (i.e. -0.002603). The reason why the coefficients will not converge for degree polynomials higher than the 5^{th} degree polynomial is because the f coefficient of the 5^{th} degree polynomial is already very close to zero. Table 7.5 presents the load-forecast results for the 2 fortnights for the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} degree polynomial-regression analysis. The program file for this is AECIPF_Linear_&_Polynomial_Regression2.xls. This file is on a disk.

Table 7.5: The load forecasting results (AECIPF) for the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} polynomial

Fortnight	**2^{nd} degree: APE (%)**	**3^{rd} degree: APE (%)**	**4^{th} degree: APE (%)**	**5^{th} degree: APE (%)**
29 May-9 June	3.19	3.11	3.13	3.19
12-23 June	1.88	1.95	1.91	2.00
MAPE (%)	**2.54**	**2.53**	**2.52**	**2.60**

Table 7.5 shows that the 4^{th} degree polynomial produces overall (i.e. MAPE) the most accurate load forecast. Table 7.6 presents the comparison between the linear regression analysis results, ANN analysis results and the 4^{th} degree polynomial results.

Table 7.6 Comparison of the Linear, ANN and polynomial analysis for AECIPF

Fortnight	**Linear: APE (%)**	**ANN: APE (%)**	**4^{th} degree: APE (%)**
29 May-9 June	2.38	3.08	3.13
12-23 June	2.92	2.03	1.91
MAPE (%)	**2.65**	**2.56**	**2.52**

The results in table 7.6 show that the 4^{th} degree polynomial produces overall (i.e. MAPE) the most accurate load forecast. The results also indicate that the accuracy overall between the linear, ANN and 4^{th} degree polynomial analysis is very close. This could be due to this customer (AECIPF) data set. The load forecasting accuracy results could be different for another customer load data set and therefore the next section performs the same analysis but on other customer load data sets to determine which analysis produces the most accurate load forecast.

7.5 Linear, Polynomial Regression and ANN Analysis for other Customer Data Sets

This section will perform linear, polynomial and ANN analysis on 3 other customer load data sets namely Corobrick, Dolphin Beach and CM Milerton Sludge.

7.5.1 Analysis for Customer data set Corobrick

The daily temperature and daily peak load data for Corobrick for the period 21 February – 26 May 2000 is presented in figure 7.6.

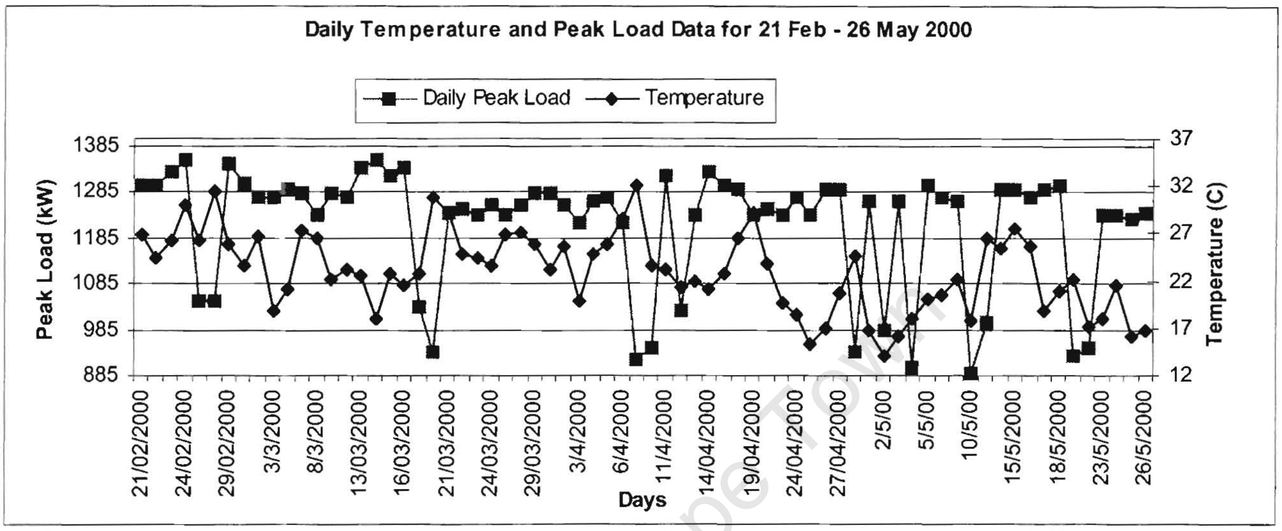


Figure 7.6 Daily Load and Temperature Data for Corobrick

7.5.1.1 Linear Regression Analysis for Corobrick

Table 7.7 presents the load-forecast results for the 2 fortnights. The program file for this analysis is called Corobrick_Linear_&_Polynomial_Regression.xls.

Table 7.7: The Linear load forecasting results for Corobrick

Fortnight to be forecasted	APE (%)
29 May-9 June	8.47
12-23 June	11.13
MAPE (%)	**9.80**

7.5.1.2 ANN Model for Corobrick

The training data set that will be used to train the ANN is presented in figure 7.6. The 1-30-1 and 1-40-1 network was tested and the 1-30-1 network will be used in the ANN model since the 1-30-1 network is large enough. The historical temperature data is used as input to the 1-30-1 network. The ANN model is run 4 times and the forecasting result is presented in table 7.8. The program file that contains this analysis is called ANNModelCorobrick2.m.

Table 7.8: The ANN load forecasting results for Corobrick

Fortnight to be forecasted	APE (%)	APE (%)	APE (%)	APE (%)
29 May-9 June	7.08	7.11	7.11	7.08
12-23 June	10.13	10.10	10.10	10.13
MAPE (%)	**8.61**	**8.61**	**8.61**	**8.61**

7.5.1.3 Polynomial Regression Analysis for Corobrick

The data presented in figure 7.6 is used in the polynomial regression analysis. Table 7.9 presents the coefficients for each of the 2nd, 3rd, 4th and 5th degree polynomials.

Table 7.9: Coefficients of the 2nd, 3rd, 4th and 5th degree polynomials for Corobrick

Degree of Polynomial	Coefficient: a	Coefficient: b	Coefficient: c	Coefficient: d	Coefficient: e	Coefficient: f
2nd	97.69734	100.7690	**-2.20228**	---	---	---
3rd	1910.203	-145.718	8.664608	**-0.155691**	---	---
4th	-12335.0	2474.607	-168.132	5.034541	**-0.055992**	---
5th	8315.822	-2354	275.1838	-14.9511	0.386700	**-0.003857**

The coefficients did not converge for the 6th, 7th, 8th etc. degree polynomials. The f coefficient of the 5th degree polynomial is very close to zero (i.e. -0.003857). The limit for this data set is also the 5th degree polynomial. Table 7.10 presents the load-forecast results for the 2 fortnights for the 2nd, 3rd, 4th and 5th degree polynomial-regression analysis. The program file for this is Corobrick_Linear_&_Polynomial_Regression.xls.

Table 7.10: The load forecasting results (Corobrick) for the 2nd, 3rd, 4th and 5th polynomial

Fortnight	2nd degree: APE (%)	3rd degree: APE (%)	4th degree: APE (%)	5th degree: APE (%)
29 May-9 June	6.48	6.89	7.30	7.12
12-23 June	9.57	9.95	10.07	10.02
MAPE (%)	**8.03**	**8.42**	**8.69**	**8.57**

Table 7.10 shows that the 2nd degree polynomial produces overall (i.e. MAPE) the most accurate load forecast. This is different from the AECIPF load forecasting results, which showed that the 4th degree polynomial produced the most accurate results. Table 7.11 presents the comparison between the linear regression analysis results, ANN analysis results and the 2nd degree polynomial results for Corobrick.

Table 7.11 Comparison of the Linear, ANN and polynomial analysis for Corobrick

Fortnight	**Linear: APE (%)**	**ANN: APE (%)**	**2nd degree: APE (%)**
29 May-9 June	8.47	7.08	6.48
12-23 June	11.13	10.13	9.57
MAPE (%)	**9.80**	**8.61**	**8.03**

The results in table 7.11 show that the 2nd degree polynomial produces overall (i.e. MAPE) the most accurate load forecast.

7.5.2 Analysis for Customer data set Dolphin Beach

The daily temperature and daily peak load data for Dolphin Beach for the period 21 February – 26 May 2000 is presented in figure 7.7.

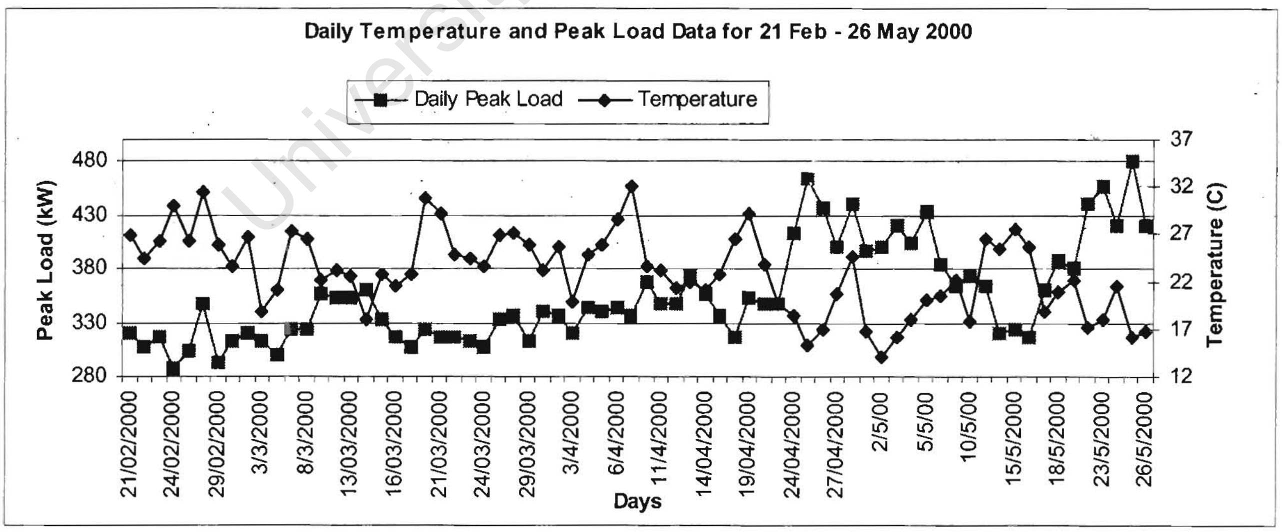


Figure 7.7 Daily Load and Temperature Data for Dolphin Beach

7.5.2.1 Linear Regression Analysis for Dolphin Beach

Table 7.12 presents the load-forecast results for the 2 fortnights. The program file that contains this analysis is called DolphinBeach_Linear_&_Polynomial_Regression.xls.

Table 7.12: The Linear load forecasting results for Dolphin Beach

Fortnight to be forecasted	APE (%)
29 May-9 June	10.89
12-23 June	4.20
MAPE (%)	**7.55**

7.5.2.2 ANN Model for Dolphin Beach

The training data set that will be used to train the ANN is presented in figure 7.7. The 1-30-1 and 1-40-1 network was tested and the 1-30-1 network will be used in the ANN model since the 1-30-1 network is large enough. The ANN model is run 4 times and the forecasting result is presented in table 7.13. The program file that contains this analysis is called ANNModelDolphinBeach2.m.

Table 7.13: The ANN load forecasting results for Dolphin Beach

Fortnight to be forecasted	APE (%)	APE (%)	APE (%)	APE (%)
29 May-9 June	6.53	6.53	6.53	6.53
12-23 June	0.813	0.813	0.813	0.813
MAPE (%)	**3.67**	**3.67**	**3.67**	**3.67**

7.5.2.3 Polynomial Regression Analysis for Dolphin Beach

The data presented in figure 7.7 is used in the polynomial regression analysis. Table 7.14 presents the coefficients for each of the 2nd, 3rd, 4th and 5th degree polynomials.

Table 7.14: Coefficients of the 2nd, 3rd, 4th and 5th degree polynomials for Dolphin Beach

Degree of Polynomial	Coefficient: a	Coefficient: b	Coefficient: c	Coefficient: d	Coefficient: e	Coefficient: f
2nd	858.8865	-37.5836	**0.660202**	---	---	---
3rd	573.5448	1.220615	-1.05057	**0.02451**	---	---
4th	-1135.91	315.6649	-22.2665	0.647349	**-0.006719**	---
5th	-13059.1	3103.553	-278.223	12.18645	-0.262316	**0.002227**

The coefficients also did not converge for the 6^{th}, 7^{th}, 8^{th} etc. degree polynomials. The limit for this data set is also the 5^{th} degree polynomial. Table 7.15 presents the load-forecast results for the 2 fortnights for the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} degree polynomial-regression analysis. The program file that contains this analysis is called DolphinBeach_Linear_&_Polynomial_Regression.xls.

Table 7.15: The load forecasting results (Dolphin Beach) for the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} polynomial

Fortnight	**2^{nd} degree: APE (%)**	**3^{rd} degree: APE (%)**	**4^{th} degree: APE (%)**	**5^{th} degree: APE (%)**
29 May-9 June	2.81	4.81	7.87	7.91
12-23 June	5.81	3.13	1.16	1.27
MAPE (%)	**4.31**	**3.97**	**4.51**	**4.59**

Table 7.15 shows that the 3^{rd} degree polynomial produces overall (i.e. MAPE) the most accurate load forecast. This is different from the AECIPF load forecasting results, which showed that the 4^{th} degree polynomial produced the most accurate results and from the Corobrick load forecasting results, which showed that the 2^{nd} degree polynomial produced the most accurate results. Table 7.16 presents the comparison between the linear regression analysis results, ANN analysis results and the 3^{rd} degree polynomial results for Dolphin Beach.

Table 7.16 Comparison of the Linear, ANN and polynomial analysis for Dolphin Beach

Fortnight	**Linear: APE (%)**	**ANN: APE (%)**	**3^{rd} degree: APE (%)**
29 May-9 June	10.89	6.53	4.81
12-23 June	4.20	0.813	3.13
MAPE (%)	**7.55**	**3.67**	**3.97**

The results in table 7.16 show that the ANN model produces overall (i.e. MAPE) the most accurate load forecast.

7.5.3 Analysis for Customer data set CM Milerton Sludge

The daily temperature and daily peak load data for CM Milerton Sludge for the period 21 February – 26 May 2000 is presented in figure 7.8.

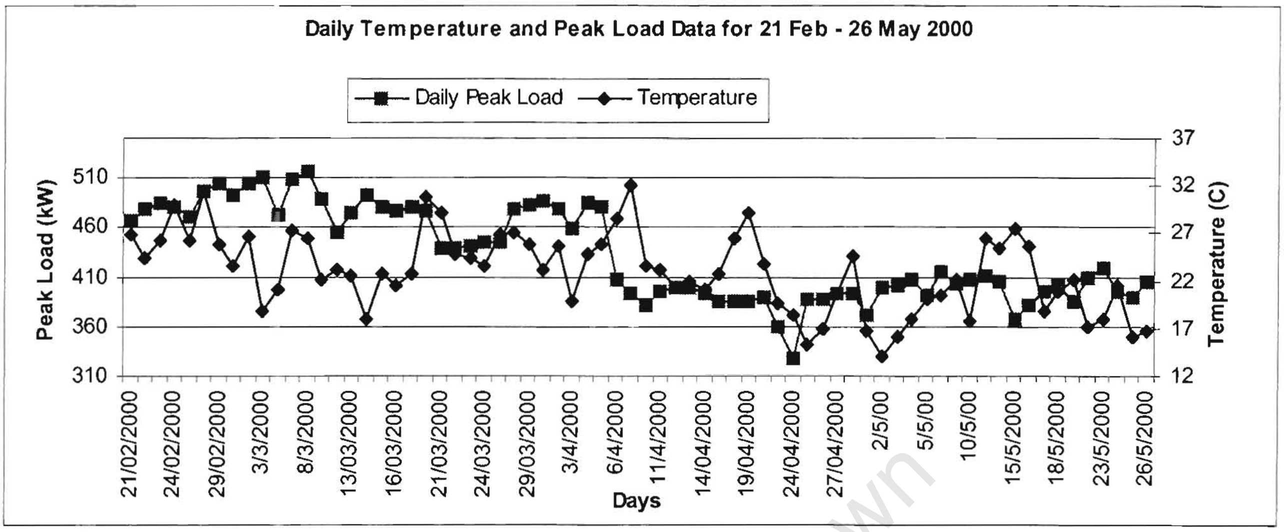


Figure 7.8 Daily Load and Temperature Data for CM Milerton Sludge

7.5.3.1 Linear Regression Analysis for CM Milerton Sludge

Table 7.17 presents the load-forecast results for the 2 fortnights. The program file that contains this analysis is called CMMilertonSludge_Linear_&_Polynomial_Regression.xls.

Table 7.17: The Linear load forecasting results for CM Milerton Sludge

Fortnight to be forecasted	APE (%)
29 May-9 June	16.11
12-23 June	17.00
MAPE (%)	**16.55**

7.5.2.2 ANN Model for CM Milerton Sludge

The training data set that will be used to train the ANN is presented in figure 7.8. The 1-30-1 and 1-40-1 network was tested and the 1-30-1 network will be used in the ANN model since the 1-30-1 network is large enough. The ANN model is run 4 times and the forecasting result is presented in table 7.18. The program file that contains this analysis is called ANNModelCMMilertonSludge2.m.

Table 7.18: The ANN load forecasting results for CM Milerton Sludge

Fortnight to be forecasted	APE (%)	APE (%)	APE (%)	APE (%)
29 May-9 June	16.08	16.08	16.08	16.08
12-23 June	16.90	16.90	16.90	16.90
MAPE (%)	**16.49**	**16.49**	**16.49**	**16.49**

7.5.2.3 Polynomial Regression Analysis for CM Milerton Sludge

The data presented in Figure 7.8 is used in the polynomial regression analysis. Table 7.19 presents the coefficients for each of the 2nd, 3rd, 4th and 5th degree polynomials.

Table 7.19: Coefficients of the 2nd, 3rd, 4th and 5th degree polynomials for CM Milerton Sludge

Degree of Polynomial	Coefficient: a	Coefficient: b	Coefficient: c	Coefficient: d	Coefficient: e	Coefficient: f
2nd	175.3669	18.66244	**-0.314711**	---	---	---
3rd	704.3392	-53.2737	2.856745	**-0.045438**	---	---
4th	814.1435	-73.4715	4.219522	-0.085445	**0.000432**	---
5th	-106.805	141.8656	-15.5507	0.805838	-0.019311	**0.000172**

The coefficients also did not converge for the 6th, 7th, 8th etc. degree polynomials. The limit for this data set is also the 5th degree polynomial. Table 7.20 presents the load-forecast results for the 2 fortnights for the 2nd, 3rd, 4th and 5th degree polynomial-regression analysis. The program file that contains this analysis is called CMMilertonSludge_Linear_&_Polynomial_Regression.xls.

Table 7.20: The load forecasting results (CM Milerton Sludge) for the 2nd, 3rd, 4th and 5th polynomial

Fortnight	2nd degree: APE (%)	3rd degree: APE (%)	4th degree: APE (%)	5th degree: APE (%)
29 May-9 June	15.71	15.80	15.77	15.84
12-23 June	16.46	16.55	16.53	16.57
MAPE (%)	**16.09**	**16.17**	**16.15**	**16.21**

Table 7.20 shows that the 2nd degree polynomial produces overall (i.e. MAPE) the most accurate load forecast. Table 7.21 presents the comparison between the linear regression

analysis results, ANN analysis results and the 2nd degree polynomial results for CM Milerton Sludge.

Table 7.21 Comparison of the Linear, ANN and polynomial analysis for CM Milerton Sludge

Fortnight	**Linear: APE (%)**	**ANN: APE (%)**	**2nd degree: APE (%)**
29 May-9 June	16.11	16.08	15.71
12-23 June	17.00	16.90	16.46
MAPE (%)	**16.55**	**16.49**	**16.09**

The results in table 7.21 show that the 2nd degree polynomial produces overall (i.e. MAPE) the most accurate load forecast.

For each of the four customer load data sets the polynomial and ANN analysis produces the most accurate load forecast. It is interesting too note that it is not the same degree polynomial that produces the most accurate load forecast for the four customers. For customer AECIPF it is the 4th degree polynomial, for customer Corobrick it is the 2nd degree polynomial, for Dolphin Beach it is the 3rd degree polynomial and for customer CM Milerton Sludge it is the 2nd degree polynomial. For three of the customers (i.e. AECIPF, Corobrick and CM Milerton Sludge) the polynomial analysis is more accurate than the ANN analysis and for the other remaining customer (i.e. Dolphin Beach) the ANN analysis is more accurate than the polynomial analysis.

From these analyses it is the conclusion that the polynomial and ANN analysis could be used to produce a more effective long-term model since they are more accurate than the linear regression analysis. It is evident that for some customer data sets the polynomial analysis will be more accurate than the ANN analysis and for other customer data sets the ANN analysis will be more accurate than the polynomial analysis. However to determine the most accurate degree polynomial for a customer data set, various degree polynomials have to be tested since the same degree polynomial will not always produce the most accurate load forecast as presented above. The advantage that the ANN analysis has is that only one ANN analysis has to be performed for a customer data set.

Chapter 8: Implementation of the Cluster Analysis Approach in Long-term Load Forecasting

8 Content

The load-forecast approach in the previous chapters does not consider classifying customers into groups. This chapter discusses load forecasting by classifying customers using cluster analysis. A cluster analysis model is developed which makes use of the customers load parameters as discussed in chapter 4.

8.1 The Implementation of the Cluster Analysis Model

This section builds on the ideas expressed in section 4.3. This section describes the choice of load parameter combinations, customer data preparation, choice of distance measure, choice of clustering algorithm, the calculation of the standardised data matrix and the interpretation of the dendograms for each cluster analysis model.

8.1.1 Choice of Distance Measure for Cluster Analysis

Consider figure 4.6 in chapter 4. In order to calculate the distance matrix a distance measure is used, Romesburg (1984). The Euclidean distance measure will be used in this thesis. There is no particular reason for choosing the Euclidean distance measure. From the literature, Everitt (1974) and Romesburg (1984), this seems to be a popular distance measure used.

8.1.2 Choice of Clustering Algorithm for Cluster Analysis

In figure 4.6 a clustering algorithm is used to transform the distance matrix to the dendogram. The nearest neighbor-clustering algorithm will be used in this thesis. The reason for choosing the algorithm is for its simplicity. The literature, Everitt (1974) and Romesburg (1984), do not suggest which method is suited for a particular problem. Nearest neighbor means that objects are fused according to the distance between their nearest members, Romesburg (1984). The objects with the smallest distance are fused together. Software called Statistica (Kernel release 5.5) will be used to perform the cluster analysis and it implements the nearest neighbor-clustering algorithm.

8.1.3 Preparation of Customer Load Parameters for the Cluster Analysis

The load parameters that will be used are max[kVA, kW, kVAr], ave[kVA, kW, kVAr], ave[Power Factor] and the Load Factor.

The objective is to forecast the peak load for the 4 fortnights 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000. This will be based on customer load data from the period 21 February to 28 April 2000. This customer load data will be used to perform the cluster analysis. The customer load data for the period 21 February to 28 April 2000 will be transformed into load parameters that will be used in the cluster analysis. Each customer has these load parameters for the period 21 February to 28 April 2000. The load parameters are prepared from the kWh and kVArh that is made available in half-hourly integrating periods for each customer. The kVAh is prepared from the load data kWh and kVArh using equation 4.1.

Table 8.1: Calculation of Load Parameters for the cluster analysis

Load Parameter	Units	Calculation of Load Parameters for ½ hourly (h) integrated period	Calculation of **Daily** Load Parameter
Max Apparent Power	kVA	kVAh / h	**Max** of ½ hourly kVA
Max Active Power	kW	kWh / h	**Max** of ½ hourly kW
Max Reactive Power	kVAr	kVArh / h	**Max** of ½ hourly kVAr
Ave Apparent Power	kVA	kVAh / h	**Ave** of ½ hourly kVA
Ave Active Power	kW	kWh / h	**Ave** of ½ hourly kW
Ave Reactive Power	kVAr	kVArh / h	**Ave** of ½ hourly kVAr
Ave Power Factor	---	---	$\frac{Daily\ Ave\ Active\ Power}{Daily\ Ave\ Apparent\ Power}$
Load Factor	---	---	---

Each of these load parameters is a single value number. The expression max[kVA, kW, kVAr] means the maximum value of the kVA, kW and kVAr for the period 21 February to 28 April 2000. Firstly the daily peak/maximum kVA, kW and kVAr is calculated for each day, as shown in table 8.1, for the period 21 February to 28 April 2000. The max kVA is determined by calculating the maximum of the daily kVA peaks/maximums for the period 21 February to 28 April 2000. The max kW is determined by calculating the maximum of the daily kW peaks/maximums for the period 21 February to 28 April 2000. The max kVAr is determined by calculating the maximum of the daily kVAr peaks/maximums for the period 21 February to 28 April 2000. The same explanation/approach applies to the ave[kVA, kW, kVAr].

The ave[Power Factor] refers to the average power factor for the period 21 February to 28 April 2000. Firstly the daily average power factor is calculated for each day, as shown in table 8.1, of the period 21 February to 28 April 2000. The ave[Power Factor] is determined by calculating the average of the daily average power factors.

The Load Factor is determined for the period 21 February to 28 April 2000. It is calculated using the total kWh over the period 21 February to 28 April 2000 and the peak load over the period 21 February to 28 April 2000 using equation 4.2.

8.1.4 Calculation of Data Matrix and Standardised Data Matrix

The process of calculating the standardised matrix is presented and the reasons why a standardised matrix is used.

There are 14 customers considered in this analysis. The 14 customers are AECIPF, CM Milnerton Sludge, Continental China, Corobrick, Crammix, D&H Industrial (Rocla), Denel Edms Bkp-Firgrove, Denel Edms Bkp-Kranzkop, Dolphin Beach, Hoechst SA Pty Ltd, JJ Bricks, Kohler Bricks, Much Asphalt and Ciolli Brothers. Of the 14 customers only 12 will be used in the cluster analysis and the load forecast will be carried out for the other 2 customer's i.e. Denel Edms Bkp-Firgrove and Much Asphalt. The data matrix has as its rows the 12 customers and the columns the customers' load parameters. The data matrix is presented in table 8.2. The 2 customers are chosen such that customer Denel Edms Bkp-Firgrove is a large power user and customer Much Asphalt is a small power user. This is hopefully not to have the two customers identify to the same cluster regularly. If the level of power use of the two customers were relatively the same then they would probably identify together all the time. It is not a problem if the two customers do identify together, however for the purpose of the investigation a variety of clustering options is best to analyze so as to test and investigate various scenarios.

Table 8.2: Data Matrix of the 12 customers and their load parameters

Customers	**Max kW**	**Ave kW**	**Max kVAr**	**Ave kVAr**	**Max kVA**	**Ave kVA**	**Ave PF**	**LF**
AECIPF	4104.0	2165.3	1404.0	553.3	4335.6	2239.9	0.967	0.528
CM Milerton Sludge	516.3	377.5	245.7	128.8	563.3	399.6	0.945	0.731
Continental China	2512.0	1032.7	608.0	278.8	2538.2	1072.3	0.961	0.411
Corobrick	1352.0	829.3	592.0	248.9	1454.0	868.3	0.956	0.613
Crammix	1296.0	789.3	976.0	612.4	1603.3	1000.0	0.789	0.609
D&H Industrial,Rocla	181.8	101.0	43.5	10.9	183.2	101.8	0.992	0.556
Denel Edms Bkp-Kra	3176.0	1599.5	2324.0	863.2	3912.9	1831.9	0.894	0.504
Dolphin Beach	464.0	224.7	132.0	103.0	479.3	248.9	0.902	0.484
Hoechst SA Pty Ltd	2154.0	1323.4	561.0	224.7	2212.7	1343.2	0.983	0.614
J J Bricks	409.2	140.2	195.0	60.5	449.6	153.4	0.914	0.343
Kohler Bricks	627.0	375.3	255.6	105.6	654.0	390.8	0.959	0.599
Ciolli Brothers	326.4	79.7	182.4	42.6	371.4	91.5	0.891	0.244

The next step is optional and that is the step of standardizing the data matrix, Romesburg (1984). Standardization the data matrix converts the original load parameters of the customers to new dimensionless load parameters. Romesburg (1984) gives two main reasons for standardizing a data matrix. Firstly the units of the attributes (load parameters) can affect the measure of similarity between the objects (customers). Standardizing the attributes removes these effects. The measure of the similarity between objects is determined in the form of a distance matrix. The distance matrix is calculated using the data matrix. The second reason is that standardizing the data matrix makes the attributes contribute more equally to the similarity between objects in the distance matrix. It is because of these two reasons that the data matrix in table 8.2 will be standardized.

A standardizing function is used to standardize the data matrix, Romesburg (1984). There are various ways to standardize a data matrix. This thesis will only implement one way and it is the one that is used most commonly in practice as described in Romesburg

(1984). The standardizing function is as follows: First consider the data matrix with X as its entries.

Data Matrix

$$
\begin{array}{cc}
 & \textbf{Attributes} \\
 & \begin{array}{ccccccc} 1 & 2 & \cdots & i & \cdots & n \end{array} \\
\textbf{Objects}\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ j \\ \vdots \\ t \end{array} &
\begin{bmatrix}
X_{11} & X_{12} & \cdots & X_{1i} & \cdots & X_{1n} \\
X_{21} & X_{22} & \cdots & X_{2i} & \cdots & X_{2n} \\
\vdots & \vdots & & \vdots & & \vdots \\
X_{j1} & X_{j2} & \cdots & X_{ji} & \cdots & X_{jn} \\
\vdots & \vdots & & \vdots & & \vdots \\
X_{t1} & X_{t2} & \cdots & X_{ti} & \cdots & X_{tn}
\end{bmatrix}
\end{array}
$$

The corresponding standardized data matrix with the standardized entries Z is as follows:

$$
\begin{array}{cc}
 & \textbf{Attributes} \\
 & \begin{array}{ccccccc} 1 & 2 & \cdots & i & \cdots & n \end{array} \\
\textbf{Objects}\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ j \\ \vdots \\ t \end{array} &
\begin{bmatrix}
Z_{11} & Z_{12} & \cdots & Z_{1i} & \cdots & Z_{1n} \\
Z_{21} & Z_{22} & \cdots & Z_{2i} & \cdots & Z_{2n} \\
\vdots & \vdots & & \vdots & & \vdots \\
Z_{j1} & Z_{j2} & \cdots & Z_{ji} & \cdots & Z_{jn} \\
\vdots & \vdots & & \vdots & & \vdots \\
Z_{t1} & Z_{t2} & \cdots & Z_{ti} & \cdots & Z_{tn}
\end{bmatrix}
\end{array}
$$

The standardizing function that is used to transform the data matrix to the standardized data matrix is as follows where Z_{ji} is the standardized value of X_{ji}:

$$Z_{ji} = \frac{X_{ji} - \overline{X_i}}{S_i} \quad \ldots 8.1$$

where

$$\overline{X_i} = \frac{\sum_{j=1}^{t} X_{ji}}{t} \quad \ldots 8.2$$

and

$$S_i = \left(\frac{\sum_{j=1}^{t} \left(X_{ji} - \overline{X_i} \right)^2}{t-1} \right)^{\frac{1}{2}} \quad \ldots 8.3$$

The value $\overline{X_i}$ represents the mean of the i th attribute and S_i represents the standard deviation of the i th attribute. Using the above standardizing function the data matrix of table 8.2 is transformed into a standardized data matrix in table 8.3. The values in table 8.3 are dimensionless since they are standardized values.

Table 8.3: Standardized Data Matrix of the 12 customers and their load parameters

Customers	**Max kW**	**Ave kW**	**Max kVAr**	**Ave kVAr**	**Max kVA**	**Ave kVA**	**Ave PF**	**LF**
AECIPF	2.084	2.108	1.169	1.060	1.953	2.003	0.669	0.059
CM Milerton Sludge	-0.709	-0.561	-0.573	-0.525	-0.704	-0.578	0.272	1.575
Continental China	0.845	0.417	-0.028	0.035	0.687	0.365	0.563	-0.808
Corobrick	-0.058	0.114	-0.052	-0.077	-0.077	0.079	0.476	0.698
Crammix	-0.102	0.054	0.525	1.280	0.028	0.264	-2.506	0.665
D&H Industrial,Rocla	-0.969	-0.973	-0.877	-0.965	-0.972	-0.996	1.119	0.268
Denel Edms Bkp-Kra	1.362	1.263	2.553	2.216	1.656	1.431	-0.640	-0.119
Dolphin Beach	-0.749	-0.789	-0.744	-0.621	-0.764	-0.790	-0.490	-0.263
Hoechst SA Pty Ltd	0.566	0.851	-0.099	-0.167	0.458	0.746	0.963	0.705
J J Bricks	-0.792	-0.915	-0.649	-0.780	-0.785	-0.924	-0.269	-1.318
Kohler Bricks	-0.622	-0.564	-0.558	-0.611	-0.641	-0.591	0.529	0.588
Ciolli Brothers	-0.857	-1.005	-0.668	-0.846	-0.840	-1.010	-0.686	-2.050

8.1.5 Choice of Load Parameter Combinations

In chapter 4 the load parameters were identified but there was no mention of how these load parameters are going to be used to test whether cluster analysis can improve the accuracy of a load forecast. To come to a conclusion of whether cluster analysis improves the accuracy of a load forecast numerous experiments/cluster analyses have to be performed. There are 8 load parameters considered in this thesis. If all the 8 load parameters together are used in one cluster analysis then this results in one output. There has to be more cluster analyses performed so as to come to a better conclusion about the use of cluster analysis in load forecasting. In order to perform more cluster analyses the author has decided to use groups of load parameters. There can be groups with 1, 2, 3, 4, 5, 6, 7 and 8-load parameter(s). The total number of combinations for each load parameter group (i.e. 1, 2, 3, 4, 5, 6, 7 and 8) is: $({}^{8}{}_{1}) = 8$, $({}^{8}{}_{2}) = 28$, $({}^{8}{}_{3}) = 56$, $({}^{8}{}_{4}) = 70$, $({}^{8}$

$_{5}) = 56$, $(^{8}{}_{6}) = 28$, $(^{8}{}_{7}) = 8$ and $(^{8}{}_{8}) = 1$. $(^{n}{}_{k})$ is read n combination k. $(^{8}{}_{2}) = 28$ means that there are a total of 28 combinations with groups of two-load parameters.

For each of the load parameter groupings (i.e. 1, 2, 3, 4, 5, 6, 7 and 8) not all the combinations will be considered for the cluster analysis. The reason is that each set of combinations do not have the load parameter max kW in each combination. The peak load is to be forecasted i.e. the maximum kW or max kW. This load parameter has to be one of the load parameters in a group because this will be forecasted. If max kW is not one of the load parameters in a group then possibly the customers in each cluster could have widely ranging maximum loads i.e. max kW. The idea is to have customers in a cluster that are similar in special respect to their maximum loads because these customers' maximum loads are used to forecast the maximum load in the future. The total number of combinations that do have the load parameter max kW in a group is presented in table 8.4.

Table 8.4 Number of combinations with load parameter max kW

Number of load parameters in a group	Number of combinations only with load parameter: max kW
1	1
2	7
3	21
4	35
5	35
6	21
7	7
8	1

The author has decided to use groups with 3 load parameters since 21 combinations are easier to manage considering the many load forecasts that have to be performed for each cluster and possibly large enough to hopefully come to a conclusion about the use of cluster analysis in load forecasting. Groups of 6 load parameters also resulted in 21 combinations and there is no particular reason for using groups of 3 load parameters. The 21 groups with 3 load parameters that do have max kW as one of the load parameters are:

$$\begin{bmatrix} \max kW \\ ave\,kVA \\ ave\,kW \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kVA \\ ave\,k\,\mathrm{var} \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kVA \\ \max kVA \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kVA \\ \max k\,\mathrm{var} \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kVA \\ ave\,PF \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kVA \\ LF \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kW \\ ave\,k\,\mathrm{var} \end{bmatrix},$$

$$\begin{bmatrix} \max kW \\ ave\,kW \\ \max kVA \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kW \\ \max k\,\mathrm{var} \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kW \\ ave\,PF \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,kW \\ LF \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,k\,\mathrm{var} \\ \max kVA \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,k\,\mathrm{var} \\ \max k\,\mathrm{var} \end{bmatrix},$$

$$\begin{bmatrix} \max kW \\ ave\,k\,\mathrm{var} \\ ave\,PF \end{bmatrix}, \begin{bmatrix} \max kW \\ ave\,k\,\mathrm{var} \\ LF \end{bmatrix}, \begin{bmatrix} \max kW \\ \max kVA \\ \max k\,\mathrm{var} \end{bmatrix}, \begin{bmatrix} \max kW \\ \max kVA \\ ave\,PF \end{bmatrix}, \begin{bmatrix} \max kW \\ \max kVA \\ LF \end{bmatrix}, \begin{bmatrix} \max kW \\ \max k\,\mathrm{var} \\ ave\,PF \end{bmatrix},$$

$$\begin{bmatrix} \max kW \\ \max k\,\mathrm{var} \\ LF \end{bmatrix} \text{ and } \begin{bmatrix} \max kW \\ ave\,PF \\ LF \end{bmatrix}.$$

8.1.6 The Interpretation of the Dendograms

Interpretation of a tree means that clusters have to be determined. To determine the clusters the dendogram has to be "cut" somewhere. Where the dendogram is cut determines the number of clusters in the classification. The method that will be used is the one that Romesburg (1984) and Everitt (1974) implement. The most popular method, Romesburg (1984) and Everitt (1974), is to cut the tree where the cluster structure remains stable for a long distance. This means that all the distances (i.e. linkage distances) between clusters in the tree are determined and the longest distance calculated is where the tree is "cut".

This section presents the dendograms for each of the 21 groups of 3 load parameters. Only the first two dendograms are presented and interpreted in this section. The other 19 dendograms are presented in the appendix. Consider the first tree with the load parameters Max kW, Ave kVA and Ave kW in figure 8.1.

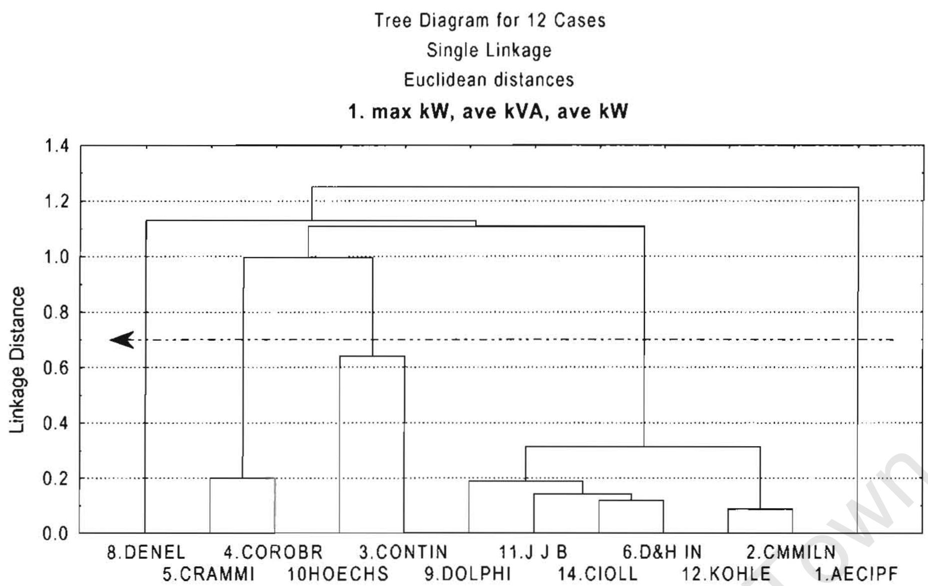


Figure 8.1 Tree of Max kW, Ave kVA and Ave kW

Table 8.5 shows the different clusters and the customers in each cluster for the dendogram in figure 8.1. The number(s) in each cluster refers to a customer. It refers to the number that precedes the customer in the dendogram.

Table 8.5: Cluster formation for max kW, ave kVA, ave kW

Cluster	1	2	3	4	5
Customers in Cluster	8	4,5	3,10	2,6,9,11,12,14	1

The second load parameter grouping is max kW, ave kVA and ave kVAr. The resultant dendogram is presented in figure 8.2 and the cluster formation is presented in table 8.6.

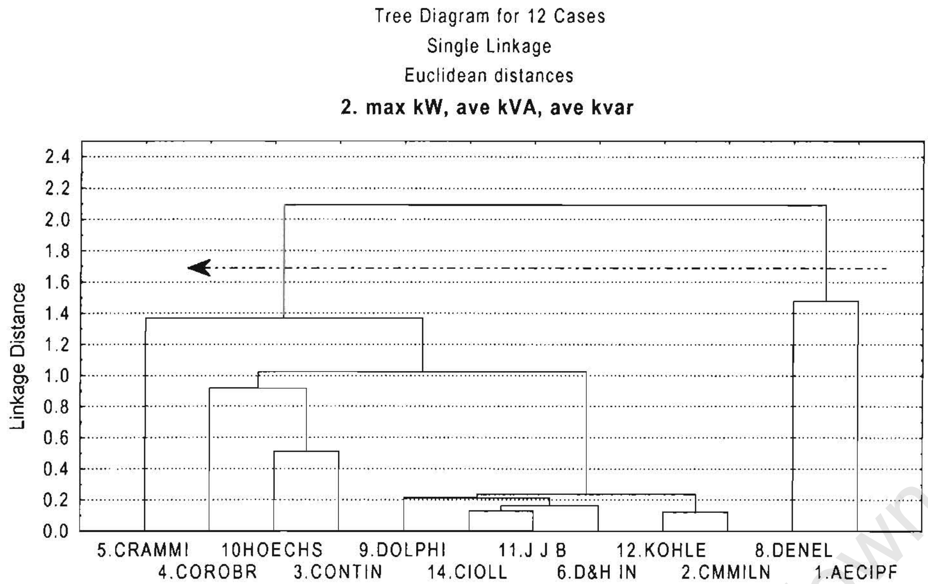


Figure 8.2 Tree of Max kW, Ave kVA and Ave kVAr

Table 8.6: Cluster formation for max kW, ave kVA, ave kVAr

Cluster	1	2
Customers in Cluster	5,4,10,3,9,14,11,6,12,2	8,1

Investigating the cluster formation from interpreting each of dendograms in figure 8.1, 8.2 and the dendograms in the appendix show that customers 8 and 1 are clustered together 5 times for different load parameter groups. Similarly customers 2, 3, 4, 5, 6, 9, 10, 11, 12 and 14 are clustered together 8 times for different load parameter groups and customers 2, 3, 4, 6, 9, 10, 11, 12 and 14 are clustered together 5 times for different load parameter groups. The analysis shows that certain customers group together.

8.2 The Classification of the Customers to be forecasted

The load is forecasted for the two customers Denel Edms Bkp-Firgrove and Much Asphalt. Each of the two customers must be separately identified to a cluster for each of the 21 groups of 3 load parameters.

The process of identifying a customer to a cluster is discussed in section 4.3.4. The customers load parameters are used to identify a customer to a cluster. In order to do this the load parameters for these two customers are calculated for the period 21 February to 28 April 2000 and are presented in table 8.7.

Table 8.7: The load parameters for Denel Edms Bkp-Firgrove and Much Asphalt

Customers	Max kW	Ave kW	Max kVAr	Ave kVAr	Max kVA	Ave kVA	Ave PF	LF
Denel Edms Bkp-Fir	3294	1886.2	2100	1013.5	3851	2126.8	0.883	0.567
Much Asphalt	505.2	183.8	46.2	5.4	505.3	184	0.999	0.364

The load parameters in table 8.7 will be standardized since the data matrix (table 8.2) was standardized (table 8.3). The load parameters are standardized according to the same variables $\overline{X_i}$ and S_i that standardized the data matrix. The standardized load parameters of the 2 customers are presented in table 8.8. The values are dimensionless.

Table 8.8: The standardized load parameters for Denel Edms Bkp-Firgrove and Much Asphalt

Customers	Max kW	Ave kW	Max kVAr	Ave kVAr	Max kVA	Ave kVA	Ave PF	LF
Denel Edms Bkp-Fir	1.454	1.661	2.216	2.778	1.612	1.845	-0.823	0.349
Much Asphalt	-0.717	-0.850	-0.873	-0.985	-0.745	-0.881	1.233	-1.159

These standardized load parameters will be used to calculate the Euclidean distance between the customers (Denel Edms Bkp-Firgrove and Much Asphalt) and each customer in each cluster for each of the 21 groups of 3 load parameters. The Euclidean distance will be calculated in 3 dimensions, Romesburg (1984) i.e.

$$\text{Euclidean Distance} = \sqrt{a^2 + b^2 + c^2} \quad \ldots\ 8.4$$

Where

a, b and c represent the distance between a load parameter

a, b and c represent the distance between 3 different load parameters

A cluster of one customer will not be considered for load forecasting and a customer identified to a cluster like this will not be considered for load forecasting. The average load forecast of a cluster, which has one customer, is the load forecast of the one customer in the cluster.

In this chapter, for the first two load parameter groups the two customers (Denel Edms Bkp-Firgrove and Much Asphalt) are identified to a cluster. The other 19 load parameter groups are presented in the appendix. Table 8.9 and 8.10 show the cluster structures for the first two load parameter groups together with the customers Denel Edms Bkp-Firgrove and Much Asphalt identified to a cluster.

Table 8.9: Cluster formation/Customer identification for max kW, ave kVA, ave kW

Cluster	1	2	3	4	5
Customers in Cluster	8	4,5	3,10	2,6,9,11,12,14	1
Customers identified	-	-	-	**Much Asphalt**	-

Table 8.10: Cluster formation/Customer identification for max kW, ave kVA, ave kVAr

Cluster	1	2
Customers in Cluster	5,4,10,3,9,14,11,6,12,2	8,1
Customers identified	**Much Asphalt**	**Denel Edms Bkp-Fir**

Analyzing table 8.9, 8.10 and the tables in the appendix shows that Denel Edms Bkp-Firgrove is always identified to the cluster that has customer 8 and 1 in it. Customer Much Asphalt identifies to a cluster with customers 2, 3, 4, 5, 6, 9, 10, 11, 12 and 14 in it and to a cluster with customers 2, 3, 4, 6, 9, 10, 11, 12 and 14 in it.

8.3 The Load Forecast Results

The load forecast process for a cluster is explained in section 4.3.3. The daily peak load data for the period 21 February to 28 April 2000 is used to forecast the future peak load for each fortnight. The daily peak load is first forecasted for each day in each fortnight. The daily peak load is forecasted for each customer in the cluster for each fortnight 1-12 May 2000, 15-26 May 2000, 29 May-9 June 2000 and 12-23 June 2000. The average daily peak load forecast of all the customers in the cluster is calculated by taking the average of the daily peak load forecasts of each customer in the cluster for each day in the fortnight. There are 10 average daily peak load forecasts in a fortnight, weekdays are only considered. The maximum of the average daily peak load forecasts for the fortnight represents the peak load forecast for the fortnight.

The load is forecasted for the two customers, Denel Edms Bkp-Firgrove and Much Asphalt. The individual load forecast of the customer is compared to the load forecast of the customer considering the cluster analysis approach. The difference between the two processes is that a load forecast of a customer using cluster analysis makes use of other customers load forecasts whereas the other process makes use of the customer's own daily peak load data using regression analysis to perform the load forecast. The load forecast results for each of the load parameter groups is presented and compared to the load forecast of the customer or the combined forecast of the two customers using its own load data in terms of the percentage error (%). The tables below show the load forecast for the cluster compared to the load forecast for the individual customer that was identified to that cluster where it is applicable.

The two groups that have the same cluster structures are **Max kW, Ave kVA, Ave kW** and **Max kW, Ave kW, Max kVA**. The difference between the two load parameter groups are the load parameters Ave kVA and Max kVA, table 8.30.

Table 8.30: Load Forecast results: **Max kW, Ave kVA, Ave kW** and **Max kW, Ave kW, Max kVA**.

	Ave kW, (**Ave kVA** or **Max kVA**)	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)
1-12 May	25.1	18.9
15-26 May	30.7	26.2
29 May-9 June	32.6	29.6
12-23 June	46.2	44.9
MAPE (%)	**33.6**	**29.9**

The five groups that have the same cluster structures are **Max kW, Ave kVA, Ave kVAr** and **Max kW, Ave kVA, Max kVA** and **Max kW, Ave kVA, Max kVAr** and **Max kW, Ave kW, Ave kVAr** and **Max kW, Ave kVAr, Max kVA**. The most prominent load parameters besides Max kW in the five groups is presented in table 8.31.

Table 8.31: Load Forecast results: **Max kW, Ave kVA, Ave kVAr** and **Max kW, Ave kVA, Max kVA** and **Max kW, Ave kVA, Max kVAr** and **Max kW, Ave kW, Ave kVAr** and **Max kW, Ave kVAr, Max kVA**

	Most Prominent: Ave **kVA & Ave kVAr**			Most Prominent: Ave **kVA & Ave kVAr**	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)		Cluster Load Forecast for Denel-Firgrove: APE (%)	Denel-Firgrove Load Forecast: APE (%)
1-12 May	57.0	18.9		20.6	40.1
15-26 May	40.7	26.2		32.2	53.5
29 May-9 June	31.8	29.6		41.7	65.0
12-23 June	1.1	44.9		48.8	74.9
MAPE (%)	**32.6**	**29.9**		**35.8**	**58.4**

The six groups that have the same cluster structures are **Max kW, Ave kVA, Ave PF** and **Max kW, Ave kW, Ave PF** and **Max kW, Ave kVAr, Max kVAr** and **Max kW, Ave kVAr, Ave PF** and **Max kW, Max kVA, Ave PF** and **Max kW, Max kVAr, Ave PF**. The most prominent load parameters in the six groups is the Ave PF, table 8.32.

Table 8.32: Load Forecast results: **Max kW, Ave kVA, Ave PF** and **Max kW, Ave kW, Ave PF** and **Max kW, Ave kVAr, Max kVAr** and **Max kW, Ave kVAr, Ave PF** and **Max kW, Max kVA, Ave PF** and **Max kW, Max kVAr, Ave PF**

	Most Prominent: **Ave PF**	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)
1-12 May	43.5	18.9
15-26 May	26.4	26.2
29 May-9 June	16.2	29.6
12-23 June	12.9	44.9
MAPE (%)	**24.8**	**29.9**

The two groups that have the same cluster structures are **Max kW, Ave kW, Max kVAr** and **Max kW, Max kVA, Max kVAr**. The difference between the two load parameter groups are the load parameters Ave kW and Max kVA, table 8.33.

Table 8.33: Load Forecast results: **Max kW, Ave kW, Max kVAr** and **Max kW, Max kVA, Max kVAr**

	Max kVAr, (**Ave kW** or **Max kVA**)	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)
1-12 May	57.0	18.9
15-26 May	40.7	26.2
29 May-9 June	31.8	29.6
12-23 June	1.1	44.9
MAPE (%)	**32.6**	**29.9**

Table 8.34 presents the forecasting results for the one group: **Max kW, Ave kVAr, LF**.

Table 8.34: Load Forecast results: **Max kW, Ave kVAr, LF**

	Max kW, Ave kVAr, LF	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)
1-12 May	22.6	18.9
15-26 May	9.7	26.2
29 May-9 June	2.7	29.6
12-23 June	21.4	44.9
MAPE (%)	**14.1**	**29.9**

Table 8.35 presents the forecasting results for the one group: **Max kW, Max kVAr, LF**.

Table 8.35: Load Forecast results: **Max kW, Max kVAr, LF**

	Max kW, Max kVAr, LF	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)
1-12 May	40.0	18.9
15-26 May	27.4	26.2
29 May-9 June	21.5	29.6
12-23 June	4.9	44.9
MAPE (%)	**23.5**	**29.9**

Table 8.36 presents the load forecasting results for the one group: **Max kW, Ave PF, LF**.

Table 8.36: Load Forecast results: **Max kW, Ave PF, LF**

	Max kW, Ave PF, LF		**Max kW, Ave PF, LF**	
Fortnight to be forecasted	Cluster Load Forecast for Much Asphalt: APE (%)	Much Asphalt Load Forecast: APE (%)	Cluster Load Forecast for Denel-Firgrove: APE (%)	Denel-Firgrove Load Forecast: APE (%)
1-12 May	117.8	18.9	68.7	40.1
15-26 May	90.0	26.2	72.9	53.5
29 May-9 June	72.6	29.6	76.4	65.0
12-23 June	27.4	44.9	78.9	74.9
MAPE (%)	**77.0**	**29.9**	**74.2**	**58.4**

The above load forecast results are comparisons between the load forecasts for the individual customers and the load forecasts for the same customers based on the cluster load forecast. Tables 8.31 and 8.36 show that both customers are identified in that group. Table 8.37 and 8.38 below show the load forecasts results that relate to tables 8.31 and 8.36. The idea is to forecast the total load. The total load is the combined load of the two customers Denel Edms Bkp-Firgrove and Much Asphalt. The total load is first forecasted for the two customers and then the total load is forecasted for the same two customers but based on the each customer's cluster load forecast. Each of the two customer's cluster load forecasts is combined to form the total load forecast. These two forms of total load forecasts are compared in tables 8.37 and 8.38.

Table 8.37: Load Forecast results: **Max kW, Ave kVA, Ave kVAr** and **Max kW, Ave kVA, Max kVA** and **Max kW, Ave kVA, Max kVAr** and **Max kW, Ave kW, Ave kVAr** and **Max kW, Ave kVAr, Max kVA**

	Most Prominent: **Ave kVA & Ave kVAr**	
Fortnight to be forecasted	Sum of Cluster Load Forecasts of Much Asphalt and Denel-Firgrove: APE (%)	Sum of Load Forecasts of Much Asphalt and Denel-Firgrove: APE (%)
1-12 May	9.7	36.7
15-26 May	22.4	49.6
29 May-9 June	32.4	60.4
12-23 June	39.3	69.4
MAPE (%)	**25.9**	**54.0**

Table 8.38: Load Forecast results: **Max kW, Ave PF, LF**

	Max kW, Ave PF, LF	
Fortnight to be forecasted	Sum of Cluster Load Forecasts of Much Asphalt and Denel-Firgrove: APE (%)	Sum of Load Forecasts of Much Asphalt and Denel-Firgrove: APE (%)
1-12 May	44.5	36.7
15-26 May	52.2	49.6
29 May-9 June	58.2	60.4
12-23 June	62.3	69.4
MAPE (%)	**54.3**	**54.0**

The load-forecast errors in the above tables are high. The large errors of around 70 % resulted are due to the poor clustering. Consider for example the load parameter group Max kW, Ave PF, LF: There are only 2 clusters formed as identified in figure 8.21 and the corresponding table 8.29 in the appendix. All the customers but one are grouped to one cluster. The other cluster only has one customer in it. The customers in the cluster that has most of the customers vary widely with respect to their power usage. The Max kW in the cluster varies from as low as 181.8 kW to as high as 4104 kW. The idea behind clustering is to have similar customers cluster together based on their Max kW. Clearly this did not occur for the load parameter group Max kW, Ave PF, LF as indicated by a Max kW of 181.8 kW and 4104 kW and therefore the load forecast for the cluster was based on customer load usage profiles that were not similar at all. This resulted in the poor load forecast for the cluster.

The MAPE in tables 8.30, 8.31, 8.33, 8.36 and 8.38 shows that the cluster analysis approach to load forecasting is worse than the normal approach to load forecasting that only uses regression analysis. On the other hand in tables 8.31, 8.32, 8.34, 8.35 and 8.37 the cluster analysis approach to load forecasting shows an improvement over the normal approach to load forecasting. The results are mixed.

This section investigates the load parameters or combinations of load parameters that were used in tables 8.31, 8.32, 8.34, 8.35 and 8.37 that resulted in the cluster analysis approach providing a more effective load forecast. From the tables, the load parameters Ave PF and Ave kVAr are the main load parameters since it appears more times than any other load parameters in the load parameter groups. There are 10 load parameter groups to which the load parameters Ave PF and Ave kVAr are apart of. The 10 load parameter groups are shown in table 8.39.

Table 8.39: The 10 Load Parameter groups with Ave PF and Ave kVAr

	Load Parameter Groups	
Max kW	**Ave kVAr**	Ave kVA
Max kW	**Ave kVAr**	Ave kW
Max kW	**Ave kVAr**	Max kVA
Max kW	**Ave PF**	Ave kVA
Max kW	**Ave PF**	Ave kW
Max kW	**Ave PF**	**Ave kVAr**
Max kW	**Ave PF**	Max kVA
Max kW	**Ave PF**	Max kVAr
Max kW	**Ave kVAr**	Max kVAr
Max kW	**Ave kVAr**	LF

Interestingly there is one load parameter group that has the Ave PF and Ave kVAr in one group i.e. Max kW, Ave PF, Ave kVAr. Instead of the 10 load parameter groups that could be used, only the one load parameter group (i.e. Max kW, Ave PF, Ave kVAr) can be used by the load forecaster for the cluster analysis model.

8.4 The Data Files for Chapter 8

In this section the files that were used in the cluster analysis process are discussed.
The first file is the file that calculated the load parameters of the 14 customers. This file is called Calculation_of_Load_Parameters.xls and is a Microsoft Excel spreadsheet. The second file is the actual calculation of the load forecasts of the customers in each cluster for each group of cluster analysis. This is a Microsoft Excel file called Cluster_Analysis_Load_Forecasting.xls. The third file shows the load parameters in the Statistica software, which will be used to determine the clusters. This is a Statistica file called Load_Parameters_of_customers.sta. In this file the unstandardized load parameters are presented for each of the 12 customers and also the standardized load parameters are presented. The standardized load parameters are used in the analysis. All these files are on a disk.

Chapter 9: Conclusions

9 Content

This chapter draw conclusions based on the results or findings of the thesis. The conclusions are linked to the objectives that are presented in chapter one.

9.1 The ANN Model for Long-term Load Forecasting

This section discusses a long-term load-forecasting model using ANN technology.

9.1.1 The ANN and Regression Model using the Data 21 February – 28 April 2000

This section discusses the ANN and Regression model using the load and weather data from 21 February – 28 April 2000 to develop the weather-load model.

The Linear Regression model proved to be more accurate than the ANN model. The MAPE of the Linear Regression model was 4.2 % while the MAPE of the ANN model ranged from 22.3 % to 23.8 %. There is a maximum difference of 19.6 % in the accuracy between the two models. This proved that ANN technology could not be used to develop an effective long-term load-forecasting model. The ANN model is tested using extra data to develop the weather-load model. This is discussed in section 9.1.2.

9.1.2 The ANN and Regression Model using the Data 21 February – 26 May 2000

The data from 21 February – 26 May 2000 is used to develop the weather-load model. The reason for using extra data is because ANN technology performs better when more data is used to train the network. The significance of the data 21 February – 26 May 2000 is that it corresponds to a seasonal period i.e. the autumn season (March, April and May). The significance of a season is that annual seasonal peak loads are forecasted in long-term load forecasting models i.e. Clayton et al (1973) and Davey et al (1973). Seasonal load and weather data is used to forecast the future annual seasonal peak load. The autumn data set is therefore the correct data size that will be used in the ANN model and the Regression model to develop the weather-load model. The MAPE of the Regression model is 2.7 % whereas the MAPE of the ANN model is 2.6 %. An effective long-term model can be developed using ANN technology based on this customer load data set. There is only a 0.1 % difference between the ANN model and the Regression model. The ANN model implements a non-iterative load forecasting methodology which can be used to forecast long-term load. The ANN model used a limited data set and little or no load growth can be assumed for the limited data set. Long-term load growth has to be taken into account for a long-term load forecast. Long-term load growth is incorporated into the ANN model by considering the annual peak loads of the years that precede the forecasting period. These annual peak loads are projected (trended) forward to future years as discussed in section 4.1.2.

Since the linear regression analysis load forecast was successful, the next step was to test how a polynomial non-linear regression analysis would succeed for the same customer (AECIPF) load data set. The results show that the 4th degree polynomial produced the most accurate load forecast compared to the linear regression and ANN analysis for customer AECIPF. The MAPE for customer AECIPF for the linear regression analysis is 2.7 %, for the ANN analysis is 2.6 % and for the 4th degree polynomial the MAPE is 2.5 %. The load-forecast accuracy is very close between the linear, polynomial and ANN analysis for this customer AECIPF load data set. Since the load-forecast accuracy is so close for this customer AECIPF, the next step is to perform the linear, polynomial and ANN analysis load forecast for another 3 customer load data sets namely Corobrick, Dolphin Beach and CM Milerton Sludge to determine, which analysis produces the most accurate load forecast that can be used for long-term load forecasting. For each of the four customer load data sets the polynomial and ANN analysis produces the most accurate load forecast. It is interesting too note that it is not the same degree polynomial that produces the most accurate load forecast for the four customers. For customer AECIPF it is the 4th degree polynomial, for customer Corobrick it is the 2nd degree polynomial, for Dolphin Beach it is the 3rd degree polynomial and for customer CM Milerton Sludge it is the 2nd degree polynomial. For three of the customers (i.e. AECIPF, Corobrick and CM Milerton Sludge) the polynomial analysis is more accurate than the ANN analysis and for the other remaining customer (i.e. Dolphin Beach) the ANN analysis is more accurate than the polynomial analysis.

From these analyses it is the conclusion that the polynomial and ANN analysis could be used to produce a more effective long-term load forecasting model than the literature, Clayton et al (1973) and Davey et al (1973), long-term model that uses linear regression since they produce more accurate load forecasts than the linear regression analysis. The ANN analysis closely matches the accuracy of the polynomial regression analysis since the ANN analysis is also a non-linear regression analysis in that it uses non-linear transfer functions. For some customer data sets the polynomial analysis will be more accurate than the ANN analysis and for other customer data sets the ANN analysis will be more accurate than the polynomial analysis. However to determine the most accurate degree polynomial for a customer data set, various degree polynomials have to be tested since the same degree polynomial will not always produce the most accurate load forecast. The advantage that the ANN analysis has is that only one ANN analysis has to be performed for a customer data set.

9.1.3 Effects of Historical Load Data in the ANN Model

The thesis did not consider historical load data in the ANN model because historical load data is outside of the scope and objectives of the thesis since historical load data can be years of data and therefore is not a limited data set. The thesis investigated how best to utilize the available limited load data to produce an effective long-term load forecasting model. The load-forecast model was designed and modified for the limited data set.

The ANN weather-load model used in this thesis only considered the load as a function of weather. A typical ANN model in the literature (e.g. Chen et al (1992) and Dash et al (1994)) uses load as a function of weather and lagged load. This suggests that there is a pattern in the load data that would assist or improve the load forecast. In long-term load forecasts seasonal peak loads are forecasted. The load pattern is unique to each season and there is therefore a pattern in the load behavior in each season. Further analyses can be performed to determine if the use of historical load data in the ANN model would improve the accuracy of the long-term load forecast.

9.2 The Application of the Cluster Analysis in Long-term Load Forecasting

This section discusses a long-term load-forecasting model using the cluster analysis tool.

The second objective is to test whether classifying or characterising customers into groups improves the load forecast in the long-term load-forecasting model. The cluster analysis tool was used to classify the customers into groups. The conclusion is that classifying customers into groups do improve the load forecast in some cases and therefore not in all cases. The cluster analysis load forecast results are mixed. In some cases the cluster analysis approach to load forecasting improves the load forecast and in other cases it makes it worse. Analysing the improved load forecasts that resulted from the cluster analysis shows that the load parameters Ave PF and Ave kVAr are prominent in the improved load forecasts and therefore the load forecaster can only use the 10 load parameter groups identified in table 8.39 to obtain an improved load forecast.

The cluster analysis load forecast was performed using a limited data set. For a limited data set little or no load growth can be assumed. Load growth has to be considered for long-term load forecasting. The cluster analysis approach can be applied to long-term load forecasting. The reason is that the load forecasting methodology used is a non-iterative process as discussed in sections 4.3.3.1 and 4.4. Long-term load growth is considered by using the annual peak loads of the years that precede the forecasting period. These annual peak loads are projected (trended) forward to future years.

Chapter 10: References and Bibliography

References

Books and Manuals

1. Demuth H and Beale M, Neural Network Toolbox User's Guide, Version 4 (The Math Works Inc., 2000)
2. Everitt B, Cluster Analysis (London: Heinemann Educational Books Ltd, 1974)
3. Glover J. D and Sarma M, Power System Analysis and Design, Second Edition, (Boston: PWS Publishing Company, 1994)
4. Patterson D.W, Artificial Neural Networks: Theory and Applications (Singapore: Prentice Hall, 1996)
5. Romesburg H. C, Cluster Analysis for Researchers (California: Lifetime Learning Publications, 1984)
6. Sullivan R. L, Power System Planning (New York: Hemisphere Publishing Corporation, 1977)

Journal Articles

1. Abou-Hussien M. S, Kandil M. S, Tantawy M. A and Farghal S. A, "An accurate Model for short-term load forecasting", IEEE Transactions on Power Apparatus and Systems, vol. PAS-100, No. 9, September 1981, pp. 4158-4164.
2. Bashir Z and El-Hawary M. E, "Short term load forecasting by using wavelet neural networks", 2000 Canadian conference on electrical and computer engineering. Conf. Proc. Navigating to a new era, IEEE, Piscataway, NJ, USA, vol. 1, pp. 163-166.
3. Chen S, Yu D. C and Moghaddamjo A. R, "Weather sensitive short-term load forecasting using nonfully connected artificial neural network", Trans. on Power Systems, vol. 7, No. 3, August 1992, pp. 1098-1105.
4. Clayton R. E, Priest K.W and Truax C. J, "Load model sensitive to weather data", Electrical World, 15 March 1973, pp. 96-98.
5. Corpening S. L, Reppen N. D and Ringlee R. J, "Experience with weather sensitive load models for short and long-term forecasting", IEEE Trans. PAS, vol. PAS-92, No. 6, Nov/Dec 1973, pp. 1966-1972.

6. Dash P. K, Liew A. C, Rahman S, "Fuzzy neural network and fuzzy expert system for load forecasting", IEE Proc.- Gener. Transm. Distrib., vol. 143, No. 1, January 1996, pp. 106-114.
7. Dash P. K, Liew A. C, Rahman S, Satpathy J. K and Ramakrishna G, "An improved neural network approach for weather sensitive short-term load forecasting", Int. Jnl. of Engineering Intelligent Systems for Elec. Engineering and Communication, vol. 2, No. 3, Sept. 1994, pp. 185-199.
8. Davey J, Saacks J. J, Cunningham G. W and Priest K. W, "Practical application of weather sensitive load forecasting to system planning", IEEE Trans. PAS, vol. PAS-92, No. 3, May/June 1973, pp. 971-977.
9. Gaunt C. T, Personal Communication
10. Gupta P. C and Yamada K, "Adaptive short-term forecasting of hourly loads using weather information", IEEE Trans. PAS, vol. PAS-91, Sept./Oct. 1972, pp. 2085-2094.
11. Heinemann G. T, Nordman D. A and Plant E. C, "The relationship between Summer weather and Summer loads – A Regression Analysis", IEEE Transactions on Power Apparatus and Systems, vol. PAS-85, No. 11, November 1966, pp. 1144-1154.
12. Khotanzad A, Afkhami-Rohani R, Lu T, Abaye A, Davis M and Maratukulam D. J, "ANNSTLF-A neural-network-based electric load forecasting system", IEEE Trans. on Neural Networks, vol. 8, No. 4, July 1997, pp. 835-846.
13. Khotanzad A, Hwang R, Abaye A and Maratukulam D, "An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities", IEEE Trans. on Power Systems, vol. 10, No. 3, August 1995, pp. 1716-1722.
14. Stanton K. N, "Medium-range weekly and seasonal peak demand forecasting by probability methods", IEEE, PAS, vol. PAS-90, No. 3, May/June 1971, pp. 1183-1189.
15. Willis H. L, Tram H.N, Northcote-Green J. E. D and Brooks C. L, "Load Forecasting data and data base development for distribution planning", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 11, November 1983, pp. 3660-3666.

16. Willis H. L and Tram H, "A cluster based V.A.I method for distribution load forecasting", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No. 8, August 1983, pp. 2677-2684.
17. Willis H. Lee and Northcote-Green James E. D, "Spatial electric load forecasting: A tutorial review", Proceedings of the IEEE, Vol. 71, No. 2, February 1983, pp. 232-253.

Bibliography

Journal Articles

1. Dash P. K, Liew A. C, Rahman S and Dash S, "Fuzzy and Neuro-fuzzy computing models for electric load forecasting", Engineering Applications of artificial intelligence, vol. 8, No. 4, August 1995, pp. 423-433.
2. Dash P. K, Liew A. C and Rahman S, "A comparative study of load forecasting models using fuzzy neural networks", ISAP '94. International Conf. on intelligent system application to power systems EC2, Nanterre, Cedex, France 1994, vol. 2, pp. 865-872.
3. Dash P. K, Liew A. C and Ramakrishna G, "Power-demand forecasting using a neural network with an adaptive learning algorithm", IEE Proc.-Gener. Transm. Distrib., vol. 142, No. 6, November 1995, pp. 560-568.
4. Elkateb M. M and Kwok-Wai M, "Comparative study for load forecasting using artificial neural networks "ANN " and a statistical package", 3rd International conference on advances in power system control, operation and management, IEEE, Hong Kong 1995, vol. 2, pp. 527-531.
5. Ho K, Hsu Y and Yang C, "Short term load forecasting using a multilayer neural network with an adaptive learning algorithm", Trans. on Power Systems, vol. 7, No. 1, Feb. 1992, pp. 141-149.
6. Johns A. T, El-Alayly A. A and Youssef M. T, "An Adaptive system for short term load forecasting using ANN", 34th Universities Power Engineering Conference. Univ. Leicester, UK, 1999, vol. 1, pp. 213-216.

7. Karaki S. H, "Weather sensitive short-term load forecasting using artificial neural networks and time series", International Journal of Power and Energy Systems, vol. 19, No. 3, 1999, pp. 251-256.
8. Karaki S. H, "A neural network weather load model for short-term load forecasting", Proceedings of the 16th IASTED Int. Conf. Modelling, Identification and Control IASTED-ACTA Press, ANAHEIM, CA, USA 1997, pp. 220-224.
9 Riad A. M, Gharraf M. K and Mahmoud A. K. M, "Short term load forecasting using artificial neural networks", Egyptian Computer Journal, vol. 22, No. 2, Dec. 1994, pp. 147-162.

Appendix

Appendix I

The Matlab code for the ANN Model

The Matlab code that is used to implement the ANN Model is presented in this appendix. The Matlab code for the ANN Model is also on the disk. The file is called ANNModel.m.

```
DT14pm = [26.8 24.5 26.2 30 26.2 31.4 25.9 23.5 26.7 18.9 21 27.3 26.5
22.1 23.1 22.5 18 22.7 21.5 22.8 30.7 29.1 24.8 24.5 23.5 26.8 27.1
25.8 23.1 25.6 19.9 24.8 25.8 28.5 32.1 23.5 23.1 21.2 21.9 21.1 22.7
26.4 29.1 23.8 19.7 18.4 15.4 17 20.7 24.6];
%% DT14pm is the Dry-bulb temperature at 14H:00pm in (°C)
%% This is data from 21 Feb. 2000 to 28 April 2000, only weekdays.

DPL = [3750 3564 3708 3894 3930 3882 3606 3750 3516 3612 3570 3864 3798
3822 3708 3450 3318 3666 3684 3696 3738 3912 3750 3576 3972 3918 3906
3888 3906 4104 3510 3678 3468 3540 3660 3702 3780 3696 3750 3666 3978
3828 3756 3762 2034 3108 3150 3204 3498 3522];
%% DPL is Daily Peak Load in (kW)
%% This is data from 21 Feb. 2000 to 28 April 2000, only weekdays.
%% The DT14pm and DPL data correspond.

%%---------------------------------------------------------------
%% NOW TO DESIGN THE ANN NETWORK FOR LOAD FORECASTING
%% INPUT:DT14PM AND OUTPUT:DPL

%% Both the input and output values are scaled for efficient training.
%% The _n represents normalised data. The function premnmx does this.
[DT14pm_n, minDT14pm, maxDT14pm, DPL_n, minDPL, maxDPL] =
premnmx(DT14pm, DPL);

%% The ANN network used is the feedforward network i.e. newff
net = newff( minmax(DT14pm_n), [30 1], {'tansig','purelin'},'trainbr');
net.trainParam.show = 50;
net.trainParam.epochs = 800;

net = init(net); %% To initialise the networks weights and biases.

%% Now the training starts with the normalised data.
%% The normalised data is in the range [-1,1].
[net, tr] = train(net, DT14pm_n, DPL_n);

%%---------------------------------------------------------------
%% Now forecast load for period 1-12 May 2000.
%% Use 20 years of historical weather data, DT14pm, for the
%% period 1-12 May as input.
May1_12_DT14pm_1980 =[24.2 18.9 20.5 20.3 20 16.4 21.8 17.3 17.6 17.8];
May1_12_DT14pm_1981 =[19.2 20.9 25.2 18.5 18 23.9 18 20.1 16.8 17.8];
May1_12_DT14pm_1982 =[19.4 21.4 20.3 20.5 19 26.2 18.5 20.5 19.3 19.2];
May1_12_DT14pm_1983 =[21.5 16.8 19.7 30.3 24 23.1 21.1 15.2 17 22.3];
May1_12_DT14pm_1984 =[19.9 20.7 18.5 19.8 15 20.4 17.1 18.6 16 19.5];
```

```
May1_12_DT14pm_1985 =[19.8 19.5 22.5 22 20 19.8 21 20.4 21.5 18.8];
May1_12_DT14pm_1986 =[21.2 22 18.1 23.1 25 18.4 20.4 20.6 20.1 19];
May1_12_DT14pm_1987 =[26.2 21.5 23.2 19 17 18.4 23.5 17.8 16.7 16.1];
May1_12_DT14pm_1988 =[24.7 30.8 17.6 18 21 25.1 18.9 29.3 27 25.3];
May1_12_DT14pm_1989 =[27.1 20.9 18.7 19.9 23 21.1 21.5 19 15 12.9];
May1_12_DT14pm_1990 =[17.6 17.5 16.8 18.8 20 23.1 18.3 19.5 25.8 21];
May1_12_DT14pm_1991 =[21.4 21.5 21.5 22.2 21 22.5 22.8 21.1 23.8 26.3];
May1_12_DT14pm_1992 =[15.5 15.4 18.5 18.3 20 19.3 19.7 21.2 19.7 27.7];
May1_12_DT14pm_1993 =[25.2 18.4 15.5 18.6 18 18.4 19.3 20.6 15.5 17.9];
May1_12_DT14pm_1994 =[17.8 19.4 22.3 14.8 18 24.2 20 18.9 16.2 20.3];
May1_12_DT14pm_1995 =[20.7 19.5 27.3 20.1 14 18 19.9 17.7 19.5 19.9];
May1_12_DT14pm_1996 =[20.9 18.1 20.1 15.6 18 26.4 20.9 18.8 18.9 17.5];
May1_12_DT14pm_1997 =[20.4 20.1 19.7 22.2 25 27.4 21.5 24.9 28.6 26.7];
May1_12_DT14pm_1998 =[20.4 19.4 18.6 21.1 18 18.7 17.7 16.5 15.7 17.7];
May1_12_DT14pm_1999 =[20 17.8 15.4 17.7 18 20.7 18.4 19 22.1 20.6];

%% The above data must be normalised.
May1_12_DT14pm_1980_n=tramnmx(May1_12_DT14pm_1980,minDT14pm,maxDT14pm);
May1_12_DT14pm_1981_n=tramnmx(May1_12_DT14pm_1981,minDT14pm,maxDT14pm);
May1_12_DT14pm_1982_n=tramnmx(May1_12_DT14pm_1982,minDT14pm,maxDT14pm);
May1_12_DT14pm_1983_n=tramnmx(May1_12_DT14pm_1983,minDT14pm,maxDT14pm);
May1_12_DT14pm_1984_n=tramnmx(May1_12_DT14pm_1984,minDT14pm,maxDT14pm);
May1_12_DT14pm_1985_n=tramnmx(May1_12_DT14pm_1985,minDT14pm,maxDT14pm);
May1_12_DT14pm_1986_n=tramnmx(May1_12_DT14pm_1986,minDT14pm,maxDT14pm);
May1_12_DT14pm_1987_n=tramnmx(May1_12_DT14pm_1987,minDT14pm,maxDT14pm);
May1_12_DT14pm_1988_n=tramnmx(May1_12_DT14pm_1988,minDT14pm,maxDT14pm);
May1_12_DT14pm_1989_n=tramnmx(May1_12_DT14pm_1989,minDT14pm,maxDT14pm);
May1_12_DT14pm_1990_n=tramnmx(May1_12_DT14pm_1990,minDT14pm,maxDT14pm);
May1_12_DT14pm_1991_n=tramnmx(May1_12_DT14pm_1991,minDT14pm,maxDT14pm);
May1_12_DT14pm_1992_n=tramnmx(May1_12_DT14pm_1992,minDT14pm,maxDT14pm);
May1_12_DT14pm_1993_n=tramnmx(May1_12_DT14pm_1993,minDT14pm,maxDT14pm);
May1_12_DT14pm_1994_n=tramnmx(May1_12_DT14pm_1994,minDT14pm,maxDT14pm);
May1_12_DT14pm_1995_n=tramnmx(May1_12_DT14pm_1995,minDT14pm,maxDT14pm);
May1_12_DT14pm_1996_n=tramnmx(May1_12_DT14pm_1996,minDT14pm,maxDT14pm);
May1_12_DT14pm_1997_n=tramnmx(May1_12_DT14pm_1997,minDT14pm,maxDT14pm);
May1_12_DT14pm_1998_n=tramnmx(May1_12_DT14pm_1998,minDT14pm,maxDT14pm);
May1_12_DT14pm_1999_n=tramnmx(May1_12_DT14pm_1999,minDT14pm,maxDT14pm);

%% Now for the simulation. The output is also normalised.
May1_12_Load_1980_n = sim(net,May1_12_DT14pm_1980_n);
May1_12_Load_1981_n = sim(net,May1_12_DT14pm_1981_n);
May1_12_Load_1982_n = sim(net,May1_12_DT14pm_1982_n);
May1_12_Load_1983_n = sim(net,May1_12_DT14pm_1983_n);
May1_12_Load_1984_n = sim(net,May1_12_DT14pm_1984_n);
May1_12_Load_1985_n = sim(net,May1_12_DT14pm_1985_n);
May1_12_Load_1986_n = sim(net,May1_12_DT14pm_1986_n);
May1_12_Load_1987_n = sim(net,May1_12_DT14pm_1987_n);
May1_12_Load_1988_n = sim(net,May1_12_DT14pm_1988_n);
May1_12_Load_1989_n = sim(net,May1_12_DT14pm_1989_n);
May1_12_Load_1990_n = sim(net,May1_12_DT14pm_1990_n);
May1_12_Load_1991_n = sim(net,May1_12_DT14pm_1991_n);
May1_12_Load_1992_n = sim(net,May1_12_DT14pm_1992_n);
May1_12_Load_1993_n = sim(net,May1_12_DT14pm_1993_n);
May1_12_Load_1994_n = sim(net,May1_12_DT14pm_1994_n);
May1_12_Load_1995_n = sim(net,May1_12_DT14pm_1995_n);
May1_12_Load_1996_n = sim(net,May1_12_DT14pm_1996_n);
May1_12_Load_1997_n = sim(net,May1_12_DT14pm_1997_n);
```

```
May1_12_Load_1998_n = sim(net,May1_12_DT14pm_1998_n);
May1_12_Load_1999_n = sim(net,May1_12_DT14pm_1999_n);

%% The output: derived loads must be un-normalised.
May1_12_Load_1980 = postmnmx(May1_12_Load_1980_n, minDPL, maxDPL);
May1_12_Load_1981 = postmnmx(May1_12_Load_1981_n, minDPL, maxDPL);
May1_12_Load_1982 = postmnmx(May1_12_Load_1982_n, minDPL, maxDPL);
May1_12_Load_1983 = postmnmx(May1_12_Load_1983_n, minDPL, maxDPL);
May1_12_Load_1984 = postmnmx(May1_12_Load_1984_n, minDPL, maxDPL);
May1_12_Load_1985 = postmnmx(May1_12_Load_1985_n, minDPL, maxDPL);
May1_12_Load_1986 = postmnmx(May1_12_Load_1986_n, minDPL, maxDPL);
May1_12_Load_1987 = postmnmx(May1_12_Load_1987_n, minDPL, maxDPL);
May1_12_Load_1988 = postmnmx(May1_12_Load_1988_n, minDPL, maxDPL);
May1_12_Load_1989 = postmnmx(May1_12_Load_1989_n, minDPL, maxDPL);
May1_12_Load_1990 = postmnmx(May1_12_Load_1990_n, minDPL, maxDPL);
May1_12_Load_1991 = postmnmx(May1_12_Load_1991_n, minDPL, maxDPL);
May1_12_Load_1992 = postmnmx(May1_12_Load_1992_n, minDPL, maxDPL);
May1_12_Load_1993 = postmnmx(May1_12_Load_1993_n, minDPL, maxDPL);
May1_12_Load_1994 = postmnmx(May1_12_Load_1994_n, minDPL, maxDPL);
May1_12_Load_1995 = postmnmx(May1_12_Load_1995_n, minDPL, maxDPL);
May1_12_Load_1996 = postmnmx(May1_12_Load_1996_n, minDPL, maxDPL);
May1_12_Load_1997 = postmnmx(May1_12_Load_1997_n, minDPL, maxDPL);
May1_12_Load_1998 = postmnmx(May1_12_Load_1998_n, minDPL, maxDPL);
May1_12_Load_1999 = postmnmx(May1_12_Load_1999_n, minDPL, maxDPL);

%% Now take the maximum of the loads for each row.
May1_12_Max_Load_1980 = max(May1_12_Load_1980);
May1_12_Max_Load_1981 = max(May1_12_Load_1981);
May1_12_Max_Load_1982 = max(May1_12_Load_1982);
May1_12_Max_Load_1983 = max(May1_12_Load_1983);
May1_12_Max_Load_1984 = max(May1_12_Load_1984);
May1_12_Max_Load_1985 = max(May1_12_Load_1985);
May1_12_Max_Load_1986 = max(May1_12_Load_1986);
May1_12_Max_Load_1987 = max(May1_12_Load_1987);
May1_12_Max_Load_1988 = max(May1_12_Load_1988);
May1_12_Max_Load_1989 = max(May1_12_Load_1989);
May1_12_Max_Load_1990 = max(May1_12_Load_1990);
May1_12_Max_Load_1991 = max(May1_12_Load_1991);
May1_12_Max_Load_1992 = max(May1_12_Load_1992);
May1_12_Max_Load_1993 = max(May1_12_Load_1993);
May1_12_Max_Load_1994 = max(May1_12_Load_1994);
May1_12_Max_Load_1995 = max(May1_12_Load_1995);
May1_12_Max_Load_1996 = max(May1_12_Load_1996);
May1_12_Max_Load_1997 = max(May1_12_Load_1997);
May1_12_Max_Load_1998 = max(May1_12_Load_1998);
May1_12_Max_Load_1999 = max(May1_12_Load_1999);


%% Now take the mean of the above. This is the load forecast.
%% This is displayed as output.

May1_12_Mean = (May1_12_Max_Load_1980 + May1_12_Max_Load_1981 +
May1_12_Max_Load_1982 + May1_12_Max_Load_1983 + May1_12_Max_Load_1984 +
May1_12_Max_Load_1985 + May1_12_Max_Load_1986 + May1_12_Max_Load_1987 +
May1_12_Max_Load_1988 + May1_12_Max_Load_1989 + May1_12_Max_Load_1990 +
May1_12_Max_Load_1991 + May1_12_Max_Load_1992 + May1_12_Max_Load_1993 +
May1_12_Max_Load_1994 + May1_12_Max_Load_1995 + May1_12_Max_Load_1996 +
```

```
May1_12_Max_Load_1997 + May1_12_Max_Load_1998 +
May1_12_Max_Load_1999)/20;

%%---------------------------------------------------------------
%% Now forecast load for period 15-26 May 2000.
%% Use 20 years of historical weather data, DT14pm, for the
%% period 15-26 May as input.
May15_26_DT14pm_1980=[17.2 14.4 18.8 17.6 22.7 24.3 13.8 12.9 14 15.4];
May15_26_DT14pm_1981=[29.2 16.7 20 19.5 19.7 23.4 29.9 28.9 17.8 17];
May15_26_DT14pm_1982=[19.4 21.2 19.4 18.5 18.7 18.2 17.5 16.8 15.2 16.2
May15_26_DT14pm_1983=[11.9 15.9 19.2 24.8 13.2 16.7 16.4 17.2 19.1 15.1
May15_26_DT14pm_1984=[16.3 12.7 13.1 14.8 19.8 25.4 23.4 25.5 14.2 16];
May15_26_DT14pm_1985=[28 19.7 22.8 21.6 20.8 18.6 18.3 18.6 22.7 20.8];
May15_26_DT14pm_1986=[17.6 18 22.6 20.2 20.4 19.3 17.8 19.3 22.8 32.7];
May15_26_DT14pm_1987=[16.5 16.6 18.7 18 19.4 25.7 19.9 30.1 20.7 22.8];
May15_26_DT14pm_1988=[23.4 17.3 16.6 17 16.4 18.7 15.4 18.3 18.6 18.1];
May15_26_DT14pm_1989=[18.4 24.7 19.6 17.9 15.4 20.9 21.6 18.4 17.3 16.4
May15_26_DT14pm_1990=[29.9 30.8 21.1 17.9 21.2 13.6 17.1 16.7 17.6 19.4
May15_26_DT14pm_1991=[28.8 23.8 29.2 26.4 18 17 17.8 13.4 17.4 18];
May15_26_DT14pm_1992=[15 15.1 17.3 18.5 18.6 19.5 16.5 16.2 15.9 15.4];
May15_26_DT14pm_1993=[18.4 19.5 19 18.5 17.1 16.4 17.7 16.5 13 14.6];
May15_26_DT14pm_1994=[18.1 19.5 21 18.8 16.7 15 17.2 19.1 17.7 16.4];
May15_26_DT14pm_1995=[28.5 19.1 24.8 25.2 23.9 14.8 16.7 17.8 16.2 18.5
May15_26_DT14pm_1996=[19.4 21.2 24.9 19.7 22.9 18.5 16.5 20.8 21.7 24.9
May15_26_DT14pm_1997=[20 21.5 20.9 18.6 19 19.1 18.7 17.1 20.3 15.9];
May15_26_DT14pm_1998=[22.2 28.3 29.7 16 18 20.9 20.1 22.5 19 18.5];
May15_26_DT14pm_1999=[22.9 26.9 26 23.3 21.4 14.9 15.9 18.7 22.6 16.8];

%% The above data must be normalised.
May15_26_DT14pm_1980_n = tramnmx(May15_26_DT14pm_1980, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1981_n = tramnmx(May15_26_DT14pm_1981, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1982_n = tramnmx(May15_26_DT14pm_1982, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1983_n = tramnmx(May15_26_DT14pm_1983, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1984_n = tramnmx(May15_26_DT14pm_1984, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1985_n = tramnmx(May15_26_DT14pm_1985, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1986_n = tramnmx(May15_26_DT14pm_1986, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1987_n = tramnmx(May15_26_DT14pm_1987, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1988_n = tramnmx(May15_26_DT14pm_1988, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1989_n = tramnmx(May15_26_DT14pm_1989, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1990_n = tramnmx(May15_26_DT14pm_1990, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1991_n = tramnmx(May15_26_DT14pm_1991, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1992_n = tramnmx(May15_26_DT14pm_1992, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1993_n = tramnmx(May15_26_DT14pm_1993, minDT14pm,
maxDT14pm);
```

```
May15_26_DT14pm_1994_n = tramnmx(May15_26_DT14pm_1994, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1995_n = tramnmx(May15_26_DT14pm_1995, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1996_n = tramnmx(May15_26_DT14pm_1996, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1997_n = tramnmx(May15_26_DT14pm_1997, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1998_n = tramnmx(May15_26_DT14pm_1998, minDT14pm,
maxDT14pm);
May15_26_DT14pm_1999_n = tramnmx(May15_26_DT14pm_1999, minDT14pm,
maxDT14pm);

%% Now for the simulation. The output is also normalised.
May15_26_Load_1980_n = sim(net,May15_26_DT14pm_1980_n);
May15_26_Load_1981_n = sim(net,May15_26_DT14pm_1981_n);
May15_26_Load_1982_n = sim(net,May15_26_DT14pm_1982_n);
May15_26_Load_1983_n = sim(net,May15_26_DT14pm_1983_n);
May15_26_Load_1984_n = sim(net,May15_26_DT14pm_1984_n);
May15_26_Load_1985_n = sim(net,May15_26_DT14pm_1985_n);
May15_26_Load_1986_n = sim(net,May15_26_DT14pm_1986_n);
May15_26_Load_1987_n = sim(net,May15_26_DT14pm_1987_n);
May15_26_Load_1988_n = sim(net,May15_26_DT14pm_1988_n);
May15_26_Load_1989_n = sim(net,May15_26_DT14pm_1989_n);
May15_26_Load_1990_n = sim(net,May15_26_DT14pm_1990_n);
May15_26_Load_1991_n = sim(net,May15_26_DT14pm_1991_n);
May15_26_Load_1992_n = sim(net,May15_26_DT14pm_1992_n);
May15_26_Load_1993_n = sim(net,May15_26_DT14pm_1993_n);
May15_26_Load_1994_n = sim(net,May15_26_DT14pm_1994_n);
May15_26_Load_1995_n = sim(net,May15_26_DT14pm_1995_n);
May15_26_Load_1996_n = sim(net,May15_26_DT14pm_1996_n);
May15_26_Load_1997_n = sim(net,May15_26_DT14pm_1997_n);
May15_26_Load_1998_n = sim(net,May15_26_DT14pm_1998_n);
May15_26_Load_1999_n = sim(net,May15_26_DT14pm_1999_n);


%% The output: derived loads must be un-normalised.
May15_26_Load_1980 = postmnmx(May15_26_Load_1980_n, minDPL, maxDPL);
May15_26_Load_1981 = postmnmx(May15_26_Load_1981_n, minDPL, maxDPL);
May15_26_Load_1982 = postmnmx(May15_26_Load_1982_n, minDPL, maxDPL);
May15_26_Load_1983 = postmnmx(May15_26_Load_1983_n, minDPL, maxDPL);
May15_26_Load_1984 = postmnmx(May15_26_Load_1984_n, minDPL, maxDPL);
May15_26_Load_1985 = postmnmx(May15_26_Load_1985_n, minDPL, maxDPL);
May15_26_Load_1986 = postmnmx(May15_26_Load_1986_n, minDPL, maxDPL);
May15_26_Load_1987 = postmnmx(May15_26_Load_1987_n, minDPL, maxDPL);
May15_26_Load_1988 = postmnmx(May15_26_Load_1988_n, minDPL, maxDPL);
May15_26_Load_1989 = postmnmx(May15_26_Load_1989_n, minDPL, maxDPL);
May15_26_Load_1990 = postmnmx(May15_26_Load_1990_n, minDPL, maxDPL);
May15_26_Load_1991 = postmnmx(May15_26_Load_1991_n, minDPL, maxDPL);
May15_26_Load_1992 = postmnmx(May15_26_Load_1992_n, minDPL, maxDPL);
May15_26_Load_1993 = postmnmx(May15_26_Load_1993_n, minDPL, maxDPL);
May15_26_Load_1994 = postmnmx(May15_26_Load_1994_n, minDPL, maxDPL);
May15_26_Load_1995 = postmnmx(May15_26_Load_1995_n, minDPL, maxDPL);
May15_26_Load_1996 = postmnmx(May15_26_Load_1996_n, minDPL, maxDPL);
May15_26_Load_1997 = postmnmx(May15_26_Load_1997_n, minDPL, maxDPL);
May15_26_Load_1998 = postmnmx(May15_26_Load_1998_n, minDPL, maxDPL);
```

```
May15_26_Load_1999 = postmnmx(May15_26_Load_1999_n, minDPL, maxDPL);

%% Now take the maximum of the loads for each row.
May15_26_Max_Load_1980 = max(May15_26_Load_1980);
May15_26_Max_Load_1981 = max(May15_26_Load_1981);
May15_26_Max_Load_1982 = max(May15_26_Load_1982);
May15_26_Max_Load_1983 = max(May15_26_Load_1983);
May15_26_Max_Load_1984 = max(May15_26_Load_1984);
May15_26_Max_Load_1985 = max(May15_26_Load_1985);
May15_26_Max_Load_1986 = max(May15_26_Load_1986);
May15_26_Max_Load_1987 = max(May15_26_Load_1987);
May15_26_Max_Load_1988 = max(May15_26_Load_1988);
May15_26_Max_Load_1989 = max(May15_26_Load_1989);
May15_26_Max_Load_1990 = max(May15_26_Load_1990);
May15_26_Max_Load_1991 = max(May15_26_Load_1991);
May15_26_Max_Load_1992 = max(May15_26_Load_1992);
May15_26_Max_Load_1993 = max(May15_26_Load_1993);
May15_26_Max_Load_1994 = max(May15_26_Load_1994);
May15_26_Max_Load_1995 = max(May15_26_Load_1995);
May15_26_Max_Load_1996 = max(May15_26_Load_1996);
May15_26_Max_Load_1997 = max(May15_26_Load_1997);
May15_26_Max_Load_1998 = max(May15_26_Load_1998);
May15_26_Max_Load_1999 = max(May15_26_Load_1999);

%% Now take the mean of the above. This is the load forecast.
%% This is displayed as output.
May15_26_Mean = (May15_26_Max_Load_1980 + May15_26_Max_Load_1981 +
May15_26_Max_Load_1982 + May15_26_Max_Load_1983 +
May15_26_Max_Load_1984 + May15_26_Max_Load_1985 +
May15_26_Max_Load_1986 + May15_26_Max_Load_1987 +
May15_26_Max_Load_1988 + May15_26_Max_Load_1989 +
May15_26_Max_Load_1990 + May15_26_Max_Load_1991 +
May15_26_Max_Load_1992 + May15_26_Max_Load_1993 +
May15_26_Max_Load_1994 + May15_26_Max_Load_1995 +
May15_26_Max_Load_1996 + May15_26_Max_Load_1997 +
May15_26_Max_Load_1998 + May15_26_Max_Load_1999)/20;

%%-------------------------------------------------------------
%% Now forecast load for period 29 May-9June 2000.
%% Use 20 years of historical weather data, DT14pm, for the
%% period 29 May-9June as input.
May29_9June_DT14pm_1980=[17.4 21.3 20.4 18.1 22.2 21.5 21.7 17.4 18
18];
May29_9June_DT14pm_1981=[17.8 19 19.3 23.3 14.7 17.8 15.8 15.3 18.8
17.9];
May29_9June_DT14pm_1982=[17.1 18.4 22.8 19.9 17.7 14.3 15.1 17.9 20
17.7];
May29_9June_DT14pm_1983=[15.4 17 16.7 18.6 15.3 12.9 15.3 17 16 15.4];
May29_9June_DT14pm_1984=[21 24.6 27 23.8 17.6 17.8 18.4 22.6 18.4
17.8];
May29_9June_DT14pm_1985=[19.3 18.7 22 26.6 26.3 17.7 18.1 18 18.6
17.2];
May29_9June_DT14pm_1986=[17.6 19.8 19 16.2 15.9 17.5 18.4 18.6 25.7
19.7];
May29_9June_DT14pm_1987=[29 24.3 24.5 18.2 14.8 17.1 16.3 17.8 19.7
18.6];
```

```
May29_9June_DT14pm_1988=[15.2 18.2 18.7 18.8 14.2 15.7 15.4 13.3 13.4
15.7];
May29_9June_DT14pm_1989=[20.9 16.3 15.8 11.8 14.4 17 20.8 16.1 16.8
18.4];
May29_9June_DT14pm_1990=[18.4 18.2 18.7 19.2 17.2 13.9 15.3 15.2 15.3
16.9];
May29_9June_DT14pm_1991=[18.3 13.5 15 18.7 20.5 25.5 15.7 13.7 16.1
17.5];
May29_9June_DT14pm_1992=[17.1 16.9 14.1 14.1 15.7 17.3 20.4 23.2 23.1
26.3];
May29_9June_DT14pm_1993=[16.8 16.3 19.6 21.9 17.8 18.1 17.7 18.2 14.9
17.1];
May29_9June_DT14pm_1994=[15.8 15.3 18.7 22.6 19.8 14.7 15.9 14.2 15.5
17.3];
May29_9June_DT14pm_1995=[22.6 17.8 18.8 17.2 22.6 12.7 15.8 20.3 13.4
17];
May29_9June_DT14pm_1996=[21 17.1 16.1 14.1 16.4 21.8 16.2 12.9 15.4
17.1];
May29_9June_DT14pm_1997=[18.3 21.7 22.2 16.9 18 19.3 16.2 17.2 14.4
13.7];
May29_9June_DT14pm_1998=[19.3 22.2 22.3 15.4 17.8 15.7 13.3 13.9 15.6
19.9];
May29_9June_DT14pm_1999=[18.2 18.9 17.5 20.5 30.4 16.5 17.7 20.7 19.5
27.3];

%% The above data must be normalised.
May29_9June_DT14pm_1980_n = tramnmx(May29_9June_DT14pm_1980, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1981_n = tramnmx(May29_9June_DT14pm_1981, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1982_n = tramnmx(May29_9June_DT14pm_1982, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1983_n = tramnmx(May29_9June_DT14pm_1983, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1984_n = tramnmx(May29_9June_DT14pm_1984, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1985_n = tramnmx(May29_9June_DT14pm_1985, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1986_n = tramnmx(May29_9June_DT14pm_1986, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1987_n = tramnmx(May29_9June_DT14pm_1987, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1988_n = tramnmx(May29_9June_DT14pm_1988, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1989_n = tramnmx(May29_9June_DT14pm_1989, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1990_n = tramnmx(May29_9June_DT14pm_1990, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1991_n = tramnmx(May29_9June_DT14pm_1991, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1992_n = tramnmx(May29_9June_DT14pm_1992, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1993_n = tramnmx(May29_9June_DT14pm_1993, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1994_n = tramnmx(May29_9June_DT14pm_1994, minDT14pm,
maxDT14pm);
```

```
May29_9June_DT14pm_1995_n = tramnmx(May29_9June_DT14pm_1995, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1996_n = tramnmx(May29_9June_DT14pm_1996, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1997_n = tramnmx(May29_9June_DT14pm_1997, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1998_n = tramnmx(May29_9June_DT14pm_1998, minDT14pm,
maxDT14pm);
May29_9June_DT14pm_1999_n = tramnmx(May29_9June_DT14pm_1999, minDT14pm,
maxDT14pm);

%% Now for the simulation. The output is also normalised.
May29_9June_Load_1980_n = sim(net,May29_9June_DT14pm_1980_n);
May29_9June_Load_1981_n = sim(net,May29_9June_DT14pm_1981_n);
May29_9June_Load_1982_n = sim(net,May29_9June_DT14pm_1982_n);
May29_9June_Load_1983_n = sim(net,May29_9June_DT14pm_1983_n);
May29_9June_Load_1984_n = sim(net,May29_9June_DT14pm_1984_n);
May29_9June_Load_1985_n = sim(net,May29_9June_DT14pm_1985_n);
May29_9June_Load_1986_n = sim(net,May29_9June_DT14pm_1986_n);
May29_9June_Load_1987_n = sim(net,May29_9June_DT14pm_1987_n);
May29_9June_Load_1988_n = sim(net,May29_9June_DT14pm_1988_n);
May29_9June_Load_1989_n = sim(net,May29_9June_DT14pm_1989_n);
May29_9June_Load_1990_n = sim(net,May29_9June_DT14pm_1990_n);
May29_9June_Load_1991_n = sim(net,May29_9June_DT14pm_1991_n);
May29_9June_Load_1992_n = sim(net,May29_9June_DT14pm_1992_n);
May29_9June_Load_1993_n = sim(net,May29_9June_DT14pm_1993_n);
May29_9June_Load_1994_n = sim(net,May29_9June_DT14pm_1994_n);
May29_9June_Load_1995_n = sim(net,May29_9June_DT14pm_1995_n);
May29_9June_Load_1996_n = sim(net,May29_9June_DT14pm_1996_n);
May29_9June_Load_1997_n = sim(net,May29_9June_DT14pm_1997_n);
May29_9June_Load_1998_n = sim(net,May29_9June_DT14pm_1998_n);
May29_9June_Load_1999_n = sim(net,May29_9June_DT14pm_1999_n);

%% The output: derived loads must be un-normalised.
May29_9June_Load_1980=postmnmx(May29_9June_Load_1980_n,minDPL,maxDPL);
May29_9June_Load_1981=postmnmx(May29_9June_Load_1981_n,minDPL,maxDPL);
May29_9June_Load_1982=postmnmx(May29_9June_Load_1982_n,minDPL,maxDPL);
May29_9June_Load_1983=postmnmx(May29_9June_Load_1983_n,minDPL,maxDPL);
May29_9June_Load_1984=postmnmx(May29_9June_Load_1984_n,minDPL,maxDPL);
May29_9June_Load_1985=postmnmx(May29_9June_Load_1985_n,minDPL,maxDPL);
May29_9June_Load_1986=postmnmx(May29_9June_Load_1986_n,minDPL,maxDPL);
May29_9June_Load_1987=postmnmx(May29_9June_Load_1987_n,minDPL,maxDPL);
May29_9June_Load_1988=postmnmx(May29_9June_Load_1988_n,minDPL,maxDPL);
May29_9June_Load_1989=postmnmx(May29_9June_Load_1989_n,minDPL,maxDPL);
May29_9June_Load_1990=postmnmx(May29_9June_Load_1990_n,minDPL,maxDPL);
May29_9June_Load_1991=postmnmx(May29_9June_Load_1991_n,minDPL,maxDPL);
May29_9June_Load_1992=postmnmx(May29_9June_Load_1992_n,minDPL,maxDPL);
May29_9June_Load_1993=postmnmx(May29_9June_Load_1993_n,minDPL,maxDPL);
May29_9June_Load_1994=postmnmx(May29_9June_Load_1994_n,minDPL,maxDPL);
May29_9June_Load_1995=postmnmx(May29_9June_Load_1995_n,minDPL,maxDPL);
May29_9June_Load_1996=postmnmx(May29_9June_Load_1996_n,minDPL,maxDPL);
May29_9June_Load_1997=postmnmx(May29_9June_Load_1997_n,minDPL,maxDPL);
May29_9June_Load_1998=postmnmx(May29_9June_Load_1998_n,minDPL,maxDPL);
May29_9June_Load_1999=postmnmx(May29_9June_Load_1999_n,minDPL,maxDPL);

%% Now take the maximum of the loads for each row.
May29_9June_Max_Load_1980 = max(May29_9June_Load_1980);
```

```
May29_9June_Max_Load_1981 = max(May29_9June_Load_1981);
May29_9June_Max_Load_1982 = max(May29_9June_Load_1982);
May29_9June_Max_Load_1983 = max(May29_9June_Load_1983);
May29_9June_Max_Load_1984 = max(May29_9June_Load_1984);
May29_9June_Max_Load_1985 = max(May29_9June_Load_1985);
May29_9June_Max_Load_1986 = max(May29_9June_Load_1986);
May29_9June_Max_Load_1987 = max(May29_9June_Load_1987);
May29_9June_Max_Load_1988 = max(May29_9June_Load_1988);
May29_9June_Max_Load_1989 = max(May29_9June_Load_1989);
May29_9June_Max_Load_1990 = max(May29_9June_Load_1990);
May29_9June_Max_Load_1991 = max(May29_9June_Load_1991);
May29_9June_Max_Load_1992 = max(May29_9June_Load_1992);
May29_9June_Max_Load_1993 = max(May29_9June_Load_1993);
May29_9June_Max_Load_1994 = max(May29_9June_Load_1994);
May29_9June_Max_Load_1995 = max(May29_9June_Load_1995);
May29_9June_Max_Load_1996 = max(May29_9June_Load_1996);
May29_9June_Max_Load_1997 = max(May29_9June_Load_1997);
May29_9June_Max_Load_1998 = max(May29_9June_Load_1998);
May29_9June_Max_Load_1999 = max(May29_9June_Load_1999);

%% Now take the mean of the above. This is the load forecast.
%% This is displayed as output.
May29_9June_Mean = (May29_9June_Max_Load_1980 +
May29_9June_Max_Load_1981 + May29_9June_Max_Load_1982 +
May29_9June_Max_Load_1983 + May29_9June_Max_Load_1984 +
May29_9June_Max_Load_1985 + May29_9June_Max_Load_1986 +
May29_9June_Max_Load_1987 + May29_9June_Max_Load_1988 +
May29_9June_Max_Load_1989 + May29_9June_Max_Load_1990 +
May29_9June_Max_Load_1991 + May29_9June_Max_Load_1992 +
May29_9June_Max_Load_1993 + May29_9June_Max_Load_1994 +
May29_9June_Max_Load_1995 + May29_9June_Max_Load_1996 +
May29_9June_Max_Load_1997 + May29_9June_Max_Load_1998 +
May29_9June_Max_Load_1999)/20;

%%------------------------------------------------------------
%% Now forecast load for period 12-23June 2000.
%% Use 20 years of historical weather data, DT14pm, for the
%% period 12-23June as input.
June12_23_DT14pm_1980=[27.6 17 16 16.4 16.3 19.6 14 14.1 13.8 16.5];
June12_23_DT14pm_1981=[19.2 18.6 16.5 15.8 16.8 18.8 13.7 13.6 15.8
19.1];
June12_23_DT14pm_1982=[14.5 14.4 14.9 14.5 15.3 21 24 15 15 16.1];
June12_23_DT14pm_1983=[17 17.4 23.5 17.7 16.7 21.4 19.1 16.6 12.4
14.7];
June12_23_DT14pm_1984=[12.7 14.5 17.1 19 21.3 11.7 20.4 20.2 16.8 26];
June12_23_DT14pm_1985=[15.5 17 18 14.6 16.8 17.3 18.9 17.1 15.2 14.6];
June12_23_DT14pm_1986=[16 15.5 18.3 23.5 17.3 17.2 14.7 15.7 16.4
21.4];
June12_23_DT14pm_1987=[20.4 18.8 18.4 16 17.3 17 19 23.3 23.4 22.2];
June12_23_DT14pm_1988=[15.7 15.4 13.4 16 17.7 17 13.4 14.3 17 17.5];
June12_23_DT14pm_1989=[18.2 21.1 16.7 25.2 26.4 15.8 20.8 16.4 15.4
15.7];
June12_23_DT14pm_1990=[18.6 22.3 19.9 15.7 15.3 19.2 13.3 13.3 14.4
14.8];
June12_23_DT14pm_1991=[17.3 14.9 17.7 20 15.1 16 19.4 17.8 13.4 12.5];
June12_23_DT14pm_1992=[15.7 18.8 22.2 17.8 22.1 23.3 16.3 15 13.6
16.9];
```

```
June12_23_DT14pm_1993=[12.9 18.4 21.9 17.6 15.9 16.3 16.6 17.7 17.9
16.6];
June12_23_DT14pm_1994=[22.6 18.3 19.3 22.6 18.7 13.4 14.8 14.5 15.4
12.9];
June12_23_DT14pm_1995=[17.1 15.2 12.8 14.9 16 24.3 21.2 15.6 16.8
12.5];
June12_23_DT14pm_1996=[15 15.5 15 16.8 16.7 20.9 19.7 25.5 24.5 16.9];
June12_23_DT14pm_1997=[18.3 17.9 17.7 16.5 13.9 15.3 19.1 15.7 14
14.1];
June12_23_DT14pm_1998=[17.6 18.8 18.6 17.7 18.4 15.7 15.5 16.4 16.2
18.5];
June12_23_DT14pm_1999=[25.9 21.5 15.2 17.2 18.9 16.6 19.8 15.8 17.6
15.7];

%% The above data must be normalised.
June12_23_DT14pm_1980_n = tramnmx(June12_23_DT14pm_1980, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1981_n = tramnmx(June12_23_DT14pm_1981, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1982_n = tramnmx(June12_23_DT14pm_1982, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1983_n = tramnmx(June12_23_DT14pm_1983, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1984_n = tramnmx(June12_23_DT14pm_1984, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1985_n = tramnmx(June12_23_DT14pm_1985, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1986_n = tramnmx(June12_23_DT14pm_1986, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1987_n = tramnmx(June12_23_DT14pm_1987, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1988_n = tramnmx(June12_23_DT14pm_1988, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1989_n = tramnmx(June12_23_DT14pm_1989, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1990_n = tramnmx(June12_23_DT14pm_1990, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1991_n = tramnmx(June12_23_DT14pm_1991, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1992_n = tramnmx(June12_23_DT14pm_1992, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1993_n = tramnmx(June12_23_DT14pm_1993, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1994_n = tramnmx(June12_23_DT14pm_1994, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1995_n = tramnmx(June12_23_DT14pm_1995, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1996_n = tramnmx(June12_23_DT14pm_1996, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1997_n = tramnmx(June12_23_DT14pm_1997, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1998_n = tramnmx(June12_23_DT14pm_1998, minDT14pm,
maxDT14pm);
June12_23_DT14pm_1999_n = tramnmx(June12_23_DT14pm_1999, minDT14pm,
maxDT14pm);

%% Now for the simulation. The output is also normalised.
```

```
June12_23_Load_1980_n = sim(net,June12_23_DT14pm_1980_n);
June12_23_Load_1981_n = sim(net,June12_23_DT14pm_1981_n);
June12_23_Load_1982_n = sim(net,June12_23_DT14pm_1982_n);
June12_23_Load_1983_n = sim(net,June12_23_DT14pm_1983_n);
June12_23_Load_1984_n = sim(net,June12_23_DT14pm_1984_n);
June12_23_Load_1985_n = sim(net,June12_23_DT14pm_1985_n);
June12_23_Load_1986_n = sim(net,June12_23_DT14pm_1986_n);
June12_23_Load_1987_n = sim(net,June12_23_DT14pm_1987_n);
June12_23_Load_1988_n = sim(net,June12_23_DT14pm_1988_n);
June12_23_Load_1989_n = sim(net,June12_23_DT14pm_1989_n);
June12_23_Load_1990_n = sim(net,June12_23_DT14pm_1990_n);
June12_23_Load_1991_n = sim(net,June12_23_DT14pm_1991_n);
June12_23_Load_1992_n = sim(net,June12_23_DT14pm_1992_n);
June12_23_Load_1993_n = sim(net,June12_23_DT14pm_1993_n);
June12_23_Load_1994_n = sim(net,June12_23_DT14pm_1994_n);
June12_23_Load_1995_n = sim(net,June12_23_DT14pm_1995_n);
June12_23_Load_1996_n = sim(net,June12_23_DT14pm_1996_n);
June12_23_Load_1997_n = sim(net,June12_23_DT14pm_1997_n);
June12_23_Load_1998_n = sim(net,June12_23_DT14pm_1998_n);
June12_23_Load_1999_n = sim(net,June12_23_DT14pm_1999_n);

%% The output: derived loads must be un-normalised.
June12_23_Load_1980 = postmnmx(June12_23_Load_1980_n, minDPL, maxDPL);
June12_23_Load_1981 = postmnmx(June12_23_Load_1981_n, minDPL, maxDPL);
June12_23_Load_1982 = postmnmx(June12_23_Load_1982_n, minDPL, maxDPL);
June12_23_Load_1983 = postmnmx(June12_23_Load_1983_n, minDPL, maxDPL);
June12_23_Load_1984 = postmnmx(June12_23_Load_1984_n, minDPL, maxDPL);
June12_23_Load_1985 = postmnmx(June12_23_Load_1985_n, minDPL, maxDPL);
June12_23_Load_1986 = postmnmx(June12_23_Load_1986_n, minDPL, maxDPL);
June12_23_Load_1987 = postmnmx(June12_23_Load_1987_n, minDPL, maxDPL);
June12_23_Load_1988 = postmnmx(June12_23_Load_1988_n, minDPL, maxDPL);
June12_23_Load_1989 = postmnmx(June12_23_Load_1989_n, minDPL, maxDPL);
June12_23_Load_1990 = postmnmx(June12_23_Load_1990_n, minDPL, maxDPL);
June12_23_Load_1991 = postmnmx(June12_23_Load_1991_n, minDPL, maxDPL);
June12_23_Load_1992 = postmnmx(June12_23_Load_1992_n, minDPL, maxDPL);
June12_23_Load_1993 = postmnmx(June12_23_Load_1993_n, minDPL, maxDPL);
June12_23_Load_1994 = postmnmx(June12_23_Load_1994_n, minDPL, maxDPL);
June12_23_Load_1995 = postmnmx(June12_23_Load_1995_n, minDPL, maxDPL);
June12_23_Load_1996 = postmnmx(June12_23_Load_1996_n, minDPL, maxDPL);
June12_23_Load_1997 = postmnmx(June12_23_Load_1997_n, minDPL, maxDPL);
June12_23_Load_1998 = postmnmx(June12_23_Load_1998_n, minDPL, maxDPL);
June12_23_Load_1999 = postmnmx(June12_23_Load_1999_n, minDPL, maxDPL);

%% Now take the maximum of the loads for each row.
June12_23_Max_Load_1980 = max(June12_23_Load_1980);
June12_23_Max_Load_1981 = max(June12_23_Load_1981);
June12_23_Max_Load_1982 = max(June12_23_Load_1982);
June12_23_Max_Load_1983 = max(June12_23_Load_1983);
June12_23_Max_Load_1984 = max(June12_23_Load_1984);
June12_23_Max_Load_1985 = max(June12_23_Load_1985);
June12_23_Max_Load_1986 = max(June12_23_Load_1986);
June12_23_Max_Load_1987 = max(June12_23_Load_1987);
June12_23_Max_Load_1988 = max(June12_23_Load_1988);
June12_23_Max_Load_1989 = max(June12_23_Load_1989);
June12_23_Max_Load_1990 = max(June12_23_Load_1990);
June12_23_Max_Load_1991 = max(June12_23_Load_1991);
June12_23_Max_Load_1992 = max(June12_23_Load_1992);
```

```
June12_23_Max_Load_1993 = max(June12_23_Load_1993);
June12_23_Max_Load_1994 = max(June12_23_Load_1994);
June12_23_Max_Load_1995 = max(June12_23_Load_1995);
June12_23_Max_Load_1996 = max(June12_23_Load_1996);
June12_23_Max_Load_1997 = max(June12_23_Load_1997);
June12_23_Max_Load_1998 = max(June12_23_Load_1998);
June12_23_Max_Load_1999 = max(June12_23_Load_1999);

%% Now take the mean of the above. This is the load forecast.
%% This is displayed as output.
June12_23_Mean = (June12_23_Max_Load_1980 + June12_23_Max_Load_1981 +
June12_23_Max_Load_1982 + June12_23_Max_Load_1983 +
June12_23_Max_Load_1984 + June12_23_Max_Load_1985 +
June12_23_Max_Load_1986 + June12_23_Max_Load_1987 +
June12_23_Max_Load_1988 + June12_23_Max_Load_1989 +
June12_23_Max_Load_1990 + June12_23_Max_Load_1991 +
June12_23_Max_Load_1992 + June12_23_Max_Load_1993 +
June12_23_Max_Load_1994 + June12_23_Max_Load_1995 +
June12_23_Max_Load_1996 + June12_23_Max_Load_1997 +
June12_23_Max_Load_1998 + June12_23_Max_Load_1999)/20;

%% This is the actual fortnight peak load for the periods above.
Actual_Fortnight_Peak_Load = [3528 3510 3474 3618];

May1_12_Mean
May15_26_Mean
May29_9June_Mean
June12_23_Mean

May1_12_Error_Percentage = abs((((May1_12_Mean-
Actual_Fortnight_Peak_Load(1,1))/Actual_Fortnight_Peak_Load(1,1))*100))

May15_26_Error_Percentage = abs((((May15_26_Mean-
Actual_Fortnight_Peak_Load(1,2))/Actual_Fortnight_Peak_Load(1,2))*100))

May29_9June_Error_Percentage = abs((((May29_9June_Mean-
Actual_Fortnight_Peak_Load(1,3))/Actual_Fortnight_Peak_Load(1,3))*100))

June12_23_Error_Percentage = abs((((June12_23_Mean-
Actual_Fortnight_Peak_Load(1,4))/Actual_Fortnight_Peak_Load(1,4))*100))

Mean_Absolute_Percentage_Error = (May1_12_Error_Percentage +
May15_26_Error_Percentage + May29_9June_Error_Percentage +
June12_23_Error_Percentage)/4

echo off
```

Appendix II

The Cluster Analysis Dendograms for the groups of 3 load parameters

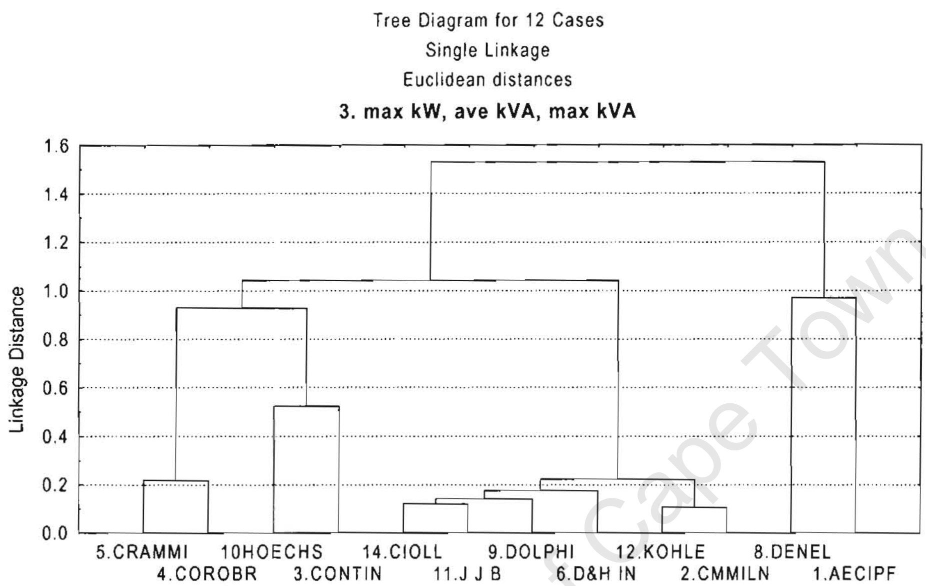


Figure 8.3: Tree of Max kW, Ave kVA and Max kVA

Table 8.11: Cluster formation/Customer identification for max kW, ave kVA, max kVA

Cluster	1	2
Customers in Cluster	5,4,10,3,14,11,9,6,12,2	8,1
Customers identified	**Much Asphalt**	**Denel Edms Bkp-Fir**

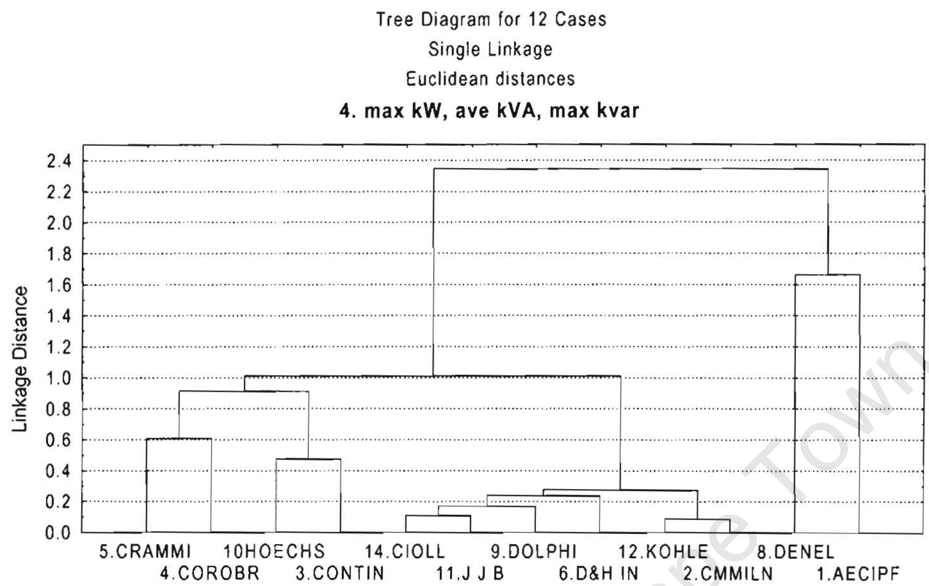


Figure 8.4: Tree of Max kW, Ave kVA and Max kVAr

Table 8.12: Cluster formation/Customer identification for max kW, ave kVA, max kVAr

Cluster	1	2
Customers in Cluster	5,4,10,3,14,11,9,6 ,12,2	8,1
Customers identified	**Much Asphalt**	**Denel Edms Bkp-Fir**

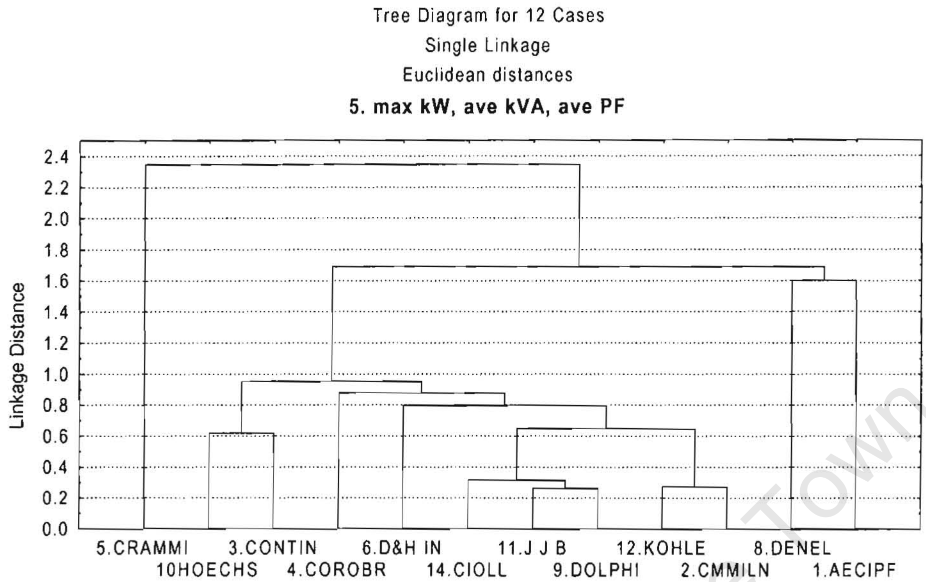


Figure 8.5: Tree of Max kW, Ave kVA and Ave PF

Table 8.13: Cluster formation/Customer identification for max kW, ave kVA, ave PF

Cluster	1	2	3	4
Customers in Cluster	5	10,3,4,6,14,11,9,12,2	8	1
Customers identified	-	**Much Asphalt**	-	-

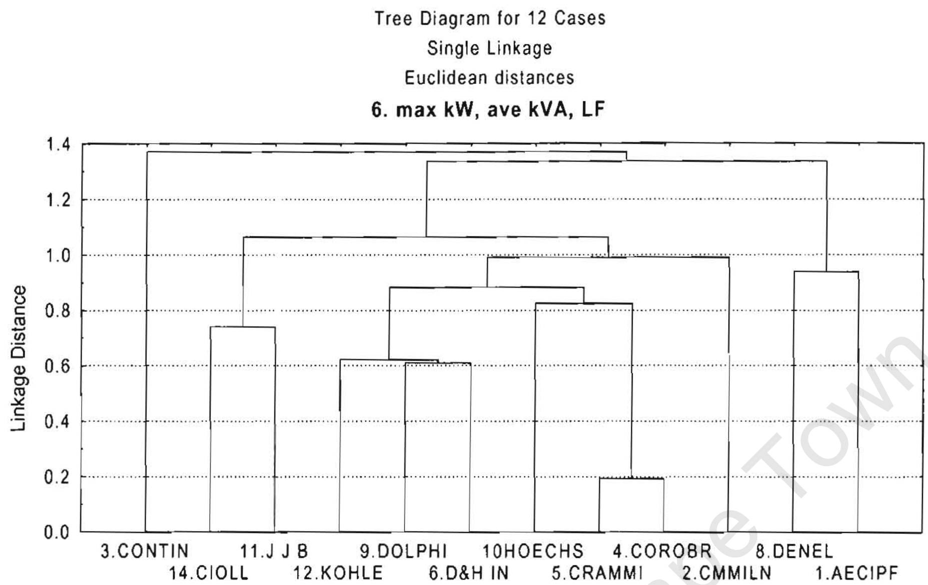


Figure 8.6: Max kW, Ave kVA and LF

Table 8.14: Cluster formation/Customer identification for max kW, ave kVA, LF

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Customers in Cluster	3	14	11	9	12	6	10	5,4	2	8	7	1
Customers identified	-	-	-	-	-	-	-	-	-	-	-	-

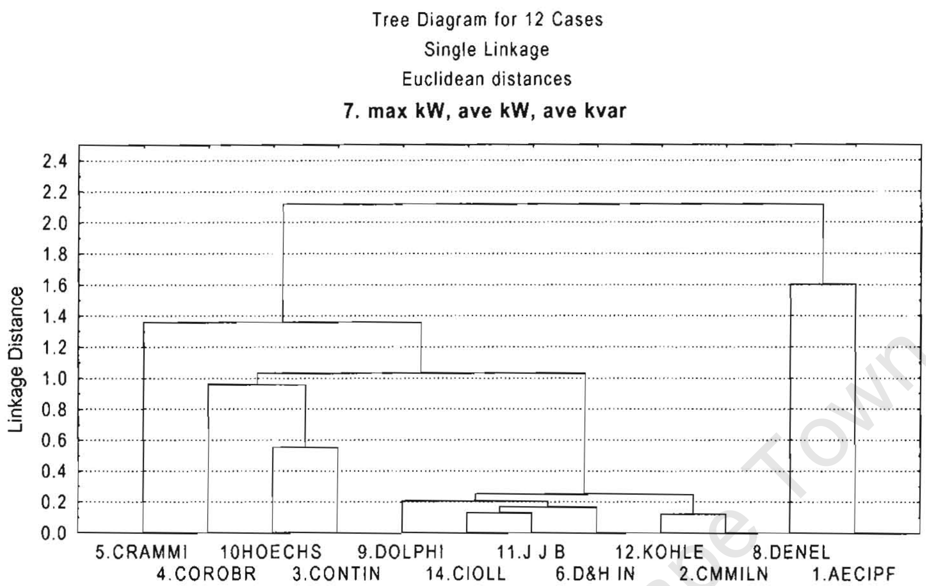


Figure 8.7: Tree of Max kW, Ave kW and Ave kVAr

Table 8.15: Cluster formation/Customer identification for max kW, ave kW, ave kVAr

Cluster	1	2
Customers in Cluster	5,4,10,3,9,14,11,6,12, 2	8,1
Customers identified	**Much Asphalt**	**Denel Edms Bkp-Fir**

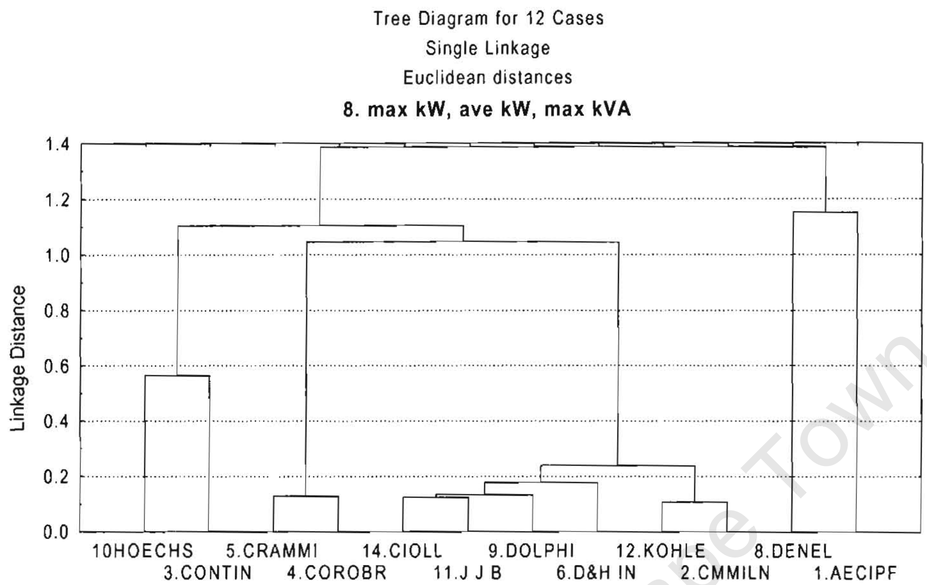


Figure 8.8: Max kW, Ave kW and Max kVA

Table 8.16: Cluster formation/Customer identification for max kW, ave kW, max kVA

Cluster	1	2	3	4	5
Customers in Cluster	10,3	5,4	14,11,9,6,12,2	8	1
Customers identified	-	-	**Much Asphalt**	-	-

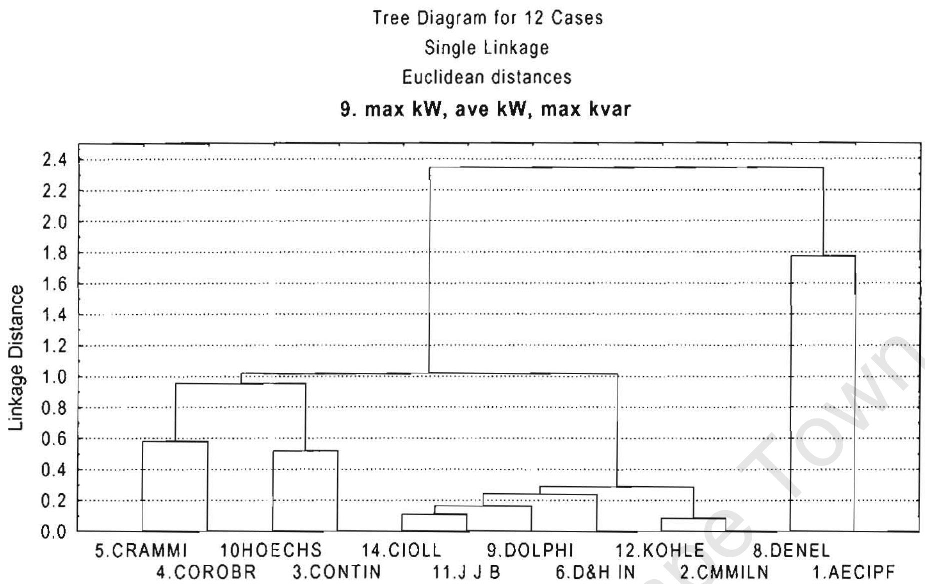


Figure 8.9: Tree of Max kW, Ave kW and Max kVAr

Table 8.17: Cluster formation/Customer identification for max kW, ave kW, max kVAr

Cluster	1	2	3
Customers in Cluster	5,4,10,3,14,11,9,6,12, 2	8	1
Customers identified	**Much Asphalt**	-	-

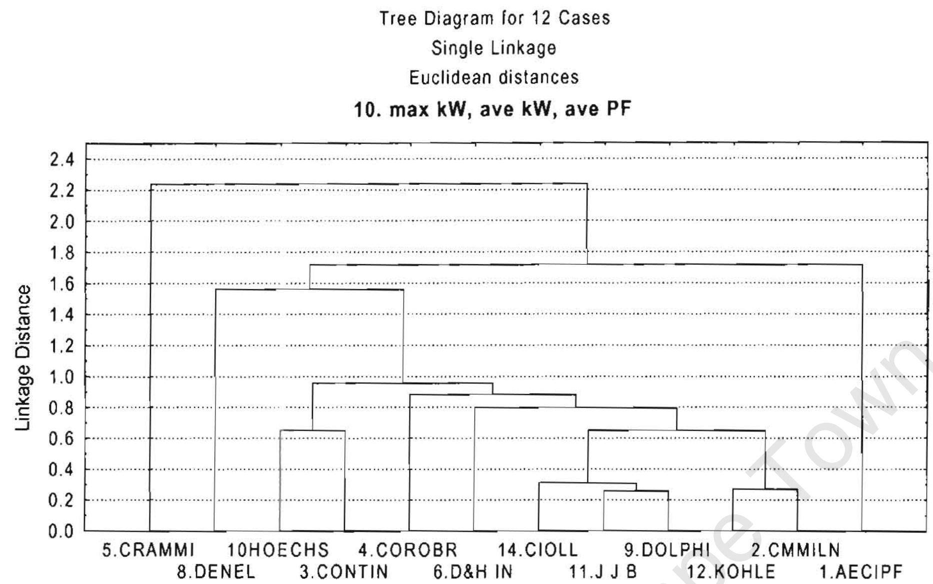


Figure 8.10: Tree of Max kW, Ave kW and Ave PF

Table 8.18: Cluster formation/Customer identification for max kW, ave kW, ave PF

Cluster	1	2	3	4
Customers in Cluster	5	8	10,3,4,6,14,11,9,12,2	1
Customers identified	-	-	**Much Asphalt**	-

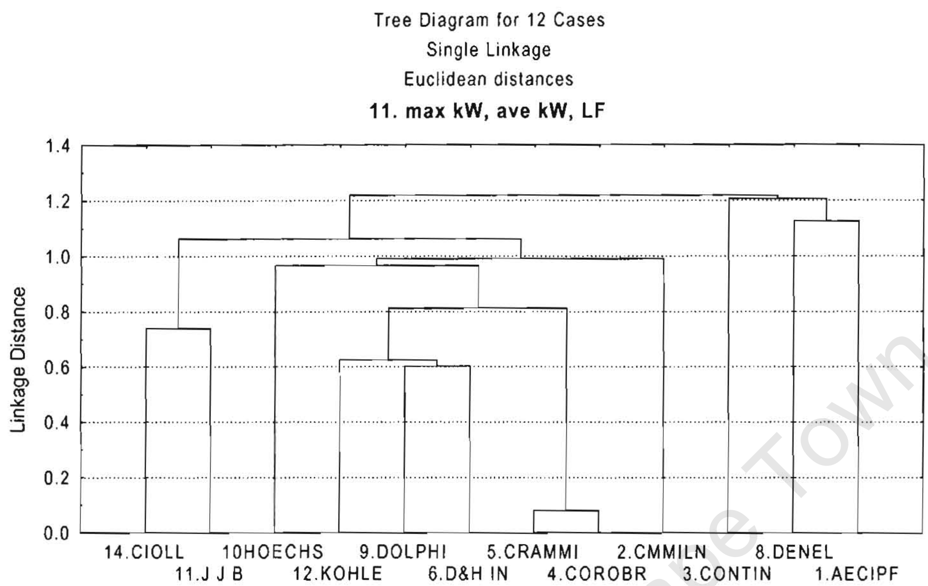


Figure 8.11: Max kW, Ave kW and LF

Table 8.19: Cluster formation/Customer identification for max kW, ave kW, LF

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Customers in Cluster	14	11	10	12	9	6	5,4	2	3	8	7	1
Customers identified	-	-	-	-	-	-	-	-	-	-	-	-

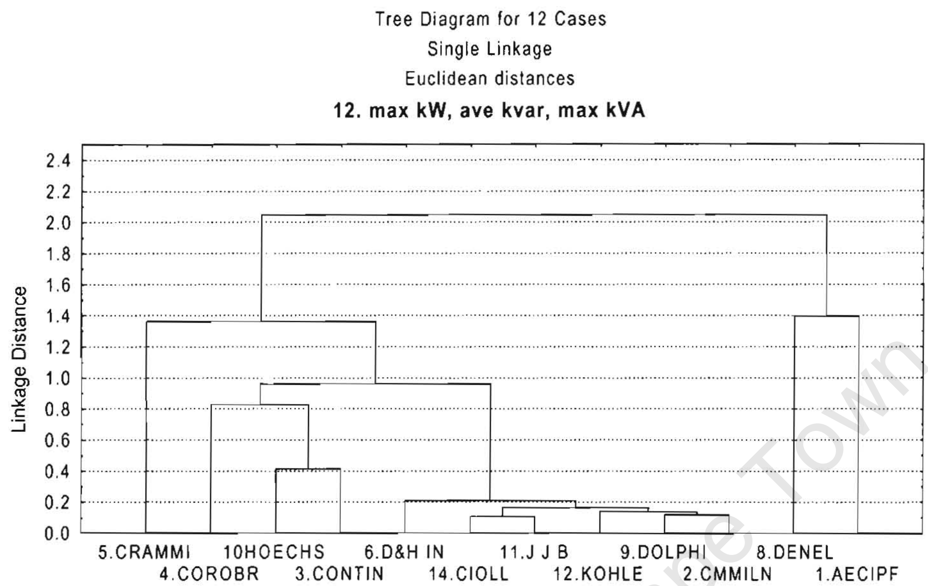


Figure 8.12: Max kW, Ave kVAr and Max kVA

Table 8.20: Cluster formation/Customer identification for max kW, ave kVAr, max kVA

Cluster	1	2
Customers in Cluster	5,4,10,3,6,14,11,12,9,2	8,1
Customers identified	**Much Asphalt**	**Denel Edms Bkp-Fir**

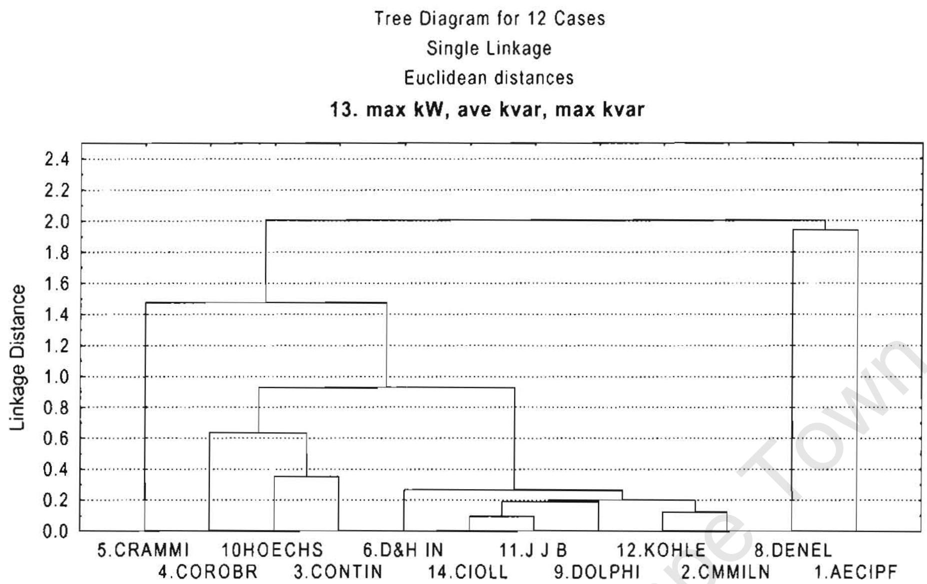


Figure 8.13: Tree of Max kW, Ave kVAr and Max kVAr

Table 8.21: Cluster formation/Customer identification for max kW, ave kVAr, max kVAr

Cluster	1	2	3
Customers in Cluster	5,4,10,3,6,14,11,9,12,2	8	1
Customers identified	**Much Asphalt**	-	-

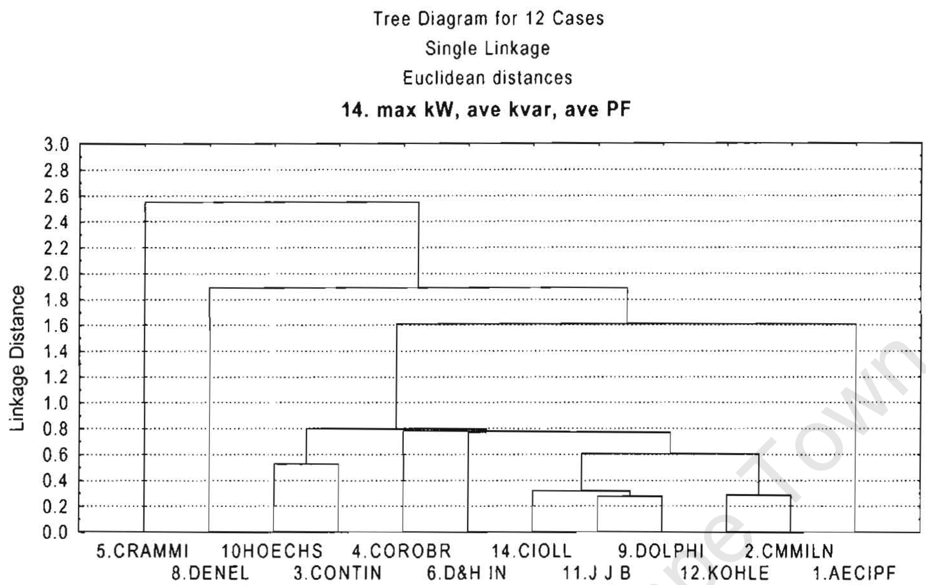


Figure 8.14: Max kW, Ave kVAr and Ave PF

Table 8.22: Cluster formation/Customer identification for max kW, ave kVAr, ave PF

Cluster	1	2	3	4
Customers in Cluster	5	8	10,3,6,4,14,11,9,12,2	1
Customers identified	-	-	**Much Asphalt**	-

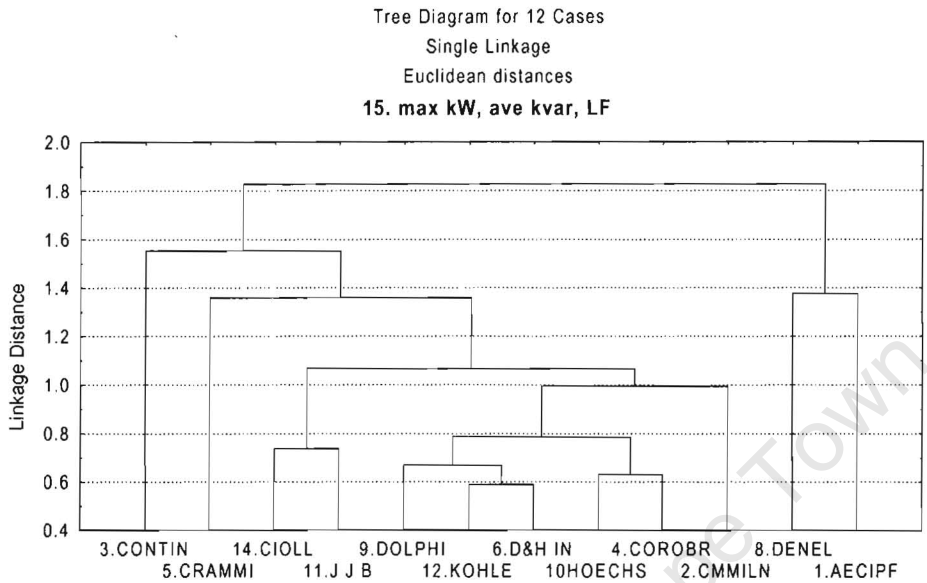


Figure 8.15: Tree of Max kW, Ave kVAr and LF

Table 8.23: Cluster formation/Customer identification for max kW, ave kVAr, LF

Cluster	1	2	2	3	4
Customers in Cluster	3	5	14,11,9,12,6,10,4,2	8	1
Customers identified	-	-	**Much Asphalt**	-	-

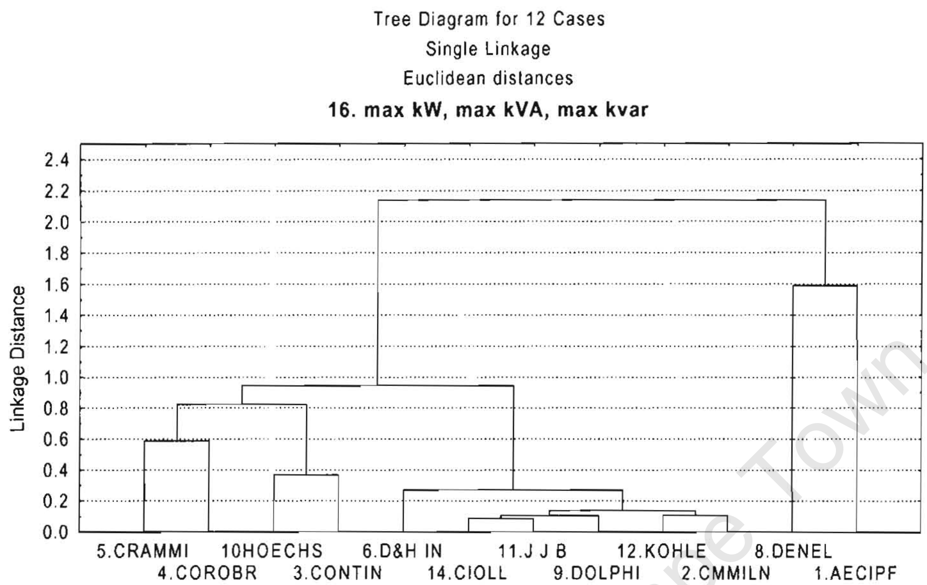


Figure 8.16: Tree of Max kW, Max kVA and Max kVAr

Table 8.24: Cluster formation/Customer identification for max kW, max kVA, max kVAr

Cluster	1	2	3
Customers in Cluster	5,4,10,3,6,14,11,9,12,2	8	1
Customers identified	**Much Asphalt**	-	-

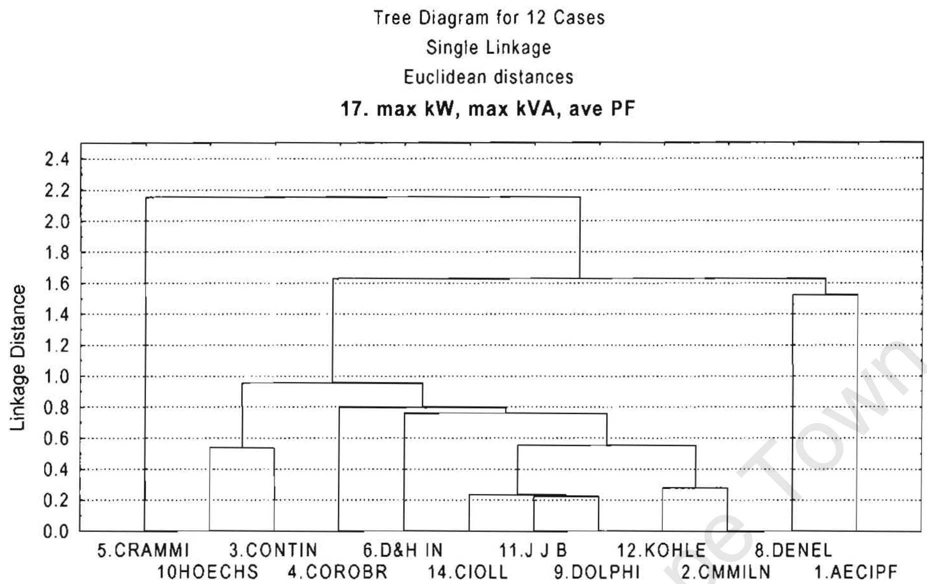


Figure 8.17: Tree of Max kW, Max kVA and Ave PF

Table 8.25: Cluster formation/Customer identification for max kW, max kVA, ave PF

Cluster	1	2	3	4
Customers in Cluster	5	10,3,6,4,11,14,9,12,2	8	1
Customers identified	-	**Much Asphalt**	-	-

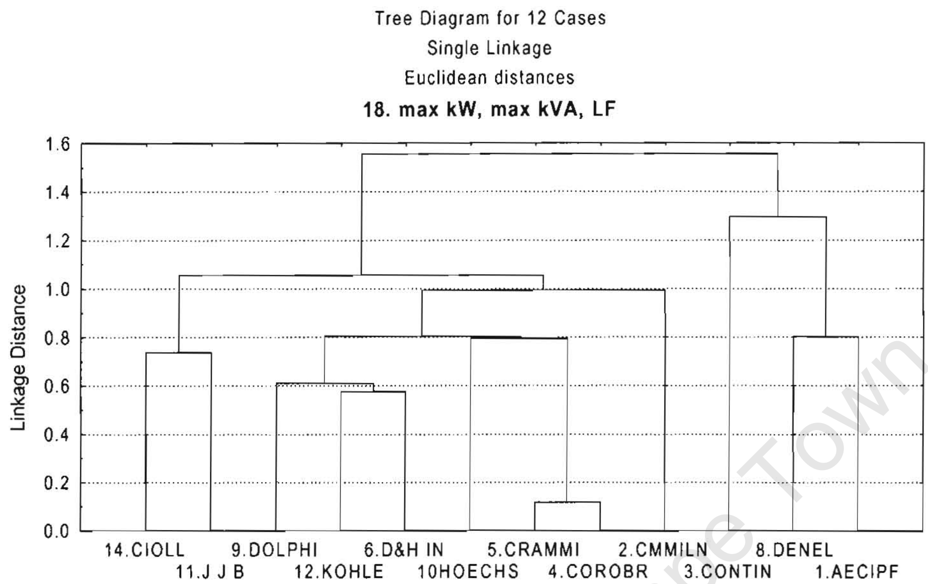


Figure 8.18: Tree of Max kW, Max kVA and LF

Table 8.26: Cluster formation/Customer identification for max kW, max kVA, LF

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Customers in Cluster	14	11	9	12	6	10	5,4	2	3	8	7	1
Customers identified	-	-	-	-	-	-	-	-	-	-	-	-

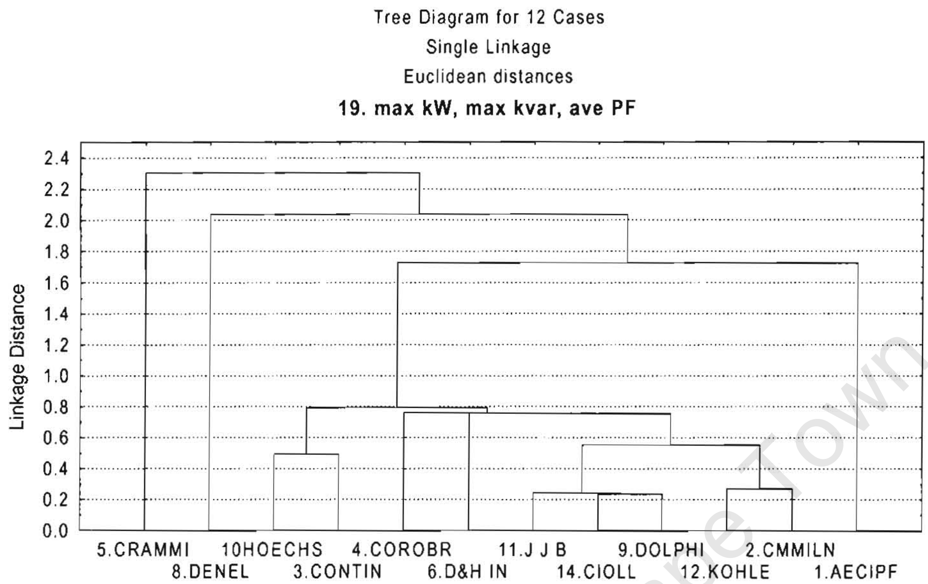


Figure 8.19: Tree of Max kW, Max kVAr and Ave PF

Table 8.27: Cluster formation/Customer identification for max kW, max kVAr, ave PF

Cluster	1	2	3	4
Customers in Cluster	5	8	10,3,6,4,11,14,9,12,2	1
Customers identified	-	-	**Much Asphalt**	-

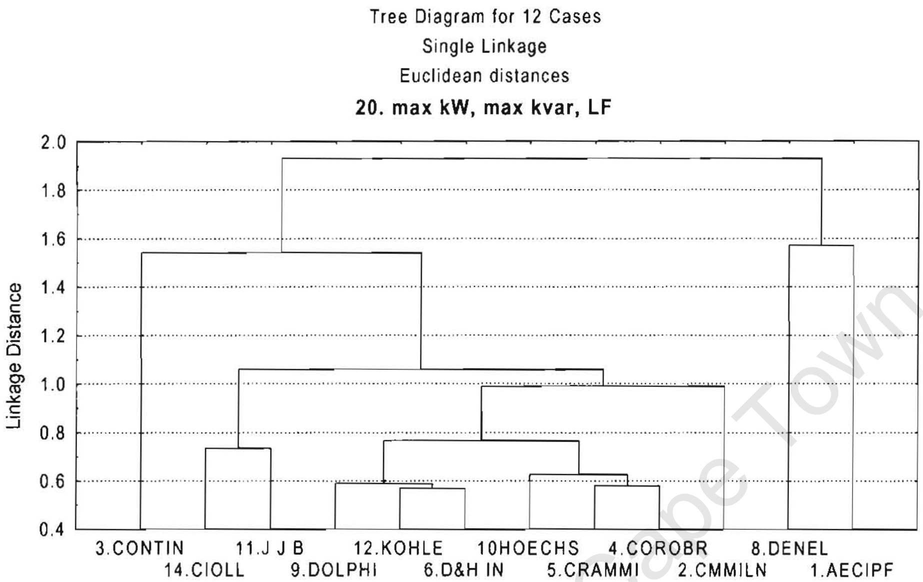


Figure 8.20: Tree of Max kW, Max kVAr and LF

Table 8.28: Cluster formation/Customer identification for max kW, max kVAr, LF

Cluster	1	2	3	4
Customers in Cluster	3	14,11,9,12,6,10,5,4,2	8	1
Customers identified	-	**Much Asphalt**	-	-

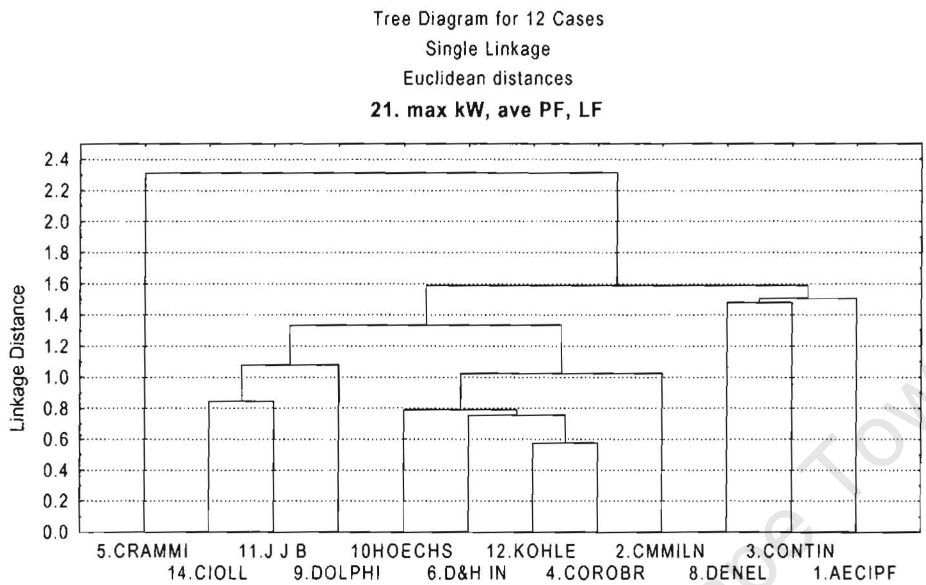


Figure 8.21: Tree of Max kW, Ave PF and LF

Table 8.29: Cluster formation/Customer identification for max kW, ave PF, LF

Cluster	1	2
Customers in Cluster	5	14,11,9,6,10,12,4,2,8,3,1
Customers identified	-	**Denel Edms Bkp-Fir and Much Asphalt**