SPATIAL ELECTRIC LOAD FORECASTING MODEL FOR SRI LANKA

Mihiri Kalpani Madawala

(139510M)

Degree of Master of Science in Electrical Engineering

Department of Electrical Engineering

University of Moratuwa
Sri Lanka

April 2018

SPATIAL ELECTRIC LOAD FORECASTING MODEL FOR SRI LANKA

Mihiri Kalpani Madawala

(139510M)

Dissertation submitted in partial fulfilment of the requirements for the Degree Master of Science in Electrical Engineering

Department of Electrical Engineering

University of Moratuwa
Sri Lanka

April 2018

DECLARATION OF THE CANDIDATE & SUPERVISOR

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the Masters dissertation under our supervision.

Signature of the supervisor: Date

Prof J. R. Lucas

Signature of the supervisor: Date

Dr P. S. N. De Silva

Abstract

With the high level of city expansion observed during the last few decades, distribution utilities currently face new challenges when planning network expansion with profitable operations. Thus distribution utilities should consider spatial electric load forecasting as the basis for the planning of the electricity distribution networks. Spatial electric load forecasting helps in determining how the increase in demand of electrical energy will be distributed geographically in the service area.

In Sri Lanka, the load forecasting in distribution planning is mainly based on trending methods which lacks the accuracy needed for present dynamic consumer market. The objective of this research is to prepare a simple yet accurate and effective spatial electric load forecasting model which can be used in the local context.

This research deals with a new method for spatial electric load forecasting using artificial neural networks. The electric load growth inside the service area of an electric utility can be expected for two reasons, natural load growth of existing consumers and addition of new loads because of new consumers. In the study, the addition of new consumers in future is regarded as the new load additions in the vacant areas. This is forecasted using the spatial electric load forecasting model implemented using artificial neural network. The growth of existing consumers is addressed with a constant growth.

The implemented model is presented and tested with data from two real midsized cities. The outcome is compared with the ones obtained from the utility planning department existing software. The results illustrate that the proposed model provides an accurate and user-friendly technique to predict yearly residential electrical load in Sri Lanka.

KEYWORDS: Spatial electric load forecasting, land use, artificial neural network, distribution planning.

Acknowledgement

It is with great pleasure that I express gratitude to those who were behind me in completing my research project.

Firstly, I would like to express my sincere gratitude to my supervisor Prof. J. R. Lucas for his great insights and perspective with much patience throughout the entire period and Dr. Narendra De Silva for his continuous support with his immense knowledge motivating me to do this research. Their guidance helped me all the time, from starting of this research, up until the end writing of this dissertation. Thank you very much for giving me the opportunity to work with you both.

In addition, I would like to thank all the Lecturers in the Department of Electrical Engineering and the Post Graduate Studies Division, Faculty of Engineering, who engaged in this MSc course in various ways to educate us and broaden our vision.

My sincere thanks go to my managers and colleagues at Ceylon Electricity Board and Lanka Electricity Company (Private) Limited, who helped me in many ways during this period.

Finally, I am thankful to my family members including my husband, parents and brother who always encouraged and help me to complete this research.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Description
ANN	Artificial Neural Network
CSC	Customer Service Centre
GDP	Gross Domestic Product
GIS	Geographical Information System
IEC	International Electro technical Commission
IEEE	Institute of Electrical and Electronic Engineers
kWh	Kilo Watt Hour
LECO	Lanka Electricity Company (pvt) Ltd
LM	Levenberg-Marquardt
MAPE	Mean Absolute Percentage Error
SLFM	Spatial Electric Load Forecasting Model

CHAPTER 1

1. INTRODUCTION

The Power System, which supplies electricity to the consumer, basically comprises of three sectors, namely generation, transmission, and distribution. All these sectors should be treated equally and developed in parallel to avoid bottlenecks in delivering the generated power to the end-consumer. However in a financial and technical sense, it is very common that the distribution sector gets reduced attention, compared to the generation and transmission, in many of the countries. Yet, it is often the most critical component since it directly takes the system power up to consumers in every corner of the country. It has the highest power quality issues, has the highest reliability issues and has the highest technical and non-technical losses, requiring precise planning an essential need. Thus, with the high level of city expansion observed in last few decades, distribution utilities currently face new challenges when planning network expansion that offers a reliable and economical service inside their service area. This situation is more difficult as distribution utilities should now run their operations profitably while maintaining existing aging electrical structures. This study is focus on such a critical yet less addressed problem in planning of the electricity distribution network in the power sector and proposes a novel, accurate user-friendly implementation methodology.

1.1. Importance of Load Forecasting

Load forecasting mainly refers to forecasting of electricity demand and energy. It is being used throughout all sectors of the electric power industry, including generation, transmission and distribution. Applications of load forecasts includes power supply planning, transmission and distribution systems planning, demand side management, power systems operations and maintenance, financial planning, preparation of tariff and so forth. Due to the fundamental role of load forecasting in the utility business operations, inaccurate load forecasts may result in financial burden of a utility company. While load forecasting provides a key input to power systems operations and planning, inaccurate load forecasts can lead to equipment failure or even system-wide blackout.

Taking everything into account, electricity storage limitations and social necessity of electricity, lead to several interesting features of load forecasting, such as identifying the complex load patterns, data collection across the grid and the need to be extremely accurate. This study is intended to simplify complex concepts, terms and statistics used in distribution system load forecasting.

1.2. Traditional Load Forecasting

Conventionally, the load forecasting technique used was the econometric approach which combines economic theory and statistical techniques for forecasting electricity demand. The approach estimates the relationships between energy consumption (dependent variables) and factors influencing consumption. Then by using the trending methods, past load growth patterns are extrapolated into the future. The most common trending method and the method most often thought of as representative of trending in general, is multiple regression used to fit a polynomial function to historic peak load data and extrapolate that function into the future. This is most widely used by many [1], due to its ease of implementation and not necessarily due to any superiority in results.

In Sri Lanka, the conventional load forecasting methods are based on national level growth factors such as national GDP, and the trending values of historical load data. Even though these methods address the transmission planning to a certain extent, they are not very much suitable for distribution planning, as these traditional load forecasting results only identifies how much demand will occur in the assumed time frame but cannot identify where the new loads takes place. Therefore, those methods are not helpful for power facilities construction location planning which is the fundamental requirement of efficient distribution network planning.

For a proper forecast it is needed to predict what, when and where the future loads will be.

1.3. Spatial Electric Load Forecasting

As already mentioned, to plan the efficient operation and economical capital expansion of the electric power delivery system, the following needs to be anticipated; a prediction of future electric demand that includes 'Where' it will be needed, as one its chief elements, in addition to magnitude (how much) and temporal (when) characteristics.

Spatial electric load forecasting is a process that tries to determine the future load growth for an electric energy distribution utility based on special aspects. This process, apart from answering the question of how much load growth is expected in a determined time frame, also identifies the locations that are most likely to get the new loads. This task includes forecasting where vacant areas will have a new development; it also includes the redevelopment as well as the reduction in land-use density.

1.3.1. Why Spatial Load Forecasting

Spatial analysis is principally, the course of mapping where things are, and identifying how they relate. Finding out where these things are, is very useful in the right context.

Once, the locations of things are known, it is vital to identify what is nearby. Waldo Tobler [20], who created the first law of geography, states: 'Everything is related to everything else, but near things are more related than distant things.' Finding out what's nearby, deepens the understanding of a location. As an example, consider, in the electricity distribution context, the importance of having an idea of the proximity of customers to retail centers.

In many cases, things are connected to other things. Therefore, understanding the interconnectedness of nearby things is of substantial importance to help make faster and more accurate decisions.

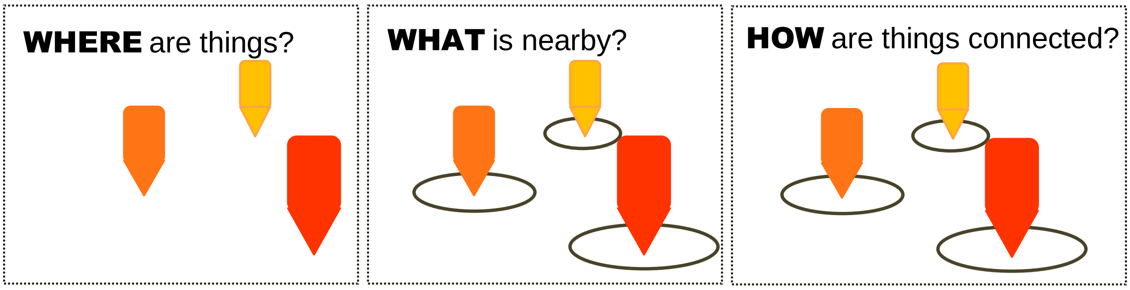

Figure 1-1: Basics of Spatial Analysis

The use of spatial analysis for electricity load forecast, provide the knowledge on the distribution of electricity load in the geography. Knowing the location of things, what is near them and how they are interconnected can help make informed decisions figure 1-1. It can predict the behavior of the interconnected parts of the electricity distribution network.

As the gist, Spatial load forecasting will help to predict large load additions to the system years in advance compared to currently implemented load forecasting methods in Sri Lanka and it helps to determine where new infrastructure should be added to the network.

1.3.2. Advantages of Spatial Load Forecasting in Sri Lanka

Traditionally, electrical engineers focused on equipment in their substations to manage load and performed trend analysis to forecast future loads. Planning engineers will now be able to use the power of Geographic Information Systems (GIS) to visualize the electricity load distribution and forecast where they are likely to see new load additions to the system. Understanding this load relative to a substation service area helps to determine where new substations might optimally be located. Spatial load forecasting produces a forecast of electric load growth inside a region of the service territory, suitable as a base for comprehensive transmission and distribution expansion planning. Forecast results are used to predict future load centers, identify substation property requirements, prioritize projects and obtain budgeting approval while minimizing financial risk.

Spatial load forecasting can:

- Predict the extent and timeline of community development,
- Identify where infrastructure investments should be directed
- Explore the impacts of new initiatives or localized development events
- Demonstrate the effects of changes in fully developed areas
- Predict general locale of new substations

1.3.3. Limitations of Spatial Load Forecasting

Spatial load forecasting does have limitations. It is not an off-the-shelf design package and does not replace the knowledge and experience of engineers. It cannot identify specific substation sites, such as exactly which land lot to purchase.

Spatial load forecasting can support the transmission and distribution system planning process but does not predict the routing of transmission lines or distribution feeders. Rather, spatial load forecasting anticipates the location of future demand in a relative timeline, assisting capital investment planning, rights of way acquisition, permitting lead times and improvement of negotiations for land acquisition. This allows a more productive interaction with the community being served.

1.4. Supporting Organisation

The research is carried out with the support of Lanka Electricity Company (Pvt) Ltd (LECO) staff including Dr. Narendra de Silva, for the electrical distribution network planning requirement of LECO. LECO provides distribution services in the coastal area from Negambo to Hikkaduwa, which contributes more than 50% of the local economy. It now operates in 32 Local Government areas and has grown from 12,000 to over 500,000 consumers. Thus special load forecasting has become an urgent need for LECO to plan their distribution.

1.5. Organization of the Dissertation

The dissertation encapsulates the most vital technical details of this study and information on literary analysis of related work and background studies. The preparation and order of the thesis is according to the chronological order of the approach to the project, hence should be referenced in the following order.

1. Introduction:

   Introduces the reader to the project and briefly discusses the traits of the initiation of the project. This discusses the background of the project, motivation for selecting the project, impact of the work and organization of the report.

2. Project Overview:

   The theoretical framework of spatial load forecasting and Artificial Neural Network is discussed in this chapter. It further describes and analyzes the previous research on the topic.

3. Methodology

   This section gives a comprehensive technical coverage on spatial electrical load forecasting technique used in the research.

4. System Modeling

   The comprehensive modeling of the Artificial Neural Network is presented. Obtaining the other related details is also described in this chapter.

5. Results

   The spatial electric load forecasting method is validated using two midsized cities in Sri Lanka and the results are presented.

6. Conclusion:

   Summarizes the study indicating how the objectives are achieved discussing advantages and disadvantages and the future direction for the research.

CHAPTER 2

2. PROJECT OVERVIEW

2.1. Evolution of Spatial Electric Load Forecasting

During the last few decades, an increasing migration of population from countryside to cities, especially in Asian cities [2], is observed. This has created a high level of city expansion and distribution utilities are currently facing new challenges when planning a network expansion that offers a reliable and economical service inside their service zone. Because of this, over the last few years, research in the distribution area has increased substantially.

To improve the distribution network expansion and planning by considering these factors, distribution utilities has considered spatial electric load forecasting as an important development. As an example, Electrical Distribution Utilities in Brazil impose the use of Spatial Electric Load Forecasting in their studies [3].

Therefore, the location of the new consumers is of vital importance for the utilities so that they can adopt all the necessary resources for meeting the needs of the future system.

Spatial electric load forecasting discussion can be traced back to Van Wormer in 1954 [4], where the need for a more detailed load forecasting considering the shape of the cities was evidenced. Most of the work developed in this area is compiled in Lee, W, H. Spatial Electric Load Forecasting, 2nd Edition [1], which is one of the fundamental sources of information in this area.

The first techniques of spatial load forecasting used pattern recognition to determine preference maps [4], where different factors are weighted to determine a specific land use type. This process was later improved using elements from fuzzy logic [5]. Recently, statistical approaches [6] are used for spatial electric load forecasting.

Later years have not shown special advances in spatial electrical load forecasting [2]. Hence, it is necessary to take advantage of new advances in algorithms to develop new methodologies especially suited to the reality of new cities in emerging countries [1], considering the requirements of new users and comply with new regulations in the sector.

The spatial electric load forecasting process begins with the data collection from different sources. These data are collected from different maps of the utility service area. Those data include different information such as land use classification, geographical setting, socioeconomic distribution, electric structure inventory, historical peak load in different equipment like substations and feeders.

One of the major problems for distribution utilities to carry out a spatial electric load forecasting is the lack of available data [2]. Hence, the method presented in this work considers that the utility have access to basic information of the geographic network.

Recent works of spatial electric load forecasting identifies the load growth using conventional methods [7] and the new load is allocated into the service area using the geographical and socioeconomic information from the spatial database.

2.2. Electric Load Forecasting Techniques

Accurate load forecasting is a difficult task [8] considering the availability of the information and the importance of the outcomes. As an example the load at a given hour is dependent not only on the load at the previous hour, but also on the load at the same hour on the previous day, and on the load at the same hour on the day with the same denomination in the previous week. It is relatively easy to get forecast with about 10% [9] mean absolute error. However, the cost of error is so high in the present scenario that research could help to reduce to a few percent.

The electricity supply industry requires forecasting electricity demand with lead times that range from the short term (1 to 5 years ahead) to the long term (up to 20 years ahead). Based on this load forecasting has three basic techniques. [1]

i. Short Term Load Forecasting – 1 to 5 years ahead
ii. Medium Term Load Forecasting – 6 to 15 years ahead
iii. Long Term Load Forecasting – 16 to 20 years ahead

The temporal features required in load forecasting vary from level to level of a power system, being roughly proportional to the capacity of the power equipment involved. In every sense, distribution planning will be more sensitive to detail about location than other sectors. Therefore, it has higher spatial resolution as shown in figure 2-1.

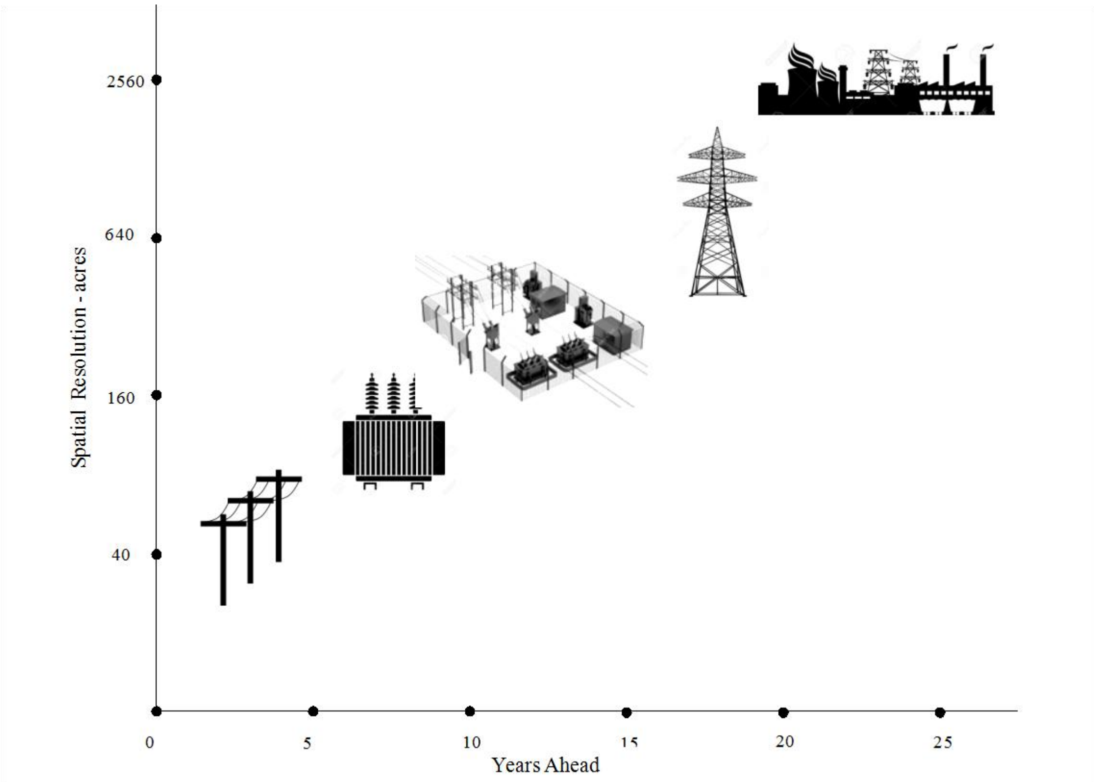

Figure 2-1: Planning Period and Spatial Resolution

It is seen from figure 2-1, that distribution planning belongs to short range period. Various forecasting techniques have been applied to short-term load forecasting to improve accuracy and efficiency. In general, these techniques can be classified as either traditional or modern. Traditional statistical load forecasting techniques, such as regression, time series, pattern recognition etc., have been used in practice for a long time, showing the forecasting accuracy that is system dependent. These traditional methods can be combined using weighted multi-model forecasting techniques, showing adequate results in practical systems. [8] [9]

However, these methods cannot properly represent the complex nonlinear relationships that exist between the load and a series of factors that influence it, which are typically dependent on system changes (eg. season or time of day).

2.2.1. Regression Based Approach

Linear regression is a technique which examines the dependent variable to estimate specified independent variable. The independent variables are where the change occurs. In load forecasting, the dependent variable is usually demand or price of the electricity, because it depends on production which on the other hand depends on the independent variables. Independent variables used in Sri Lanka are usually macro economic factors such as gross domestic product (GDP), population, income, efficiency etc. Essentially, regression analysis attempts to measure the degree of correlation between the dependent and independent variables, thereby establishing the latter's predicted values.

Regression is one of the most widely used statistical techniques [1]. For electric load forecasting, regression methods are usually used to model the relationship of load consumption and other dependant factors for each consumer class.

2.2.2. Time Series Analysis

Time series forecasting is based on the idea that reliable predictions can be achieved by modeling patterns in a time series plot, and then extrapolating those patterns to the future. Using historical data as input, time series analysis fits a model according to seasonality and trend.

Time series models can be accurate in some situations [1], but are especially complex and require large amounts of historical data. Additionally, careful efforts must be made to ensure an accurate time line throughout data collection, filtering, modeling and recall processes. Time series analysis is widely used for forecasting of consumer demand for goods and services [8]. It is not widely used for energy industry forecasting.

2.2.3. Artificial Neural Networks

Artificial Neural Networks (ANN), still at research stage, is electronic model based on the neural structure of the brain. It is designed to learn from the experience just like the brain. The biological inspired methods are thought to be the major advancement in the computational industry. In a neural network, the basic processing element is the neuron. These neurons get input from some source, combine them, perform all necessary operations and put the final results on the output. Artificial neural networks are developed since mid-1980 and extensively applied. They have very successful applications in pattern recognition and many other problems.

Forecasting using ANN is based on the pattern observed from the past event and estimates the values for the future. ANN is well suited to forecasting for two reasons. First, it has been demonstrated that ANN are able to approximate numerically any continuous function to the desired accuracy. In this case the ANN is seen as a multivariate, nonlinear and nonparametric method. Secondly, ANNs are date-driven methods, in the sense that it is not necessary for the researcher to use tentative models and then estimate their parameters. ANNs are able to automatically map the relationship between input and output. They learn this relationship and store this learning into their parameters. [9] [10]

2.3. Fuzzy Logic

Fuzzy Logic is a form of multi-valued logic that attempts to solve problems by assigning values to an imprecise spectrum of data. In fuzzy logic the truth values may be any real number between 0 and 1, so as to arrive at the most accurate conclusion possible. It is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false. Fuzzy logic is designed to solve problems in the same way that humans do, by considering all available information and making the best possible decision given the input.

Therefore, Fuzzy Logic approach, just as ANN, is considered as modern forecasting technique proposed for short term load forecasting. General problem with fuzzy logic approach for energy forecasting is inaccuracy of prediction and numerical instability.

This is due to the fact that with the same parameters this is difficult to find an accurate and comprehensive linear regression curve for fuzzy logic method.

On the contrary, ANN considers load varieties during load forecasting. With ANN one can model complex and nonlinear relationships between input and output thereby giving better accuracy in complex modeling.

2.4. Statement of Problem

It is evident from aforementioned particulars that Spatial load Forecasting is vital for an accurate load forecast in electricity demand thereby accurate planning of the distribution system. In fact, the spatial load forecasting is the basis [1] of the distribution network planning.

There are numerous countries and utilities that employ spatial load forecasting models for the distribution planning. It is difficult to adopt the same model in different contexts as there are various factors that need to be considered, such as requirements from the model, the environment the model will be utilized in, and availability of resources. As an example, it will be difficult to use a load forecasting model used in a European country to be used in a developing Asian country like Sri Lanka, as being a tropical country the energy utilization will not depend heavily on weather patterns as in European countries. Furthermore, the resources will be limited compared to a developed country.

However, despite the fact that spatial electric load forecasting is important for distribution planning, such a spatial electric load forecasting model is still lacking in Sri Lanka.

2.5. Objectives of the Study

Spatial electric load forecasting is a novel concept for Sri Lanka even though similar models have been developed in various developed countries.

The main objective of this research is to develop a spatial electric load forecasting model which can be used in the Sri Lankan context to determine yearly feeder load. The goal of the model is that it should be,

- User-friendly
- Accurate

The spatial electric load forecasting model developed in this research will be used in the Sri Lankan context for distribution planning purposes of Lanka Electricity Company (pvt) Ltd.

2.6. Originality

The originality of this research is developing an easy to use, accurate spatial electric load forecasting model, to be implemented in the Sri Lankan distribution utilities to estimate residential electric load, using Artificial Neural Network as the load forecasting technique.

CHAPTER 3

3. METHODOLOGY

Spatial electric load forecasting is a process that tries to determine the future load growth for an electric energy distribution utility. This process, apart from answering the question of how much load growth is expected in a determined time frame, also identifies the locations that are most likely to receive the new loads.

This research deals with a new method for spatial electric load forecasting, by considering new developed zones and redevelopment of the existing ones. Concepts of development of load in vacant areas and elements from artificial neural network algorithms are applied with special emphasis on developing a simple but powerful method with a reasonable use of available data.

3.1. Research Design

Electric load growth inside the service area of an electric utility can be expected for two reasons. First reason is the natural growth because of the natural behavior of existing consumers. The second reason is addition of new loads because of new consumers. The natural behavior of existing consumers is stationary, with low expected growth, thus, the main reason for load growth is the new consumers, inside and outside the actual service zone.

Most of these new consumers will be located outside the service area, continuing the natural growth of the cities, but others will be located inside the service area in zones that are redeveloped.

In this study, the addition of new consumers is regarded as the new load additions to the vacant areas which are forecasted using the spatial electric load forecasting model. The growth of existing consumers is addressed as an annual constant growth.

Since, the factors affecting load development in vacant areas are not linearly related [1] to the load developed in the vacant area, traditional techniques for load forecasting cannot be used. Artificial neural network models, which have proven characteristics to determine non-linear relationships in recent research is used for modelling purposes of this study.

The method is presented and tested with data from two real midsized cities, Kelaniya and Galle. The results present the applicability and accuracy of proposed method. The validation process of the results is made by comparing the results with the ones obtained in the utility planning department using the existing procedure.

3.2. Customer Class Identification

The type of load (customer class) is often an important factor in distribution planning. Traditionally, basic distinctions of customer class, residential, commercial, industrial etc., have been used by distribution planners because they broadly identify the expected load density, load factor, equipment types and power quality issues on an area basis. Spatial electric load forecasting methods based on forecasting customer type and density have been used since 1930s [1].

Today many spatial forecasts distinguish among subclasses within residential (apartment, small homes, large homes), commercial (retail, low-rise, high-rise, institutional) and industrial (various classes and purposes).

For the scope of this study the residential customer class is employed for modeling of the spatial electric load forecasting model.

3.3. Forecasting Time Period

The transmission and distribution planning procedure consists of five stages as shown in figure 3-1 [1], including a load forecast, followed by coordinated steps of transmission, substation, distribution and customer level planning.

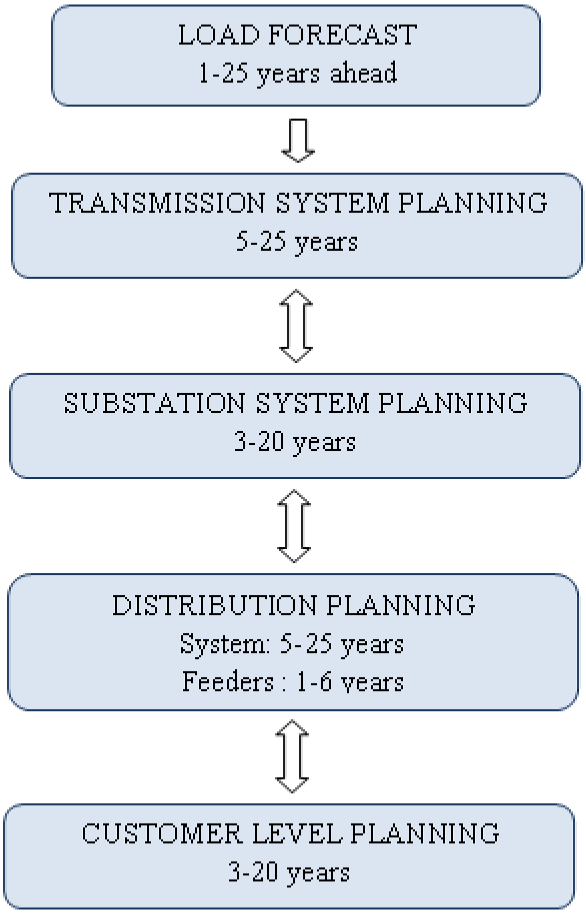

Figure 3-1: The Basic Transmission and Distribution Planning Process

In particular, distribution planning involves detailing the route of each feeder and identifying its equipment to engineering precision. Complete feeder plans include specification of equipment, route maps, pole locations and all construction details. Generally, feeder planning is done by ‘a feeder at a time’ basis. [1]

Taking that into account, the proposed method forecasts the residential load of the selected feeders. Therefore, a time period of one year is considered for the forecast.

3.4. Load Growth Behavior

Peak demand and energy usage within a distribution utility system grow for only two reasons. All demand and energy growth is due to one or a combination of both of these factors. Those two factors are discussed below.

3.4.1. New Consumer Additions

Load will increase if more customers are buying the utility's product that is electricity. New construction and a net population in migration to the utility area will add new customers and increase peak load. With more people purchasing electricity, the peak load and annual energy sales will most likely increase.

3.4.2. New Uses of Electricity

Existing customers may add new appliances or replace existing equipment with improved devices that require more power. With every customer purchasing more electricity, the peak load and annual energy sales will most likely increase. There are no other causes of load growth. Similarly, any decrease in electric demand is due to reductions in either or both of these two factors. Regardless of what happens to the load or how one looks at the load growth or decline, change in one or both of these two factors is what causes any increase or decrease in energy demand.

The bulk of load growth on most power systems is due to the additions of new uses of electricity. Load growth caused by new consumers who are located in previously vacant areas is usually the focus of distribution planning, because this growth occurs where planner has little or no distribution facilities. Therefore, the load in the vacant areas should be derived using the spatial load forecasting model.

The changes in usage among existing customers are more or less stationary with low expected growth. Therefore, it can be addressed with a constant percentage growth.

3.5. Spatial Modeling of Land Use Change

As already mentioned, the task of forecasting the location where changes in land use and customers will occur in the future includes, forecasting which vacant areas will have a new development.

Many factors affect the land use preferences in the utility service area and they differ from context to context they are addressed in. Therefore, it is of vital importance to identify the proper factors that affect the load development of vacant areas of the particular context.

Further, the availability of data is the most important factor in any methodologies dealing with land use determination in spatial electric load forecasting. The residential consumer database available at the distribution utility, LECO, identified all the consumers by name, address, consumer code and class. This information can be cross-referenced with the technical databases using the consumer code. With both databases, it is possible to allocate energy consumption spatially in the service area.

Generally, from the technical database, it is possible to identify the location of the different structures of interest such as schools, hospitals, and commercial zones, among others. The major roads and zones without potential can further be identified from aerial maps, and other information is obtained from the utility planning department, like activity centers, and special urban interest zones, to name a few. The information used in this work is readily available in most of the electric utilities around the world.

In the point of view of this research, identifying the correct factors affecting residential electric load development in vacant areas in Sri Lanka is of vital importance. Further, the requirements of a technical database are fulfilled by the Geographical Information System (GIS) maps available at LECO which are updated annually.

From the information readily available, a spatial database has been implemented to be used for the spatial electric load forecasting model.
Spatial data base includes the following data,

- Current spatial land use data
- Geographic data such as major streets, highways, railroads and urban poles etc.
- Proximity and surrounding factors such as distance to nearest main road, residential load density within the area etc.

3.6. Division of Service Utility Area

In order to realize the forecasts of the planning year, objective service utility area should be divided into a set of small areas. So the spatial load forecast is based on the forecast of every small area.

Generally there are two ways to divide the service area for spatial load forecasting requirements, regular (square or hexagon) and irregular. Theoretically both are possible based on the electric power platform. However, there are many other, not only spatial, factors that decide the division of the service utility area.

Basically, different ways may be used with different planning areas. For a larger area, such as cities, countryside, mountains and rivers, regular division is suitable. However, for urban power system planning, the irregular way is better. That is to divide according to the real streets and functional areas. [1] [11]

In this study, the planning area is divided according to the supply range of each feeder. It is convenient for data collection, flexible and high precision. By this, it is not only more suitable for urban electric power planning but also consistent with data preparation requirement for load forecasting model analysis.

Based on GIS, objective service utility area of LECO can be divided into a set of small areas according to different feeders. Hence, feeder-wise vacant areas and feeder-wise load growth of existing consumers can be forecasted through spatial load forecasting model.

3.7. Artificial Intelligence Based Technique

Several conventional techniques have been used for the load forecasting [12]. However, the disadvantages of these techniques have led to the use of the artificial intelligence based technique. Artificial Neural Network (ANN) is the most proved for forecasting application and cited among the most powerful computational tools ever developed [10] [13] [14]. The reason is ANN methodology solves the complex relationships between the independent and dependent variables by a mathematical mapping algorithm without detailed mathematical modeling.

ANN models can handle large and complex systems with many interrelated parameters. They simply seem to ignore excess data that are of minimal significance and concentrate instead on the more important inputs. [14]

Several types of neural architectures are available, among which the multi-layer back propagation (BP) neural network is the most widely used. A back propagation network typically employs three or more layers for the architecture.

- Input layer
- Output layer
- At least one hidden layer

The BP algorithm is a gradient descent algorithm. It tries to improve the performance of the neural network by reducing the total error by changing the weights along its gradient. The objective of the back propagation algorithm is to minimize the square errors of the system.

In brief, the procedure to set up a back propagation network is,

- Select input and define output variables

- Determine number of layer and the number of neurons in hidden layers. – No rule is available for determining them, and may depend on trial and error.

- Learning (or training) from historical data. – A neural network modifies its weights with respect to the response to inputs to minimize the error. The equation that specifies this change is called the learning rule.

- Testing – When a neural network is trained after learning, it is processed via a test set with historical data that the network has never seen. If the testing results are in an acceptable range, the network can be considered as fully trained.

CHAPTER 4

4. SYSTEM MODELING

The step by step implementation of the spatial electric load forecasting model is discussed in this Chapter. The suggested methodology of this study is simulated using MATLAB toolbox.

4.1. Load Growth due to Existing Customers

The load growth due to existing customers occurs because of the discovery of new uses of electricity consumption. Existing customers may add new appliances or replace existing equipment with improved devices that require more power. With every existing customer buying more electricity, the electricity demand as well as peak load and annual energy sales will most likely increase.

This increase is fairly constant over the years and can be addressed by introducing a constant percentage growth to the existing load. From the data obtained from LECO and referring to values introduced by H. Lee Willis [1], this value is taken as 5% annually.

4.2. Load developed in the Vacant Lands

As discussed in Chapter 3, there are many factors that affect the land use preferences in the utility service area. Those factors can be spatial or generic, macro-economic or micro-economic. The importance is identifying the appropriate and correct factors that affect electric demand. Further, the availability of data should also be taken into consideration when deciding factors that will implement the forecasting model.

Considering all aforementioned requirements, following factors were identified as factors that influence residential load development in vacant lands. [1] [4] [14]

The factors belong to different sectors discussed as in sections 4.2.1, 4.2.2 and 4.2.3.

4.2.1. Geographical Factors

Among the factors that affect electricity load development in vacant lands, the most influencing factors are the geographical factors. These factors attract the new consumers to new lands expanding the city growth. The geographical factors identified based on previous work and experiences of planning engineers are as follows.

1. Distance to nearest school
2. Distance to nearest hospital
3. Distance to nearest major road
4. Distance to main activity centre
5. Local highest temperature
6. Area of the land
7. Entrance road width
8. Distance from feeder to the service pole
9. Geometry of the land
10. Prone to natural disasters

4.2.2. Neighboring Factors

Apart from the geographical factors, the neighboring factors also influence the fact whether there will be a load developed in the given land plot or not. What is adjoining the vacant lands plays a major role in the psychology of the consumers. Though all the geographical factors are in favor, the neighboring factors may influence the final decision.

The neighboring factors that may affect spatial load forecasting are listed as follows.

1. Whether area is residential or not
2. Local cultural background
3. Number of adjacent houses
4. Per person area of the nearby houses

4.2.3. Socio Economic Factors

Socio economic factors are as important as geographical and neighboring factors discussed above. These factors bring the influence of economy to the spatial aspects of new load additions. Therefore it is evident that they will be taken into account to implement the spatial electric load forecasting model.

1. Local GDP
2. Local population
3. Per perch land value
4. Typical electricity consumption of the adjacent houses
5. Typical water consumption of the adjacent houses
6. Per person area of the house

Applying all these factors to the spatial electric load forecasting model will be tedious and inefficient as a means of historical data collection and analysis.

One of the objectives of this study is to develop an easy to use forecasting model. Therefore, from the above factors, the factors that affect electric load development will be identified. Further, there should be a convenient way of collecting data related to these factors, from either or both the databases already present at LECO or GIS maps already available.

Thus, based on availability of data at hand, as the GIS maps and databases at LECO and heuristics of planning engineers at LECO, the factors were filtered out and following parameters are taken into account for the modeling of spatial electric load forecasting. When selecting factors to be used for the model, much attention was paid to select factors from all three sectors discussed, namely, geographical factors, neighboring factors and socio economic factors.

The following factors are confirmed to be used for modelling.

1. Distance to nearest major road
2. Whether area is residential or not
3. Number of adjacent houses
4. Area of the land
5. Entrance road width
6. Distance from feeder to the service pole
7. Typical electricity consumption of the adjacent houses

Once the factors are finalized, the data should be collected to design the spatial electric load forecasting model.

Though the above factors were selected considering the availability of data, taking factors such as local GDP, inflation of the economy into account will give better understanding of the probability of load growth of the locality.

4.3. Implementing the Spatial Database

Once the factors that affect residential load development of vacant lands are finalized, the data related to those factors is collected. These collected data is used to train the ANN from which the load development is forecasted.

An area was identified where there is a possibility to further develop residential load. The area first chosen was Gampaha district. Gampaha being a highly residential area there were almost no vacant lands where data could be collected. Therefore, Maharagama area was selected. From LECO Maharagama Branch Office, information on new electricity connections given for the period of from September 2013 to February 2015 was obtained.

Figure 4-1 shows a part of the pole database extracted to obtain data on new electricity connections. This data is from Maharagama Udahamulla area which belongs to the utility service area of LECO Maharagama Branch Office, Maharagama Customer Service Centre.

		Kandawala			WGS84		
OBJECTID	Name	X	Y	Z	POINT_X	POINT_Y	POINT_Z
1	NR34	104628.892	184392.788	23.781	79.91080514	6.858902837	23.781
2	NR35	104577.054	184397.91	26.209	79.91033606	6.85894831	26.209
3	NR35/3C						
4	NR35/1	104561.751	184404.74	27.772	79.9101975	6.859009819	27.772
5	NR35/2	104530.168	184409.752	26.994	79.90991168	6.859054626	26.994
6	NR35/3	104500.396	184414.058	29.904	79.90964225	6.859093078	29.904
7	NR35/4	104487.384	184417.597	31.099	79.90952447	6.859124867	31.099
8	NR35/5	104467.215	184426.715	28.885	79.90934185	6.859206986	28.885
9	NR35/5/1	104487.336	184465.031	29.248	79.90952327	6.859553773	29.248
10	NR35/5/2	104500.608	184488.835	28.42	79.90964296	6.859769229	28.42
11	NR35/5/2/1	104537.558	184485.504	25.142	79.90997731	6.85973971	25.142
12	NR35/5/2/3						
13	NR35/5/3	104505.02	184499.595	27.86	79.9096827	6.859866595	27.86
14	NR35/5/3A	104463.487	184503.109	22.776	79.90930688	6.859897694	22.776
15	NR35/5/3A1	104474.929	184543.787	24.202	79.90940974	6.860265698	24.202
16	NR35/5/3B	104430.076	184506.383	24.004	79.90900455	6.859926755	24.004
17	NR35/5/3C	104388.156	184511.007	25.094	79.90862521	6.859967884	25.094
18	NR35/5/4	104519.333	184521.42	26.637	79.90981184	6.860064173	26.637
19	NR35/5/5	104530.349	184542.742	24.837	79.90991115	6.86025715	24.837
20	NR35/5/5/1	104572.96	184537.664	21.004	79.91029675	6.860211925	21.004
21	NR35/5/5/2	104604.412	184535.614	20.301	79.91058134	6.8601939	20.301
22	NR35/5/6	104536.556	184555.601	25.674	79.9099671	6.860373524	25.674
23	NR35/5/7	104553.889	184590.265	22.885	79.91012335	6.860687244	22.885
24	NR35/5/8	104564.868	184609.036	21.864	79.91022238	6.860857153	21.864
25	NR35/5/9	104576.217	184631.107	24.545	79.9103247	6.861056908	24.545
26	NR35/5/9-1	104589.162	184655.125	22.264	79.91044143	6.861274293	22.264
27	NR35/5/10	104596.044	184670.871	22.79	79.91050343	6.861416783	22.79
28	NR35/5/11	104596.718	184703.075	22.72	79.91050901	6.861707989	22.72

Figure 4-1: Part of the Pole Database of Maharagama

The pole locations and transformers are then marked in the particular GIS map of that area which will provide the details of the geographic location of the new consumers who have obtained the new electricity supply.

Figure 4-2 shows the GIS map of Maharagama area which depicts the major equipment, feeder routes and poles. The pole locations of considered new connections were marked in this map and that information was made use of to collect data in the site visit.

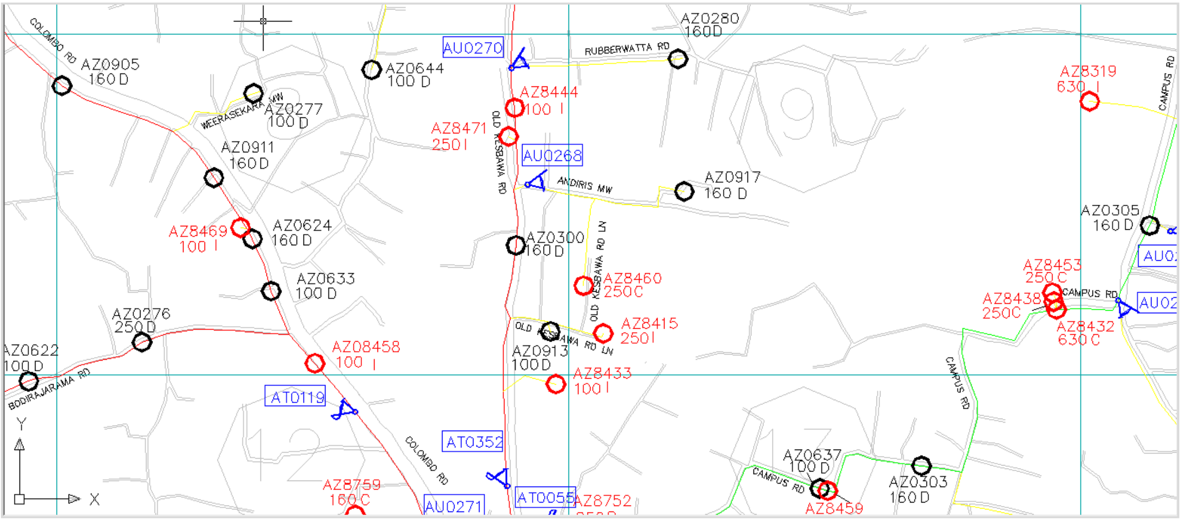

Figure 4-2: GIS Map of Maharagama Area

From the GIS map, the data related to the geographical factors can be obtained without any stress. However, a site visit was needed to obtain data associated with neighboring and socio economic factors. Hence, that information was obtained from a site visit to the area on June 2015. A sample of the data obtained is shown in figure 4-3.

MAHARAGAMA DATA 2							
Land No.	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Nearest Pole No.	**NR34**	**NR35**	**NR35/2**	**NR35/5**	**NR35/5**	**NR35/5**	**NR35/5/**
Distance from road	370	322	265	215	215	218	390
Adjacent Houses	3	3	2	2	3	3	3
perches	21.5	20	22	12	8	15	30
area residential	1	1	1	1	1	1	1
Road width	20	20	20	12	12	12	12
feeder to service point distanc	20	10	10	5	5	3	30
Residential Load (Target)	0	1	0	0	0	0	0

Figure 4-3: Part of the Spatial Database

From the data collected from the GIS map and through site visit, an ANN was developed. The developed ANN model predicted which vacant lands would have a new load developed within the next year.

4.4. Artificial Neural Network Model

The key element of ANN model is the novel structure of its information processing system. An ANN model is composed of a large number of highly interconnected processing elements called neurons, working together to solve specific problems. ANN models, like people, learn by example.

Neural networks are essentially nonlinear circuits that have the demonstrated capability to do non-linear curve fitting. This adaptive information processing system is much suited for this application due to the nature of the relationship between inputs and output of this study.

The ANN for this application is developed using MATLAB Toolbox.

Section 4.4.1 describes the process of modeling the ANN for this application.

4.4.1. Neuron Model

A neuron with a single scalar input and no bias appears on the figure 4-4 below.

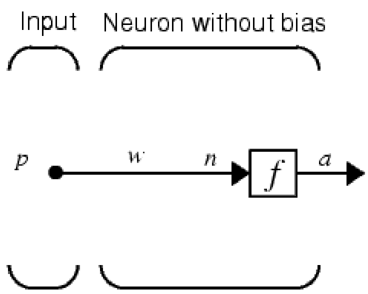

Figure 4-4: A Neuron with a Single Scalar Input

$p = \ Scalar\ input$

$w = \ Scalar\ weight$

$f \ = \ Transfer\ function$

$a = Scalar\ output$

$b = \ Bias$

The scalar input p is transmitted through a connection that multiplies its strength by the scalar weight w, to form the product wp, again a scalar. Here, the weighted input wp is the only argument of the transfer function f, which produces the scalar output a.

$$a = f(wp) \qquad \text{Equation (4.1)}$$

The neuron shown in figure 4-5 has a bias, “b”. The bias is simply being added to the product wp as shown by the summing junction.

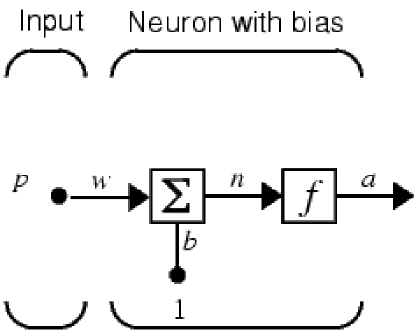

Figure 4-5: A Neuron with a Single Scalar Input and Bias

$$a = f(wp + b) \qquad \text{Equation (4.2)}$$

The transfer function net input “n”, again a scalar, is the sum of the weighted input wp and the bias “b”. This sum is the argument of the transfer function f. The parameters “w” and “b” are both adjustable scalar values of neurons. The basic idea of neural networks is that such parameters can be adjusted so that the network exhibits desired behavior. Thus, the network can be trained to do a particular job by adjusting the weight and/or bias parameters.

4.4.2. Network with Multiple Neurons

A network with multiple neurons is shown in figure 4-6.

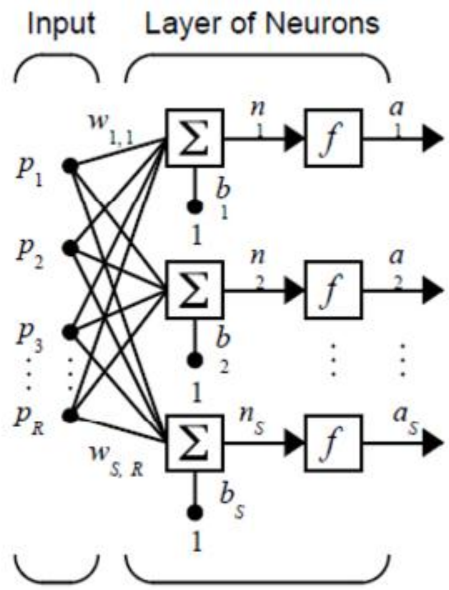

Figure 4-6: Network with Multiple Neurons

A layer includes the combination of the weights, the multiplication and summing operation, realized as a vector product Wp, the bias "b" and the transfer function f. The array of inputs, vector p, is not called a layer. It is called the input. The output can be written as in equation (4.3) to equation (4.5).

$a_i \quad = Output\ of\ node\ i$

$w_{i,j} = Connection\ weight\ between\ node\ i\ and\ j$

$p_i \quad = Input\ signal\ from\ node\ i$

$b_i \quad = Bias\ of\ node\ i$

$$n_1 = p_1 w_{1,1} + p_2 w_{1,2} + p_3 w_{1,3} + \cdots + p_R w_{1,R} + b_1 \qquad \text{Equation (4.3)}$$

$$a_1 = f(n_1) \qquad \text{Equation (4.4)}$$

$$\therefore\ a_i = f(\sum_{j=1}^{R} (p_i w_{i,j} + b_i)) \qquad \text{Equation (4.5)}$$

4.4.3. Transfer Function

There are many transfer functions available to be used for neural network modeling in MATLAB which are shown in figure 4-7.

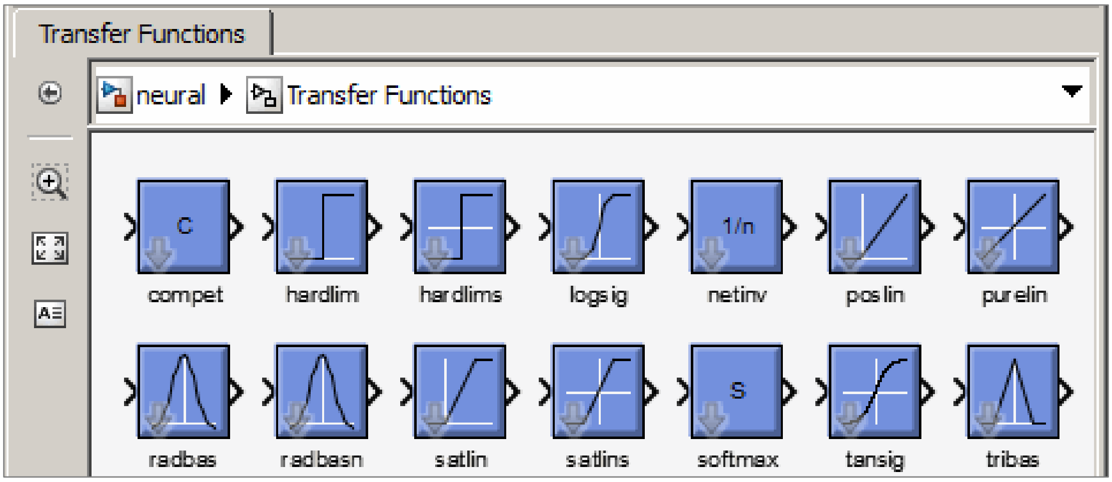

Figure 4-7: Transfer Functions Available in MATLAB

The three most commonly used transfer functions are

- Hardlimit transfer function
- Linear transfer function
- Sigmoid transfer function

The hard-limit transfer function, shown in figure 4-8 and equation (4.6), limits the output of the neuron to either “0”, if the net input argument, n, is less than zero, or “1”, if n is greater than or equal to zero.

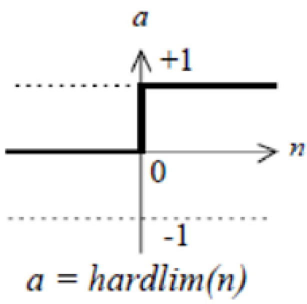

Figure 4-8: Hard-Limit Transfer Function

$$\begin{aligned} &If\ n \leq 0, \quad a = 0 \\ &If\ n > 0, \quad a = 1 \end{aligned} \qquad \text{Equation (4.6)}$$

The linear transfer function characteristic is shown in figure 4-9.

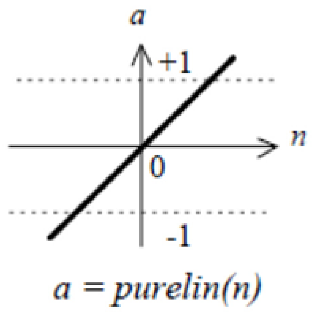

Figure 4-9: Linear Transfer Function

$$a = c.n \quad \text{Equation (4.7)}$$

$c = Gradient\ of\ the\ function$

The sigmoid transfer functions, equation (4.7), take the input, which may have a value between plus and minus infinity, and squashes the output into the range either 0 to 1, for log-sigmoid transfer function, shown in figure 4-10, or -1 to 1, for tan-sigmoid transfer function, shown in figure 4-11.

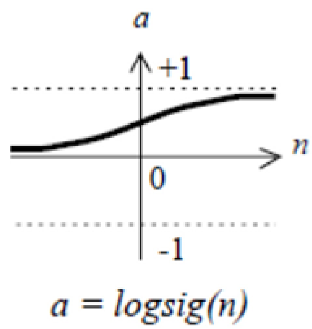

Figure 4-10: Log-Sigmoid Transfer Function

$$a = \frac{1}{1 + e^{-n}} \quad \text{Equation (4.8)}$$

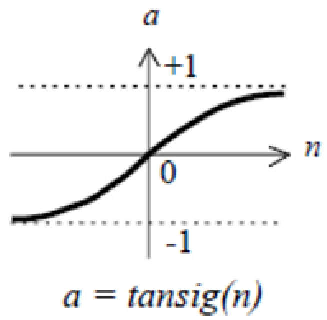

Figure 4-11: Tan-Sigmoid Transfer Function

$$a = \frac{2}{1 + e^{-2n}} - 1$$ Equation (4.9)

The objective of this model is to predict whether the vacant land would develop a load or not next year, once the spatial factors are given. The output of the model should be either yes (1) or no (0). Therefore, the log-sigmoid transfer function is used for this application.

4.4.4. Learning Rule

Learning rule is defined as the procedure for modifying the weights and biases of a network. It is also sometimes referred to as training algorithm. The learning rule is applied to train the network to perform a particular task.

In supervised learning, the learning rule is provided with a set of examples of proper network behavior which is called the training data set.

$$\{p_1 t_1\}, \{p_2 t_2\}, \dots, \{p_m t_m\}$$ Equation (4.10)

$p_i \ = Input\ to\ the\ network$

$t_i \ = Desired\ output\ (Target)$

The learning rule is then used to adjust the weights and biases of the network in order to move the network outputs closer to the targets.

It is very difficult to know which training algorithm will be the fastest for a given problem. It depends on many factors, including the complexity of the problem, the number of data points in the training set, the number of weights and biases in the network, the error goal, and whether the network is being used for pattern recognition (discriminate analysis) or function approximation (regression).

The table 4-1 lists the algorithms that are tested and the acronyms used to identify them.

Table 4-1: Training Algorithms

Acronym	**Algorithm**	**Description**
LM	trainlm	Levenberg-Marquardt
BFG	trainbfg	BFGS Quasi-Newton
RP	trainrp	Resilient Backpropagation
SCG	trainscg	Scaled Conjugate Gradient
CGB	traincgb	Conjugate Gradient with Powell/Beale Restarts
CGF	traincgf	Fletcher-Powell Conjugate Gradient
CGP	traincgp	Polak-Ribiére Conjugate Gradient
OSS	trainoss	One Step Secant
GDX	traingdx	Variable Learning Rate Backpropagation

There are several algorithm characteristics that can be deduced. In general, on function approximation problems, for networks that contain up to a few hundred weights, the Levenberg-Marquardt algorithm will have the fastest convergence [21]. This advantage is especially noticeable if very accurate training is required. In many cases, Levenberg-Marquardt algorithm is able to obtain lower mean square errors than any of the other algorithms tested. However, as the number of weights in the network increases the advantage of Levenberg-Marquardt algorithm decreases.

The Resilient Backpropagation algorithm is the fastest algorithm on pattern recognition [21] problems. However, it does not perform well on function approximation problems. Its performance also degrades as the error goal is reduced.

The conjugate gradient algorithms, in particular Scaled Conjugate Gradient algorithm, perform well, particularly for networks with a large number of weights. The Scaled Conjugate Gradient algorithm is almost as fast as the Levenberg-Marquardt algorithm on function approximation problems [21] (faster for large networks) and is almost as fast as Resilient Backpropagation algorithm on pattern recognition problems. Its performance does not degrade as quickly as Resilient Backpropagation algorithm performance does when the error is reduced.

The performance of BFGS Quasi-Newton algorithm is similar to that of Levenberg-Marquardt algorithm. But the computation required increase with the size of the network [21], because the equivalent of a matrix inverse must be computed at each iteration.

The Variable Learning Rate Backpropagation is usually much slower [21] than the other methods. It is useful when there are situations in which it is better to converge more slowly. For example, when using early stopping, obtaining inconsistent results are possible if an algorithm that converges too quickly is used.

Therefore considering the number of parameters and accuracy requirements, Levenberg-Marquardt algorithm is used as the training algorithm for this application.

4.4.5. Back Propagation

The Back Propagation (BP) algorithm calculates the weight change for a given neuron by comparing the error between the desired output and the ANN model output. If it does not match the error requirement, each weight is revised and backpropagated layer by layer from output layer to hidden layer and to input layer.

The basic structure of an ANN is shown in figure 4-12.

$i1, i2 \quad = Input\ layer\ neurons$

$h1, h2 = Hidden\ layer\ neurons$

$o1, o2 = Output\ layer\ neurons$

$b1, b2 = Bias\ values$

$w1 - \ w8 = Weights\ between\ nodes$

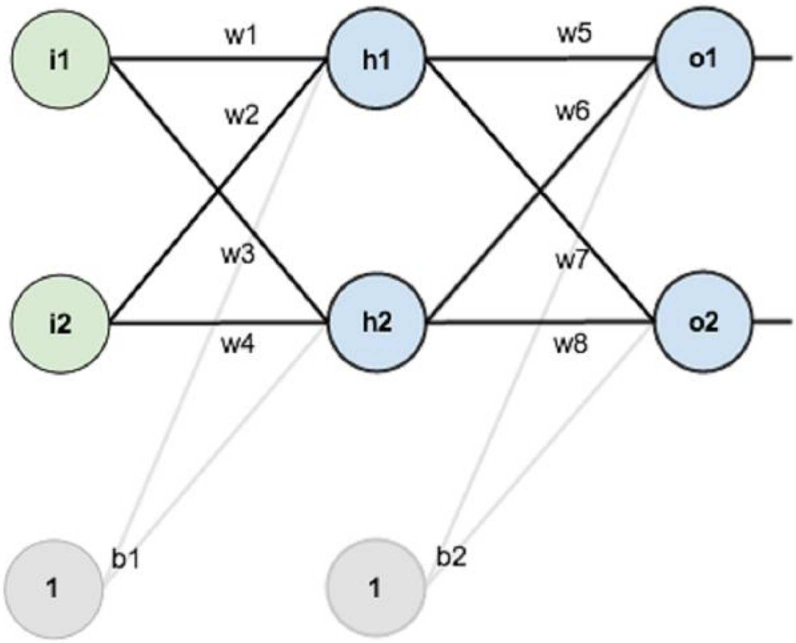

Figure 4-12: The Basic Structure of an ANN

The goal of the backproagation is to optimize the weights so that the neural network can learn how to correctly map arbitrary inputs to outputs.

Following equations derive the process of backpropagation and change of weights and biases of the network model.

The total net input for node h_1 and h_2:

$$net_{h1} = \ w_1 \times i_1 + w_2 \times i_2 + b_1 \times 1 \qquad \text{Equation (4.11)}$$

$$net_{h2} = \ w_3 \times i_1 + w_4 \times i_2 + b_1 \times 1 \qquad \text{Equation (4.12)}$$

Assuming the transfer function of node h_1 and h_2 to be logsid transfer function, the output of h_1 and h_2 can be written as follows.

$$out_{h1} = \frac{1}{1 + e^{-net_{h1}}}$$ Equation (4.13)

$$out_{h2} = \frac{1}{1 + e^{-net_{h2}}}$$ Equation (4.14)

Repeat this process for the output layer neurons, using the output from the hidden layer neurons as inputs.

$$net_{o1} = w_5 \times out_{h1} + w_6 \times out_{h2} + b_2 \times 1$$ Equation (4.15)

$$net_{o2} = w_7 \times out_{h1} + w_8 \times out_{h2} + b_2 \times 1$$ Equation (4.16)

Assuming the transfer function of the output nodes as logsig transfer function, the output for o_1 and $o_2$is,

$$out_{o1} = \frac{1}{1 + e^{-net_{o1}}}$$ Equation (4.17)

$$out_{o2} = \frac{1}{1 + e^{-net_{o2}}}$$ Equation (4.18)

Calculating the total error:

Calculate the error for each output neuron using the squared error function and sum them to get the total error.

$$E_{Total} = \sum \frac{1}{2}(Target - Output)^2$$ Equation (4.19)

$$E_{Total} = E_{o1} + E_{o2}$$ Equation (4.20)

$$E_{Total} = \sum \frac{1}{2}(Target_{o1} - Output_{o1})^2 + \sum \frac{1}{2}(Target_{o2} - Output_{o2})^2$$ Equation (4.21)

$$E_{Total} = \sum \frac{1}{2}(Target_{o1} - Out_{o1})^2 + \sum \frac{1}{2}(Target_{o2} - Out_{o2})^2$$ Equation (4.22)

To update each of the weights in the network, how much change from each weight affects the output will be calculated. This is calculated by taking partial derivatives.

To update the weight w_5 following will be calculated.

$$\frac{\partial E_{Total}}{\partial w_5} \qquad \text{Equation (4.23)}$$

By applying chain rule,

$$\frac{\partial E_{Total}}{\partial w_5} = \frac{\partial E_{Total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5} \qquad \text{Equation (4.24)}$$

The chain rule can be visually shown as figure 4-13.

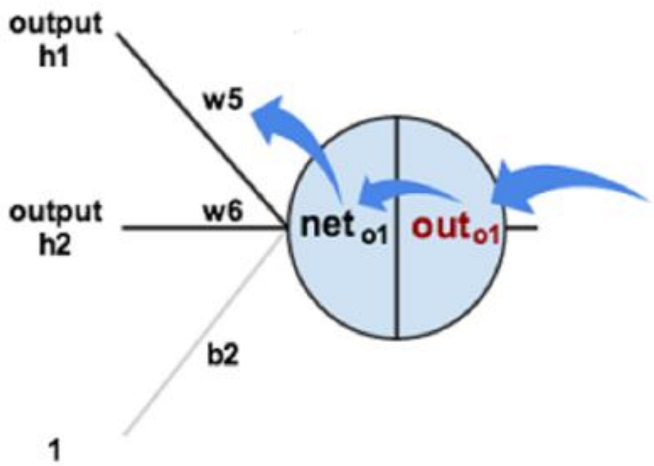

Figure 4-13: Chain Rule

Referring the figure 4-13 and equation (4.22) it is observed that, to calculate total error with respect to the weight value the following will be identified.

- How much does the total error change with respect to the output

  From equation (4.22),

$$\frac{\partial E_{Total}}{\partial out_{o1}} = 2 \times \frac{1}{2}(Target_{o1} - Out_{o1})^{2-1} \times -1 + 0$$
$$= -(Target_{o1} - Out_{o1}) \qquad \text{Equation (4.25)}$$

- How much does the output change with respect to the net input

  From equation (4.17),

$$\frac{\partial out_{o1}}{\partial net_{o1}} = out_{o1} \times (1 - out_{o1})$$ Equation (4.26)

- How much does the total error change with respect to the output

  From equation (4.15),

$$\frac{\partial net_{o1}}{\partial w_5} = 1 \times out_{h1} + 0 + 0 = out_{h1}$$ Equation (4.27)

Putting all together,

$$\frac{\partial E_{Total}}{\partial w_5} = \frac{\partial E_{Total}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial w_5}$$

$$\frac{\partial E_{Total}}{\partial w_5} = -(Target_{o1} - Out_{o1}) \times out_{o1}(1 - out_{o1}) \times out_{h1}$$ Equation (4.28)

The effect of change of other weights on total error can similarly be computed.

To decrease the total error, the change of error due to change of weight value is subtracted from the current weight value. Optionally this value is multiplied by a constant named, learning rate.

The new weight value can be calculated as,

$$w_5^{+} = w_5 - \eta \times \frac{\partial E_{Total}}{\partial w_5}$$ Equation (4.29)

$\eta = Learning\ rate$

The actual update in the neural network is done after calculating the weights of hidden layer neurons.

The backward pass is extended to the hidden layer as follows.

$$\frac{\partial E_{Total}}{\partial w_1} = \frac{\partial E_{Total}}{\partial out_{h1}} \times \frac{\partial out_{h1}}{\partial net_{h1}} \times \frac{\partial net_{h1}}{\partial w_1}$$ Equation (4.30)

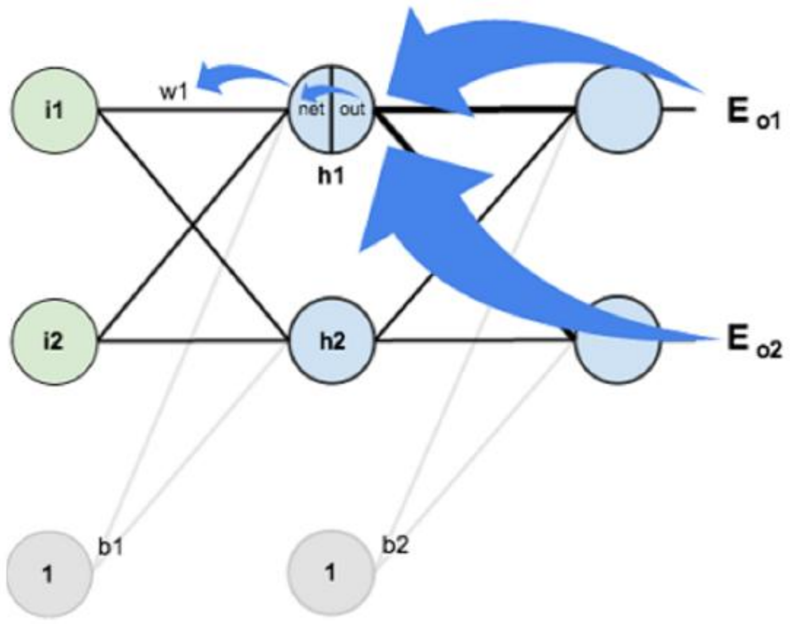

Figure 4-14: Backward Pass to the Hidden Layer

Output of each hidden layer neuron contributes to the output and therefore error of multiple output neurons. Therefore,

$$\frac{\partial E_{Total}}{\partial out_{h1}} = \frac{\partial E_{01}}{\partial out_{h1}} + \frac{\partial E_{02}}{\partial out_{h1}} \quad \text{Equation (4.31)}$$

Starting with,

$$\frac{\partial E_{01}}{\partial out_{h1}} = \frac{\partial E_{01}}{\partial net_{o1}} \times \frac{\partial net_{01}}{\partial out_{h1}} = \frac{\partial E_{01}}{\partial out_{o1}} \times \frac{\partial out_{o1}}{\partial net_{o1}} \times \frac{\partial net_{o1}}{\partial out_{h1}} \quad \text{Equation (4.32)}$$

Substituting values for equation (4.32),

$$\frac{\partial E_{01}}{\partial out_{h1}} = -(Target_{o1} - Out_{o1}) \times out_{o1}(1 - out_{o1}) \times w_5 \quad \text{Equation (4.33)}$$

Similarly,

$$\frac{\partial E_{02}}{\partial out_{h1}} = \frac{\partial E_{02}}{\partial net_{o2}} \times \frac{\partial net_{02}}{\partial out_{h1}} = \frac{\partial E_{02}}{\partial out_{o2}} \times \frac{\partial out_{o2}}{\partial net_{o2}} \times \frac{\partial net_{o2}}{\partial out_{h1}} \quad \text{Equation (4.34)}$$

$$\frac{\partial E_{02}}{\partial out_{h1}} = -(Target_{o2} - Out_{o2}) \times out_{o2}(1 - out_{o2}) \times w_7$$

Therefore,

$$\frac{\partial E_{Total}}{\partial out_{h1}} = -(Target_{o1} - Out_{o1}) \times out_{o1}(1 - out_{o1}) \times w_5 - (Target_{o2} - Out_{o2}) \times out_{o2}(1 - out_{o2}) \times w_7$$ Equation (4.35)

Then from equation (4.11) and equation (4.13),

$$\frac{\partial out_{h1}}{\partial net_{h1}} = out_{h1}(1 - out_{h1})$$ Equation (4.36)

$$\frac{\partial net_{h1}}{\partial w_1} = i_1$$ Equation (4.37

Putting it all together,

$$\frac{\partial E_{Total}}{\partial w_1} = [-(Target_{o1} - Out_{o1}) \times out_{o1}(1 - out_{o1}) \times w_5 - (Target_{o2} - Out_{o2}) \times out_{o2}(1 - out_{o2}) \times w_7 \times \frac{\partial out_{h1}}{\partial net_{h1}} \times \frac{\partial net_{h1}}{\partial w_1}] \times out_{h1}(1 - out_{h1}) \times i_1$$ Equation (4.38)

The weight w_1 can now be updated as,

$$w_1^{+} = w_1 - \eta \times \frac{\partial E_{Total}}{\partial w_1}$$ Equation (4.39)

All the weights and bias values can be updated in this manner. This process may repeat several times before the required error requirements are reached.

4.4.6. Training and Validation

After defining the structure of the ANN, the network model needs to be trained properly in order to successfully detect possible outlying events.

The data used for training the network is subdivided into three sets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases as discussed in section 4.4.5.

The second subset is the validation set and the third set is the testing set. 60% of the samples are assigned to the training set, 20% to the validation set, and 20% to the test set. This is a process performed by the network, where the data are randomly divided. Samples for each of these three sets are constructed with a fair representation of all data set classes.

The criteria to define the performance of ANN model when trained, is the Mean Absolute Percentage Error (MAPE) value.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Target_i - Output_i}{Target_i}\right| \qquad \text{Equation (4.40)}$$

The required error value was set as 5% for this study. The variation of the MAPE during training for different values of network parameters is observed. The parameters which give the minimum deviation with the set MAPE are considered to be optimum for a network.

CHAPTER 5

5. ANALYSIS OF PROBLEM

The spatial data with respect to the spatial factors identified, were collected from Maharagama and Galle regions using site visits and GIS maps at LECO. Data were collected feederwise. The ANN model trained was then feed-forwarded by these data and the output was taken as to whether a load will be developed in the vacant areas of that feeder. The load addition from existing consumers was addressed as a constant percentage growth. Addition of these two was taken as the final load of the feeder.

The neural fitting tool within the MATLAB neural network toolbox was used to conduct all modeling and analysis. It was used to test the accuracy of load prediction on data from Maharagama and Galle regions. The correctness of the results was evaluated with the mean squared error of the predicted values with respect to the actual load values and values obtained from the existing load forecasting technique.

5.1. Simulation of ANN Model

In ANN, multi-layer networks are most commonly used for the forecasting application [12]. In this work 3-layered network is used. The number of neurons in the input layer is equal to the vector elements of input vector. Inputs to the network are as follows and therefore the number of neurons in the input layer is six.

1. Distance to nearest major road
2. Whether area is residential or not
3. Number of adjacent houses
4. Area of the land
5. Entrance road width
6. Distance from feeder to the service pole

The output layer consists of one neuron giving the decision of whether a load will be developed in the particular vacant lands whose spatial data mentioned above, are inputted to the network. If load will be developed in the next year, the output of the network would be one, otherwise 0.

The number of neurons in the hidden layer of the network was taken to be 25. The number of neurons in the hidden layer affects the speed of the training and how efficiently the neural network can process data, for a given number of data points and the algorithm used. The default number of neurons in MATLAB's neural network tool box is set to be 20. For the purpose of comparison, 15, 20, 25 and 30 neurons were tested in order to find which would yield the best results.

Log-sigmoid (logsid) transfer function which squashes the output between 0 and 1 is used for the hidden layer and linear activation function is used in the output layer.

The two most accurate back propagation training algorithms [21] in MATLAB are Levenberg-Marquardt algorithm and Bayesian Regularization algorithm. The Levenberg-Marquardt algorithm is used for this application as it gives the optimum results for medium sized NNs.

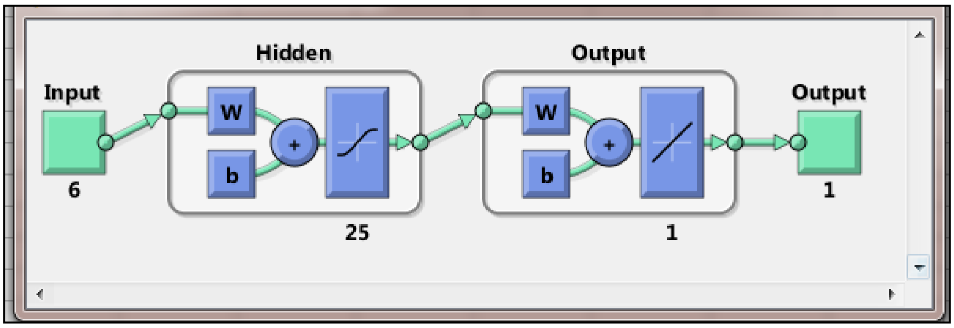

Figure 5-1: Artificial Neural Network Model

Figure 5-1 illustrates the neural network model implemented in MATLAB. It has an input layer with six input parameters, a hidden layer with 25 hidden neurons and an output layer with one output parameter. The input parameters are multiplied by the weight matrix of the hidden layer and the bias matrix is added. The summation is fed into the hidden layer transfer function, Log-sigmoid (logsid) transfer function, to get the output from the hidden layer.

The output from the hidden layer is then multiplied by the weight matrix of the output layer and the bias matrix is added. The summation is fed into the transfer function of the output layer, linear activation function, to get the output of the model.

The MATLAB simulation window is shown in figure 5-2. The information of input matrix, output matrix, hidden layer size and validation matrices are available at the workspace to right hand side of the window. The selected information appears in the matrix view at the middle of the window.

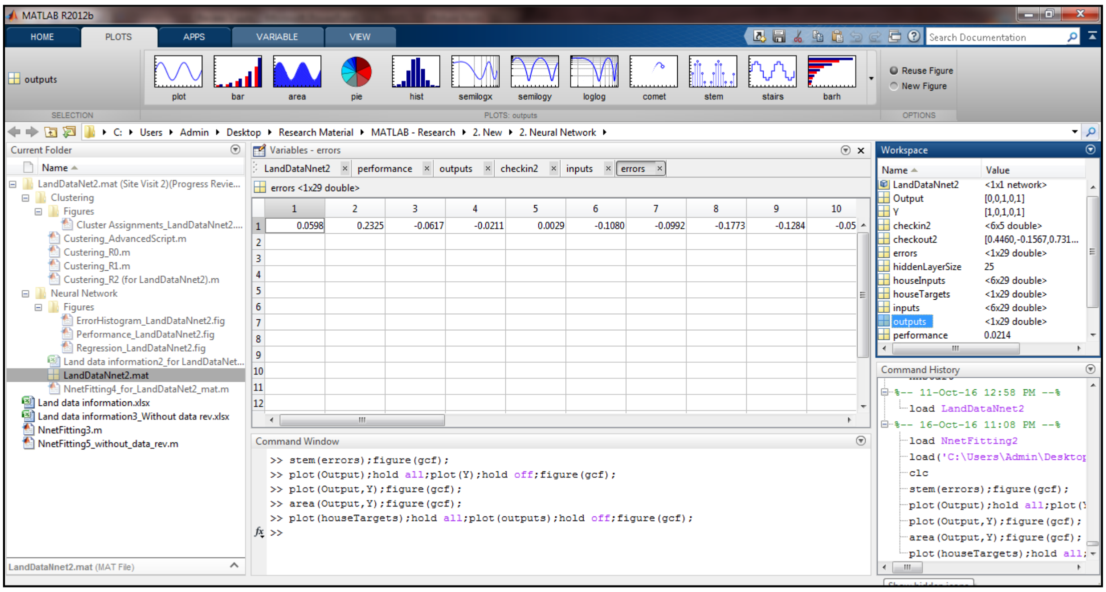

Figure 5-2: MATLAB Simulation

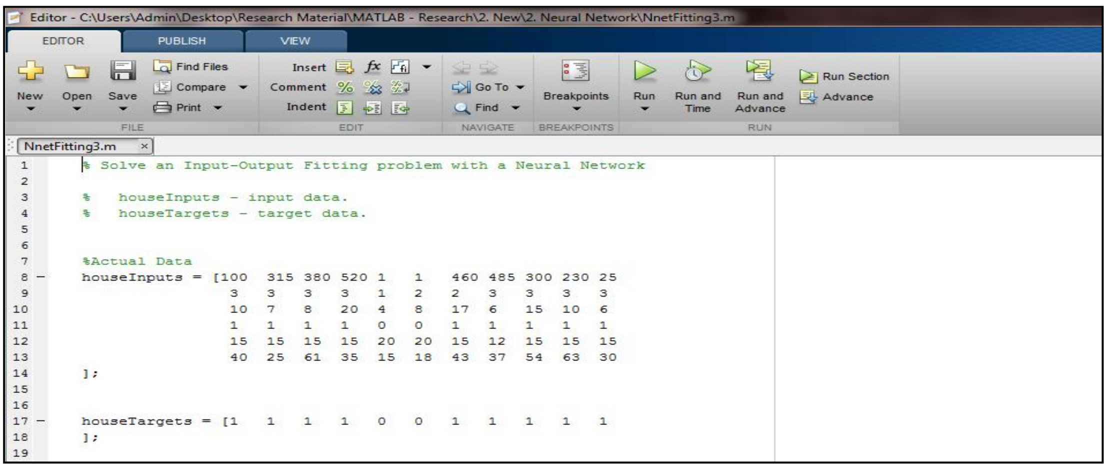

Figure 5-3: Part of the MATLAB m-file

The data to the NN model is fed in, in the form of an m-file. M-files are convenient in introducing user specific data and processes into the model. Further, it facilitates trouble-free developments of the parameters. Figure 5-3 is a part of the m-file used in this application.

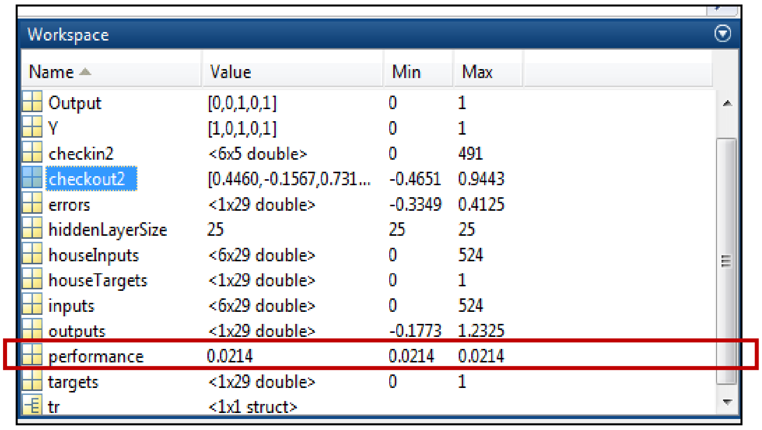

Figure 5-4: MATLAB Performance Values

The criteria to define the performance of ANN model when trained, is the Mean Absolute Percentage Error (MAPE) value. This value is set for 5% or this application. However, according to figure 5-4, the performance is 2%.

The regression plots show the relationship between the outputs of the network and the targets. If the training was perfect, the network outputs and the targets would be equal and therefore give an R value equals to 1. The dashed lines present the perfect results. The solid line represents the best fit linear regression line between outputs and targets.

Figure 5-5 shows the regression plots representing the training, validation, and testing data of this study. All the R values are greater than 0.8 reflecting good relationship between outputs of the network and the targets.

The scatter plot is helpful in showing any outliers. If there are any outliers, additional data should be collected on these points and used in the test set. When the number of outliers increases, the performance of the network deteriorates. This is further discussed in section 5.2.

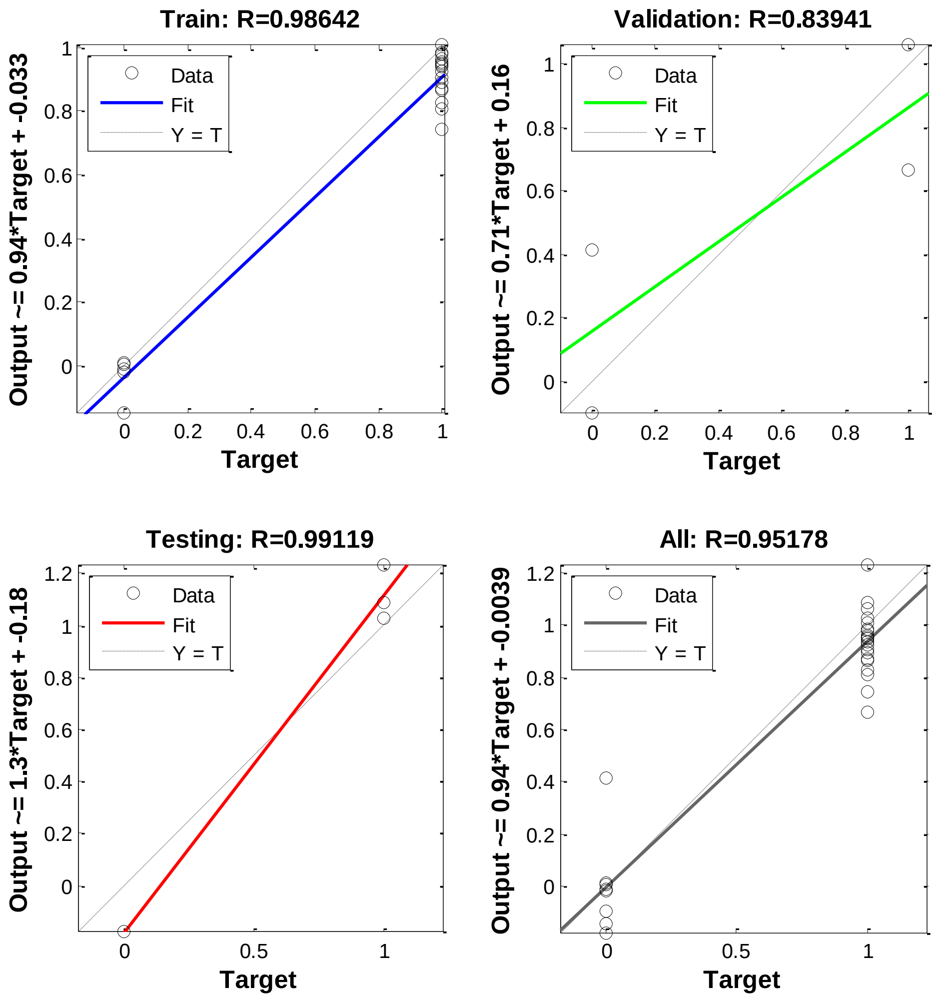

Figure 5-5: Regression Plot of the Model

5.2. Cluster Analysis

Clustering is a physical or abstract process of grouping and collecting of objects. [17] A cluster analysis was done on the target values and the outputs of neural network model. The objective was to identify any outliers of the data set. If there were outliers, spatial data sets of similar characteristics would further be incorporated to the neural network model to get an optimum output.

5.2.1. K-means Clustering

K-means clustering in MATLAB Toolbox is used to perform clustering analysis of this study. It uses a two-phase iterative algorithm to minimize the sum of point-to-centroid distances, summed over all K number of clusters.

The first phase uses batch updates, where iteration consists of reassigning points to their nearest cluster centroid, all at once, followed by recalculation of cluster centroids. This phase may be thought of as providing a fast but potentially only approximate solution as a starting point for the second phase.

The second phase uses on-line updates, where points are individually reassigned if doing so will reduce the sum of distances, and cluster centroids are recomputed after each reassignment. Iteration during this second phase consists of one pass though all the points. K-means clustering can converge to a local optimum, which is a partition of points in which moving any single point to a different cluster increases the total sum of distances.

As seen in the figure 5-6 the output data points are divided to two clusters. They are on either side of the diagonal line of the plot which means the data points are well clustered. If any data point lied near to the diagonal line of the plot, it will be regarded as an outlier.

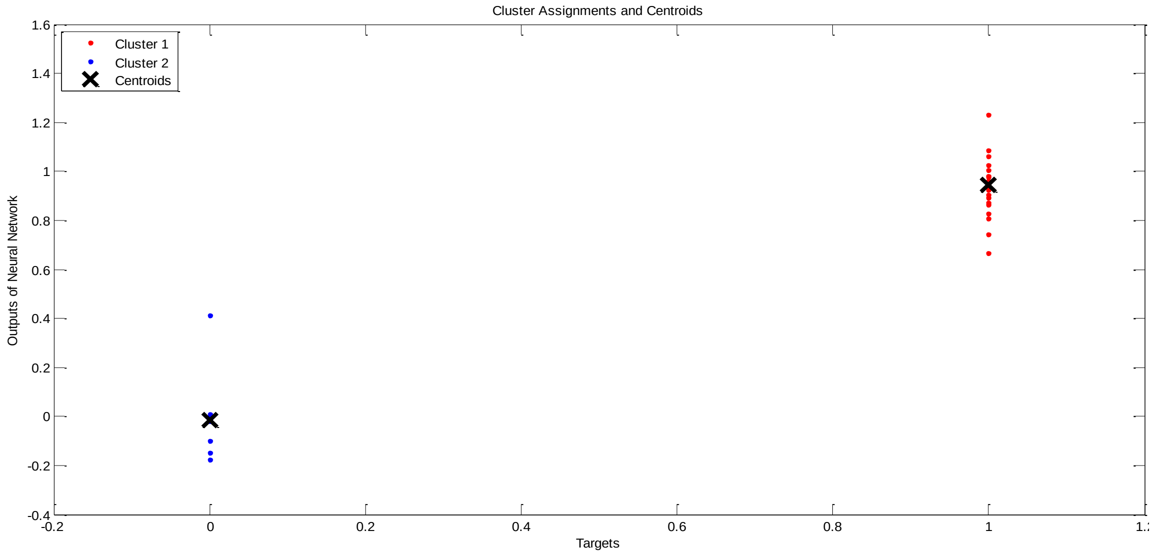

Figure 5-6: Cluster Plot of the Target and Output

5.3. Average Consumption of Electricity

The average consumption of electricity depends on the area and the customer class. It is statistically derived that average consumption of adjacent houses is rather same.

Therefore, the average consumption of residential customers in each feeder can be calculated and this value can be used to simply multiply the output of the neural network model. i.e. We assume, if a new load will develop in the vacant land, the new load that is developed will be equal to the average electricity consumption of that particular feeder. This is fairly a justifiable assumption [1] [11] according to the references and heuristics.

This study validates the model on data from Kelaniya and Galle regions. The average electricity of each feeder is calculated from the data available at the LECO EMR Reports. Figure 5-7 shows a Part of EMR Report LECO Kelaniya Area Dalugama CSC.

Dalugama	JANUARY										
	FDS Measurments				R		Y		B		
Transformer No	R	Y	B	Feeder No	Curren	End Voltag	Curren	End Voltag	Current	End Voltage	Status
AZ-043	234	239	235	1	72	230	32	235	62	228	0
AZ-043	234	239	235	2	145	228	93	210	150	205	1
AZ-043	234	239	235	3	35	220	36	224	44	215	1
AZ-043	234	239	235	4	44	232	30	234	31	231	0
AZ-050	235	233	235	1	46	233	48	230	48	229	0
AZ-050	235	233	235	2	20	234	28	230	55	228	0
AZ-050	235	233	235	3	53	231	50	232	24	231	0
AZ-050	235	233	235	4	52	230	49	225	54	226	0
AZ-051	248	249	249	1	34	244	14	245	39	242	0
AZ-051	248	249	249	2	58	230	54	233	87	232	1
AZ-051	248	249	249	3	125	229	118	229	113	230	1
AZ-052	232	234	234	1	52	230	51	232	60	230	0
AZ-052	232	234	234	2	46	231	36	233	47	233	0
AZ-052	232	234	234	3	55	230	33	231	41	229	0
AZ-052	232	234	234	4	81	230	53	232	41	232	0
AZ-053	222	223	223	1	46	220	24	220	33	221	0

Figure 5-7: Part of EMR Report LECO Kelaniya Area Dalugama CSC

5.4. Feeder Residential Load

Putting all together the feeder residential load can be calculated as follows.

$$\begin{aligned}&Feeder\ Residential\ Load\\&= Occupancy\ of\ the\ land\ derived\ from\ the\ ANN\ model\\&\times Average\ electricy\ consumption\ of\ adjacent\ houses\\&+ Existing\ customer\ load\ of\ the\ feeder \times 105\%\end{aligned}$$ Equation (4.41)

According to the records 2014-2016, the average electricity consumption at Kelaniya (selected) feeders is 3.7A and the average electricity consumption at Galle (selected) feeders is 2.6A.

CHAPTER 6

6. RESULTS

The ANN model is developed using the data from Udahamulla area belong to Maharagama CSC of Maharagama Area Office. To validate the spatial electric load forecasting model's functionality it is being implemented on data of Kelaniya and Galle region. The results are compared with actual values and output obtained from LECO Regression Analysis Software. The obtained results are presented in tabulated form in table 6-1 and table 6-2 and in plots for better understanding.

Table 6-1: Kelaniya Area Feeder Residential Load

KELANIYA BRANCH OFFICE – LECO									
	2014 values in A			2015 values in A			2016 values in A		
	Real *	SLFM **	Regre. ***	Real *	SLFM **	Regre. ***	Real *	SLFM **	Regre. ***
Dalugama									
AZ-043 F1	31.59	32.23	33.55	43.60	44.60	46.65	55.30	57.80	59.89
AZ-051 F2	42.66	41.55	40.01	54.60	52.96	50.78	66.30	64.10	60.80
Wattala									
AZ-0083 F1	58.57	57.05	54.94	65.60	63.30	61.01	70.30	67.50	64.47
AZ-0048 F2	29.23	29.99	31.04	38.00	39.94	40.66	46.00	48.20	49.82
MAPE (%)		2.45%	6.20%		3.48%	7.00%		4.15%	8.30%

* Data from LECO EMR Reports 2014-2016
** Data for ANN was extracted from GIS maps 2013-2015
*** Calculated from LECO Regression Analysis Software

Table 6-1 demonstrates the results obtained when the model is tested with the data from Kelaniya area. Two customer service centres (CSC) are considered, two feeders from each. From Dalugama CSC, feeder 1 (F1) connected to transformer AZ-043 and feeder 2 (F2) connected to transformer AZ-051 are considered. Similarly, data from Wattala CSC, feeder 1 (F1) connected to transformer AZ-0083 and feeder 2 (F2) connected to transformer AZ-0048 are taken into account to validate the results.

The current from each feeder, Dalugama AZ-043 F1, AZ-051 F2 and Wattala AZ-0083 F1, AZ-0048 F2, for years 2014, 2015 and 2016 are estimated using the model proposed in this research. The results obtained from the model are listed under the column SLFM (Spatial Electric Load Forecasting Model) in the table 6-1.

The actual current values for the feeders, for years 2014, 2015 and 2016, are obtained from the LECO Energy Management Reports (EMR). They are tabulated under the column Real in the table 6-1.

The estimated current values for the Dalugama AZ-043 F1, AZ-051 F2 and Wattala AZ-0083 F1, AZ-0048 F2 feeders, for years 2014, 2015 and 2016, from the Regression Analysis Software currently used at LECO is mentioned under the column Regre. in the table 6-1.

The mean absolute percentage error (MAPE) between the real current values and the output from the spatial electric load forecasting model proposed is presented for each feeder in each year in the table 6-1. It is observed that the MAPE between the real current values and the output from the SLFM proposed, averaged for the four feeders for year 2014 is 2.45%. The MAPE averaged between real values and values obtained from LECO Regression Analysis Software for year 2014 is 6.2%. For year 2015, MAPE between the real current values and the output from the SLFM proposed is 3.48% and MAPE averaged between real values and values obtained from LECO Regression Analysis Software for year 2014 is 7.0%. For year 2016, they are 4.15% and 8.3% respectively. Therefore, from these results it is clear that the proposed model has better accuracy than the conventional method used.

The graphs depicted in figure 6-1, figure 6-2, figure 6-3 and figure 6-4 gives a better representation of the tabulated results in table 6-1. It is seen from the graphs that the plot of SLFM fits the real values plot better than the regression model plot. Therefore, the figure 6-1 to figure 6-4 proves that the SLFM model shows a higher level of performance than the regression model.

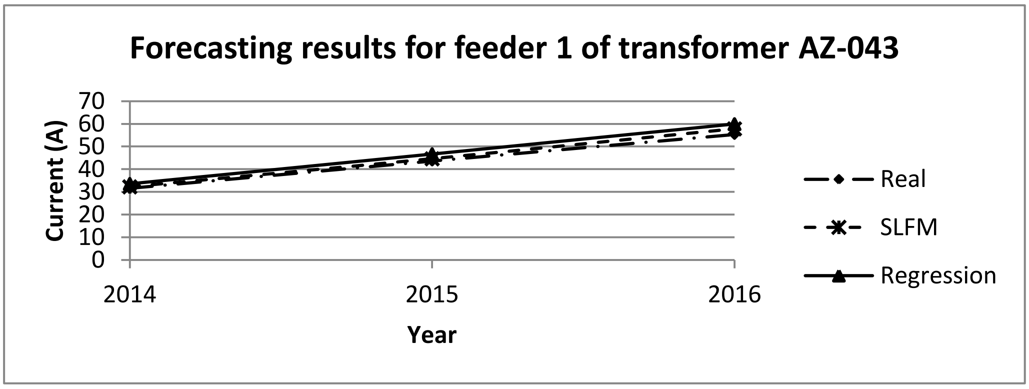

Figure 6-1: Forecast for feeder 1 of transformer AZ-043, Dalugama CSC

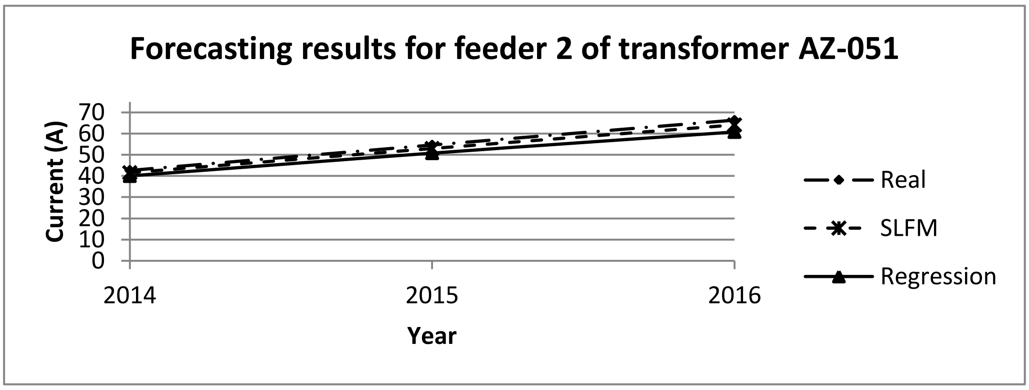

Figure 6-2: Forecast for feeder 2 of transformer AZ-051, Dalugama CSC

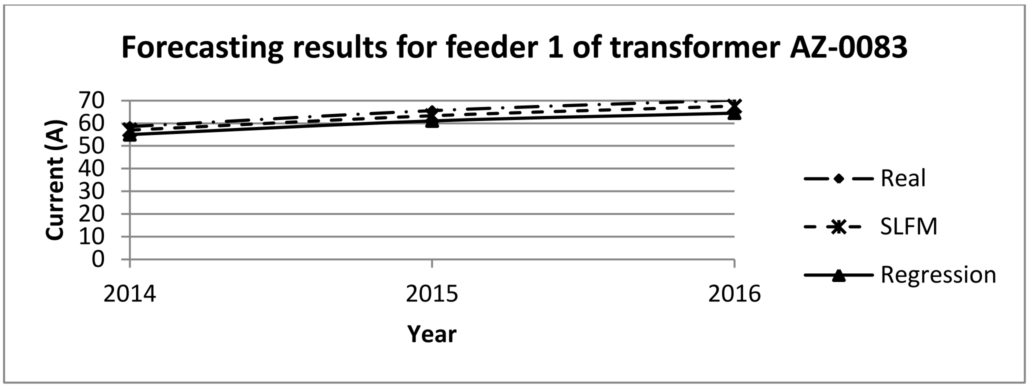

Figure 6-3: Forecast for feeder 1 of transformer AZ-0083, Wattala CSC

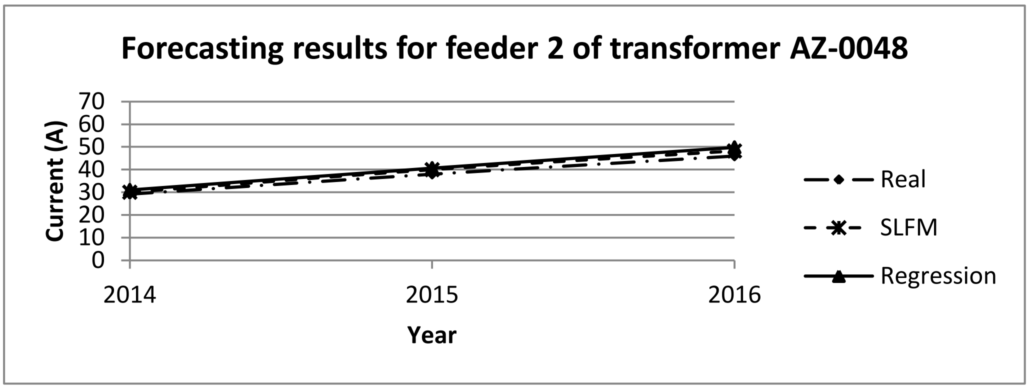

Figure 6-4: Forecast for feeder 2 of transformer AZ-0048, Wattala CSC

Table 6-2 demonstrates the results obtained when the model is tested with the data from Galle area. Two customer service centres (CSC) are considered, two feeders from each. From Hikkaduwa CSC, feeder 1 (F1) connected to transformer AZ-5612 and feeder 1 (F1) connected to transformer AZ-5517 are considered. Similarly, data from Ambalangoda CSC, feeder 2 (F2) connected to transformer AZ-5005 and feeder 1 (F1) connected to transformer AZ-5006 are taken into account to validate the results.

The current from each feeder, Hikkaduwa AZ-5612 F1, AZ-5517 F1 and Ambalangoda AZ-5005 F2, AZ-5006 F1, for years 2014, 2015 and 2016 are estimated using the model proposed in this research. The results obtained from the model are listed under the column SLFM (Spatial Electric Load Forecasting Model) in the table 6-2.

The actual current values for the feeders, for years 2014, 2015 and 2016, are obtained from the LECO Energy Management Reports (EMR). They are tabulated under the column Real in the table 6-2.

The estimated current values for the Hikkaduwa AZ-5612 F1, AZ-5517 F1 and Ambalangoda AZ-5005 F2, AZ-5006 F1 feeders, for years 2014, 2015 and 2016, from the Regression Analysis Software currently used at LECO is mentioned under the column Regre. in the table 6-2.

The mean absolute percentage error (MAPE) between the real current values and the output from the spatial electric load forecasting model proposed is presented for each feeder in each year in the table 6-2. It is observed that the MAPE between the real current values and the output from the SLFM proposed, averaged for the four feeders for year 2014 is 4.75%. The MAPE averaged between real values and values obtained from LECO Regression Analysis Software for year 2014 is 8.1%. For year 2015, MAPE between the real current values and the output from the SLFM proposed is 3.78% and MAPE averaged between real values and values obtained from LECO Regression Analysis Software for year 2014 is 6.9%. For year 2016, they are 4.90% and 8.4% respectively. Therefore, from these results it is clear that the proposed model has better accuracy than the conventional method used.

Table 6-2: Galle Area Feeder Residential Load

GALLE BRANCH OFFICE – LECO									
	2014 values in A			2015 values in A			2016 values in A		
	Real *	SLFM **	Regre. ***	Real *	SLFM **	Regre. ***	Real *	SLFM **	Regre. ***
Hikkaduwa									
AZ5612 F1	16.36	16.65	17.68	22.57	23.21	24.13	28.67	30.45	31.08
AZ5517 F1	16.60	15.80	15.25	24.56	23.80	22.87	32.67	30.94	29.93
Ambalangoda									
AZ5005 F2	14.30	15.30	15.46	21.60	22.46	23.09	30.67	31.80	33.25
AZ5006 F1	15.62	16.47	16.88	19.21	20.21	20.54	22.67	23.67	24.57
MAPE (%)		4.75%	8.10%		3.78%	6.90%		4.90%	8.40%

*	Data from LECO EMR Reports 2014-2016
**	Data for ANN was extracted from GIS maps 2013-2015
***	Calculated from LECO Regression Analysis Software

The graphs depicted in figure 6-5, figure 6-6, figure 6-7 and figure 6-8 gives a better representation of the tabulated results in table 6-2. It is seen from the graphs that the plot of SLFM fits the real values plot better than the regression model plot. Therefore, the figure 6-5 to figure 6-8 proves that the SLFM model shows a higher level of performance than the regression model.

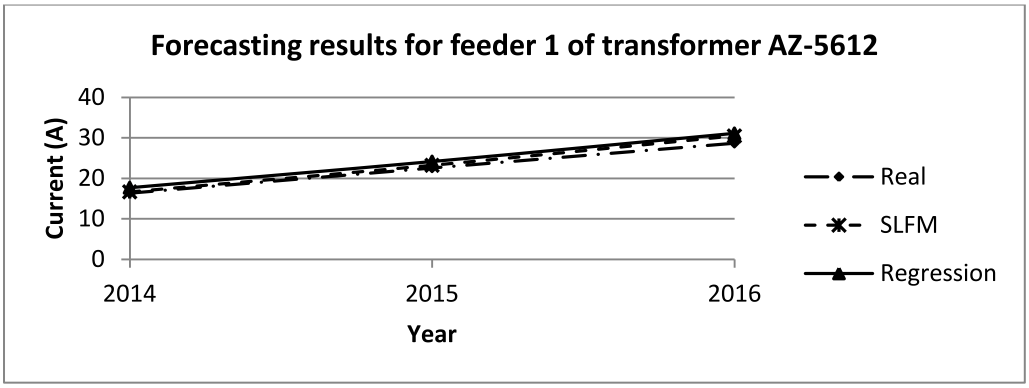

Figure 6-5: Forecast for feeder 1 of transformer AZ-5612, Hikkaduwa CSC

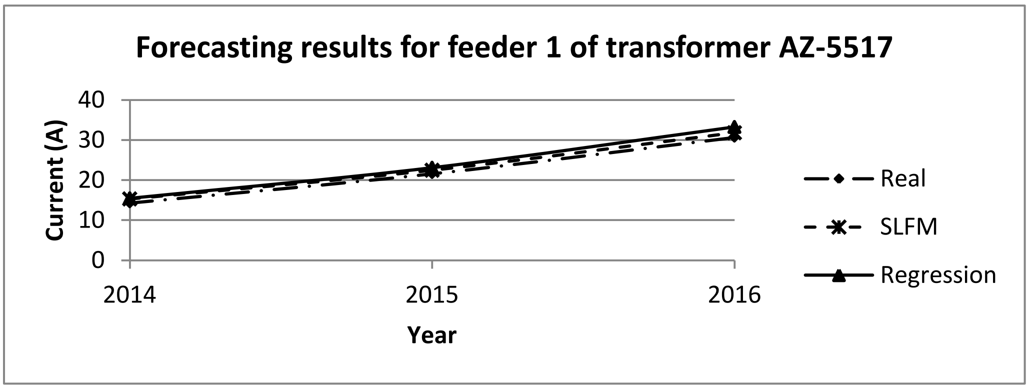

Figure 6-6: Forecast for feeder 1 of transformer AZ-5517, Hikkaduwa CSC

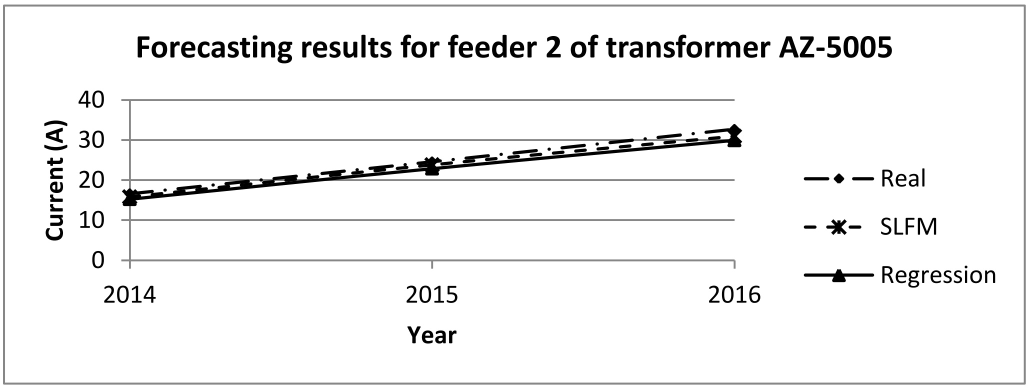

Figure 6-7: Forecast for feeder 2 of transformer AZ-5005, Ambalangoda CSC

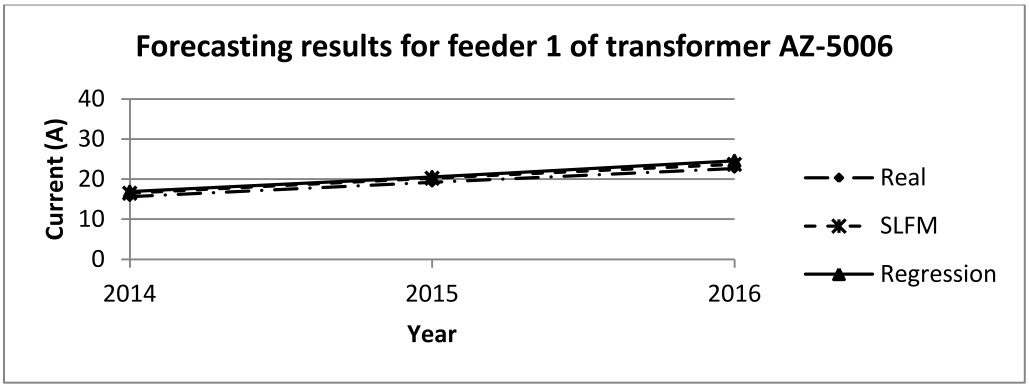

Figure 6-8: Forecast for feeder 1 of transformer AZ-5006, Ambalangoda CSC

From the results it can be concluded that the spatial electric load forecasting model has better accuracy than the regression based methods. It can be used for spatial areas with different load behaviours, such as Kelaniya and Galle, without losing accuracy. The MAPE of the spatial electric load forecasting model is less than 5% while the MAPE of the regression model is about 10%, proving better accuracy in the proposed method.

CHAPTER 7

7. Conclusions

Most utilities have in the past undertaken aggregate load forecasting, which are suitable for planning the generation and planning networks. Most load forecasting emphasized aggregate load forecasting. Unfortunately, the results of such load forecasting does not identify where the power loads takes place and are also are not helpful in the location and construction planning of distribution power facilities. To plan the efficient operation and economical capital expansion of the electric power delivery system, a prediction of future electric demand that includes 'Where' will it be needed as one its chief elements, in addition to magnitude (how much) and temporal (when) characteristics must be anticipated.

In this study, a simple yet accurate and efficient algorithm to determine future land use and occupation in the spatial electric load forecasting process has been presented by considering the identification of vacant lands with potential growth and the redevelopment of the existing ones.

Electric load growth inside the service area of an electric utility can be expected for two reasons, natural growth because of the natural behaviour of existing consumers and addition of new loads because of new consumers. The natural behaviour of existing consumers is stationary, with low expected growth. Thus, the main reason for load growth is the new consumers inside and outside the actual service zone.

The addition of new consumers is regarded as the new load additions to the vacant lands which are forecasted using the spatial electric load forecasting model. The user preferences about land use have been taken into consideration, when determining the potential load growth of vacant lands. The growth of existing consumers is addressed as an annual constant growth.

Since, the factors affecting load development in vacant areas are not linearly related to the load developed, traditional techniques for load forecasting cannot be used. Artificial neural network model, which have proven characteristics to model nonlinear relationships in recent researches, is used to implement the spatial electric load forecasting model of this study. The output of Artificial Neural Network (ANN) model is combined with the load growth of the existing consumers and arrived at the final load.

Today many spatial forecasts distinguish among residential, commercial and industrial subclasses of customers. For the scope of this study the residential customer class is employed.

Distribution planning basically involves detailing the route of each feeder and identifying the capacity of equipment. Therefore, the proposed method forecasts the residential load of the selected feeders for a period of one year.

The spatial electric load forecasting model is developed using the data from Maharagama area. Validation is done on data from Kelaniya and Galle areas for years 2014, 2015 and 2016. The outcome is compared with actual values and output obtained from Lanka Electricity Company (pvt) Ltd. (LECO) Regression Analysis Software. The results show that the Mean Absolute Percentage Error (MAPE) between actual current values and output from Spatial Load Forecasting Model (SLFM) for feeders in Kelaniya area in year 2014 is 2.45%. The MAPE between actual current values and values obtained from regression model currently used at LECO, for feeders in Kelaniya area in year 2014 is 6.20%. For year 2015, MAPE between actual values and SLFM output is 3.48% and MAPE between actual values and regression model output is 7.00%. In year 2016, theses two values are 4.15% and 8.30% respectively. Therefore, it is seen that the spatial electric load forecasting model proposed in this research has better accuracy than the regression based methods currently used.

Comparing the results in Galle area, the MAPE between actual values and the output from SLFM in year 2014 is 4.75%, MAPE between actual values and output from regression model is 8.10%. In year 2015, the MAPE values are 3.78% and 6.90% respectively and in year 2016, the MAPE values are 4.90% and 8.40% respectively. Hence, the spatial electric load forecasting model has better accuracy than the regression based methods. Further, the spatial electric load forecasting model developed can be used for spatial areas with different load behaviours, such as Kelaniya and Galle, without losing accuracy. The MAPE of the spatial electric load forecasting model is less than 5% while the MAPE of the regression model lies in the range of 10% to 15%, proving better accuracy in the proposed method.

The main advantage of this method is that the input data the model requires, such as distance to nearest major road, area of the land, entrance road width and distance from feeder to the service pole, can easily be extracted from the GIS maps already available at LECO and typical electricity consumption of the adjacent houses can be obtained from the customer database.

Although the proposed tool has provided useful results, it has some limitations. The ANN model should be trained at least once a year with the new data of the past year. A drawback is that the model may require new data surveys for better accuracy in future as spatial factors that affect load forecasting may change in the future.

Another limitation would be, since this model was developed focusing to be used in LECO distribution areas which are urbanized or semi urbanized, it may not give promising results for rural areas. As an example, a new model will need to be developed to address spatial electricity load development in an area where many of the land areas are used for agriculture and houses are situated far away from each other.

Similarly, when considering growth of electricity load in highly developed areas, old low rise buildings are continuously torn down and are replaced by very high rise buildings. This vertical expansion may also be taken into account for an accurate load forecast in a developed area.

Additionally, to develop this model further, studies may be used to pre-process the input data of the model to capture more spatiality to the model. It may be possible to combine the historical data available to do trending for known loads and the simulation techniques will take care for new (unknown) loads. Further, it may be necessary have access to information from other utilities, to help establish better forecasting techniques.

It can be concluded, as the results verify, that the objective of this research - to develop an easy to use, accurate spatial electric load forecasting model to be implemented in Sri Lankan distribution utilities to estimate residential electric load - has been successfully achieved. This model will prove to be an important tool in assisting distribution network planning engineers to identify important zones with the possibility of growth inside the service area.

References

[1] Lee, W, H. *Spatial Electric Load Forecasting*, 2nd Edition. New York, Marcel Drekker Inc, 2002.

[2] Carreno, E.M.; Padilha-Feltrin, A., "Evolutionary heuristic to determine future land use," in *Proc. Power and Energy Society General Meeting*, Jul. 2008, pp.1-6.

[3] ANEEL, *Procedimentos de Distribuição - PRODIST 2008*. Available online http://www.aneel.gov.br

[4] Edgar Manuel Carreno, Rodrigo Mazo Rocha and Antonio Padilha-Feltrin, "A Cellular Automaton Approach to Spatial Electric Load Forecasting," in *IEEE Transactions on power systems*, vol.26, no. 2, May 2011, pp. 532 – 540

[5] Jingmin Wang, Xiaodong Liu and Ruiqing Cui, "Spatial Load Forecasting Method for The New Urban Area Based on Information Classification Approach," in *International Journal of Digital Content Technology and its Applications (JDCTA),* Volume7, Number 4, February 2013, pp. 527 – 534

[6] Vasquez-Arnez, R.L.; Jardini, J.A.; Casolari, R.; Magrini, L.C.; Semolini, R.; Pascon, J.R., "A methodology for electrical energy forecast and its spatial allocation over developing boroughs," in *Proc Transmission and Distribution Conference and Exposition*, 2008. pp.1-6.

[7] Alfares H. K., Nazeeruddin M. "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science*, v.33, no 1, p. 23-24, Jan. 2002

[8] Badar Ul Islam, "Comparison of Conventional and Modern Load Forecasting Techniques Based on Artificial Intelligence and Expert Systems", in *IJCSI International Journal of Computer Science Issues*, Vol. 8, Issue 5, No 3, September 2011.

[9] I. Mogharm, S. Rehaman, "Analysis and Evaluation of Five Short Term Load Forecasting Techniques", in *IEEE Trans on Power Systems*, Vol.4, Oct 1999

[10] Ashton Webberley and David Wenzhong Gao, "Study of Artificial Neural Network Based Short Term Load Forecasting," in *Power and Energy Society General Meeting, Vancouver*, July 2013, pp. 1-4

[11] Xiao Bai, Mu Gang and Li Ping, “A Method of Spatial Load Forecasting Based on Feeder,” in *Electric Utility Deregulation and Restructuring and Power Technologies, Nanjing*, May 2008, pp. 1548-1553

[12] Faiazy and Ebtehaj, “Short Term Load Prediction of a Distribution Network based on an Artificial Intelligent Method,” in *22nd International Conference on Electricity Distribution, Stockholm,* June 2013

[13] Mohd Fatimie Irzaq Khamis, Zuhairi Baharudin, Nor Hisham Hamid et al, "Electricity Forecasting for Small Scale Power System using Fuzzy Logic", *IEEE Trans., The 9th International Power and Energy Conference (IPEC),* Oct. 2010, pp:1040-1045

[14] Che-Chiang Hsu, Chia-Yon Chen, “Regional load forecasting in Taiwan—applications of artificial neural networks,” in *Energy Conversion and Management 44*, 2003, pp. 1941 – 1949

[15] C. Bhurtun, I. Jahmeerbacus and C. Jeewooth, “Short Term Load Forecasting In Mauritius Using Neural Network,” in *8th Conference on the Industrial and Commercial Use of Energy*, 2011, pp. 184 – 191

[16] Matt Mazur, *A Step by Step Backpropagation*: https://mattmazur.com

[17] Xiao Bai, Guo Peng-wei et al, “A Spatial Load Forecasting Method Based on the Theory of Clustering Analysis,” in *International Conference on Applied Physics and Industrial Engineering*, 2012, pp. 176 – 183

[18] A. Badri, Z. Ameli and A.Motie Birjandi, “Application of Artificial Neural Networks and Fuzzy logic Methods for Short Term Load Forecasting,” in *Energy Procedia 14*, 2012, pp. 1883 – 1888

[19] Mohsen, “Enhancing reliability in passive anti-islanding protection schemes for distribution systems with distributed generation,” M.S. thesis, Electrical and Computer Engineering, The University of Western Ontario, 2012

[20] Tobler's First Law and Spatial Analysis, http://onlinelibrary.wiley.com

[21] Neural Network Toolbox, http://in.mathworks.com