NON-NEGATIVE LEAST SQUARE OPTIMIZATION MODEL FOR INDUSTRIAL PEAK LOAD ESTIMATION

Hari Priya Moda

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

In

Electrical Engineering

APPROVED:

Robert P. Broadwater, Chair

Kwa Sur Tam

Amos L. Abbott

December 04, 2009

Blacksburg, Virginia

Keywords: Load Research, Non-Negative Least Square Optimization, NNLS, Peak Load

NON-NEGATIVE LEAST SQUARE OPTIMIZATION MODEL FOR INDUSTRIAL PEAK LOAD ESTIMATION

Hari Priya Moda

(ABSTRACT)

Load research is the study of load characteristics on a power distribution system which helps planning engineer make decisions about equipment ratings and future expansion decisions. As it is expensive to collect and maintain data across the entire system, data is collected only for a sample of customers, where the sample is divided into groups based upon the customer class. These sample measurements are used to calculate the load research factors like kWHr-to-peak kW conversion factors, diversity factors and 24 hour average consumption as a function of class, month and day type. These factors are applied to the commonly available monthly billing kW data to estimate load on the system.

Among various customers on a power system, industrial customers form an important group for study as their annual kWHr consumption is among the highest. Also the errors with which the estimates are calculated are also highest for this class. Hence we choose the industrial class to demonstrate the Lawson-Hanson Non-Negative Least Square (NNLS) optimization technique to minimize the residual squared error between the estimated loads and the SCADA currents on the system. Five feeders with industrial dominant customers are chosen to demonstrate the improvement provided by the NNLS model. The results showed significant improvement over the Nonlinear Load Research Estimation (NLRE) method.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Robert P. Broadwater for his guidance, support and help. His knowledge and encouragement has shaped this research work. I am also greatly indebted to him for providing me financial support through my graduate studies. I wish to thank him for having faith and confidence in me. It's been a pleasure working with him.

I am indebted to Max Zhong for being my mentor and guide and for helping me with this thesis. Also I would like to thank Electrical Distribution Design Inc., and all my co-workers for providing me with all the technical and infrastructure support required to carry out the work.

I would also like to thank my parents Sripada Rao Moda and Sujatha Moda, and to my sister Preeti Moda for their love and support. Also, my special thanks to Raghuram for his encouraging words.

Last but not least I would like to express my love and thanks to Sujit for always being there for me. He has always pushed me to achieve and supported me in every way possible.

Contents

Chapter 1 1
Introduction 1
1. Problem Definition and Motivation 1
1.1 Overview 1
1.2 Peak Load Estimation 4
1.3 Load Research 4
1.4 Load Estimation Factors 7
1.5 Motivation 7
Chapter 2 9
2 Literature Review 9
2.1 Load Research History 9
2.2 Electricity consumption model based on consumption 9
2.3 Load Research Factors 14
2.4 Load estimation 18
2.5 Peak Demand and Forecasting 20
Chapter 3 24
Non-Negative Least Square Optimization Model 24
3.1 Load Research Process 24
3.2 DEW Workstation 24
3.3 Load Research Statistics Calculation 25
3.4 Peak load estimation using NLRE 31
3.5 Division of customers into classes 33
3.6 Customer penetration 36
3.7 Results for industrial customer dominated circuit 37
3.8 Non-Negative Linear Least Square Error correction for kWhr-kW conversion factor 38
Chapter 4 48
Results and Discussions 48
4.1 Load Research Calculation 48
4.2 Results discussion 54
Chapter 5 79
Future work and conclusions 79

5.1 Conclusions 79
5.2 Contributions 80
5.3 Future work 81
BIBLIOGRAPHY 82

LIST OF FIGURES

Figure 1-1 Representative daily load curve for residential customer on a weekday 2
Figure 1-2 Representative daily load for commercial customer on a weekday 3
Figure 1-3 Illustration-Load Research 6
Figure 3-1 Representative Diversity factor curve 31
Figure 3-2 Flowchart describing the division of customer classes 35
Figure 3-3 July weekday Phase A current estimates for industrial circuits 42
Figure 3-4 July weekday Phase B current estimates for industrial circuits 43
Figure 3-5 July weekday Phase C current estimates for industrial circuits 44
Figure 3-6 November weekday Phase A current estimates for industrial customers 45
Figure 3-7 November weekday Phase B current estimates for industrial customers 46
Figure 3-8 November weekday Phase C current estimates for industrial customers 47
Figure 4-1 Calculation of LDR Statistics:Login window 50
Figure 4-2 Calculation of LDR Statistics: Statistics calculation window 51
Figure 4-3 Calculation of LDR Statistics: Chart window- Diversity Load curve for class 9 for January and July weekdays 52
Figure 4-4 Calculation of LDR Statistics: Chart window- C-factor curves for class 9 for weekday and weekend 53
Figure 4-5 Calculation of LDR Statistics: Chart window- Diversity factor curve for class 9 for January and July weekday 54
Figure 4-6 July weekday Phase A current estimation for a chosen circuit 55
Figure 4-7 July weekday Phase B current estimation for a chosen circuit 56
Figure 4-8 July weekday Phase C current estimation for a chosen circuit 56
Figure 4-9 November weekday Phase A current estimation for a chosen circuit 57
Figure 4-10 November weekday Phase B current estimation for a chosen circuit 57
Figure 4-11 November weekday Phase C current estimation for a chosen circuit 58
Figure 4-12 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 35-6-13 63
Figure 4-13 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 35-6-13 63
Figure 4-14 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 35-6-13 64
Figure 4-15 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 35-9-13 64
Figure 4-16 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 35-9-13 65
Figure 4-17 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 35-9-13 65
Figure 4-18 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 51-8-13 66

Figure 4-19 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 51-8-13 66
Figure 4-20 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 51-8-13 67
Figure 4-21 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 15-4-13 67
Figure 4-22 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 15-4-13 68
Figure 4-23 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 15-4-13 68
Figure 4-24 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 67-2-13 69
Figure 4-25 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 67-2-13 69
Figure 4-26 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 67-2-13 70
Figure 4-27 Monthly peak comparison for phase A:35-6-13 71
Figure 4-28 Monthly peak comparison for phase B:35-6-13 72
Figure 4-29 Monthly peak comparison for phase C:35-6-13 72
Figure 4-30 Monthly peak comparison for phase A: 35-9-13 73
Figure 4-31 Monthly peak comparison for phase B: 35-9-13 73
Figure 4-32 Monthly peak comparison for phase C: 35-9-13 74
Figure 4-33 Monthly peak comparison for phase A:51-8-13 74
Figure 4-34 Monthly peak comparison for phase B: 51-8-13 75
Figure 4-35 Monthly peak comparison for phase C: 51-8-13 75
Figure 4-36 Monthly peak comparison for phase A:15-4-13 76
Figure 4-37 Monthly peak comparison for phase B:15-4-13 76
Figure 4-38 Monthly peak comparison for phase C:15-4-13 77
Figure 4-39 Monthly peak comparison for phase A:67-2-13 77
Figure 4-40 Monthly peak comparison for phase B:67-2-13 78
Figure 4-41 Monthly peak comparison for phase C:67-2-13 78

LIST OF TABLES

Table 3-1 Customer class division based on annual consumption 36
Table 3-2 Customer penetration for industrial circuits 37
Table 3-3 New and old C-factors 41
Table 4-1 Customer class division for the chosen feeders 59
Table 4-2 Error comparison for all 3 phases for feeder 35-6-13 60
Table 4-3 Error comparison for all 3 phases for feeder 35-9-13 60
Table 4-4 Error comparison for all 3 phases for feeder 51-8-13 61
Table 4-5 Error comparison for all 3 phases for feeder 15-4-13 61
Table 4-6 Error comparison for all 3 phases for feeder 67-2-13 62

Chapter 1

Introduction

1. Problem Definition and Motivation

1.1 Overview

To ensure uninterrupted power supply to consumers, utility companies have to manage the generation schedules and also need to plan the distribution system. However, to accomplish the above tasks it is imperative to understand the consumption patterns based on season, day type, such as a weekday or a weekend, and the geographical location since the usage patterns varies based on these factors [1] [2]. The entire system loading information can solve this problem, but load is constantly changing and growing and we need to predict. However, the costs associated with collecting and maintaining such large data stores across the entire system is enormous. Various attempts to reduce the above mentioned costs has generated significant interest in load research among utilities. Though load research has been ongoing for many years it is still of interest due to the costs associated with the planning of the distribution system and the many factors especially the human and demographic factors influencing the usage [10]. These factors also make load research partly marketing based as the customer usage pattern and demand study is based on how people behave in their daily lives and on different days in different locations as far as the electricity consumption is concerned. To illustrate the above factors, an apt example would be that almost everyone would be taking a bath in the morning. That means the water heaters draw the power and the use of cooking appliances around the same time affects the peak loads in the morning hours. Figure 1-1 is an illustration of the usage pattern for the residential circuits showing that the utilization peaks during morning and in the evening hours. This is expected for a residential customer as the consumption is high in the morning for hot water, cooking and in the evenings for TV's, cooking appliances, light bulbs etc.

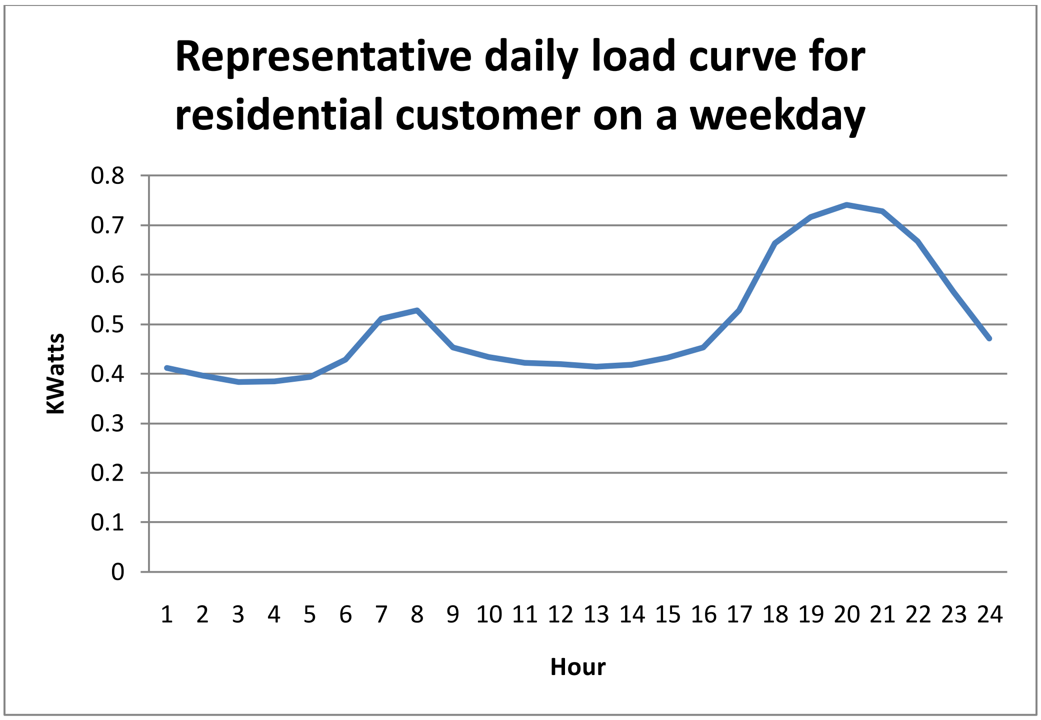

Figure 1-1 Representative daily load curve for residential customer on a weekday

This pattern varies for a commercial or an industrial customer where the load starts peaking around 8 AM in the morning and stays around the same value till 5 PM in the evening. Figure 1-2 illustrates the same in case of a commercial customer.

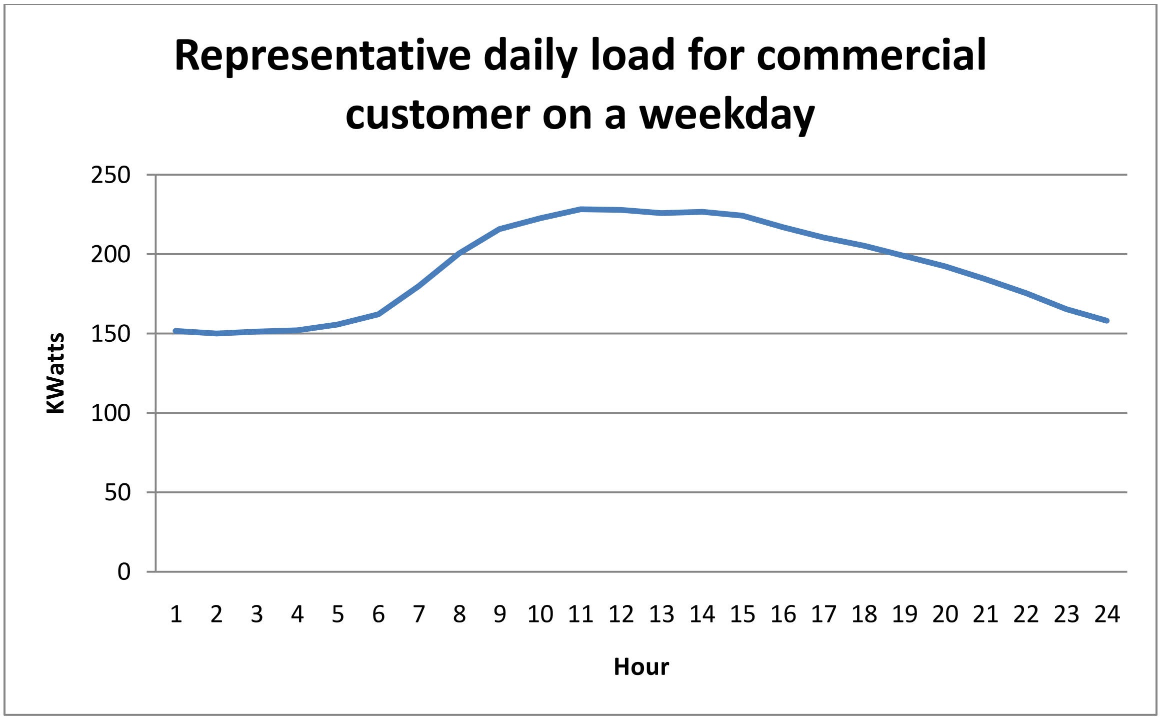

Figure 1-2 Representative daily load for commercial customer on a weekday

Electric load supply in a circuit with many different types of customers is a very complex problem and estimating the demand peak on such a system has been of primary interest for many years. The utility companies have billing data in addition to the sample measurements which are used in the load research to estimate the system peak and load curves. Various statistical factors representing the customer usage pattern are estimated and studied for residential and industrial customers to come up with more accurate results for the system.

1.2 Peak Load Estimation

Load estimation is used to schedule and plan the distribution system by planning engineers. This is essential, as the load usage varies based on type of customer, time and weather and also on the geographical location [1] [2]. Hence the planning engineer has to act in accordance to these changes and meet the demand at any given point of time.

It has been found that the most efficient way of estimating the peak load is by dividing the customers into classes based on their usage patterns [5]. Typically, the samples are collected every one hour or every half an hour at distribution substations and major equipment installations [3]. In addition to the above mentioned method one other method that is employed is to use the customer billing data to calculate the daily peak [4]. Sample customers are selected for each class and the readings are collected. The readings are collected from anywhere between 145 days to 480 days [1].

A circuit, especially in a metropolitan area is composed of different types of customers connected to distribution transformers. Hence it is essential to study the characteristics of each of these loads attached, as they form the load for the entire circuit. Hence it becomes imperative to study the load patterns for these different customer classes.

1.3 Load Research

Load research has gained significant importance in many countries and among private and public electricity boards, as it is found to reduce the equipment cost and the data processing costs, as only samples of the customer classes which form the main part of the research program are collected. Many electricity boards have been participating in these programs and have benefited from the same [1] [3].

Load research programs have been used in many applications other than estimating the peak load [1] [2] [4] [17]. This along with load forecasts can help in establishing the following:

1. Load curves for different customer classes for different weather conditions
2. Response to weather conditions.
3. Network, tariff and production planning.
4. Operating decisions, like dispatch scheduling and maintenance planning.
5. Planning the geographical location of installing the equipment, ratings of capacitor banks, transformer ratings etc.
6. Design of energy conservation to improve operation efficiency.

Load research is mainly about recording customer kWHr or demand readings for the month and analyzing the data for demand peaks, which provides essential information for planning the distribution system. As data collection for each hour for all the customers involves huge costs associated with data collection and storage, load research programs have been designed to collect around 3000 samples for each customer class for a year [2]. The raw kWHr data is aggregated based on customer classes to perform the statistical analysis to estimate load curves and peak demand. Figure 1-3 illustrates the data that goes into load research and the results obtained.

The main inputs to the Load Research process are the raw hourly load data for the entire year for all the customer classes, SCADA data collected at substation/feeders and commonly available monthly billing data. Load Research process calculates the kWHr-to-peak-kW conversion factors, diversity factors, and load diversity curves from the annual loading information and SCADA data and applies these factors to the monthly billing data to estimate the system loading information.

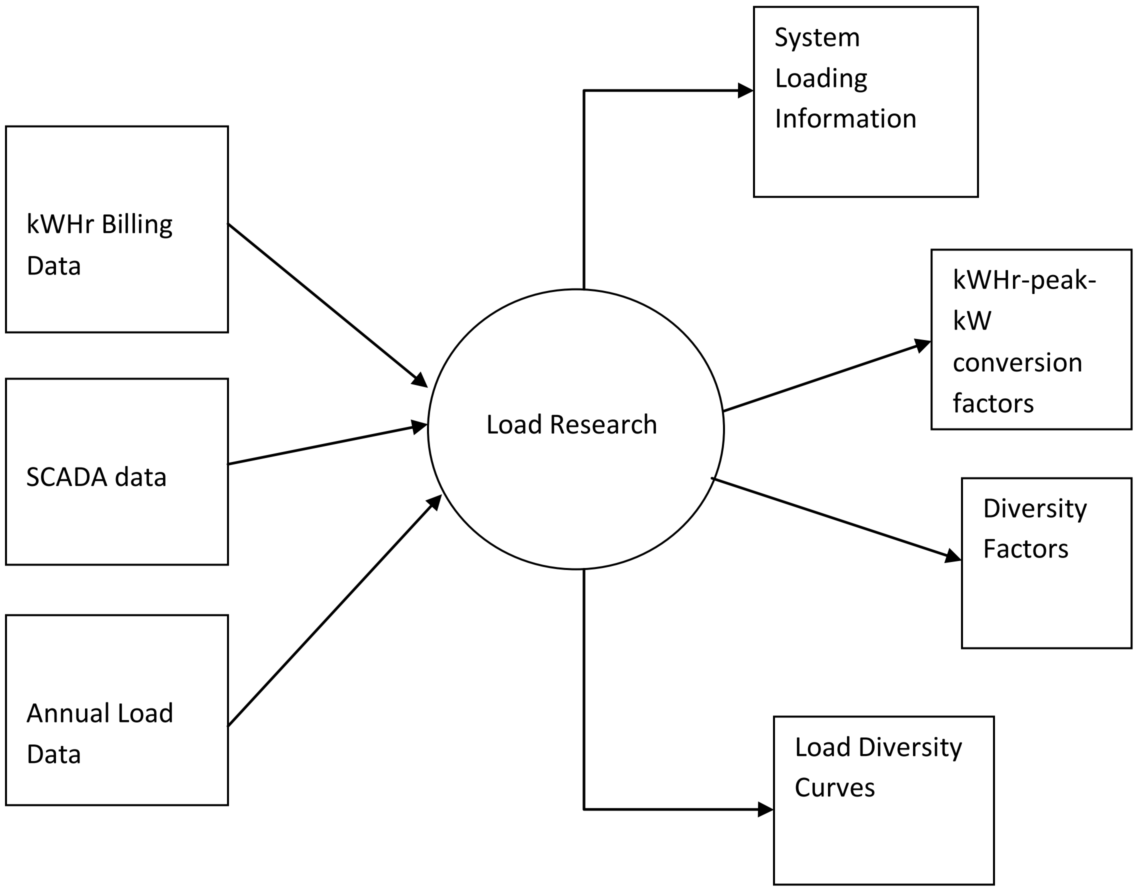

Figure 1-3 Illustration-Load Research

Load research was started by several utility companies to comply with regulatory requirements, like PURPA, to better allocate the costs among rate classes in a regulated rate setting. Demand side management (DSM) is a program where the customer usage patterns are influenced to match the present or extrapolated generation. [AEIC load research manual]. DSM is an area in which load research has been applied to advantage. It has been used to measure and verify the advantages of DSM programs [12].

Methods like Automated Metering Infrastructure (AMI) are being implemented, which collect samples at every customer, and hence have the possibility to replace statistical analysis of load research program. But the fact that AMI does not cover the entire population and ignoring one area can bias the estimates reinstates the importance of load research [11].

1.4 Load Estimation Factors

Load research factors are defined per customer class, type of day, weather condition, month and geographical location [5]. Monthly billing data with the load research statistics like kWHr-kW peak conversion factor or C-factor and diversity factor are used in estimating peak load demand [5]. The availability of data for large number of customer classes enables the class based C-factors applied to monthly billing data to estimate the peak loads give better estimates. For non-homogeneous customers connected to a single transformer, diversity factors are applied to take into account the diversity of energy use in the group. Also, parsing factors are another important factor used to organize the monthly billing data into calendar month cycles as the billing cycles for each customer are not same.

1.5 Motivation

In order to accurately analyze and plan the power distribution system, correct loading information across the system is required. The customer billing information is used and the conversion factors and diversity factors are applied to the billing information to accurately estimate the time and the peak for a circuit. The past work, based on data supplied by Arkansas Power and Light Company [5], has studied and applied the various load research statistics to estimate the time and peak load for residential heat and non-heat customers.

Demand meters were employed and the readings are collected for each test customer for a year, with a total of 8760 points for each customer. Load data for 299 residential customers was used in this study and the residential customers are divided into electric and non-electric heat. The peak power is then estimated for the control group and the results are found to be accurate within 13%.

The results on the industrial predominant circuits has been analyzed in this thesis and has been found to give the results which are over estimated by more than 100%. To provide improved results in case of industrial dominated circuits a new method has been proposed which utilizes the Supervisory Control and Data Acquisition (SCADA) data to provide better results. The load research statistics are calculated [5] and the loads for particular customers are estimated. Circuits are chosen which are predominant in industrial class and the SCADA measurements for the particular circuits are obtained from the utility companies. The C-factors are then adjusted using the Non-Negative Least Squares (NNLS) algorithm and the new estimated values have been compared to the SCADA and have been found to provide improved results.

The remainder of this thesis is organized as follows. Chapter 2 provides an in-depth discussion on the work related to load research and various methods that have been employed to estimate the peak load. Chapter 3 describes the NNLS algorithm and its application to the load research data. Chapter 4 demonstrates the experimental results. Finally, the conclusions and recommendations are provided in Chapter 5.

Chapter 2

Problem Background and Related Work

2 Literature Review

2.1 Load Research History

As discussed, load research is being used by electric boards and has been found to be very effective in reducing costs. AEIC, the Association of Edison Illuminating Companies is an organization formed in 1885 in the United States and has been involved in guiding the electric companies in load research and many other power related areas. The load research committee organizes annual seminars to discuss and train the industry with the latest load research trends.

Load Research programs were started to comply with utility acts like the Public Utility Regulatory Act which address such things as cost of service rates based on the time of day, resource planning, and load management [13]. Deregulation by government agencies and changing requirements for load research, due to increasing consumption and changing usage patterns, high costs associated with equipment, and planning distribution systems, makes load research still relevant today. Utility companies are interested in estimating the peak load and system loading information.

2.2 Electricity consumption model based on consumption

Load research is conducted by dividing the customers into homogeneous groups, since the loading information based on smaller groups of the population is of interest. Accurate estimation of load profiles and its importance has been stressed by many researchers. Its importance in the areas of demand side management and network planning has been researched [7] [8] [14].

Load curves primarily depend on the mix of customers. Few of the relevant methods proposed for coming up with homogeneous customer groups are discussed here.

Customers are broadly classified into residential, industrial and commercial customers [7]. Also proposed are classifications based on environment or geographical location like urban, rural etc, electric service rate classes like residential, commercial power, commercial lightning etc. [16]. This high level classification does not cover the sub-groups within the groups having different consumption patterns, though belonging to the same group.

Irvin et al [7] have shown that one group representing the public housing sector required two models to describe the group. Hence, this method can be applied to identify sub-groups within a homogeneous group with some confidence.

The consumer monthly billing data is used in most of the models in coming up with the demand model. A bottom up approach has been used here, where all of the consumer consumption patterns are used in the development of the model. However, since each customer cannot be modeled, they are aggregated into classes based on similar consumption patterns.

A statistical framework based on the Weibull model [35] has been employed to come up with accurate consumer groups. Demand modeling, which is responsive to the parameters within each consumer group, is the next step. The public housing sector group was chosen in this paper to demonstrate that the Weibull distribution provides the required model.

Rural and urban data for the public sector housing data was collected and the relative frequency of consumption demonstrated that the data could be classified based on the histograms into a high-skew and skew-normal modeled by:

$$\frac{f}{N} = a \exp(-aC_r) \qquad (2.1)$$

where

f= frequency density

N=sample size

C_r=relative consumption

a=shape parameter

A skew normal histogram can be modeled by:

$$\frac{f}{N} = 2aC_r \exp(-aC_r^2) \qquad (2.2)$$

The upper and lower limit polygon from SN and HS is developed and the upper and lower confidence limits established. Though these models give a satisfactory fit, they do not provide a model for the entire population of the housing sector.

A two-parameter Weibull estimation has been applied to the entire public housing population using the equation modeled by:

$$\frac{f}{N} = amC_r^{m-1}\exp(-aC_r^m) \tag{2.3}$$

where

f= frequency density

N=sample size

C_r=relative consumption

m, a= model parameters

Here ‘a’ and ‘m’ are the control parameters. By selecting accurate ‘a’ and ‘m’ values, the distribution model for a required demand group can be established. These parameters were estimated using the method of moments and method of maximum likelihood. By plotting contours of constant log likelihood and observing the behavior of log likelihood in the neighborhood it was observed that by varying one parameter ‘m’, the second parameter ‘a’ remained mostly constant. It was also observed that the frequency polygon obtained by applying these values produced a better fit. Also it was established that the frequency polygons for these two models are significantly different.

Subjective customer load classification based on usage patterns into residential, industrial and commercial customers was observed to have differences in usage patterns among the groups as shown in the above research, which would lead to errors in load estimation. Hence, objective customer classification based on artificial neural networks approach has been discussed by Nazarko et al [14] that is used to cluster the n+1th load curve when the placement for the n load curve in a cluster map is already known.

The load model considered in [14] is based on statistical methods, where the 24-hour load pattern is categorized into columnar values based on the day time into night, morning, afternoon and evening. The load pattern is divided into layers as base load, intermediate and peak load layer. Customers are classified and grouped based on the average load of 24 hour alignment for each column

Neural networks are trained based on the time and the customer composition. Peak load with the neural network approach was found to be better than the statistical method by a factor of two. But the results reported did not take into account the seasonal and week day type.

Ansii Sepalla et al [8] as part of the Finnish load research tried to establish the distribution of the customer load profiles. The confidence limits for the log-normal distribution were too high and the confidence intervals for the normal distribution were too low. As these distributions did not give accurate and reliable estimates for confidence intervals, which are required for distribution planning, the authors have proposed a modified log-normal distribution which was found to give better estimates with higher confidence. This was calculated using mean and standard deviations that are available from the distribution. The estimators for modified log-normal distribution are:

$$\xi' = \ln(E\{P\}) \tag{2.4}$$

$$\sigma' = \ln(E\{P\} + \sigma\{P\} - \xi') \quad (2.5)$$

where E{P} is the mean and $\sigma\{P\}$ is the standard deviation of the normal distribution

2.3 Load Research Factors

Load curves on the system give information about loading patterns on the whole system. Peak load estimation involves applying the information obtained from load curves and the load research factors derived from the load research data. Diversity factors, load factors and coincidence factors are some of the prime factors used to describe load characteristics [5] [15]. Some definitions pertaining to load research are presented below:

System load is the load on the system at any given time, and the time interval needs to be defined for this definition to be complete. It could be defined over a day, week, month or a year. This is an important factor to analyze the usage trends among the consumers.

System peak is defined as the maximum load required or consumed in any given time interval. It is normally defined for a day, week, month or year. This is an important factor, as this is used to plan the distribution system and to plan the capacity of the plant.

Demand is defined as the average load consumed over an interval of time. The variation of demand with the demand interval is represented as [15]

$$Z_i(t) = C_i \times X_i(t) \quad (2.6)$$

where

$Z_i(t)$ - demand at time t,

C_i - capacity of the load equipment

and $X_i(t)$ - random variable defining the usage pattern of the equipment.

The load research data does not use the instantaneous demand unless the spikes in the load need to be studied. In such a case the readings taken for 1 minute interval are used in the study. Load research studies generally use half-an-hour or hourly data.

When a variety of loads are connected to a transformer, X_i varies according to the usage pattern of the loads. X_i varies from 0 to 1, to represent the off and on conditions respectively. Using this above, system peak load for time period T can be defined as:

$$Peak = \max_{t \epsilon T} Z_i(t) \quad (2.7)$$

$$= \max_{t \epsilon T} C_i \times X_i(t) \quad (2.8)$$

If Z_1, Z_2,.. ,Z_n represent the n loads connected to the system, the sum of these loads represent the system load at time t.

$$Z(t) = \sum_{i=1}^{n} Z_i(t) \quad (2.9)$$

$$= \sum_{i=1}^{n} C_i \times X_i(t) \quad (2.10)$$

Load factor is defined as the ratio of average load to the peak load in a particular period [16]. If the load that varies with time is Z(t)

Load Factor for load i can be represented as;

$$LF_i = \frac{\frac{1}{M}\sum_{i=1}^{n} X_i(t)}{\max\limits_{t \in T} Z_i(t)} \qquad (2.11)$$

where M is the time period with M intervals of size t.

Load factor has values between 0 and 1. If the load is constant, load factor is 1.0, but the normal loads are observed to have load factors less than 1, due to varying nature of loads. Load factor is a measure of the extent to which peak load is sustained during the period [16]

American Standard Definition for Electric terms has defined **Diversity factor** as "the ratio of the sum of individual maximum demands of the various sub-divisions of the system to the maximum demand of the whole system" [18]. It represents the relation between the group peak and customer's individual peaks [15]. The diversity factor will be 1 when all the peaks for the loads in the group occur at the same time. It will be greater than 1 when the demands are non-coincident or do not occur at the same time for all the customers. The main contributing factor to the diversity factor is the size of the group. But it was observed that the diversity factor remains unchanged for groups larger than 30. The diversity factor increases with a large slope in the beginning and flattens out as the number of customers increase.

Diversity Factor for maximum demand loads D_1, D_2,… ,D_n is defined as

$$F_D = \frac{(D_1 + D_2 + .. + D_n)}{D_{1+2+3..+n}} \qquad (2.12)$$

where $D_{1+2+3..+n}$ is the maximum demand of the group of loads.

The inverse of diversity factor is the **coincidence factor**. Its value is less than 1 and is defined as the ratio of maximum coincidence demand for a group to the sum of maximum demands of individual customers in the group considered at the same point of time.

$$F_c = \frac{1}{F_d} \tag{2.13}$$

$$= \frac{D_{1+2+3..+n}}{(D_1 + D_2 + .. + D_n)} \tag{2.14}$$

Load diversity is the difference between sum of peaks of individual customers and the peak of the combined load. It is defined as follows;

$$LD = (D_1 + D_2 + .. + D_n) - (D_{1+2+3..+n}) \tag{2.15}$$

Before defining the coincidence factor the contribution factor is explained.

C_n= contribution factor for the load n which is the contribution of the nth load to the group peak. Now the coincidence factor can be represented as follows;

$$F_c = \frac{C_1 D_1 + C_2 D_2 + \cdots + C_n D_n}{D1 + D2 + .. + Dn} \tag{2.16}$$

When individual demands are same, D_1=D2=… =Dn , for example, most of the residential appliances within a certain range have the same maximum demands. We have,

$$F_c = \frac{(C_1 + C_2 + .. + C_n)}{n} \tag{2.17}$$

which means the coincidence factor is the average of contribution factors

When the contribution factors are same and demands are different, like in the case of air conditioners, we have,

$$F_c = \frac{C_1(D_1 + D_2 + .. + D_n)}{(D_1 + D_2 + .. + D_n)} = C_1 \quad (2.18)$$

the coincidence factor is same as contribution factor.

In addition to the diversity factor, coincidence factor also is a function of the number of customers in the group.

The load research factors are employed in load estimation of the peak load of a system. Lack of this information leads to incorrect estimation of ratings of transformer or capacitor banks that result in equipment damage for being operated beyond their rated capacities. In addition to equipment protection, this information is also used in planning various other aspects of distribution systems.

2.4 Load estimation

Load estimation is useful for managing the distribution system. It has found use also in monitoring loads at the distribution sub-stations. This has become important in the deregulated environment where the online load monitoring becomes important for those who buy power from the market. Fuzzy load models, state estimation techniques and statistical methods are discussed here to estimate the load.

Fuzzy load models are used to consider the uncertainties in the load models. Chang et al [21] have come up with a model to construct load profiles and to define confidence limits. Random sampling of customers is conducted using Neymunn stratification sampling and meters installed to collect the measurements of customer power demand. Bad data is detected and load patterns are derived. Using the "possibility-probability" function a fuzzy membership function is derived. Feeder measurements, kWHr data and the fuzzy load profiles are used and estimations of transformer load profiles can be obtained.

Another fuzzy neural approach based on case based reasoning is presented by Wu et al to estimate the nodal load on distribution systems [22]. A connectionism-based case-based reasoning (CBCBR) method is proposed for load estimation which was found to give better estimates. State estimation is started by real-time measurements and load profiles or customer kWHr consumption. The output of state estimation is used as input for learning by the load estimation module. Lastly, this will provide state estimation better load estimates based on stored cases. The results have shown that the CBCBR has a better self-adaptive learning ability and good resistance against bad data.

Peak load estimation using nonlinear fuzzy regression was proposed by Cartina et al[6]. Fuzzy set theory is used to model the data and the triangular fuzzy numbers are used to represent the uncertainties in the loads. This method has produced results within the range of -6 to 14% error for class of customers with residential, industrial and commercial types and two types of days (weekday and weekend).

State estimation involves estimating the state based on the measurements available [23]. Multiple load parameter Weighted Least Squares method is applied for distribution state estimation (DSE) [24] and reinstated with operating and loading constraints was proposed by Wan et al[24][25].

A specific case has been studied by Baran et al [19] where the load is estimated in the absence of SCADA data due to metering problems or bad data. Here the "regression-based prediction" method is used where the meters that have strong correlation are used to obtain data from a failed meter and the results were reliably estimated up to a week. Bad data to be estimated is calculated using regression by applying the regression coefficients calculated by eliminating the outliers. However, some problems with outliers were observed.

2.5 Peak Demand and Forecasting

Peak load is the most important load characteristic that is used in planning the distribution system. This is because the plant needs to be designed to meet the peak load at all the times. Also the plant is seldom designed to meet the present or near future demand. It is designed to meet the demand for a long term and demand forecasting helps in planning the system and equipment. The resources are synchronized to meet the demand using the least cost. Also, peak load helps in assigning the rates to particular class of customers who are responsible for the peak load, as they are more responsible for the peak capacity of the plant and the investments made to establish that capacity. Peak load estimation also helps in demand side management where the utilization companies can regulate the use during the peak periods by understanding the components contributing to the peak load and period of occurrence of the peak load.

Many short-term and long-term load forecasting methods have been proposed over the years. Short term load forecasts normally deal with hourly data and forecasts from one hour ahead to a week. Long term forecasts deal with one year to ten year forecasts.

Time series approaches have been used extensively in load forecasting, where future loads are predicted based on past data. However, the problem with the time series approach is that the results are inaccurate when it is applied for long term forecasting [20]. The time series approach has been widely discussed by many researchers [26] [27].

Short term forecasting for up to one hour ahead based on auto regressive integrated moving average (ARIMA) model, which is simpler to model than the time series approach, was proposed by Lu and Zhang [28].The method uses 48 data points, which is very less data in training. Mean absolute Percentage Error (MAPE), which is a performance indicator for the predicting process, was used to evaluate the performance and is computed as follows:

$$MAPE = \frac{1}{N}\sum_{t=1}^{n}\left|\frac{w(t) - \overline{w}(t)}{w(t)}\right| \qquad (2.19)$$

where

N- Sample number

$w(t)$- Original series

$\overline{w}(t)$-mean of $w(t)$

MAPE results indicated that the algorithm has good adaptability and good performance over weekdays and weekends.

In contrast to the other load forecasting algorithms that require learning all similar days data, Senjyu et al proposed a neural network based one-hour ahead load forecasting method [29] using the correction of similar days data. Forecasted load power is obtained by adding correction to similar day’s data, where the correction data is obtained from the neural network.

Long term forecasts are significantly different than the short-term forecasting, in that they look at ten years to about thirty years ahead. Long-term forecasting is used in deciding the capacity of the plant, its expansion plans, and investment in the plant. The long timeline involved makes uncertainties large and results error prone.

The extrapolation method is one of the statistical methods proposed for forecasting loads in the long term. Here the past data is used and the future load is extended in time. This was analyzed and found to work well for one year to three years [30]. The future peak is estimated by using the previous year's peak and extrapolating it to the future assuming the same weather conditions in the future. However, the changes in weather conditions in the recent past have resulted in high errors using this method.

Parlos et al [31] proposed an intelligent artificial neural network (ANN) based long-term forecasting. A historical database is used for training and testing the ANN. Under training or over training of the network is achieved using training set error. The ANN is based on the adaptive back propagation algorithm.

As the ANN architecture is designed for solving the nonlinear optimization problem, many ANN architectures could solve the forecasting problem. The best one is chosen based on the testing error indicators. The case study performed shows that the neuro-forecast gives a 3% error whereas the utility forecast had 13% for the Bryan Rural area. And for the city of Bryan, the results gave an error ranging from 4.3% to 9.3%.

Econometric modeling for load forecasting is another long term forecasting method, where the correlation between consumption and economic parameters, like Gross Domestic Product, inflation, etc., is established. Use of the above economic parameters makes large scale econometric models unsuitable for small area load forecasting [33].

A comparison between econometric models and neural networks was carried out in Singapore by Liu et al [34]. It was found that both the methods forecasted the historical consumption for 1960-1984 equally well but for 1985-1984, the results from both vary. The performance of neural network model is observed to be inferior to the econometric model. The elasticity estimates demonstrate that for the neural network model the elasticity changes substantially over time.

Chapter 3

Non-Negative Least Square Optimization Model

3.1 Load Research Process

The load research process begins with collecting the sample kWHr measurements and involves evaluating load curves and calculating the system peak. Data is collected at various points on the circuit to cover all the sample customers. A total of 8760 data points were collected for each customer, each hour for an entire year. The following steps detail the work flow involved in carrying out the load research:

1. Raw customer billing kWHr data is imported and stored in the database. The data is divided into homogeneous customer groups to carry out further calculations.
2. Calculating the load research factors.
3. Parsing the billing kWHr into calendar cycles.
4. Applying the load research factors to parsed billing data to estimate the load on the system.
5. Adjusting the load research factors based on SCADA and estimated values using the Non-Negative Least Squares optimization algorithm.
6. Updating the conversion factors and re-estimating the load on the system.

3.2 DEW Workstation

The Distribution Engineering Workstation or DEW is a software application used for distribution planning and design by providing an open architecture with integrated data and applications. External data can be imported into the workstation and made available to the applications. Power flow, load estimation and fault analysis are some of the analyses that can be performed in the DEW [37]. DEW provides GUI functions that allow interaction with users and the open architecture allows addition, deletion or replacement of applications [38].

The integrated model allows the standardized database to be shared by all the applications. This data includes corporate data, SCADA, engineering, financial, operations, geographical and statistical data [39]. For the load estimation we are interested in statistical data and particularly tables related to load research. These tables include customer billing data, conversion factors, and diversity factors and 24 hour average consumption which are stored as a function of customer class, month and type of day.

The Customer Information System (CIS) provides energy use, load class, peak demand and geographical location of the customers. The above load research statistics and the CIS data are used for improved estimates of the peak load [38].

3.3 Load Research Statistics Calculation

Load research statistics are a set of factors which are functions of customer class, day type and month. Since it is expensive to collect and maintain the loading information across the entire system, the readings are collected at a few distribution transformers to sample each customer class for the entire year. Using these sample readings and billing data, loading across the entire system may be estimated for any hour of the year. Since the usage patterns, peaks and peak time occurrences are different, there exists a need to study each customer group separately. Hence these readings are grouped based on class, month and day type.

The following is a description of various factors used in estimating the currents on the system.

kWHr to kW peak Conversion Factor or C-Factor is applied to monthly billing data to estimate the peak for the month [3] . The conversion factor is a function of customer class, type of day, and month.

$$conversion\ factor\ or\ cfactor = f(cls, tyd, mon) \qquad (3.1)$$

where cls = customer class

tyd = type of day

mon = month

Utility companies employ demand recorders to collect the samples for a wide variety of customers which helps in calculating class based C-factors, which when applied to billing data give better improved results. Class based C-factors are estimated from the sample kilowatt measurements using the hourly load collected. These large number of customer readings for a large variety of customer classes help estimate the peak load more accurately. The data is organized based on the customer class, month and type of day such as weekday and weekend. The daily load is aggregated for each month for all the customers and the monthly peak is estimated for each.

The individual peaks are summed and divided by the total energy usage for the entire group to estimate the C-factor for each group [12]

$$C_f = \frac{\sum(\text{Individual Peaks})}{Total\ Load} \quad (3.2)$$

The following procedure describes the steps to calculate the class based C-factors:

1. Use the raw kWHr data and calculate the 24-hour consumption for each group for each hour in a month by summing up the kWHr of all the customers by grouping the data based on class, month, type of date as follows

$$PAVGW_g(cls, mon, tyd, hr) = \sum_{i=1}^{N} P_i(hr, cls, mon, tyd) \quad (3.3)$$

where

N is the number of customers in the group.

P_i is the kWHr measurement for individual account *i*.

2. The peak for the group is calculated by finding the maximum among the 24 hour values calculated above.

$$Peak_g(cls, mon, tyd) = \max_{hr\epsilon 24} PAVGW_g(cls, mon, tyd, hr) \quad (3.4)$$

3. The total consumption for the group in a month is calculated by summing up the 24 hour kWHr for the entire month for all the customers grouped based on the class and type of day

$$MonthlyKW_g(cls, mon, tyd) = \sum_{hr=1}^{24} \sum_{i=1}^{N} P_i(cls, mon, tyd, hr) \quad (3.5)$$

4. The conversion factor for the group is calculated as

$$c_f(cls, mon, tyd) = \frac{Peak_g(cls, mon, tyd)}{MonthlyKW_g(cls, mon, tyd)} \quad (3.6)$$

If all the customers peaked at the same time and there is no diversity in the group of customers attached to a distribution transformer, the above calculated C-factor can be applied to the monthly billing kWHr data to estimate the peak load. But all the customers in the group do not peak at the same time and hence the diversity factor needs to be calculated to estimate the peak.

Diversity factor or d-factor is a function of customer class, number of customers in the group, day type, and month.

$$Diversity\ Factor\ or\ dfactor = f(N, cls, mon, tyd) \quad (3.7)$$

where

N= number of customers in the group

mon= month

tyd=type of day (weekend or weekday)

cls= customer class

The diversity for a group containing N number of customers can be calculated by following algorithm:

1. Pick 2 to N_g random customers with non-coincident peaks where N_g is the maximum number of customers for a group beyond which the diversity factor remains almost constant.
2. Diversity factor is calculated from a number of sample groups.
3. Individual daily peak for a customer on a particular day is calculated as follows

$$Peak_i(cls, mon, day) = \max_{hr} P_i(cls, mon, day, hr) \qquad (3.8)$$

4. Unique random groups are selected and the sum of individual peaks for a group where the day matches the day type, is calculated as follows

$$= \sum_{i=1}^{N_g} Peak_i(cls, mon, tyd) \qquad (3.9)$$

5. The total group peak for a sample x on a day is calculated using the total group consumption for 24 hours and finding the maximum to estimate the peak for the group where the day matches the type of day.

$$Total\ _{gx}(cls, mon, day, hour) = \sum_{i=1}^{N_g} P_j(cls, mon, day, hour) \qquad (3.10)$$

$$Peak_{gx}(cls, mon, tyd) = \max_{hr} Total_{gx}(cls, mon, tyd, hour) \quad (3.11)$$

6. Diversity factor for a sample x in the group is calculated by dividing the sum of individual peaks for the group by the group peak as given by

$$DF_x(cls, mon, day, N_g) = \frac{\sum_{j=1}^{N_g} Peak_j(cls, mon, tyd)}{Peak_{gx}(cls, mon, tyd)} \quad (3.12)$$

7. Diversity factor for a group can be calculated by taking the average of the diversity factors for all the samples as given by

$$DF\left(cls, mon, tyd, N_g\right) = \frac{\sum_{nSamples} DF_x(cls, mon, day, N_g)}{nSamples} \quad (3.13)$$

It was observed that the diversity factor for a group of customers does not vary for larger number of customers in a group. Figure 3-1 demonstrates a representative diversity factor curve and we can observe that the d-factor varies significantly for lower number of customers.

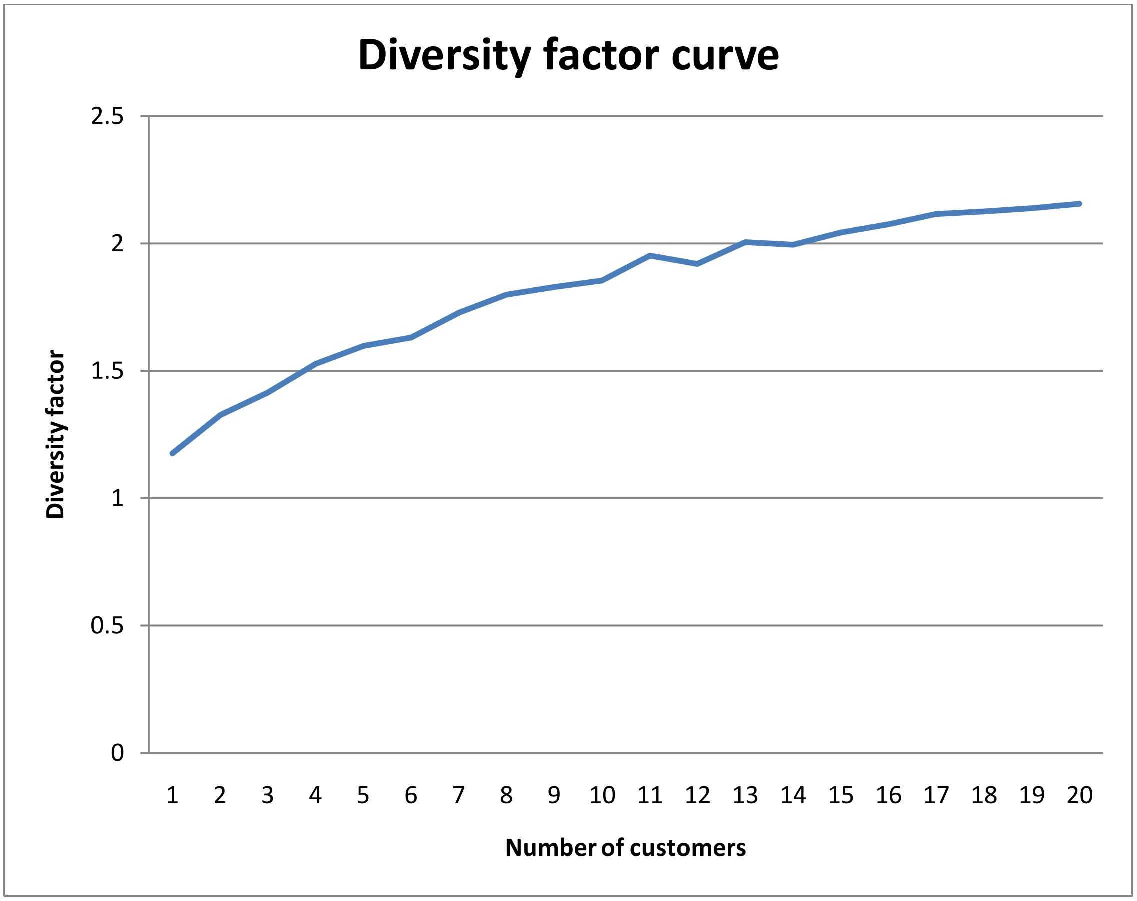

Figure 3-1 Representative Diversity factor curve

3.4 Peak load estimation using NLRE

kWHr-to-peak kW conversion factor and the diversity factors calculated can be applied to the billing kWHr parsed into the monthly billing cycles based on Nonlinear Load Research (NLRE) based algorithm[15]. The billing kWhr is divided into calendar monthly cycles based on average accumulated consumption.

Using the non linear parsed billing kWHr data and the load research factors, peak load is estimated as follows:

1. The daily peak for an account can be estimated by using the C-factor matching for that class, day type, and month and monthly kWHr billing data as given by

$$D_{Peak}\,(cls, account, mon, tyd) = c_f\,(cls, tyd, mon) \times Monthly_{kW\ billing}\ (account, class, mon, tyd) \tag{3.14}$$

2. From the 24 hour load profile and the daily peak calculated from the load research statistics, the 24 hour load profile for an account is calculated as

$$kWHr_{estimate}\ (cls, account, hour, month, tyd) = \frac{PAVGW_g(hr, mon, tyd, cls)}{Peak_g(mon, tyd, cls)} \times D_{Peak}\,(account, mon, tyd) \tag{3.15}$$

3. For a group of accounts on a distribution transformer, the total kWHr for the group can be estimated by applying the diversity factor for the group as

$$kWHr_{g,estimate}\ (cls, hr, month, tyd) = \frac{\sum_{account\ =1}^{N_g} kWHr_{estimate}\ (account, cls, hr, mon, tyd)}{DF(cls, mon, tyd, N_g)} \tag{3.16}$$

The peak for the group can be estimated for a month as

$$Peak_{g,estimate}\ (cls, mon, tyd) = \max_{hr\epsilon 24} kWHr_{g,estimate}\ \ (mon, tyd, hour) \tag{3.17}$$

3.5 Division of customers into classes

Customers are categorized into different groups based on various criteria. Rate class assigned by the utility companies is one of the methods that is used to classify the customers. Rate class is an identifier used by the utility companies to classify customers based on the class by which the rates are assigned for billing. What follows describes one of the procedures used to classify the customers in a metropolitan area with mixed residential and commercial customers.

Initially the customers are divided into residential and non-residential customers based on the annual kWHr consumption. Residential customers are further divided into 3 sub-groups by based on annual consumption. Some other means to identify residential customers which might have been missed in the above procedure can be identified based on the name of the customer, address, building type such as apartment or condominium.

Non-residential customers are further divided into identical user groups based on the rate class assigned. Another technique employed for grouping the industrial customers is based on the North American Industrial Classification System (NAICS codes). These are the codes used to classify the customers based on their industry type. Customers belonging to the same industry are expected to have similar load patterns and can be grouped into a class based on the first two digits of the NAICS code. It has been observed several times that the NAICS codes are assigned incorrectly by utility companies, or the management for the industrial company receives the bill on the behalf of the industry and hence has the incorrect grouping. In such cases, the name of the customer and address can be used. Some of the industrial customers with different usage patterns are education, restaurants, night clubs and hospital etc.

The anomalies in the groups can also be identified based on the individual account C-factor distributions. The following algorithm of Figure 3-2 shows the steps.

1. Group and assign the class type for the non-residential customers based on the first two digits of the NAICS code.
2. If there are any anomalies in the grouping, reorganize based on the service address, part supplied.
3. Further reclassification done based on the diversified load curves.
4. Final grouping done based on the individual C-factor distributions.

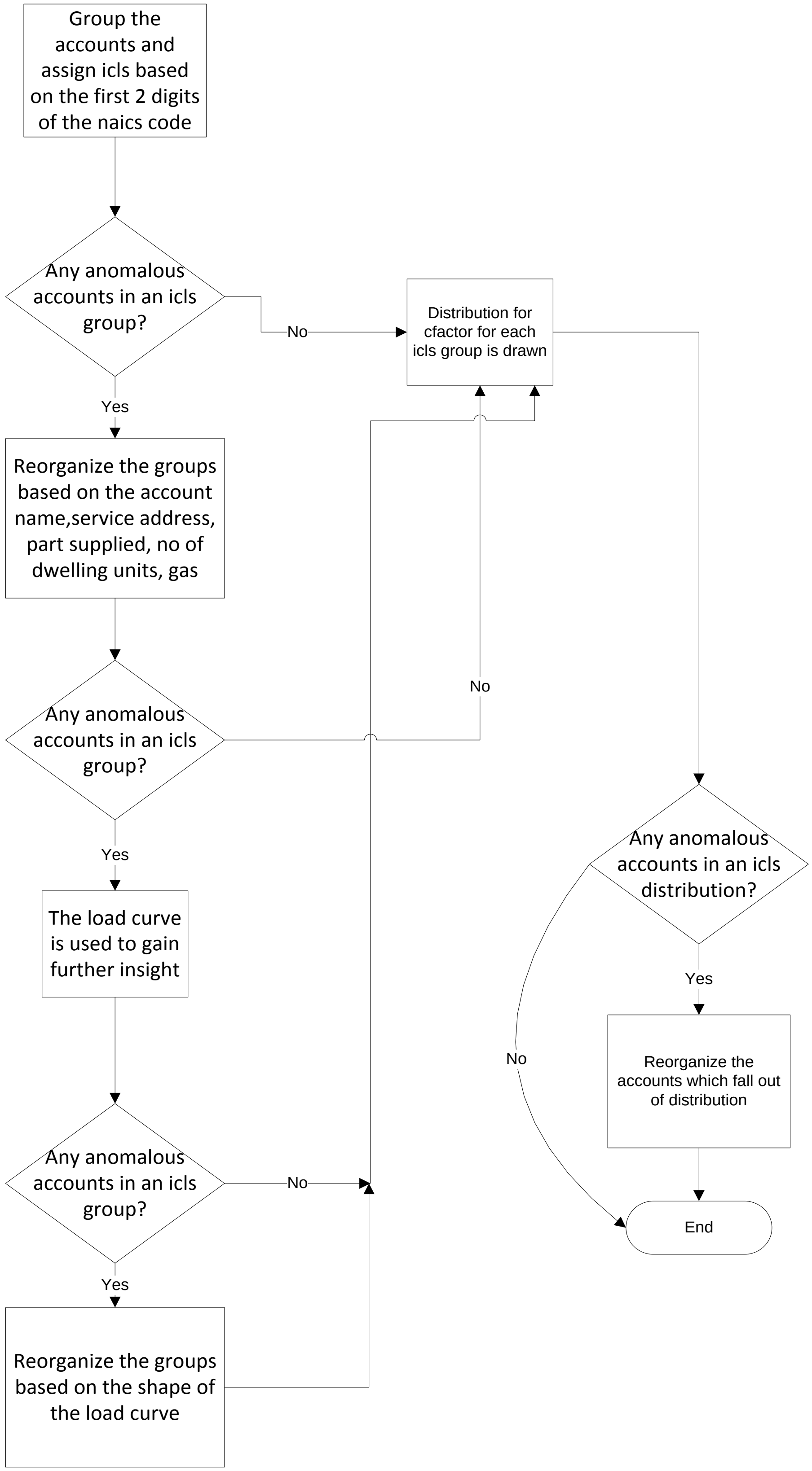

Figure 3-2 Flowchart describing the division of customer classes

In this research, annual kWHr consumption is used to divide the customers. Industrial and commercial customer class division based on annual kWHr consumption is shown in Table 3-1.

Table 3-1 Customer class division based on annual consumption

Class	Description	Annual kWHr Consumption
3	Small Commercial Type 1	0KWHR to 15,000KWHR
5	Large Commercial Type 1	15,000KWHR to 100,000KWHR
8	Single Large Load Type 1	100,000KWHR to 500,000KWHR
9	Single Large Load Type 2	500,000KWHR+

3.6 Customer penetration

The number of customers on the circuit is not proportional to the load that they contribute. As can be observed from Table 3-2, a total of sixteen residential customers of class 0 customers contribute only 5% of the total load, whereas one industrial customer of class 9 contributes to 98% of the load. Hence, a customer class is dominant on a particular circuit based on the load contributed and not based on the number of customers on the circuit.

The customer penetration is calculated by estimating the proportional kWHr accumulated by the customers from each class.

Table 3-2 Customer penetration for industrial circuits

Class	Number of customers	Contribution (%)
0	16	5.39367
3	5	0.0090675
5	2	0.0021003
9	1	98.8961

3.7 Results for industrial customer dominated circuit

The method described above when applied to residential and commercial circuits was reported to give results with a maximum error of 13% [15].Here we study the results for industrial class 9 dominated circuits.

The estimated KW is given as

$$3\ phase\ KW = \sqrt{3} \times VI \cos\theta \qquad (3.18)$$

where V is the line-line nominal voltage in Kilovolts

I is the RMS line current in amperes

$\cos\theta$ is the power factor.

As the voltage and power factor is approximately constant, we compare the currents on the system to compare the peaks during winter and summer.

Circuits with dominant industrial class 9 customers are chosen where the annual kWHr consumption is greater than 500,000 KWHR for each customer.

The estimated currents and the SCADA currents are compared for these and they show that the estimated peaks are higher than the actual as shown in Chapter 4 discussion. Hence we propose the Non-Negative Least Square estimation to reduce the errors and improve the estimates.

3.8 Non-Negative Linear Least Square Error correction for kWhr-kW conversion factor

Utility companies collect SCADA data at substations and/or feeders. In this thesis this data is used to adjust the conversion factors or C-factors in order to better estimate the peak loads.

Non-Negative Least Square fitting

The Least Squares method has been used historically in regression analysis in such fields as astronomy and engineering. This method has also been used in load estimation in the field of distributed state estimation [24], [25]. It is used to find a better fit for the data available. Estimated values are fitted to the actual data by minimizing the square of the error or the residual between the two.

A linear least squares algorithm can be used to solve the function $y = f(A, x_i)$ of the form $y_i = Ax_i + B$ where $x\epsilon R^n$ and $y\epsilon R^m$ and A is an $m \times n$ matrix.

The residual is calculated as

$$r = y_i - (Ax_i - B) \tag{3.19}$$

The least squares involves finding the minimum of the sum of squares of the residuals

$$SS = \sum_{i=1}^{n} |y_i - f(A, x_i)|^2 = min \tag{3.20}$$

Here, the industrial customer’s currents and the SCADA currents are used and the fitting is done based on Lawson-Hanson’s non-negative least squares [36].

Here the constraint $x \geq 0$ holds good.

This is essential as the C-factors cannot be negative and hence the multiplication factor x, cannot be negative.

SCADA currents are used to adjust the estimated currents which were estimated based on NLRE, as SCADA is the real data measured at the substation and expected to be more reliable and accurate than estimated currents.

A group of circuits with predominant industrial or class 9 customers are chosen and the currents are estimated using the power flow.

The estimated currents are adjusted to match the SCADA currents and the solution to the NNLS, x, is multiplied with the C-factors to obtain new updated C-factors

Hence the new C-factor is given as

$$c_{f,new}(cls,tyd,mon) = x(cls,tyd,mon) * c_f(cls,tyd,mon) \tag{3.21}$$

And x is calculated as follows

$$EC(cls,tyd,mon,hr)x(cls,tyd,mon) = SC(cls,tyd,mon,hr)$$

Let ckt1, ckt2, ckt3… cktn be the n circuits with dominant class 9 circuits which have 24 hour readings.

Let us suppose that we are adjusting for the C-factors for a particular month and day type.

We choose the SCADA and estimated currents for all the available 24 hours matching the month and day type and run the NNLS algorithm to find the x value as

$$\begin{bmatrix} EC(ckt1,hr1) \\ EC(ckt1,hr2) \\ \vdots \\ EC(ckt1,hr24) \\ \vdots \\ EC(cktn,hr1) \\ EC(cktn,hr2) \\ \vdots \\ EC(cktn,hr24) \end{bmatrix} \times X = \begin{bmatrix} SC(ckt1,hr1) \\ SC(ckt1,hr2) \\ \vdots \\ SC(ckt1,hr24) \\ \vdots \\ SC(cktn,hr1) \\ SC(cktn,hr2) \\ \vdots \\ SC(cktn,hr24) \end{bmatrix} \tag{3.22}$$

Applying (EC, SC) , we find x that minimizes the least square error for the equation (3.23).

Table 3-3 gives the values calculated of x for various months and day types for class 9.

Table 3-3 New and old C-factors

MON	TYD	X	OLD CFACTOR	NEW CFACTOR
1	0	0.6071	0.0028	0.0017
2	0	0.5500	0.0040	0.0022
3	0	0.5333	0.0045	0.0024
4	0	0.8000	0.0030	0.0024
5	0	0.8519	0.0027	0.0023
6	0	0.9259	0.0027	0.0025
7	0	0.8462	0.0026	0.0022
8	0	0.8462	0.0026	0.0022
9	0	0.8929	0.0028	0.0025
10	0	0.8519	0.0027	0.0023
11	0	0.6552	0.0029	0.0019
12	0	0.7273	0.0022	0.0016
1	1	1.0000	0.0023	0.0023
2	1	0.7500	0.0024	0.0018
3	1	0.7727	0.0022	0.0017
4	1	0.8000	0.0020	0.0016
5	1	0.9444	0.0018	0.0017
6	1	1.0000	0.0019	0.0019
7	1	1.1250	0.0016	0.0018
8	1	0.8947	0.0019	0.0017
9	1	0.9524	0.0021	0.0020
10	1	0.8182	0.0022	0.0018
11	1	0.7500	0.0020	0.0015
12	1	0.7778	0.0018	0.0014

Here EC and SC are the 3-phase estimated currents based on NLRE method and SCADA currents on the industrial dominated circuit.

The updated C-factor is then applied in a similar way to obtain new currents and we see that the new estimations follow the SACDA currents better.

Figures 3-3 to 3-5 illustrate the estimated and actual currents for a July weekday when the summer peak occurs and Figures 3-6 to 3-8 illustrate November weekday estimates and it can be seen that the NNLS method tracks the SCADA measurements better for a large number of circuits.

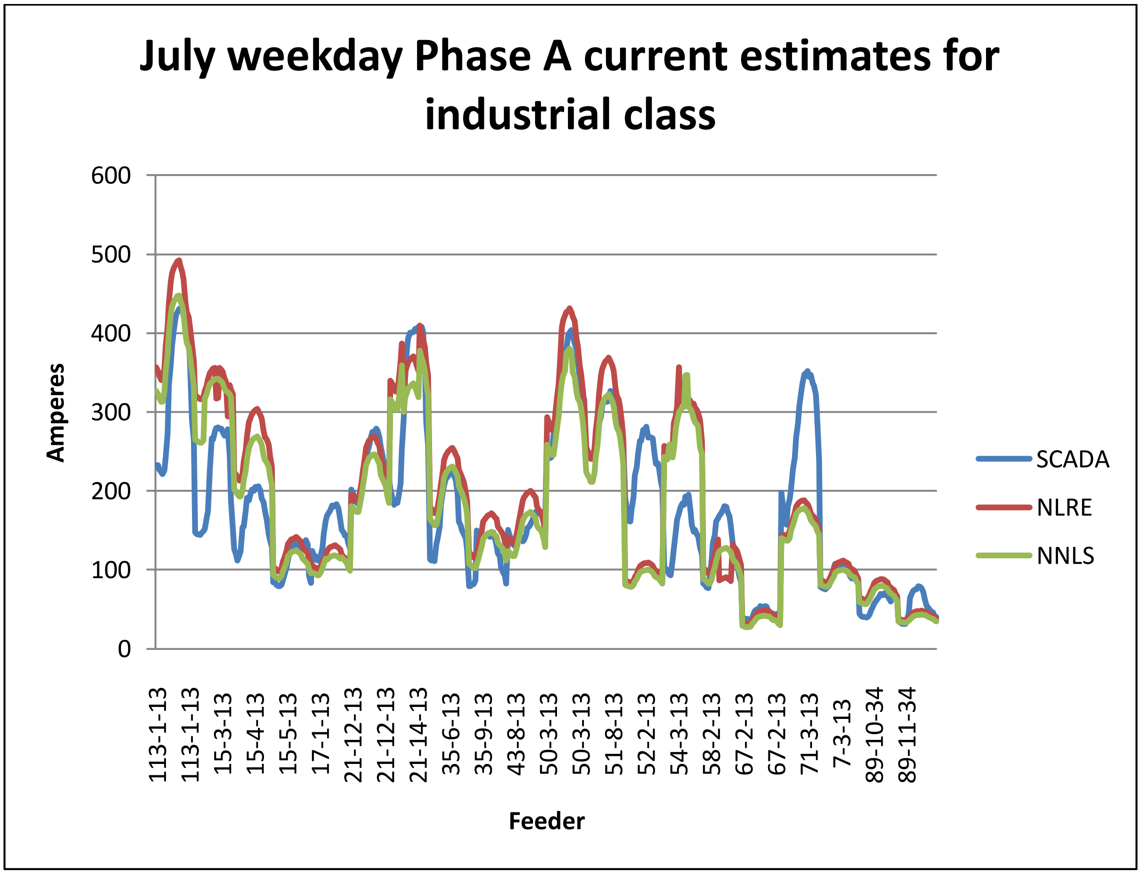

Figure 3-3 July weekday Phase A current estimates for industrial circuits

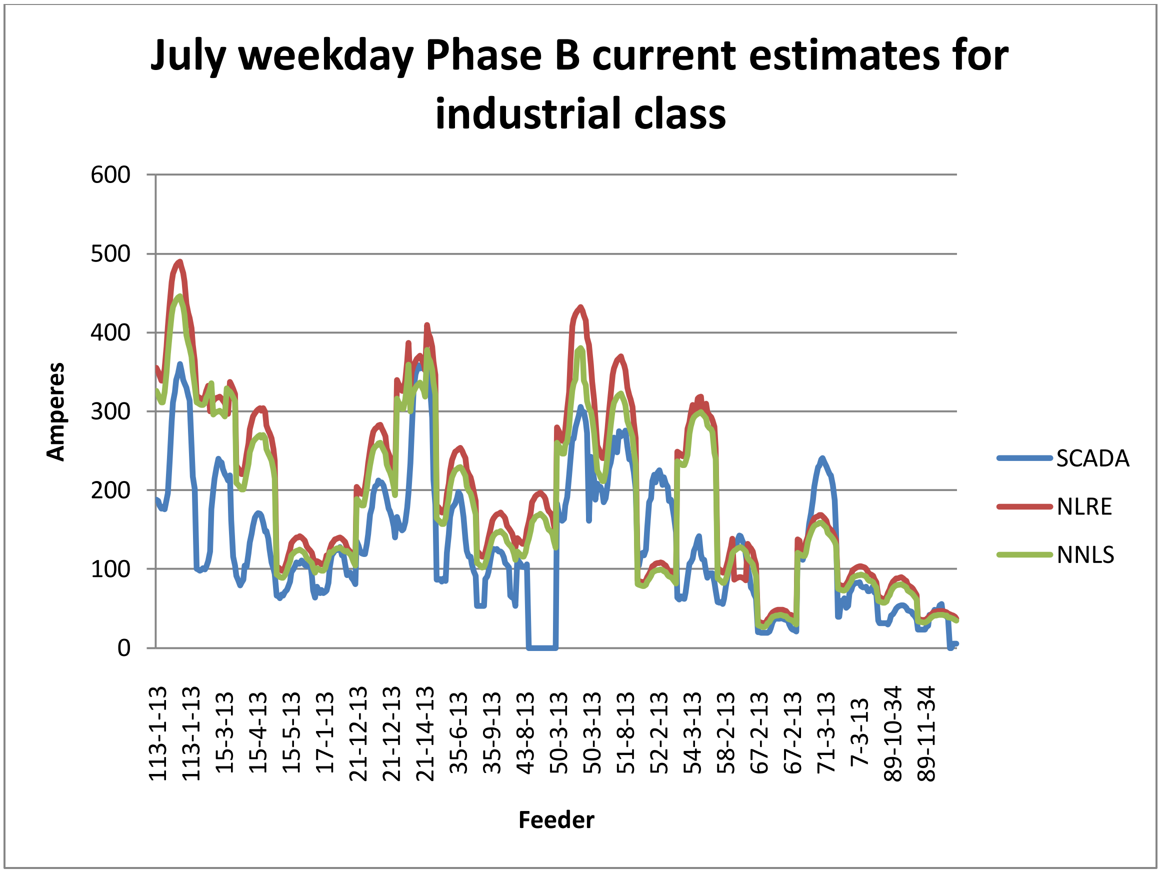

Figure 3-4 July weekday Phase B current estimates for industrial circuits

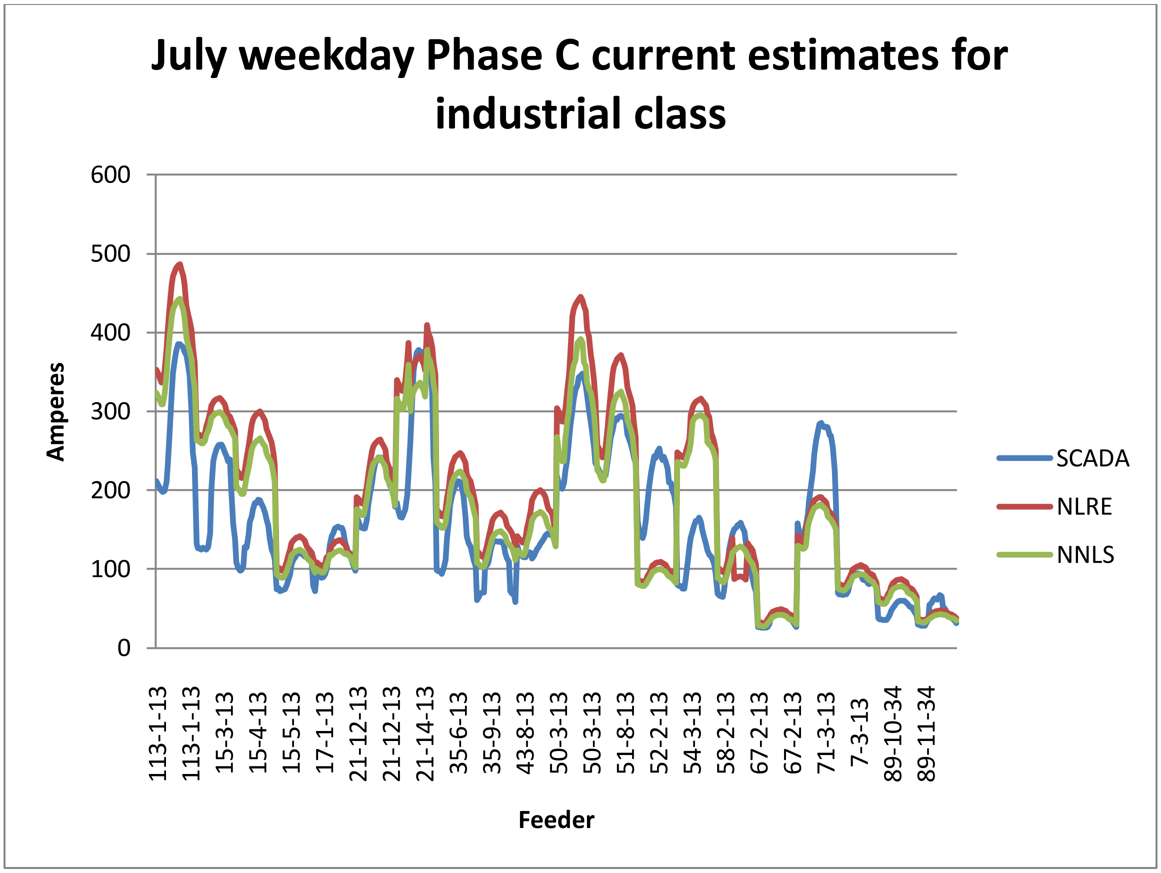

Figure 3-5 July weekday Phase C current estimates for industrial circuits

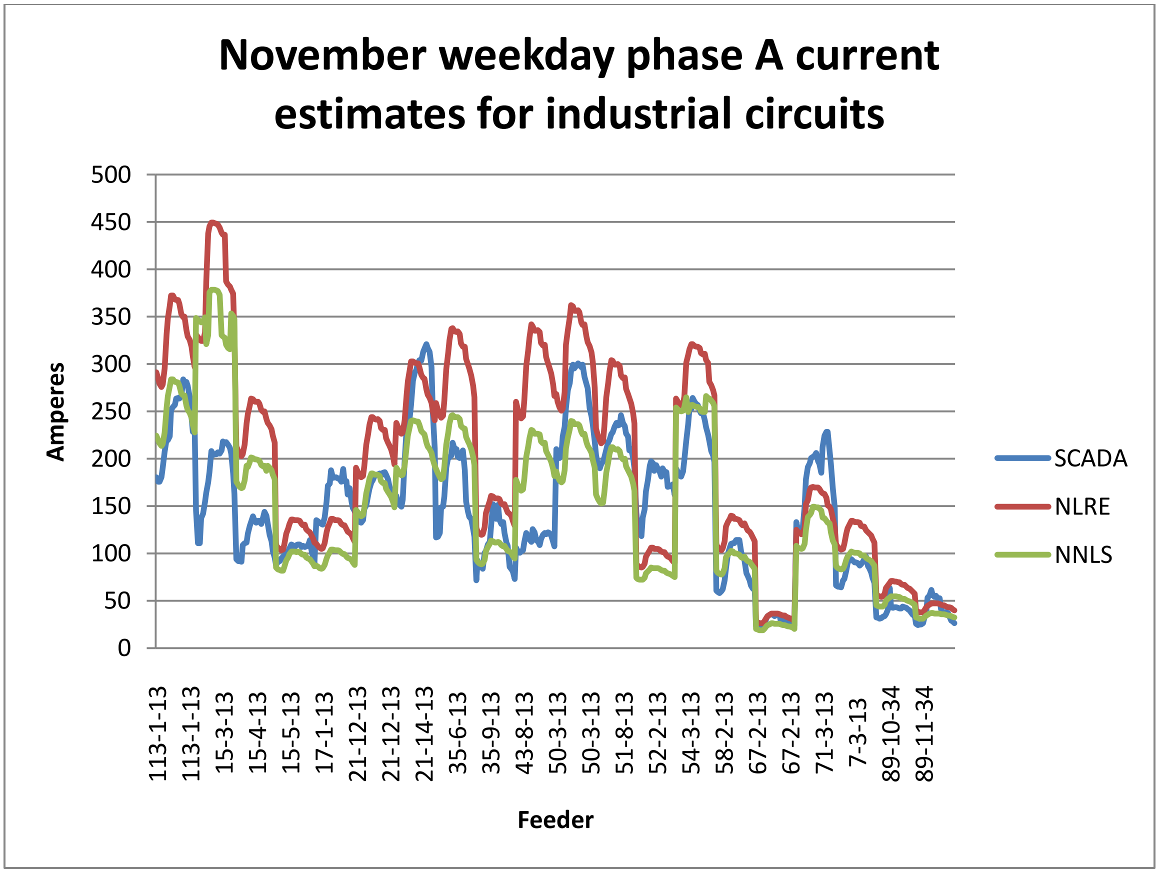

Figure 3-6 November weekday Phase A current estimates for industrial customers

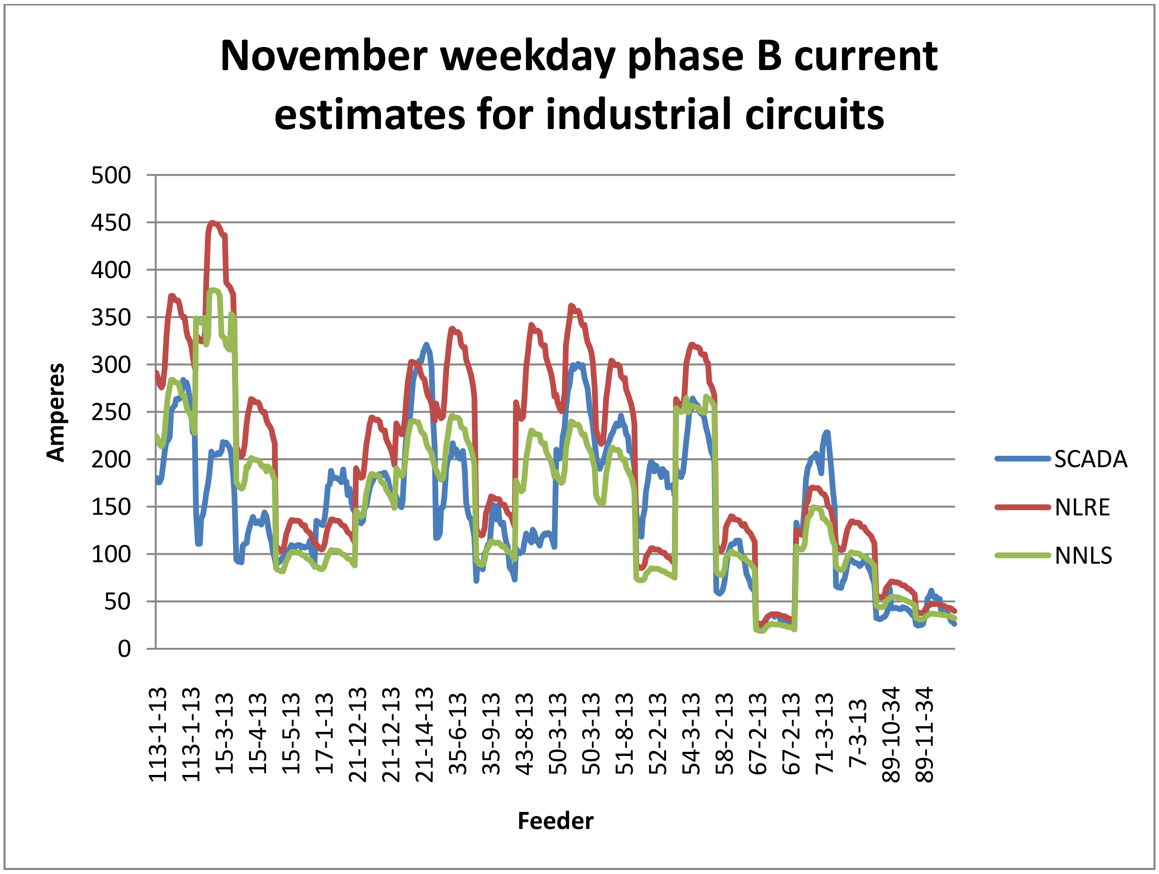

Figure 3-7 November weekday Phase B current estimates for industrial customers

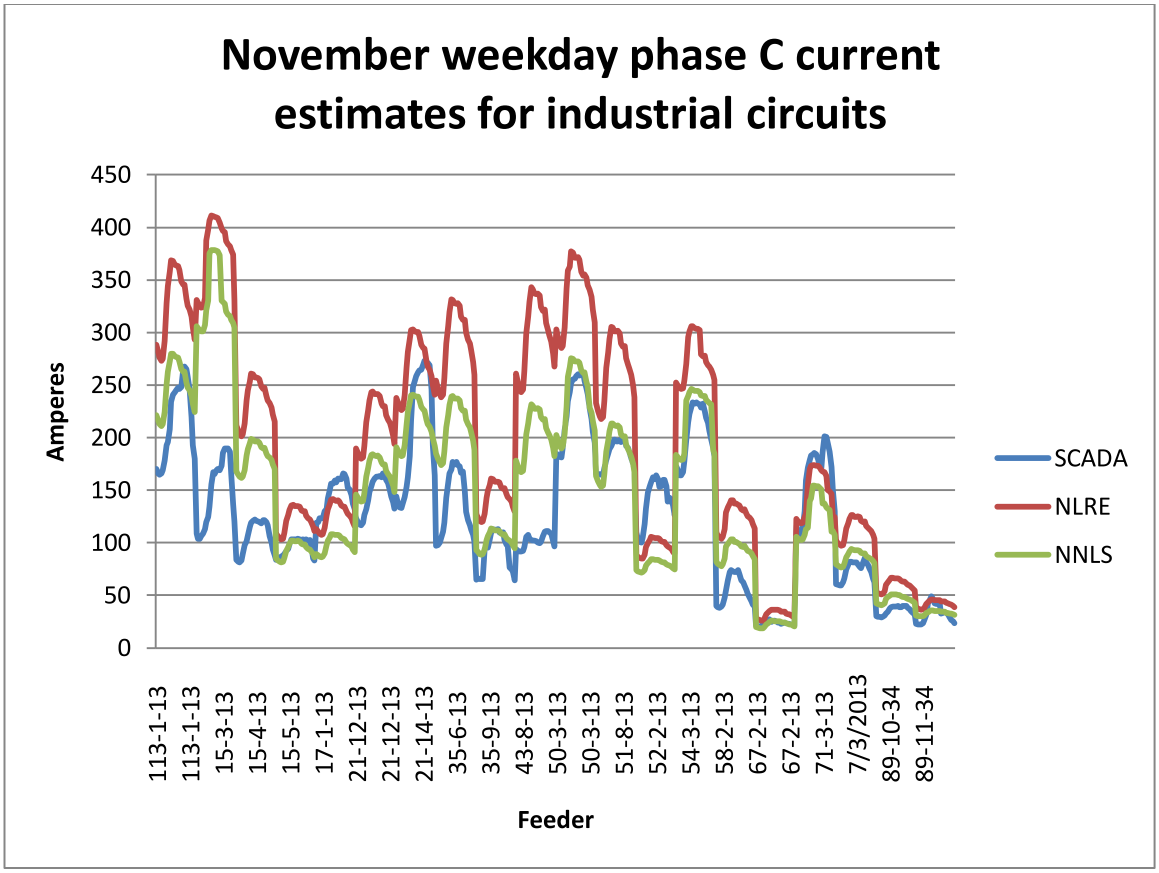

Figure 3-8 November weekday Phase C current estimates for industrial customers

Chapter 4

Results and Discussions

4.1 Load Research Calculation

The load research statistics calculation is implemented as a standalone application using C#, ADO.NET and SQL procedures. C# provides reporting services for report generation, viewing and exporting reports. Also, SQL routines provide a cleaner way to connect and execute SQL routines and data binding.

A load research statistics application calculates 24 hour kWHr consumption, daily peak, diversity factors and C-factors. It is also used to these factors. Hourly kWHr sample readings are provided by utility companies and the calculated load research factors are imported into DEW to estimate the load on the system. The application provides a GUI to enter the login information to connect to the SQL server database as shown in Figure 4-1.

The next form as shown in Figure 4-2 provides the options to calculate selected load research statistics which are inserted into the database tables. The following are main components of the form.

1. Class, type of the day, month and weather condition list boxes. One or more values can be chosen from these list boxes to select the values over which the load research factors will be calculated.
2. Check boxes to calculate all or a part of the following:
    a. 24 hour kWHr load profile for each customer class, day type, and month and weather condition group.

b. Calculating the daily peak for each group which is the maximum among the 24 hour kWHr calculations performed above.
c. Calculating kWHr-Peak-kW conversion or C-factors for the group chosen.
d. Calculating diversity factors for the group chosen.

All of the above calculations are performed by invoking the SQL stored procedures which work against the database to group the data and calculate the load research factors.

3. The start button begins the calculation and a message box indicates the end of calculation.
4. The chart button, redirects to the next form as shown in Figure 4-3 which generates the report.

The chart form has the following options and reports to choose from:

1. Class, type of the day, month and weather condition list boxes. One or more values can be chosen from these list boxes to select the values over which the load research factors will be calculated.
2. Radio buttons to chose one of the following reports:
    a. C-factor report is an overlay plot against 12 months. Class, type of day, and weather condition can be chosen for which the report has to be generated.
    b. The diversity load curve report which plots the kWHr consumption for 24 hours. The plots can be chosen for a group of class, month, day type, weather condition.
    c. Diversity factor curves against the number of customers in a group.

3. Refresh button to start report generation which can be viewed in the report area.

Figures 4-3 to 4-5 shows the reports for the above discussed load research factors.

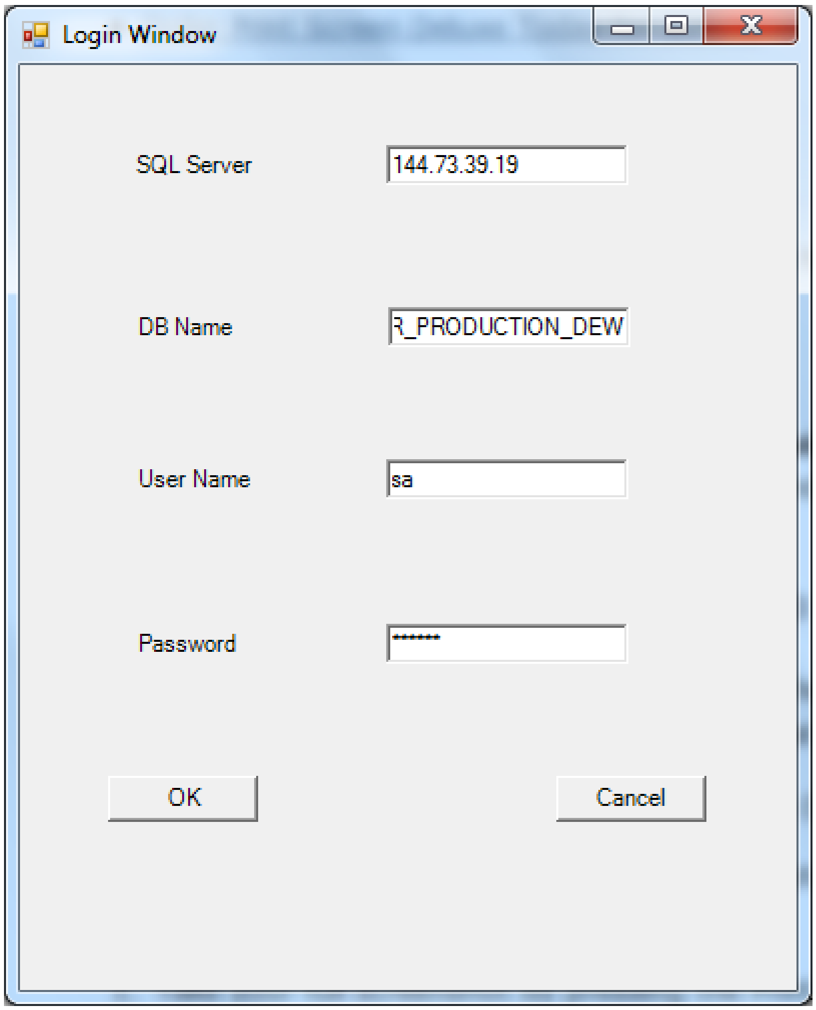

Figure 4-1 Calculation of LDR Statistics:Login window

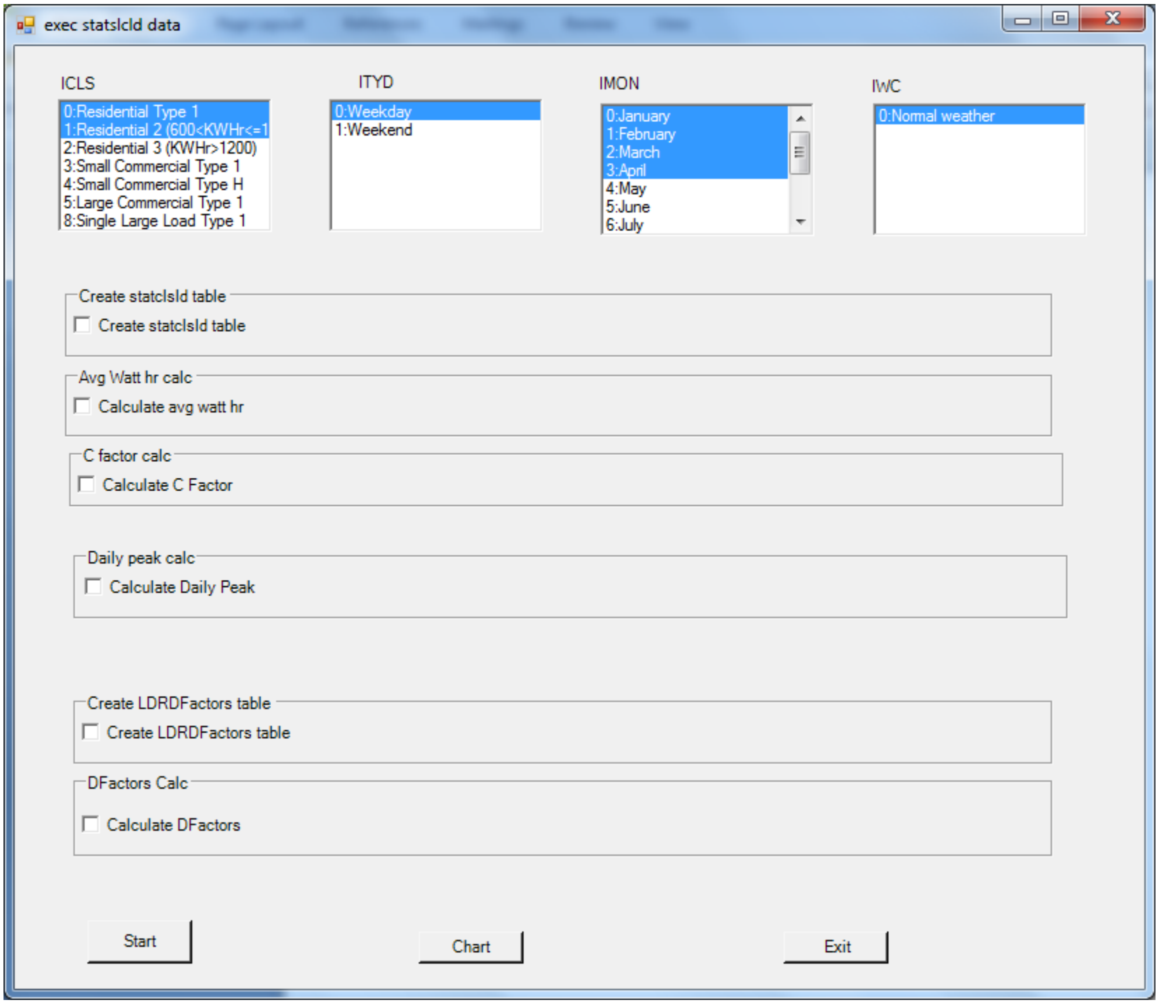

Figure 4-2 Calculation of LDR Statistics: Statistics calculation window

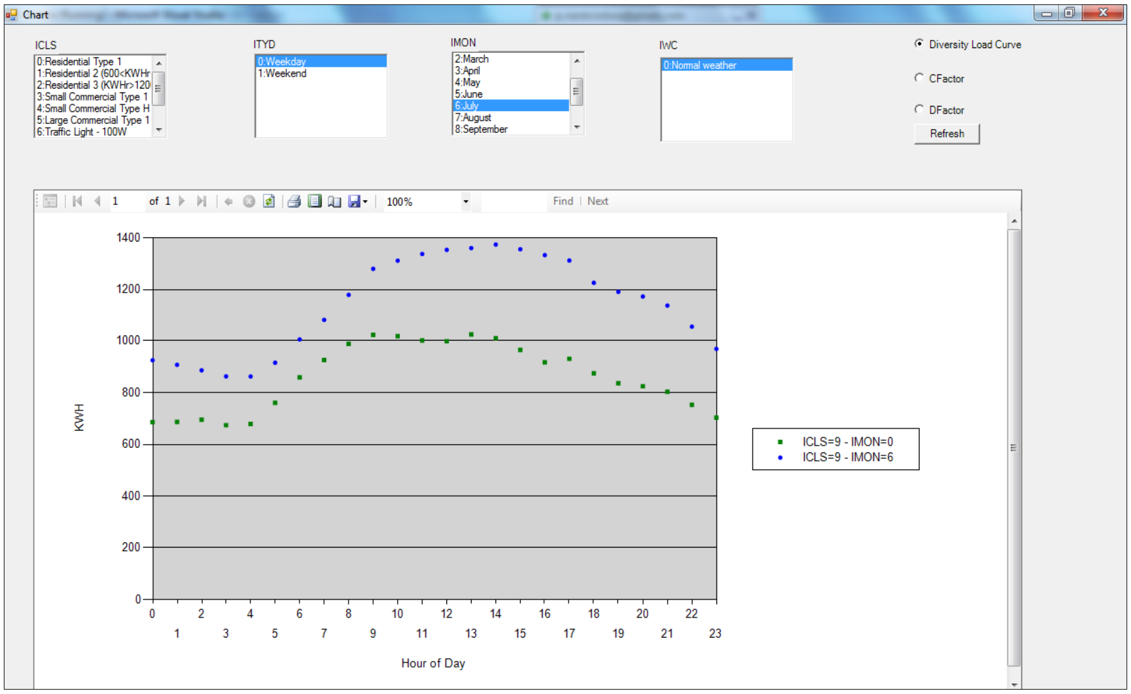

Figure 4-3 Calculation of LDR Statistics: Chart window- Diversity Load curve for class 9 for January and July weekdays

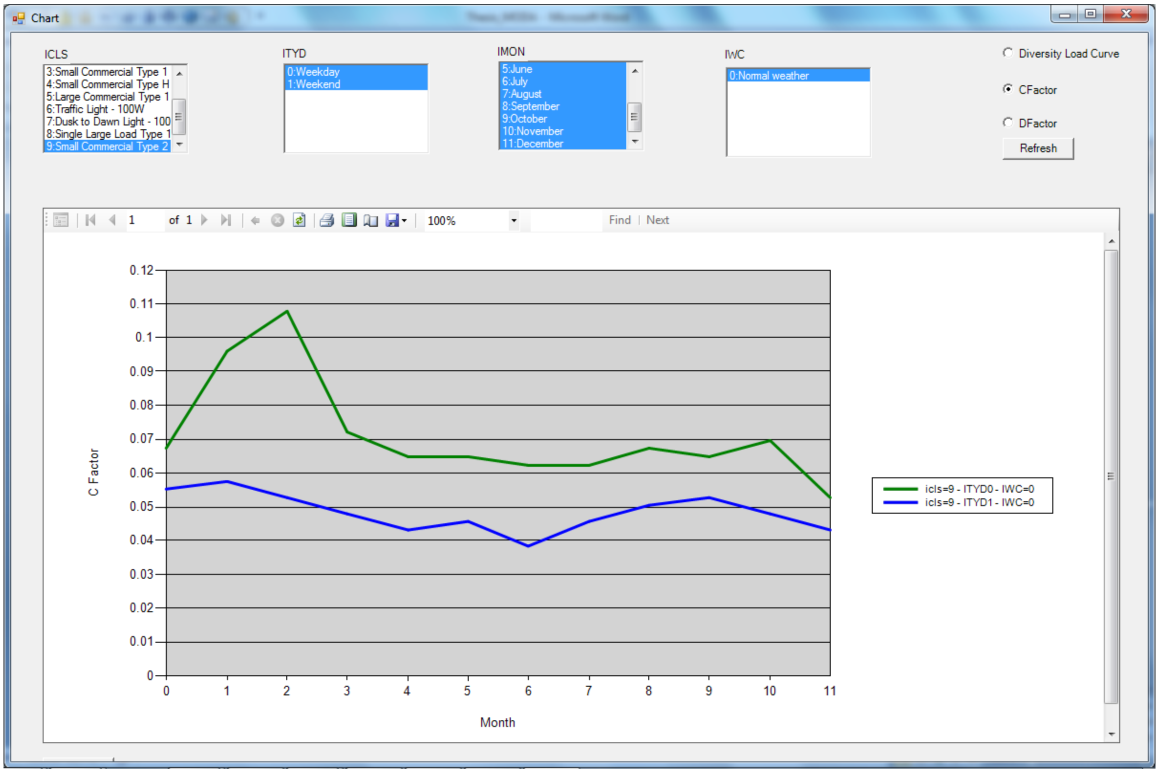

Figure 4-4 Calculation of LDR Statistics: Chart window- C-factor curves for class 9 for weekday and weekend

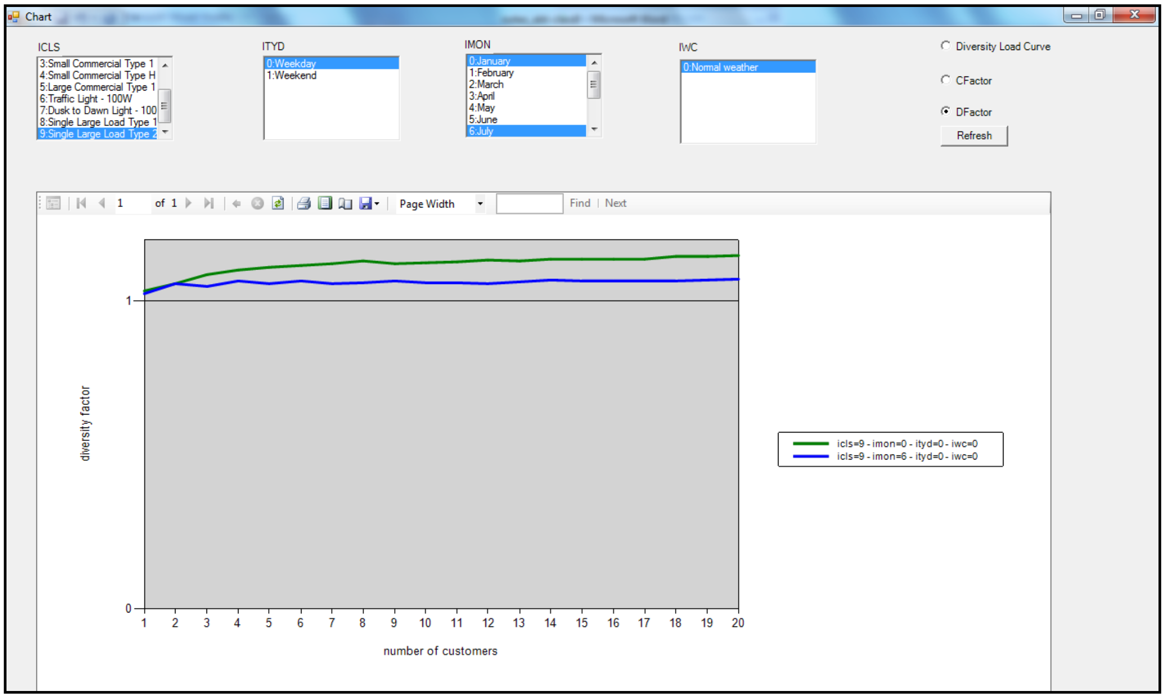

Figure 4-5 Calculation of LDR Statistics: Chart window- Diversity factor curve for class 9 for January and July weekday

The above calculated factors are applied to the monthly billing data parsed using NLRE algorithm and is used to estimate the currents on the system. The NNLS package in the R-statistical package [9] is used to minimize the residual error between SCADA and current estimates to find the new C-factors.

4.2 Results discussion

The NNLS method is applied and the estimates are compared against SCADA measurements and also a comparison is provided between the NNLS and NLRE method. To demonstrate the efficiency of the model presented, five industrial dominated circuits are chosen and the results are shown.

Figures 4-6 to 4-11 show a comparison between the estimates during July and November weekday for chosen circuits.

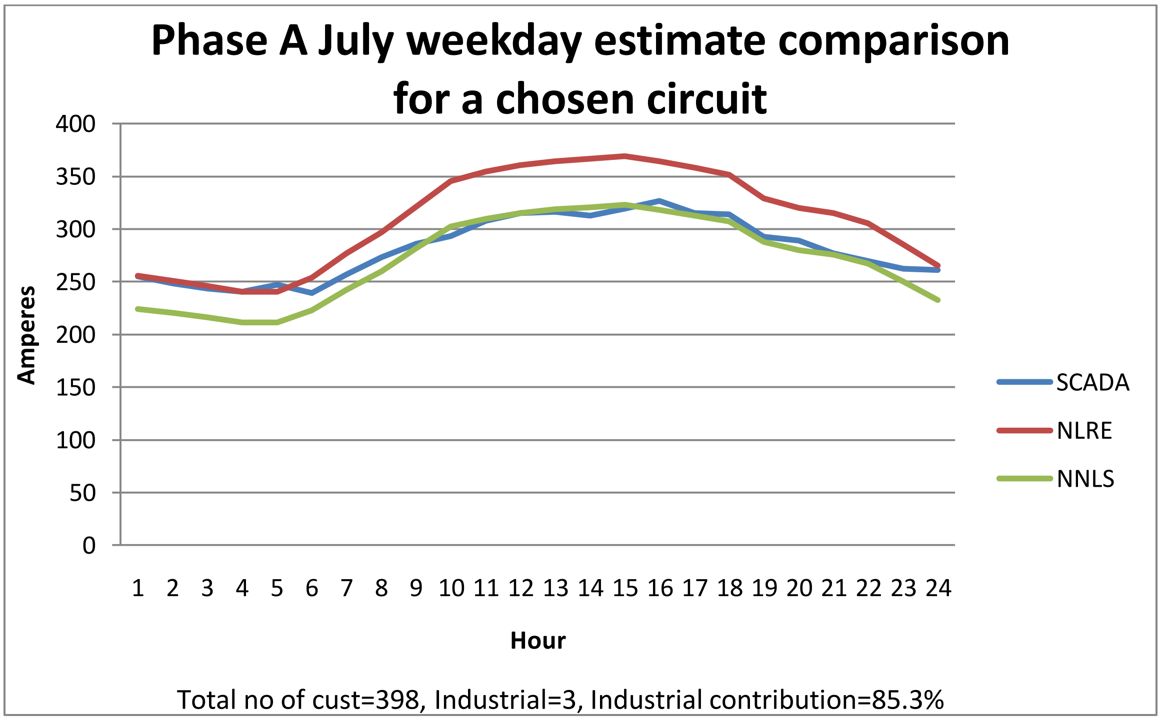

Figure 4-6 July weekday Phase A current estimation for a chosen circuit

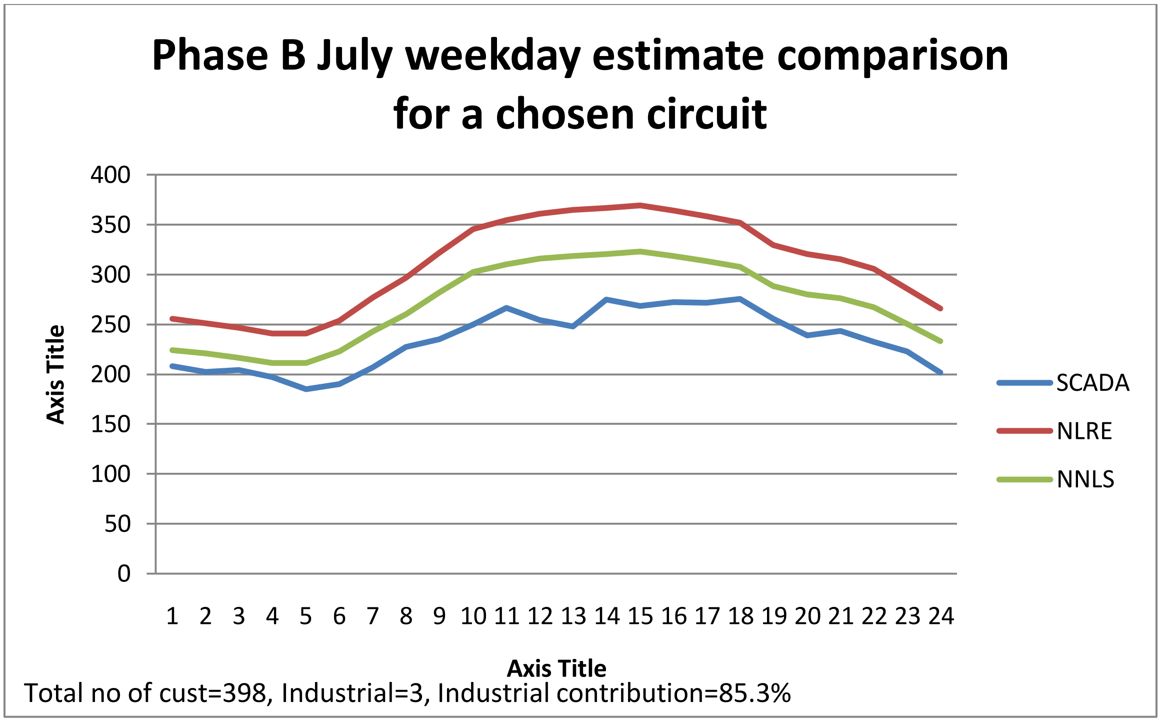

Figure 4-7 July weekday Phase B current estimation for a chosen circuit

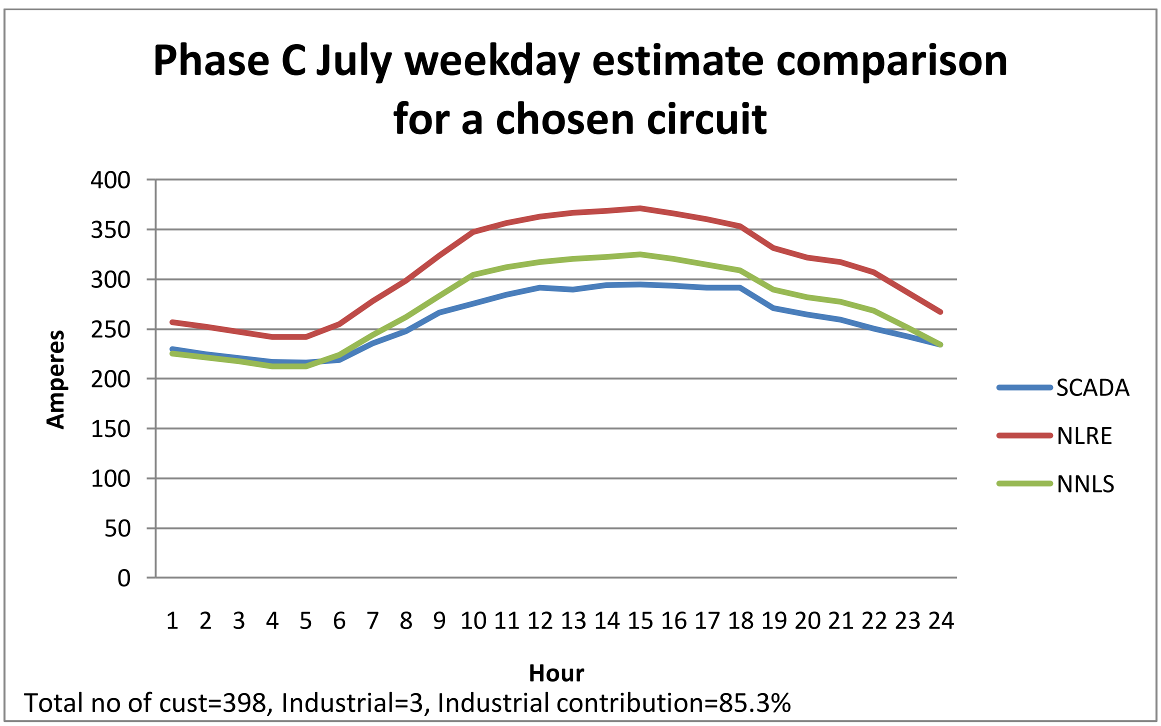

Figure 4-8 July weekday Phase C current estimation for a chosen circuit

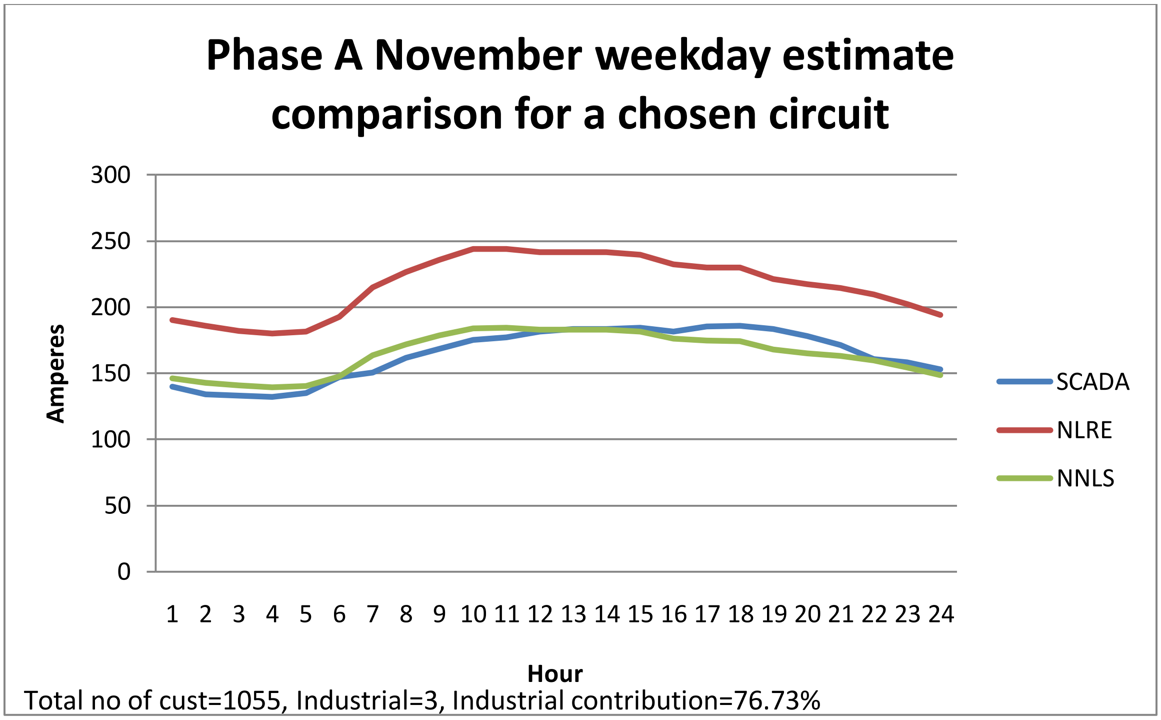

Figure 4-9 November weekday Phase A current estimation for a chosen circuit

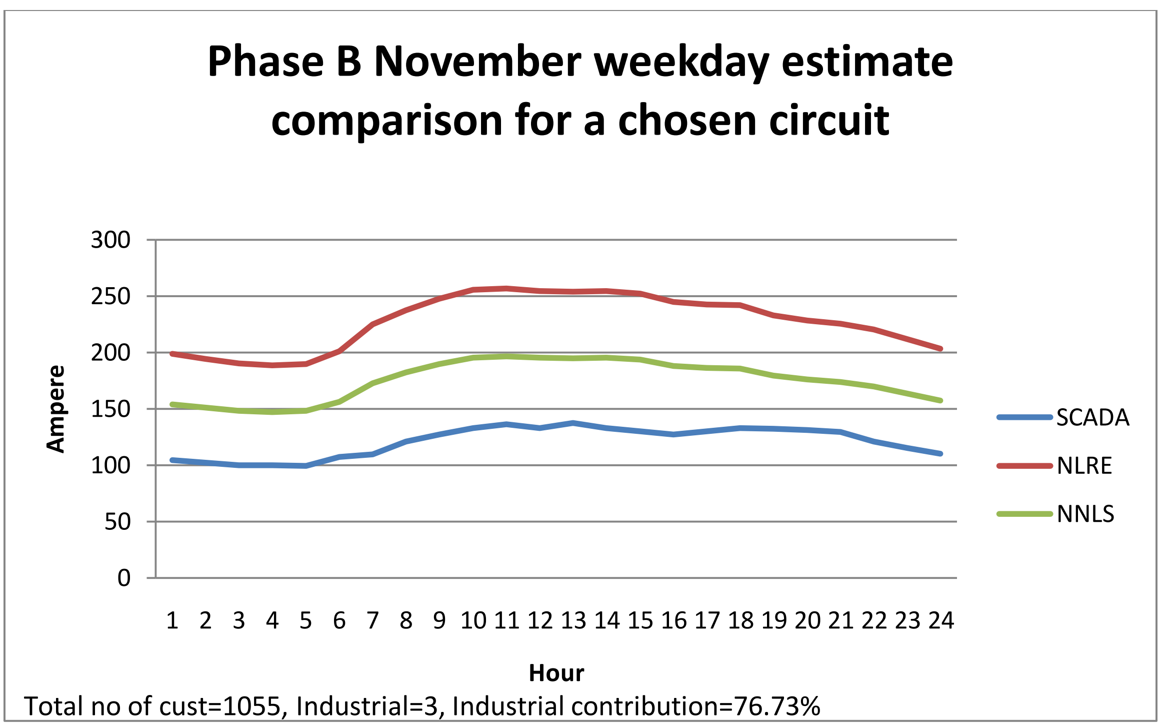

Figure 4-10 November weekday Phase B current estimation for a chosen circuit

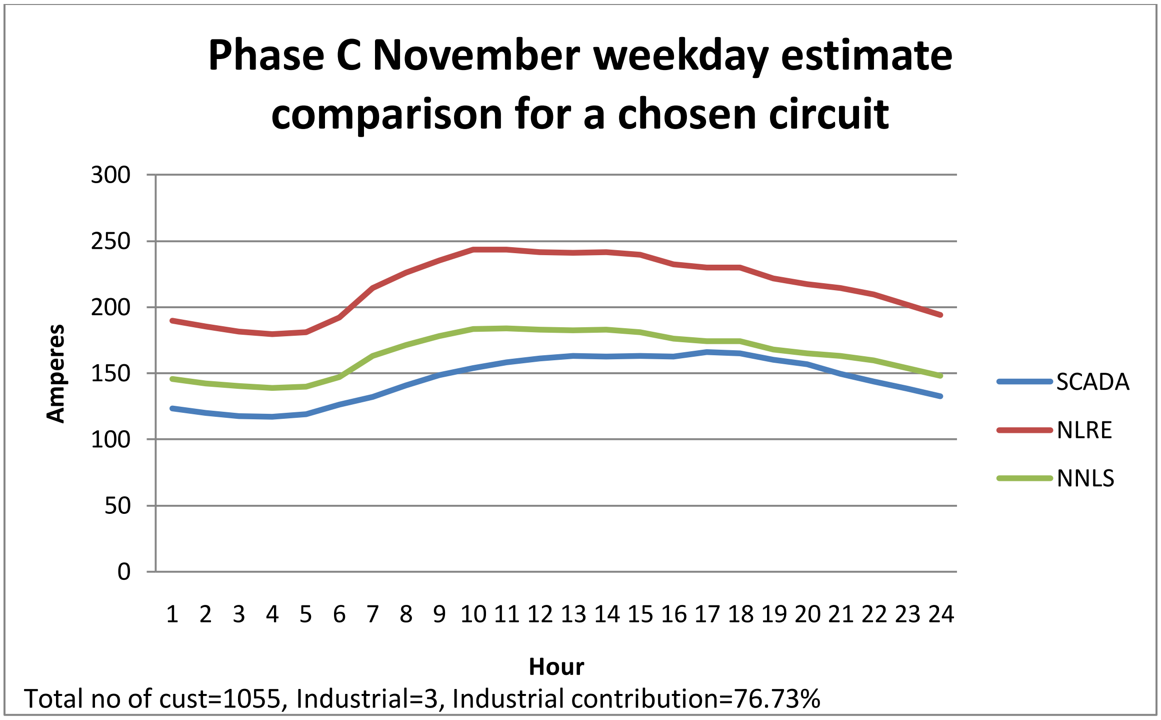

Figure 4-11 November weekday Phase C current estimation for a chosen circuit

A comparison between SCADA peak currents, currents estimated from the NNLS and the NLRE method is presented below. Five feeders dominated by industrial class 9 circuits are chosen as shown in Table 4-1 to demonstrate the accuracy of the NNLS optimization model.

Table 4-1 Customer class division for the chosen feeders

Feeder name	Total number of customers	Industrial Class 9	Industrial Class 9 contribution (%)
35-6-13	359	7	86.62
35-9-13	25	1	98.89
51-8-13	398	3	86.96
15-4-13	174	14	86.62
67-2-13	58	1	80.87

For the month of July during which the system peaks, Table 4-2 to Table 4-6 show that the NNLS has proven to track the peak better than NLRE and has reduced the error by about 10% for 35-6-13, 14% for 35-9-13, 12% for 51-8-13 and 17% for 15-4-13.

Significant improvement can be observed for the months of February and November for 35-6-13 where an improvement of over 90% is observed. For 35-9-13, 51-8-13, 15-4-13 there is an improvement of 50%-70% for the months of February and November.

The NLRE method underestimates currents for feeder 67-2-13 for Phase A. Since the entire system is over estimated, the x value is calculated to be less than 1 for most of the months to reduce the estimates across all the circuits. This leads the NNLS model to further underestimate the currents.

Table 4-2 Error comparison for all 3 phases for feeder 35-6-13

MON	**Phase A**		**Phase B**		**Phase C**	
	NLRE Error(%)	**NNLS Error(%)**	**NLRE Error(%)**	**NNLS Error(%)**	**NLRE Error(%)**	**NNLS Error(%)**
1	-13.15	-11.15	-14.26	-11.75	-11.00	-8.13
2	-95.25	-34.78	-246.49	-134.99	-142.06	-64.26
3	-77.89	-8.64	-175.20	-67.91	-125.12	-36.28
4	-13.25	3.64	-51.20	-28.30	-26.08	-6.76
5	1.46	11.78	-52.02	-35.86	-19.27	-6.42
6	-7.5	-1.98	-45.23	-37.63	-27.02	-20.26
7	**-11.61**	**-1.21**	**-28.13**	**-16.14**	**-16.80**	**-5.59**
8	-12.33	-0.33	-33.65	-19.13	-19.20	-6.03
9	11.11	17.67	-33.78	-23.78	-15.39	-6.60
10	-21.58	-8.61	-160.00	-131.89	-41.69	-26.15
11	-55.55	-13.32	-185.21	-106.67	-86.94	-35.11
12	-34.88	-8.94	-121.02	-78.03	-64.91	-32.25

Table 4-3 Error comparison for all 3 phases for feeder 35-9-13

MON	**Phase A**		**Phase B**		**Phase C**	
	NLRE Error(%)	**NNLS Error(%)**	**NLRE Error(%)**	**NNLS Error(%)**	**NLRE Error(%)**	**NNLS Error(%)**
1	6.58	14.76	6.58	14.75	6.54	14.71
2	-42.59	14.79	-109.90	-25.45	-83.23	-9.57
3	-76.53	-0.21	-142.60	-37.73	-119.22	-24.51
4	-6.81	11.75	-51.87	-25.49	-36.18	-12.55
5	6.63	18.47	-57.73	-37.74	-24.60	-8.83
6	-16.25	-8.55	-45.04	-35.43	-33.75	-24.90
7	**-14.66**	**0.99**	**-36.58**	**-17.94**	**-26.69**	**-9.41**
8	-18.37	-2.25	-43.25	-23.75	-32.05	-14.09
9	-7.34	2.87	-51.19	-36.81	-33.97	-21.24
10	-16.91	-1.91	-75.58	-53.05	-42.24	-24.01
11	-6.1	25.24	-84.78	-30.21	-42.42	-0.40
12	-11.84	12.36	-36.17	-6.71	-21.64	4.63

Table 4-4 Error comparison for all 3 phases for feeder 51-8-13

MON	Phase A		Phase B		Phase C	
	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)
1	-13.62	-7.5	-15.57	-9.36	-16.36	-10.53
2	-78.26	-13.27	-127.22	-44.65	-108.41	-33.04
3	-99.99	-14.66	-143.41	-39.42	-128.43	-31.22
4	-22.38	-0.46	-60.56	-31.72	-43.60	-17.90
5	-3.24	10.23	-57.95	-37.32	-27.76	-11.10
6	-11.19	-4.06	-44.60	-35.33	-28.31	-20.12
7	**-12.94**	**1.29**	**-34.07**	**-17.19**	**-26.09**	**-10.30**
8	-12.9	2.38	-49.25	-29.01	-25.21	-8.31
9	-7.51	2.55	-47.36	-33.52	-30.07	-17.94
10	-14.49	0.42	-62.09	-41.03	-38.44	-20.51
11	-23.49	13.77	-98.08	-38.29	-52.56	-6.75
12	-15.45	10.19	-54.34	-20.12	-37.42	-7.16

Table 4-5 Error comparison for all 3 phases for feeder 15-4-13

MON	Phase A		Phase B		Phase C	
	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)
1	-25.97	-22.77	-29.09	-25.60	-23.77	-19.60
2	-123.52	-54.75	-147.38	-70.38	-121.33	-49.83
3	-180.03	-74.32	-232.38	-108.89	-199.25	-85.91
4	-41.98	-22.96	-95.88	-70.41	-66.30	-44.06
5	-44.8	-29.06	-137.39	-112.22	-86.90	-66.65
6	-13.51	-7.62	-79.54	-70.49	-46.96	-39.36
7	**-47.96**	**-31.04**	**-77.62**	**-57.89**	**-59.66**	**-41.25**
8	-32.78	-17.87	-80.72	-60.93	-55.54	-38.12
9	-21.48	-12.34	-87.49	-73.75	-55.63	-43.96
10	-53.66	-37.53	-127.79	-104.52	-94.06	-73.77
11	-83.65	-40.16	-175.70	-111.93	-113.48	-62.77
12	-51.28	-27.78	-84.40	-56.60	-62.59	-37.24

Table 4-6 Error comparison for all 3 phases for feeder 67-2-13

MON	Phase A		Phase B		Phase C	
	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)	NLRE Error(%)	NNLS Error(%)
1	17.39749	20.23852	10.97	14.36	14.17	17.28
2	-28.2911	17.30584	-114.10	-37.47	-67.85	-8.04
3	-48.9707	14.63325	-122.05	-26.64	-88.68	-7.86
4	13.48289	29.18688	-66.03	-35.71	-27.24	-4.09
5	23.01543	33.27429	-51.27	-31.03	-14.28	0.95
6	18.65384	24.19797	-35.46	-26.20	-17.62	-9.63
7	**9.587578**	**22.33638**	**-27.08**	**-9.13**	**-10.91**	**4.64**
8	17.94104	29.62058	-19.49	-2.43	-0.28	13.97
9	15.91311	24.26289	-77.15	-59.49	-30.03	-17.14
10	5.904054	18.31674	-44.54	-25.37	-23.80	-7.46
11	-3.23362	26.93736	-66.95	-17.72	-34.80	4.62
12	24.43573	40.81912	-21.56	5.06	-0.34	21.50

Figures 4-12 to 4-26 show the absolute error comparison for the peak currents for the five feeders and for three phase currents.

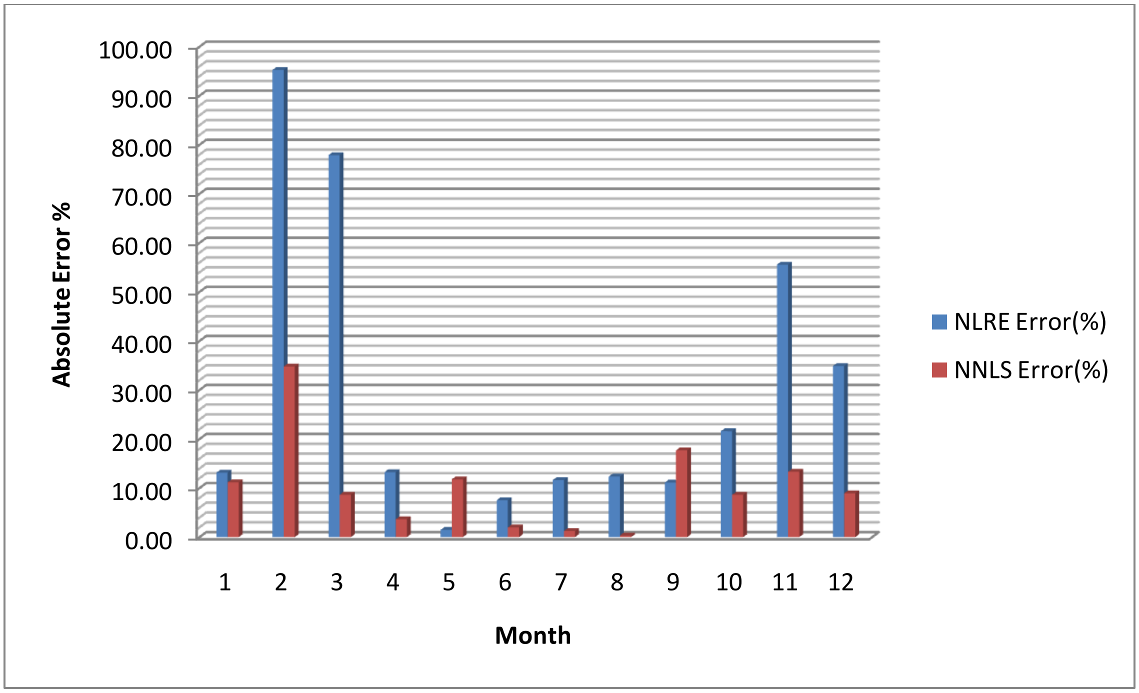

Figure 4-12 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 35-6-13

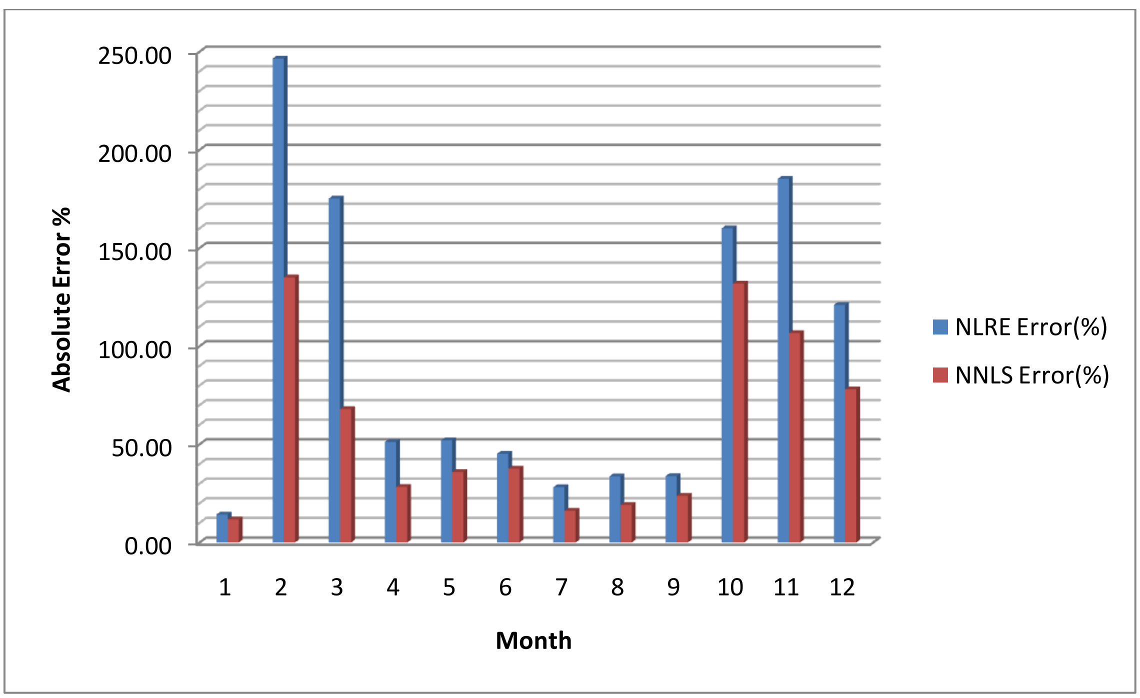

Figure 4-13 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 35-6-13

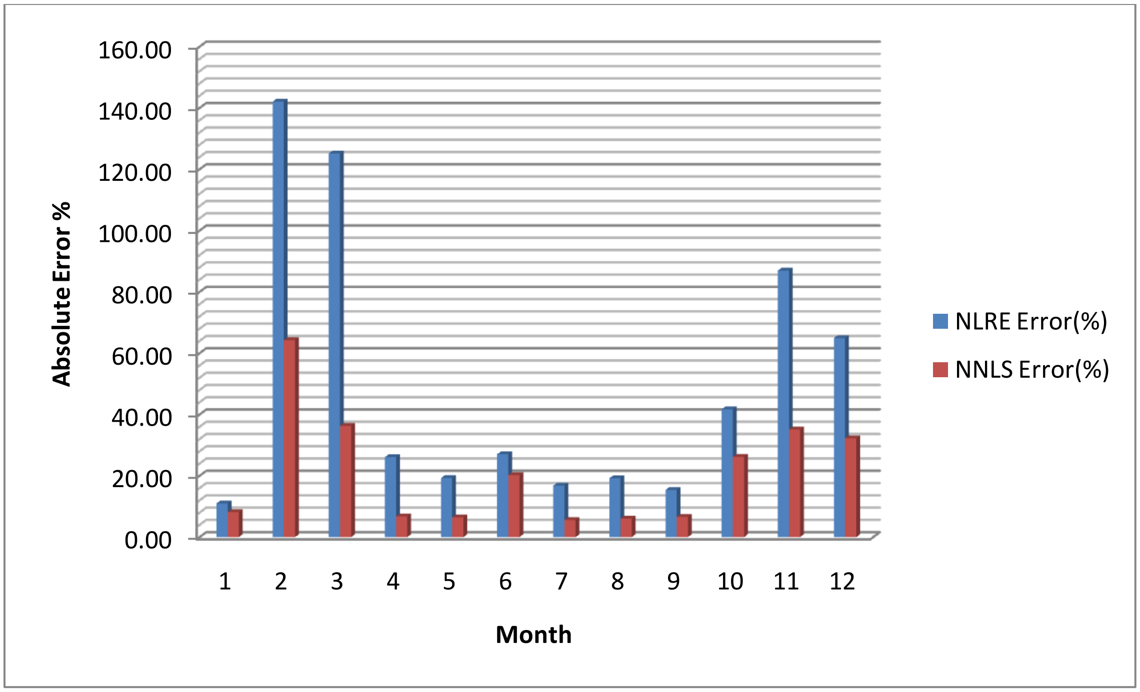

Figure 4-14 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 35-6-13

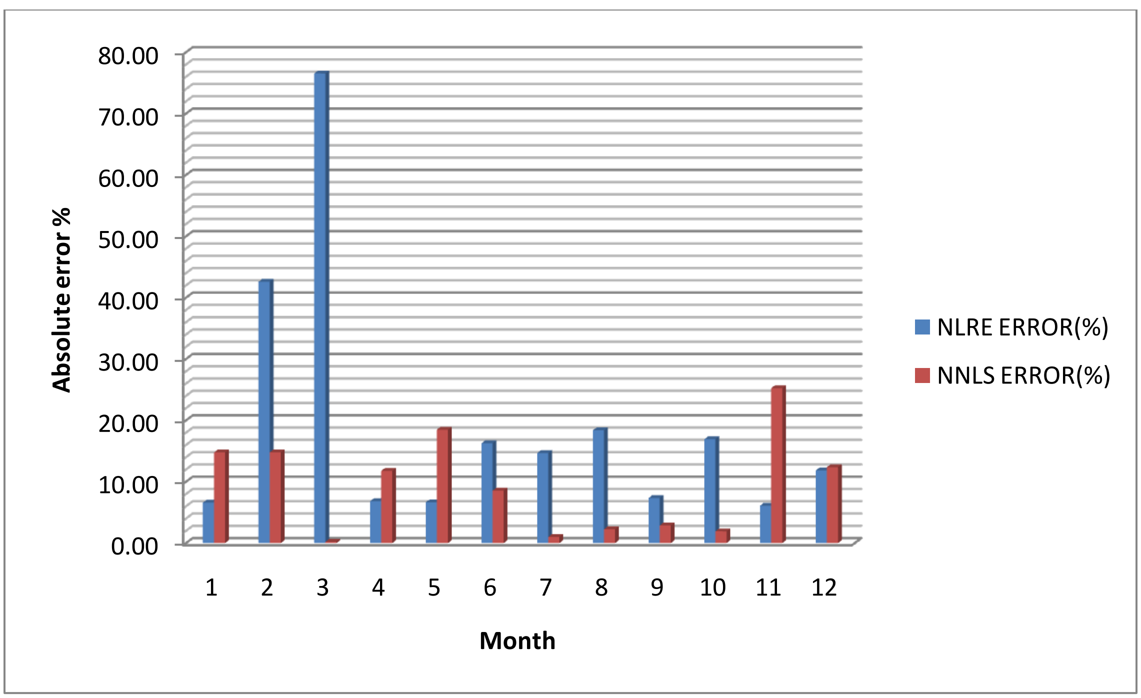

Figure 4-15 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 35-9-13

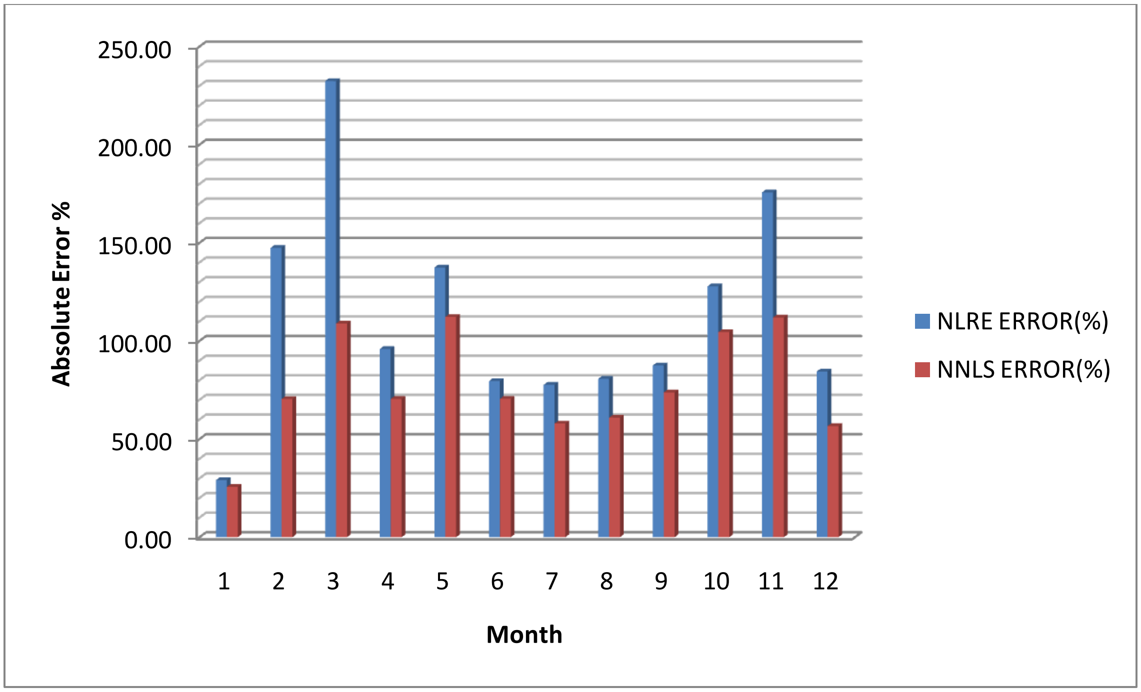

Figure 4-16 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 35-9-13

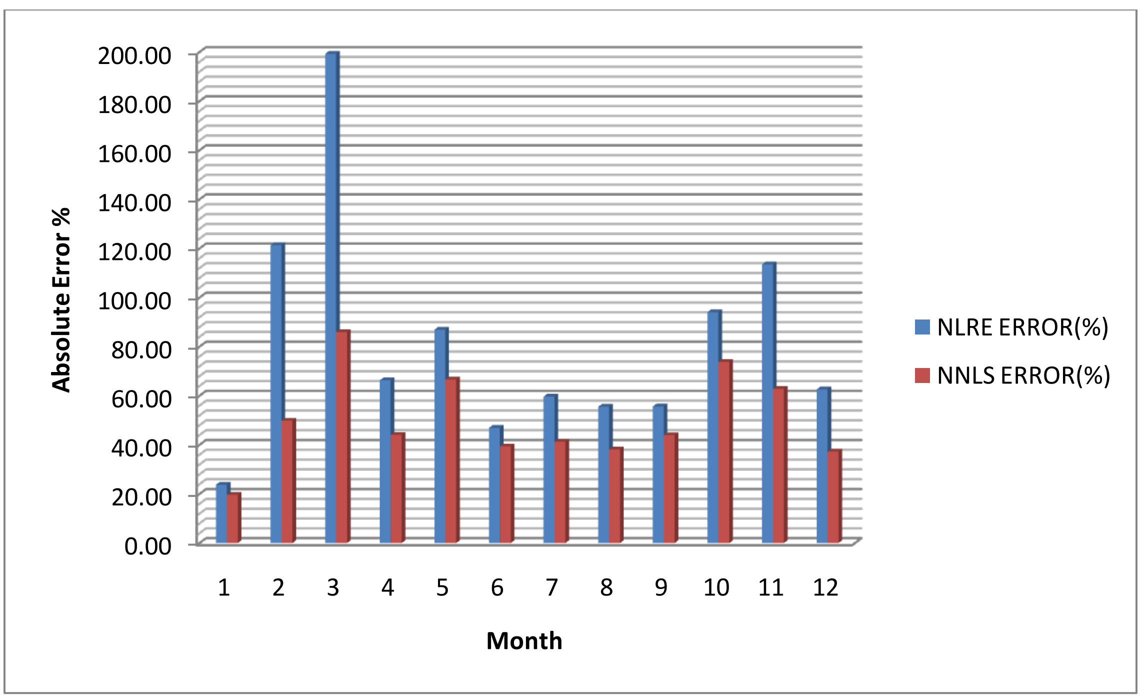

Figure 4-17 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 35-9-13

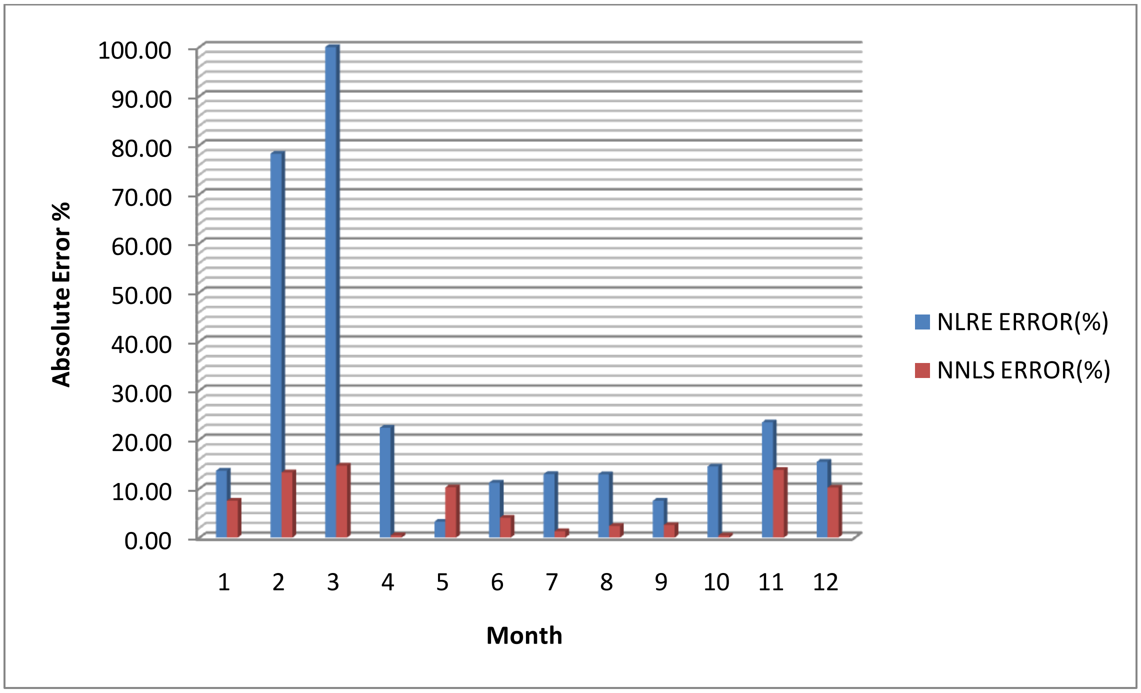

Figure 4-18 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 51-8-13

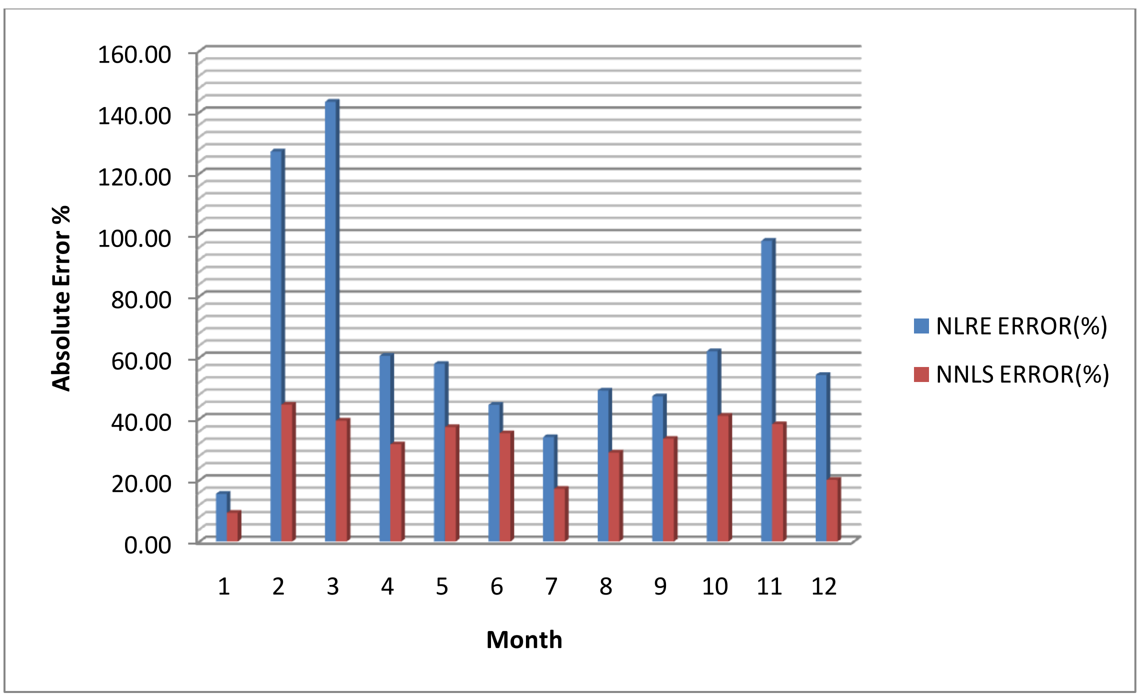

Figure 4-19 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 51-8-13

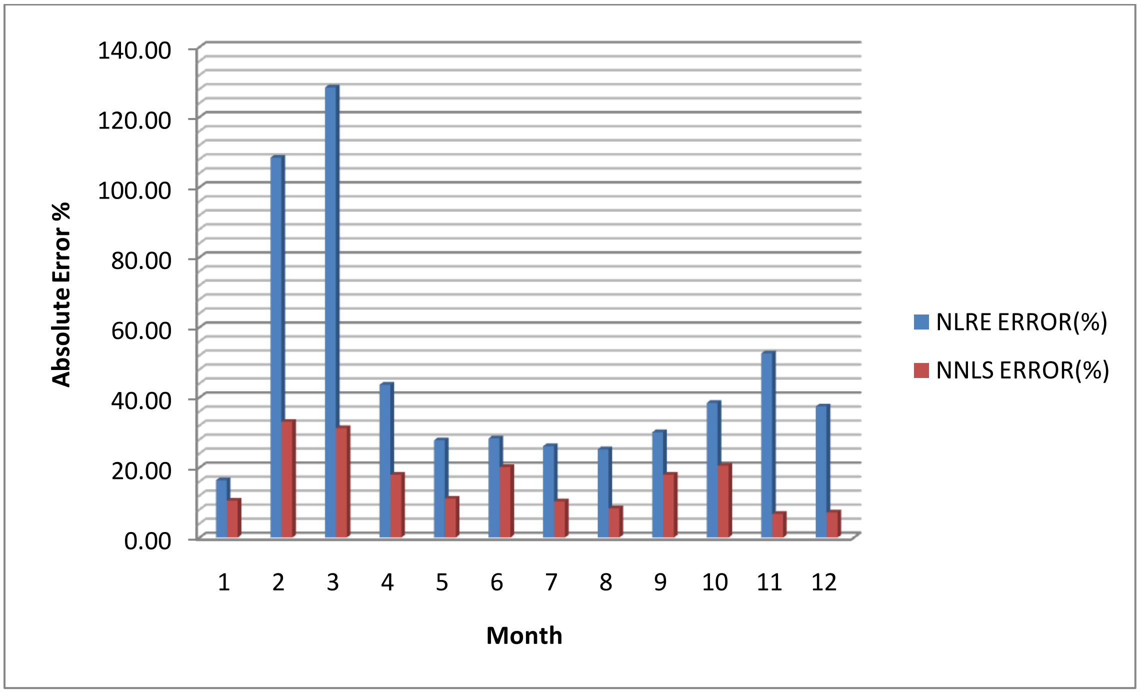

Figure 4-20 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 51-8-13

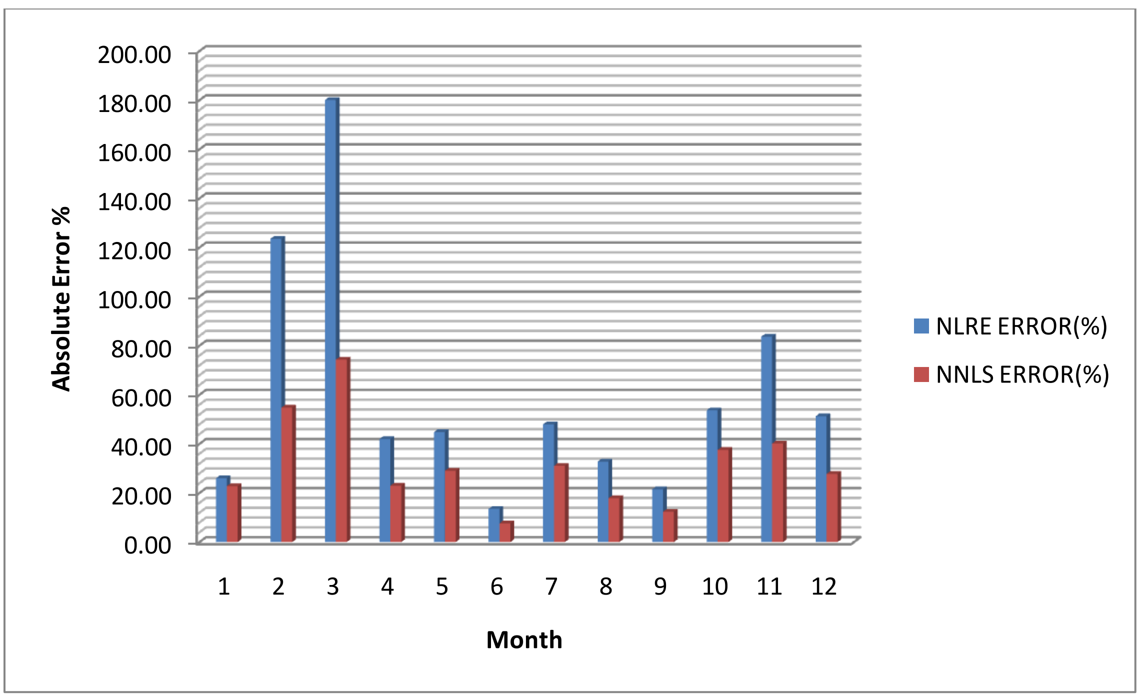

Figure 4-21 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 15-4-13

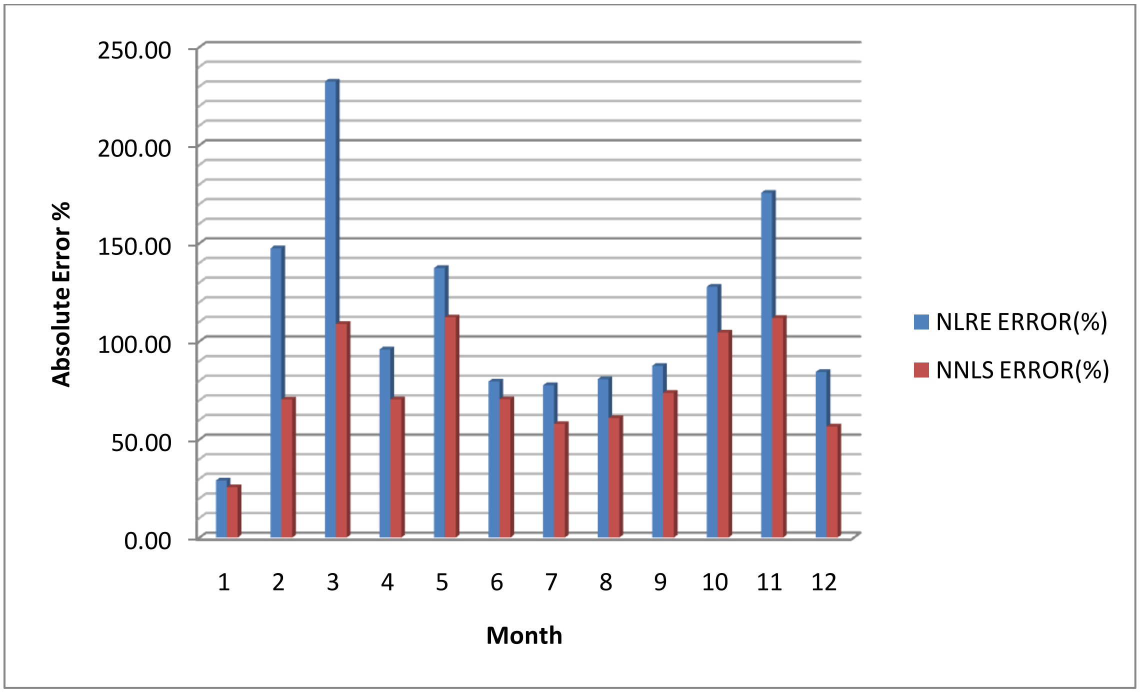

Figure 4-22 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 15-4-13

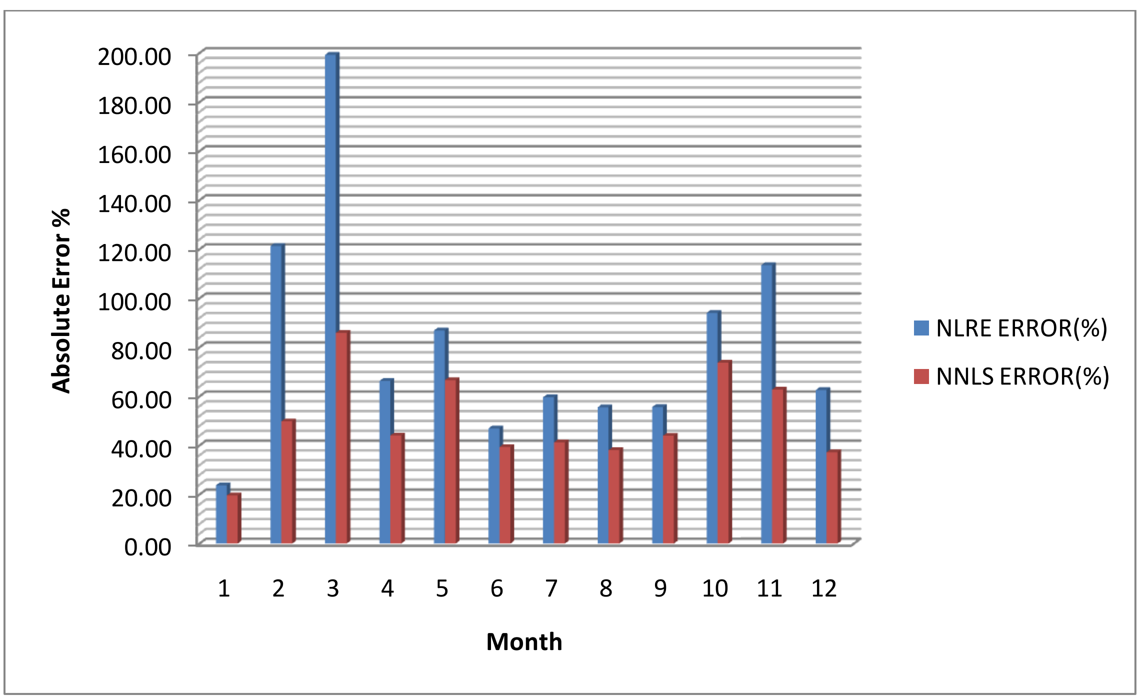

Figure 4-23 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 15-4-13

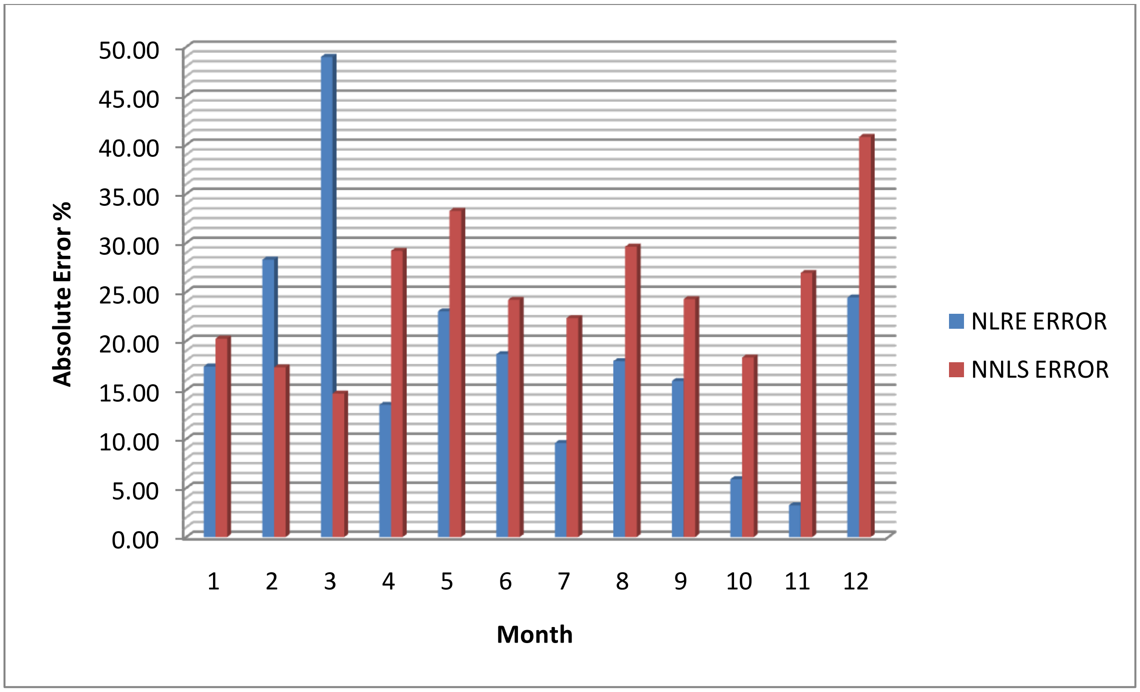

Figure 4-24 Absolute error percentage comparison between NLRE and NNLS for phase A peak currents for feeder 67-2-13

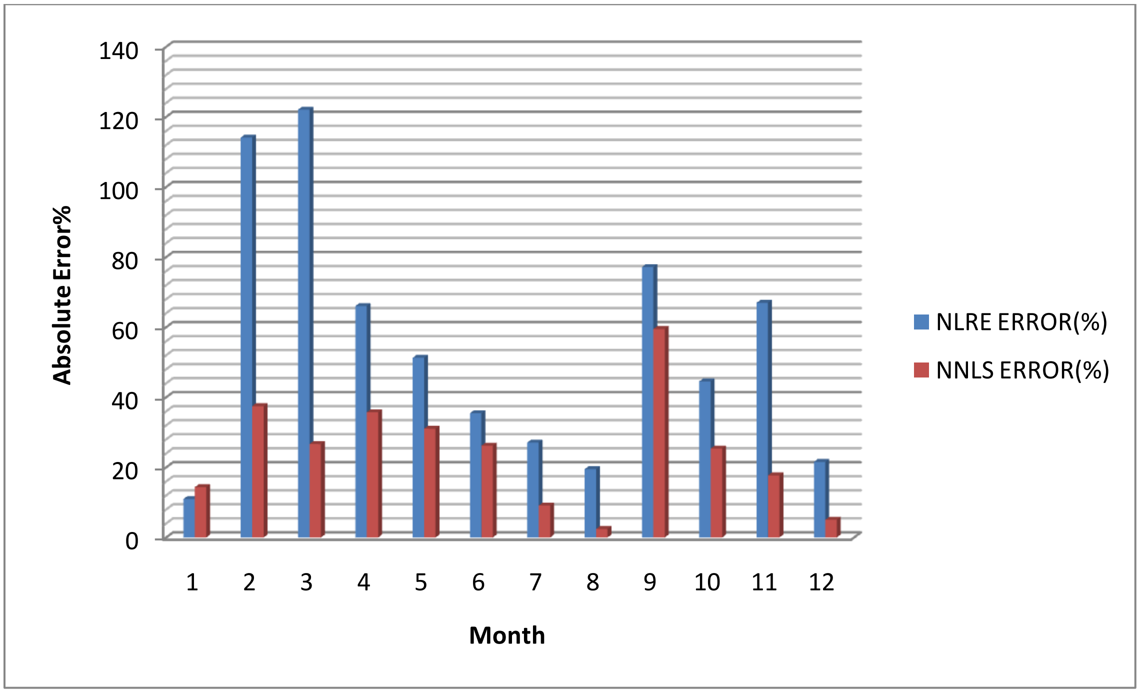

Figure 4-25 Absolute error percentage comparison between NLRE and NNLS for phase B peak currents for feeder 67-2-13

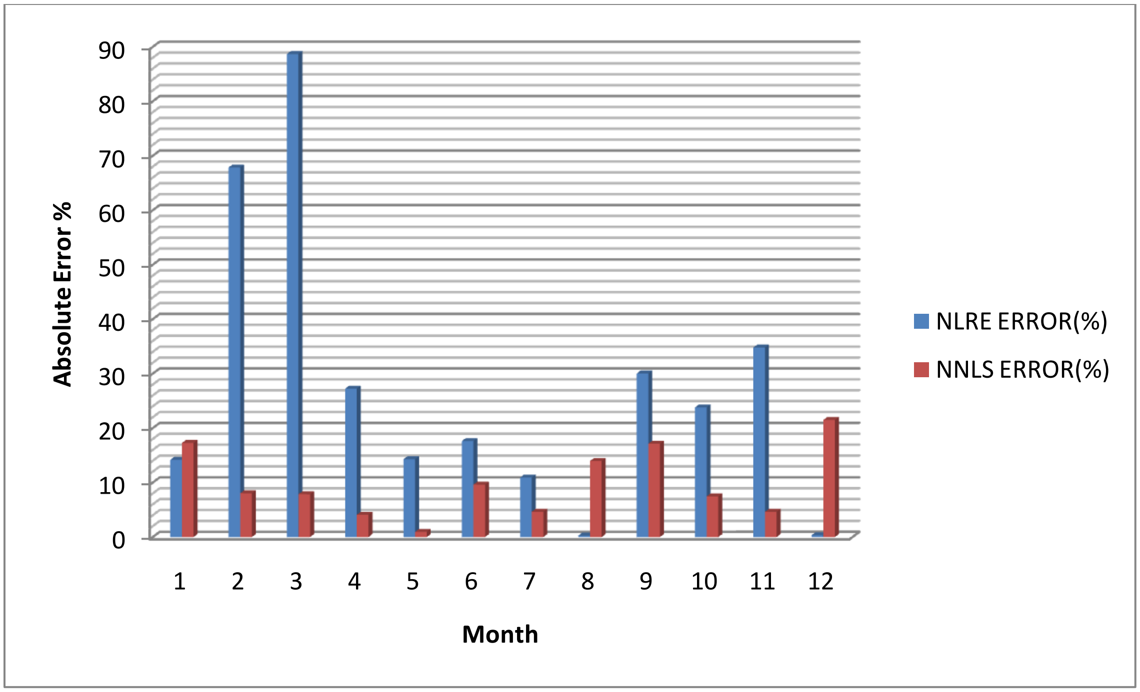

Figure 4-26 Absolute error percentage comparison between NLRE and NNLS for phase C peak currents for feeder 67-2-13

Figures 4-27 to 4-41 provides a comparison for the monthly peak currents for Phase A, Phase B and Phase C currents for the above five feeders. It can be seen that for all the feeders, the NNLS estimates the peaks more accurately than the NLRE.

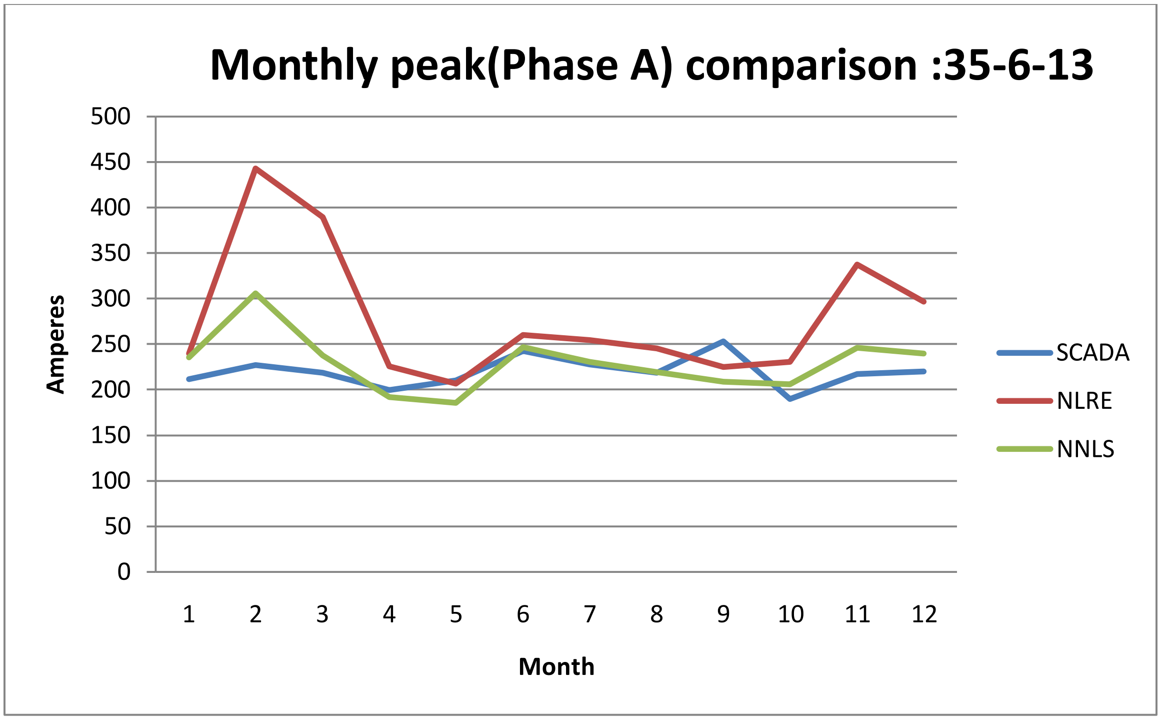

Figure 4-27 Monthly peak comparison for phase A:35-6-13

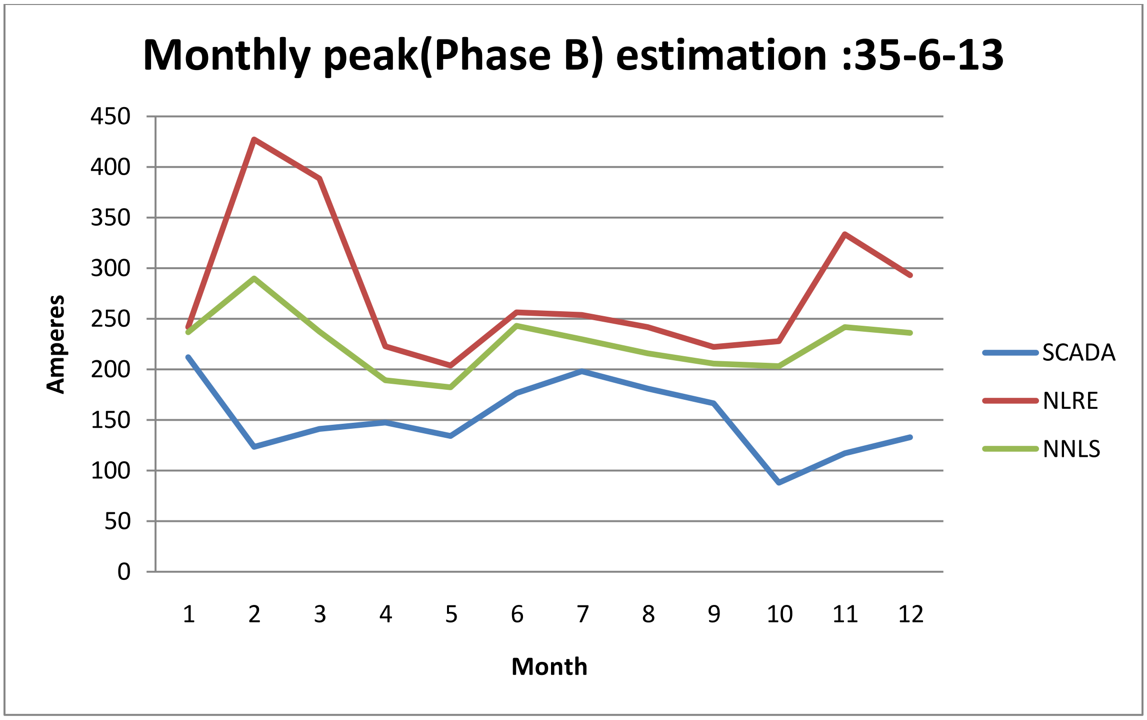

Figure 4-28 Monthly peak comparison for phase B:35-6-13

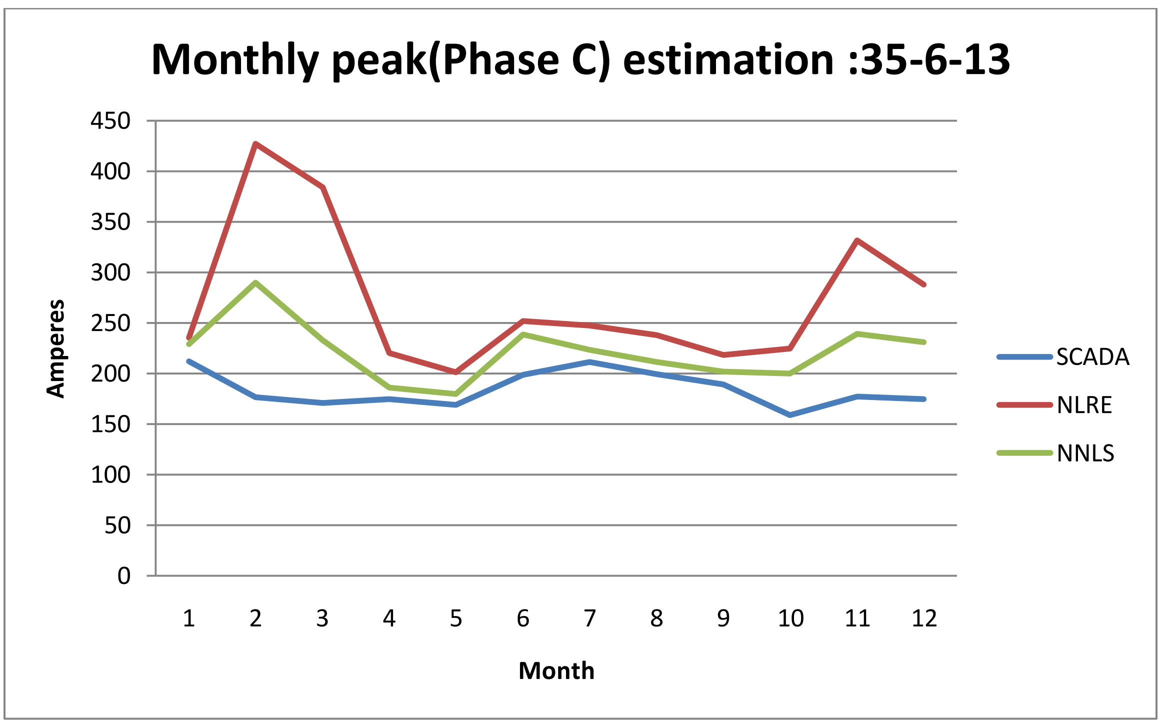

Figure 4-29 Monthly peak comparison for phase C:35-6-13

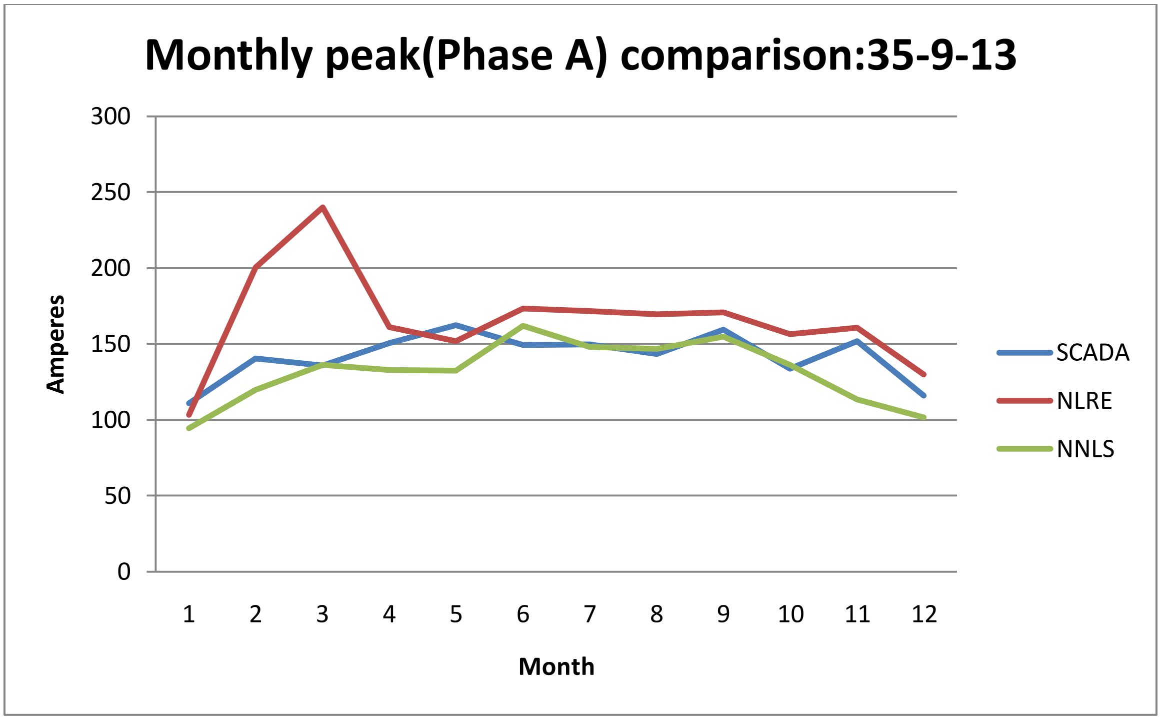

Figure 4-30 Monthly peak comparison for phase A: 35-9-13

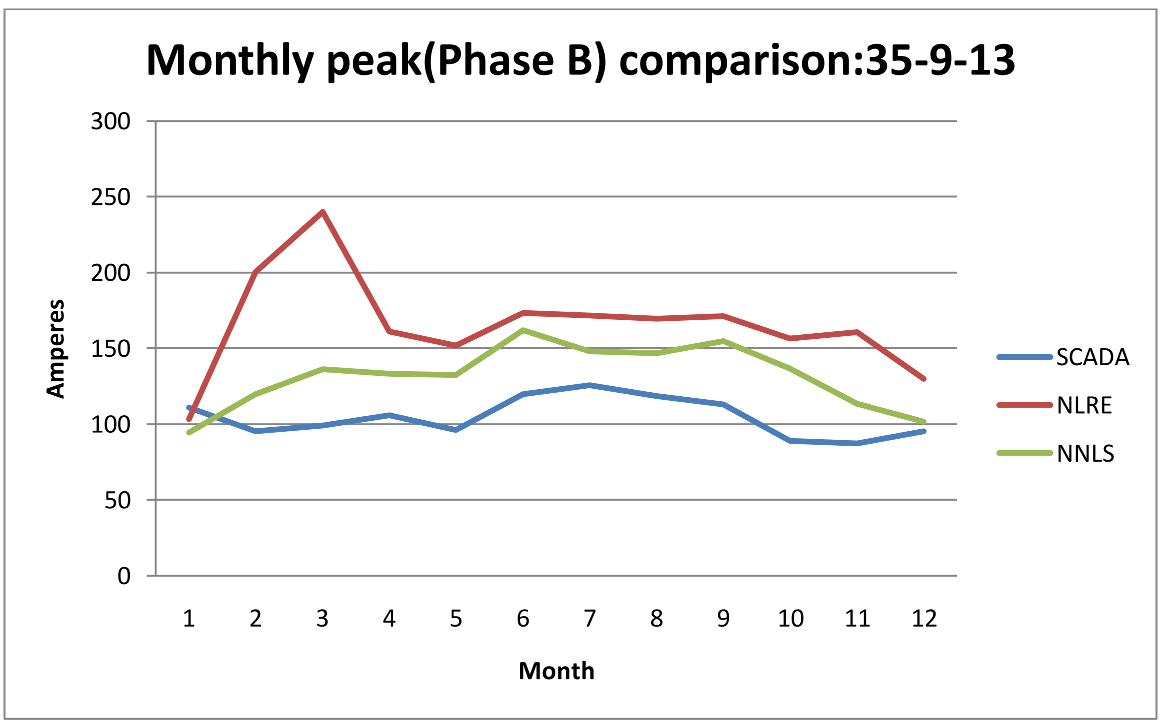

Figure 4-31 Monthly peak comparison for phase B: 35-9-13

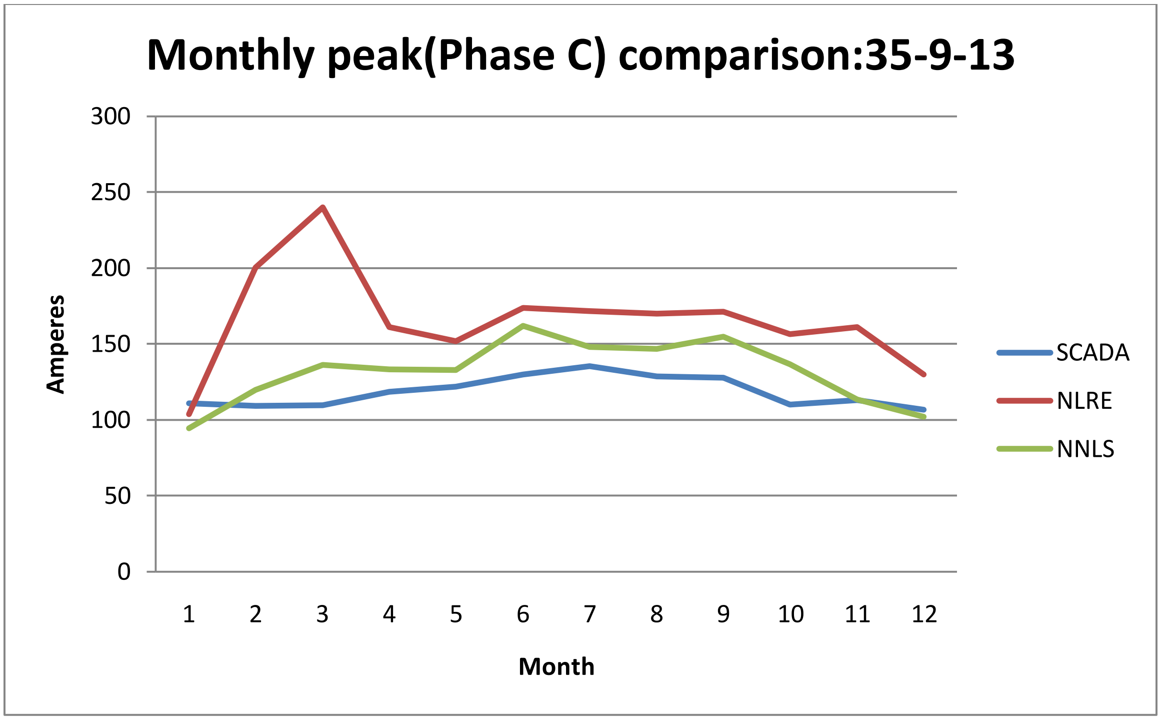

Figure 4-32 Monthly peak comparison for phase C: 35-9-13

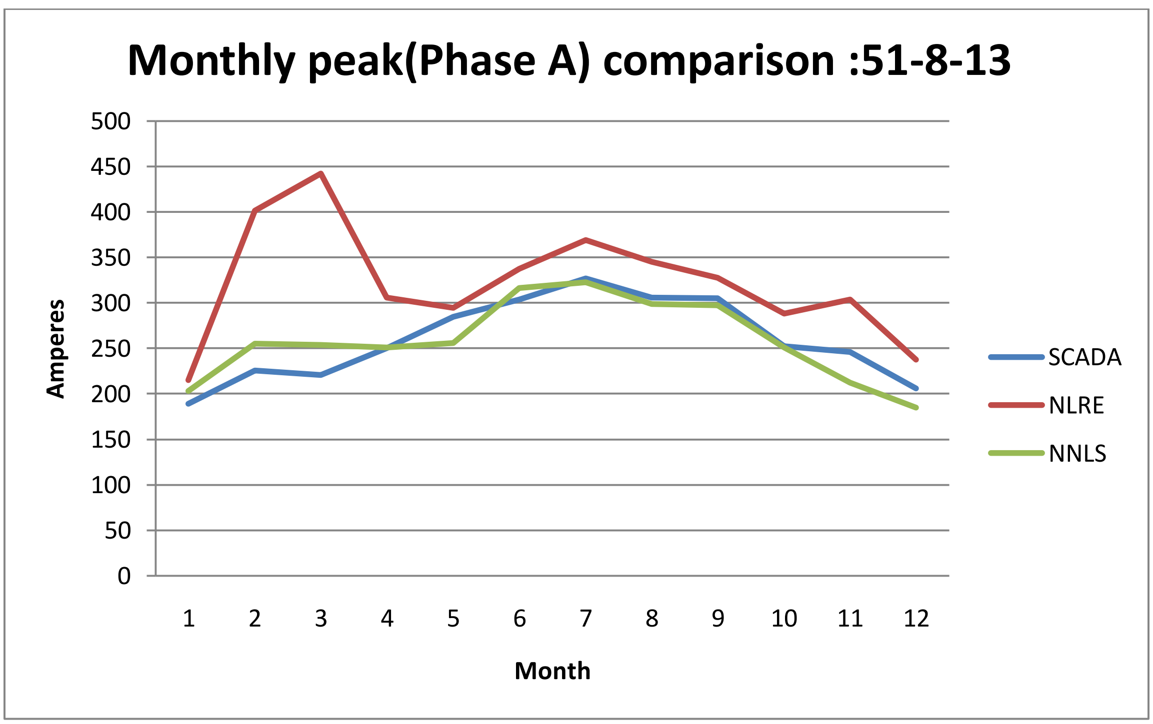

Figure 4-33 Monthly peak comparison for phase A:51-8-13

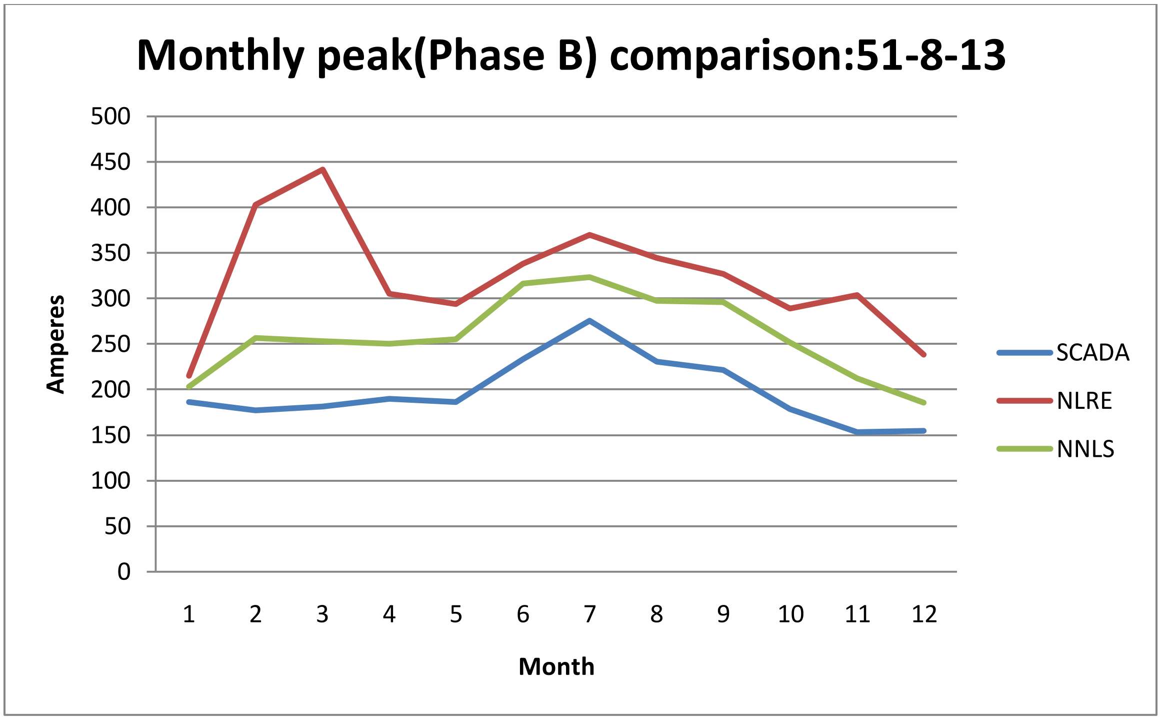

Figure 4-34 Monthly peak comparison for phase B: 51-8-13

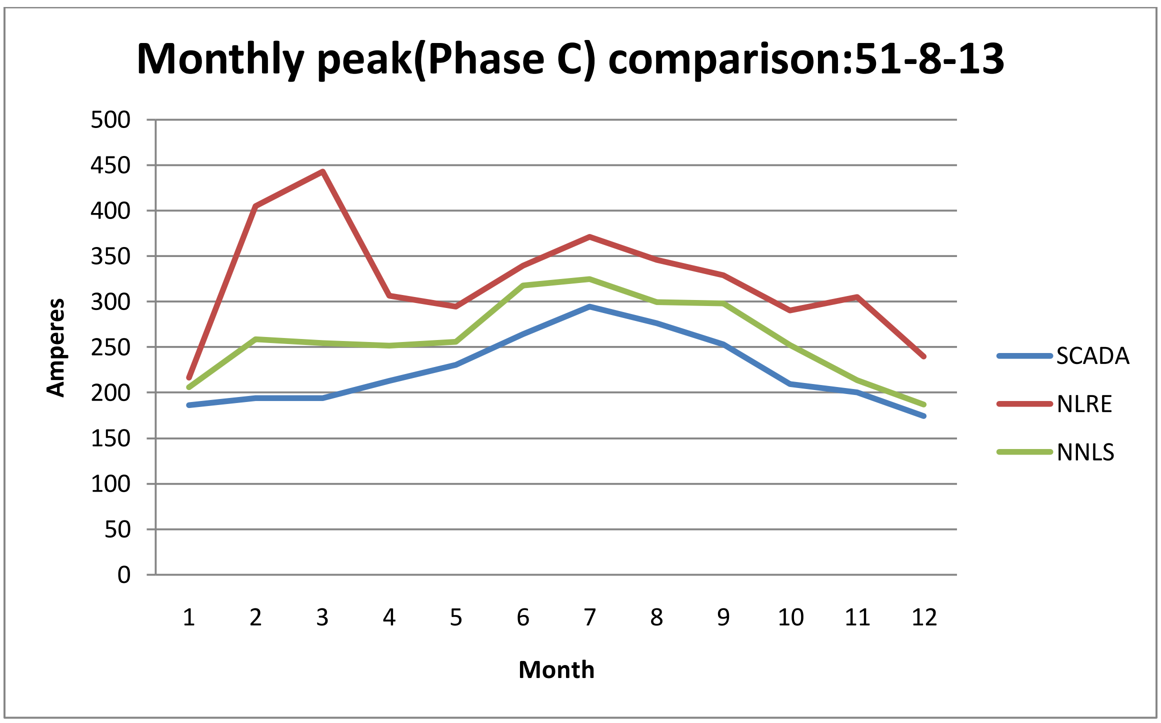

Figure 4-35 Monthly peak comparison for phase C: 51-8-13

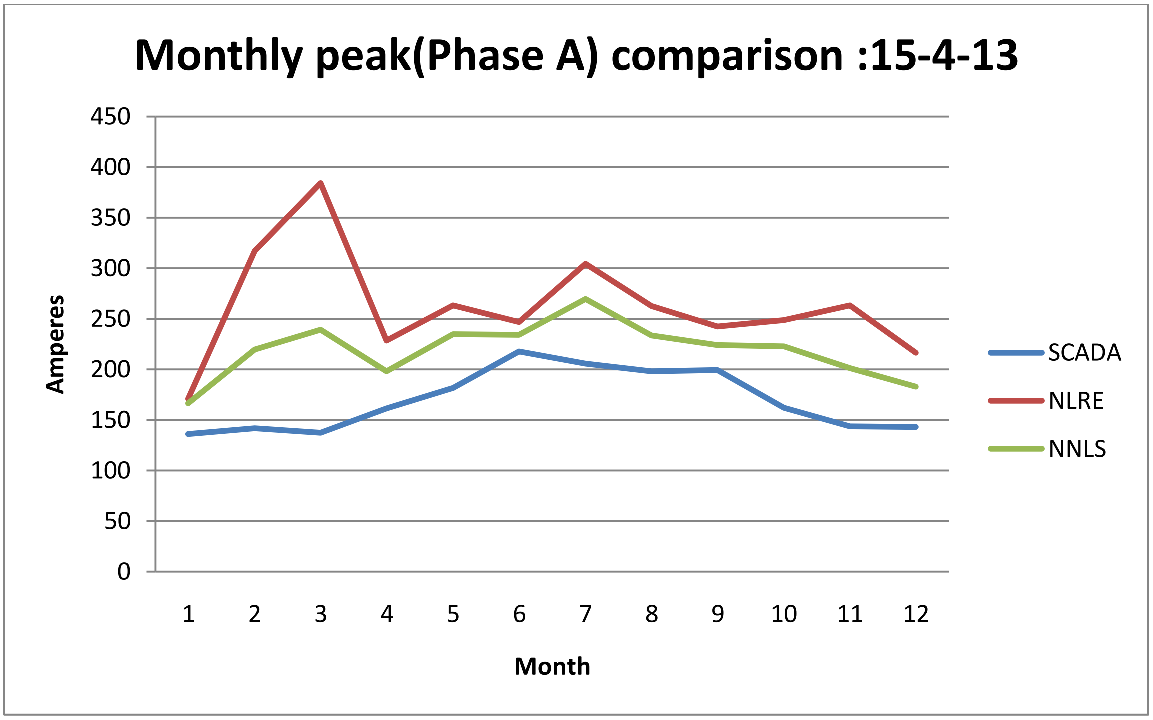

Figure 4-36 Monthly peak comparison for phase A:15-4-13

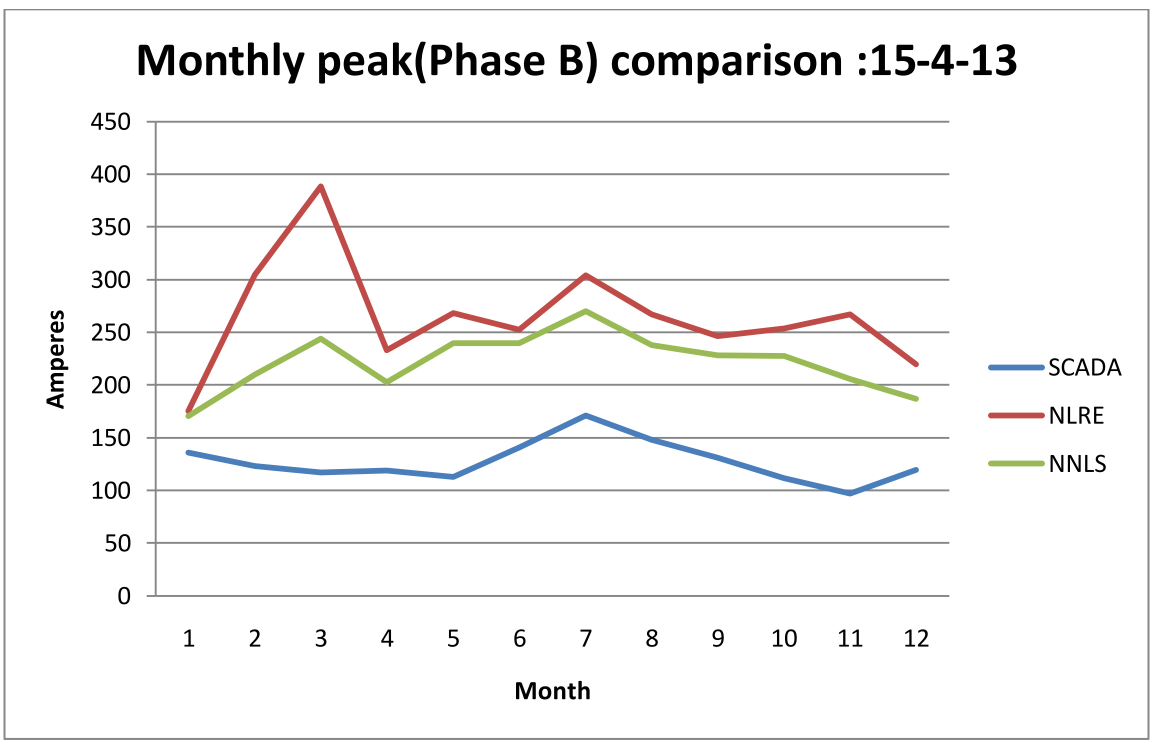

Figure 4-37 Monthly peak comparison for phase B:15-4-13

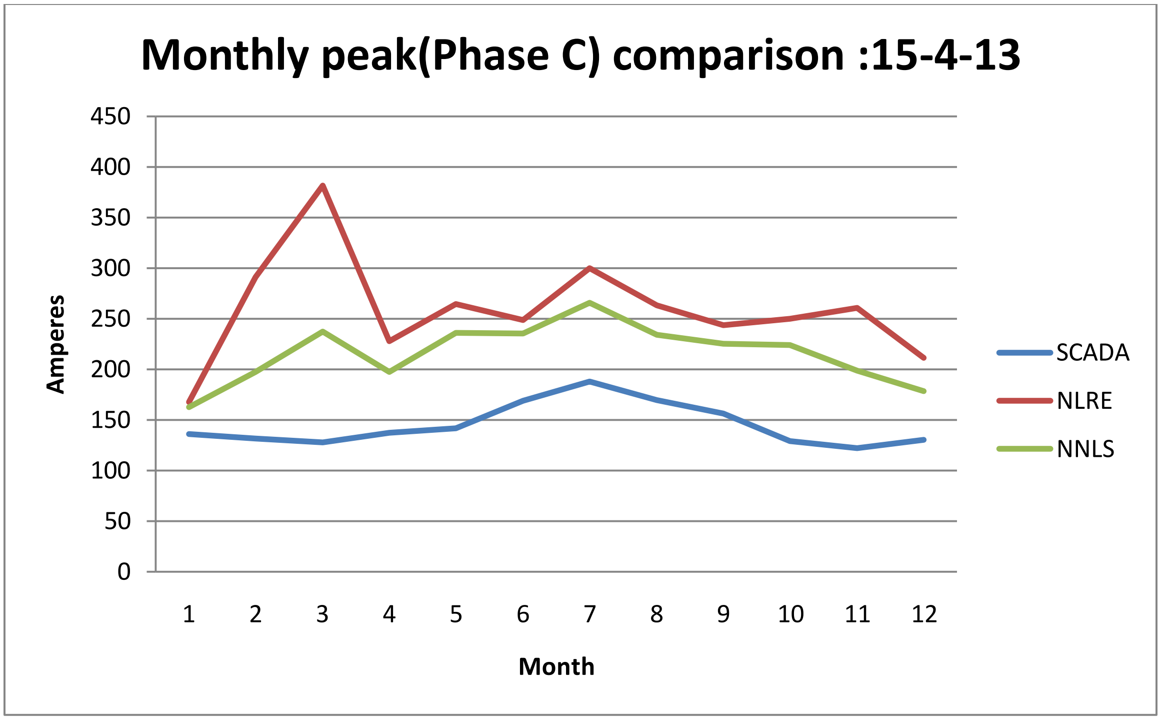

Figure 4-38 Monthly peak comparison for phase C:15-4-13

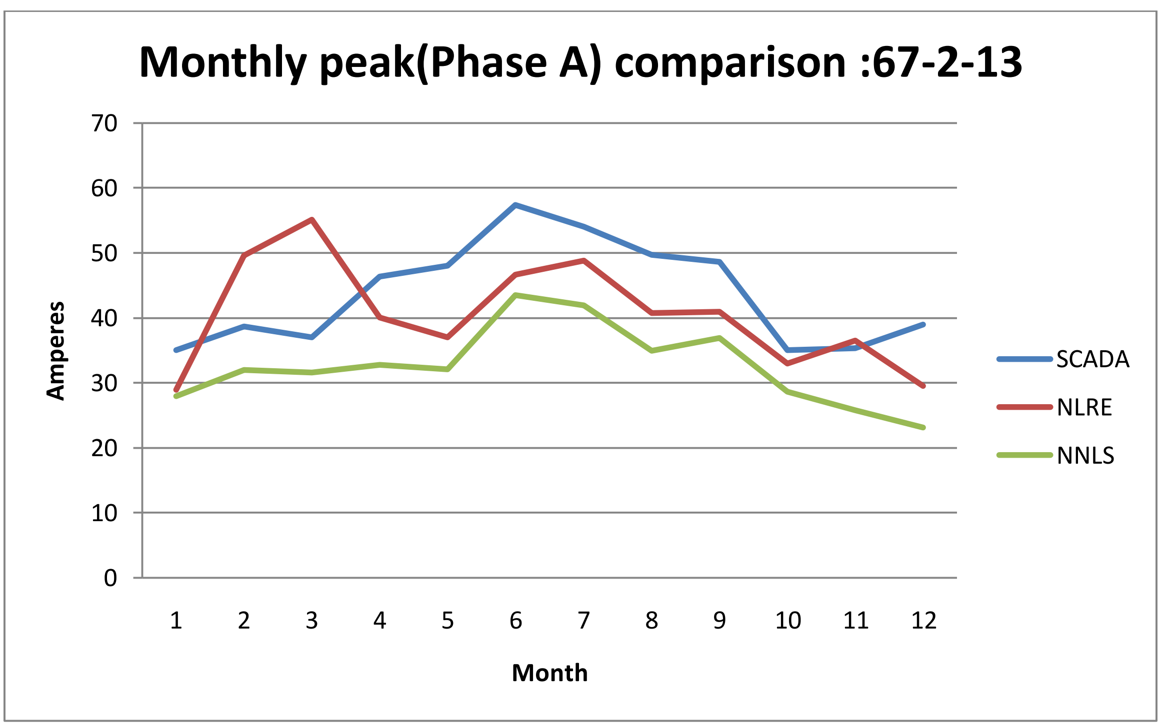

Figure 4-39 Monthly peak comparison for phase A:67-2-13

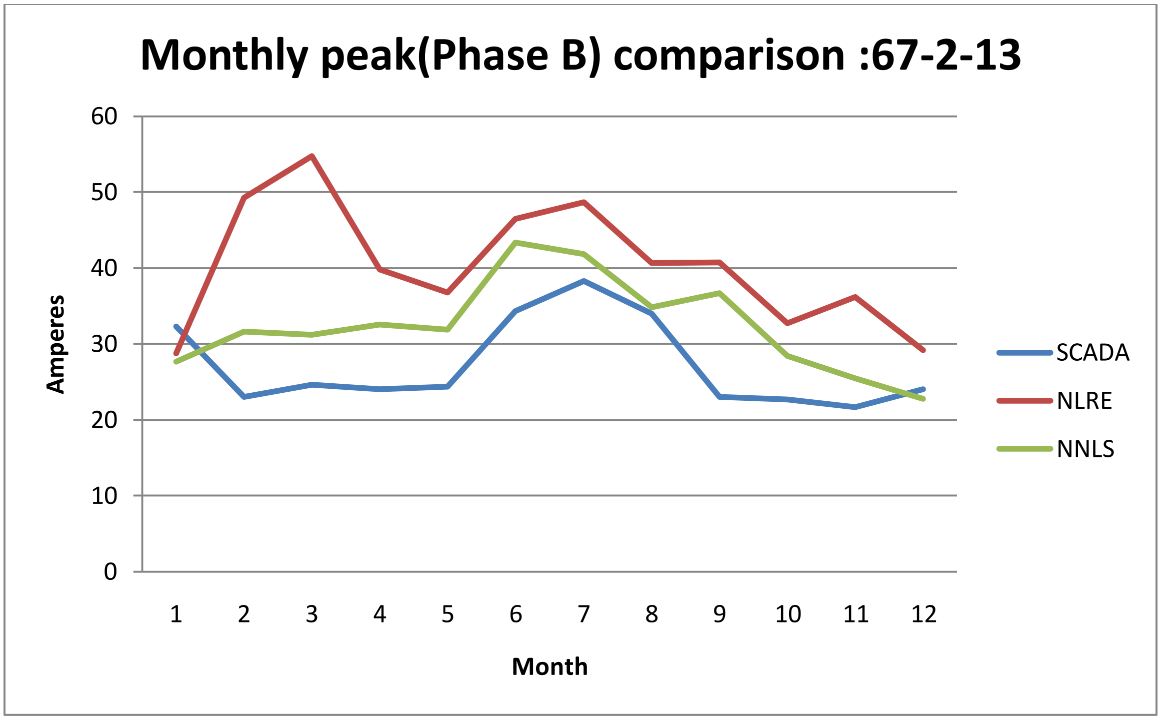

Figure 4-40 Monthly peak comparison for phase B:67-2-13

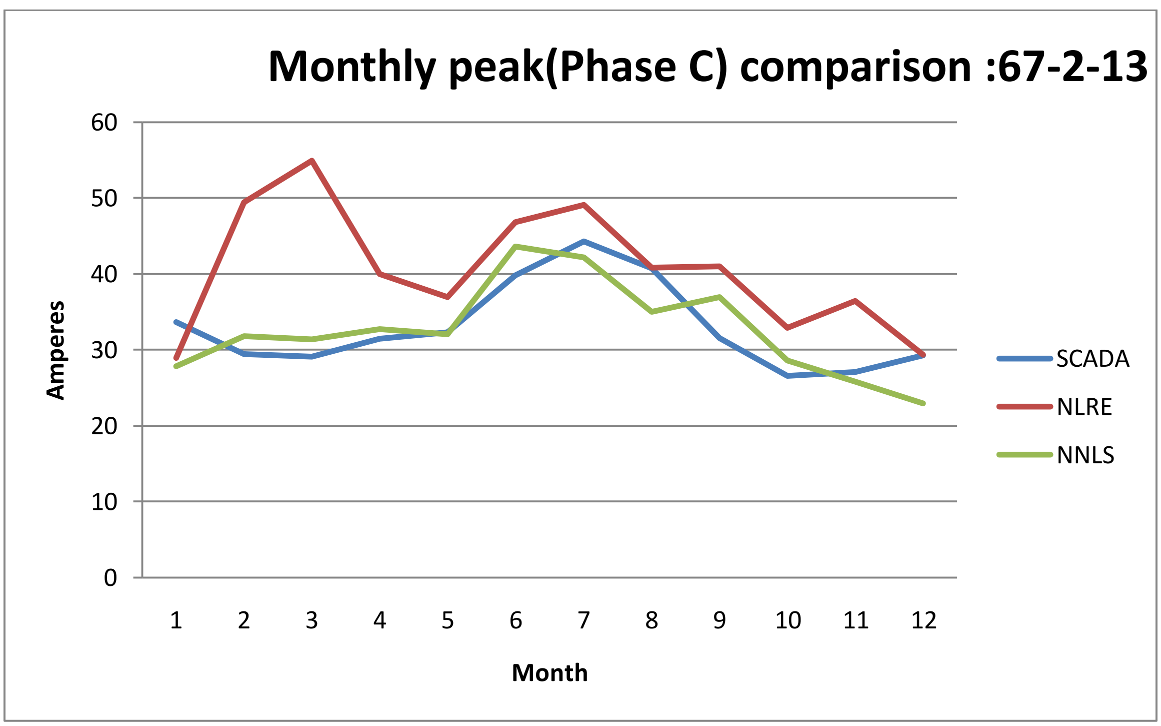

Figure 4-41 Monthly peak comparison for phase C:67-2-13

Chapter 5

Future work and conclusions

5.1 Conclusions

The work presented here is an extension of NLRE [15]. It has been optimized for industrial circuits using the SCADA data and by utilizing the NNLS Optimization technique.

Load estimation has been used over the years by utility companies to plan the distribution system and to choose the equipment ratings. The accuracy with which the load research estimates the peak and loading on the system has an effect on these decisions.

Industrial customers are huge customers and are defined here with annual consumption ranging from above 500,000KWHR. These are the key customers to any utility company as these customers are responsible for the peak loads on the system. This study has been focused on industrial customers.

In this research class based conversion factors, diversity factors and diversity load curves are estimated based on the kWHr data available for the sample customers for a complete year. Customer sample kWHr data and monthly data is imported into the database and sorted according to the class, month and day type. Load research statistics are calculated and applied to non-linear parsed billing data. Currents are estimated on the feeders and are optimized to match the SCADA data using NNLS.

Five industrial dominated feeders are chosen and the estimates are compared with the SCADA values and the error values demonstrate that the NNLS method provides an improvement of 10% to 90%.

5.2 Contributions

The NNLS model to optimize the load research factors using the SCADA measurements has been proposed and demonstrated. The goal was to provide an improved method to estimate the peaks and the load characteristics for industrial circuits. This method can be used to improve the estimates on any circuit for which the monthly billing data and the SCADA data exists. With the peak and loading information available, the planning engineers are equipped with the necessary data to plan the distribution system, generation schedules and future expansion. The prime contributions can be described as follows:

The customer billing data can be used to estimate the peak on the system by applying the load research factors. A model has been presented to minimize the error between the estimates and SCADA data for industrial circuits. The optimization is achieved by applying the NNLS algorithm to estimate the new conversion factors. The optimization is performed for each month, day type and class. The updated conversion factors are applied to find the new estimates and the SCADA currents are compared against the estimated currents and were found to reduce the errors for over estimated systems. Also a comparison between the load research based estimates using the NLRE method and the estimates based on the NNLS method was performed.

This work has shown that by applying NNLS Optimization to minimize the difference between SCADA measurements and load research model based estimates of the SCADA measurements, significant improvements in the load research model based estimates can be achieved.

5.3 Future work

Emphasis of the thesis was to present a model to optimize the estimates at the feeder level using Lawson-Hanson NNLS for industrial customers. Future recommendations are as follows.

The method can be studied on various other classes and can be generalized to apply for all the circuits by studying the weighted NNLS method to optimize estimates for circuits with customers from different classes.

Another area which can be studied as an extension of load research is the effect on peak load by changing the transformer type and/or phase change.

BIBLIOGRAPHY

[1] S V Allera,N D Alcock, A A Cook, “ Load Research in a Privatised Electric Supply Industry,” In *Proceedings of the Sixth International Conference, Metering Apparatus and Tariffs for electricity Supply*, April 1990.

[2] Xiao Jiang Ye, Zhao Xia Zhou, Qing Er, Feng Wu,” Analysis of electric loads of residential buildings during summer in Wuhan” In *Proceedings of Power and Energy Engineering Conference*, pp. 1-4, March 2009.

[3] Robert P. Broadwater, Al Sargent, Abdulrahman Yarali, Hesham E.Shaalan, Joanicjusz Nazarko,”Estimating Substation Peaks from Load Research Data”, *IEEE Transactions on Power Delivery*, Vol. 12, Issue 1, pp. 451-456, Jan 1997.

[4] Nazarko.J, Zalewski.W, “ The fuzzy regression approach to peak load estimation in power distribution systems”, *IEEE Transactions on Power Delivery*, Vol. 14, Issue 3, pp. 814-809, Aug 1999.

[5] Sargent,A., Broadwater,R.P., Thompson, J.C., Nazarko, J., “ Estimation of diversity and kWHR-to-peak-kW factors from load research data”, *IEEE Transactions on Power Systems*, Vol. 9, Issue 3, pp. 1450-1456, Aug 1994.

[6] Cartina, G., Alexandrescu, V., Grigoras, G., Moshe, M., “Peak Load Estimation in Distribution Networks by Fuzzy Regression Approach”, In *Proceedings of Electrotechnical Conference,* MELECON 2000, Vol. 3, pp. 907-910, May 2000.

[7] Irwin, G. W., Monteith, W., Beattie, W. C., “Statistical Electricity Demand Modeling from Consumer Billing Data”, *IEE Proceedings*, Vol. 133, September 1986.

[8] Seppala, A., “Statistical Distribution of Customer Load Profiles”, In *Proceedings of International Conference on Energy management and Power Delivery*, Vol. 2, pp. 696-701, November 1995.

[9] R-Project for Statistical Computing, CRAN.

[10] Fawwaz Elkarmi, ”Load Research as a tool in Electric Power System Planning, Operation and Control – The Case of Jordan”, *Energy Policy, Elseiver Ltd.*, 2008.

[11] AEIC’s Load Research Committee, “The Role of Load Research in Automated Meter Infrastructure/Meter Data Management Initiatives”, In *Proceedings of AEIC*, September 2008.

[12] AEIC’s Load Research Committee, “Demand Response Measurement and Verification”, In *Proceedings of AEIC,* March 2009.

[13] Public Utility Regulatory Policies Act of 1978. Public Law 95-617.

[14] Nazarko, J., Styczynski, Z.A., “Application of Statistical and Neural Approaches to the Daily Load Profiles Modeling in Power Distribution Systems”, In *Proceedings of the IEEE Transmission and Distribution Conference,* Vol. 1, pp. 320-325, April 1999.

[15] Yarali, A., “Peak Load Estimation using Load Research Data”, Ph.D. Thesis. Virginia Polytechnic Institute and State University, 1995.

[16] Electric Utility Engineering Reference Book - Distribution Systems, East Pittsburgh, Westinghouse Electric Corporation, Vol. 3, 1965.

[17] Chen, C.S., Hwang, J.C.; Tzeng, Y.M., Huang, C.W., Cho, M.Y., “Determination of Customer Load Characteristics by Load Survey at Taipower”, ”, *IEEE Transactions on Power Delivery,* Vol. 11, Issue 3, pp. 1430-1436, July 1996.

[18] American Standard Definitions of Electric Terms, American institute of Electrical Engineers, 1942.

[19] Baran, M.E., Freeman, L.A.A., Hanson, F., Ayers, V., "Load Estimation for Load Monitoring at Distribution Substations", *IEEE Transactions on Power Systems,* Vol. 20, Issue 1, pp. 164-170, February 2005.

[20] Liu, K., Subbarayan, S., Shoults, R.R., Manry, M.T., Kwan, C., Lewis, F.I., Naccarino, J., "Comparison of very Short-term Load Forecasting Techniques", *IEEE Transactions on Power Systems,* Vol. 11, Issue 2, pp. 877-882, May 1996.

[21] Chang, R. F., Leou, C., Lu, C. N., "Distribution Transformer Load Modeling using Load Research Data", *IEEE Power Engineering Review,* Vol.22, Issue 3, pp. 62-62, March 2002.

[22] Jianzhong Wu, Yixin Yu, "CBR-based Load Estimation for Distribution Networks", In *Proceedings of Electrotechnical Conference,* MELECON 2006, pp. 952-955, May 2006.

[23] Baran, M., McDermott, T.E., "State Estimation for Real time Monitoring of Distribution Feeders", In *Proceedings of IEEE Power and Energy Society General Meeting*, pp. 1-4, July 2009.

[24] Jie Wan, Nan Miu, K., "WLS Method for Load Estimation in Unbalanced Distribution Networks", In *Proceedings of IEEE Power Engineering Society Winter Meeting*, Vol. 2, pp. 969-974, January 2002.

[25] Jie Wan, Miu, K.N., "Weighted Least Square Methods for Load Estimation in Distribution Networks", In *Proceedings of IEEE Power Systems*, Vol. 18,Issue 4, pp. 1338-1345, November 2003.

[26] Hagan, Martin T., Behr, Suzanne M., "Time series approach to short term load forecasting", *IEEE Transactions on Power Systems*, Vol. 2,Issue 3, pp. 785-791, August 1987.

[27] Grady, W.M., Groce, L.A., Huebner, T.M., Lu, Q.C., Crawford, M.M., "Enhancement, Implementation, and Performance of an Adaptive Short-Term Load Forecasting Algorithm" *IEEE Transactions on Power Systems*, Vol. 6, Issue 4, pp. 1401-1410, November 1991.

[28] Jian-chang Lu, Xingping Zhang, Wei Sun, " A Real-time Adaptive Forecasting Algorithm for Electric Power Load", In *Proceedings of Transmission and Distribution Conference and Exhibition*, pp. 1-5, 2005.

[29] Senjyu, T., Takara, H., Uezato, K., Funabashi, T., "One-hour-ahead Load Forecasting using Neural Network", *IEEE Transactions on Power Systems*, Vol. 17, Issue 1, pp. 113-118, February 2002.

[30] B.M.Sander, "Forecasting Residential Energy Demand: A Key to Distribution Planning", In *Proceedings of the IEEE PES Summer Meeting*, Mexico City, Mex., April 1977, A 77 642-2.

[31] Parlos, A.G., Oufi, E., Muthusami, J., Patton, A.D., Atiya, A.F., "Development of an Intelligent Long-term Electric Load Forecasting System", In *Proceedings of the International Conference on Intelligent Systems Applications to Power Systems*, pp. 288-292, February 1996.

[32] Wenjin Dai, Ping Wang, "Application of Pattern Recognition and Artificial Neural Network to Load Forecasting in Electric Power System", In *Proceedings of the*

Third International Conference on Natural Computing, Vol. 1, pp. 381-385, August 2007.

[33] Villacis, E., Grossman, R., McElroy, R., Musa, D., "An Econometric Approach for Forecasting Power Peak Demands", In *Proceedings of the American Power Conference, 1981.*

[34] Liu, X.Q., Ang, B.W., Goh, T.N., "Forecasting of Electricity Consumption: A Comparison between an Econometric Model and a Neural Network Model", In Proceedings of the IEEE International Joint Conference on Neural Networks, Vol. 2, pp. 1254-1259, November 1991.

[35] Shaprio, S., and Gross, A.J., "Statistical modeling techniques", Marcel Dekker, New York and Basel, 1981.

[36] Lawson, C.L., "Solving Least Squares Problems" , Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

[37] Shaalan, H., Thompson, J., Broadwater, R., Ellis, M., Ng, H., "Distribution Engineering Tool Features a Flexible Framework", *IEEE Journal, Computer Applications in Power*, Vol. 8, Issue 3, pp. 21-24, July 1995.

[38] Shaalan, H.E., Thompson, J., Broadwater, R.P., Ng, H., "DEWorkstation: A Comprehensive Distribution Modeling Tool", In *Proceedings of Transmission and Distribution Conference and Exhibition*, pp. 7-12, September 1996.

[39] Broadwater, R., Thompson, J., Ellis, M., Ng, H., Singh, N., Loyd, D., "Application Programmer Interface for the EPRI Distribution Engineering Workstation", *IEEE Transactions on Power Systems,* Vol. 10, Issue 1, pp. 499-505, February 1995.