Scrutiny of Electricity Data for Consumption and Load Forecasting

A THESIS

Submitted to the School of Engineering and Computer Science,
BRAC University, in partial fulfilment of the requirements for the degree of
Master of Science in Computer Science and Engineering

By

Samiul Islam

CERTIFICATE OF APPROVAL

The thesis titled "Scrutiny of Electricity Data for Consumption and Load Forecasting" is completed under my supervision, meets acceptable presentation standard and can be submitted in partial fulfilment of the requirement for the degree of Master of Science in Computer Science and Engineering from the Department of Computer Science and Engineering, BRAC University.

Dr. Moinul Islam Zaber
Associate Professor
Department of Computer Science and Engineering
University of Dhaka

DECLARATION

I hereby certify that this material, which I now submit for assessment on the program Master of Science in Computer Science and Engineering is entirely my own work (consumption forecasting modeling, demand forecasting survey and load forecasting modeling part), that I have exercised reasonable care to ensure that the work is original, does not to the best of my knowledge breach any law of copyright and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work. This report also showcasing an extended theoretical and technical review of several previous works around the globe of the same field and all of these are cited and referred appropriately. All the permission and authorization for data collection have been administered by Data and Design Lab, University of Dhaka.

Samiul Islam

ID number: 15166006

Committee Members

Dr Amitabha Chakrabarty, Committee Chair

Dr Jia Uddin, Committee Member

Dr Amin Ahsan Ali, Committee Member

Chairperson
Department of Computer Science and Engineering
BRAC University, Dhaka, Bangladesh

ACKNOWLEDGEMENTS

All praises are due to the merciful Almighty Allah, the most exalted, the most Beneficent, the most Merciful who blessed me to be here at BRAC University for pursuing the master degree.

My journey towards the Master degree would not have been possible without the help of many people. It is my great pleasure to take this opportunity to thank them for the support and advice that I received.

My deepest gratitude should go to my Supervisor, Dr Moinul Islam Zaber, PhD, Department of Computer Science and Engineering. His comprehensive and constructive comments, his significant support throughout this work gave me strength during the work. I attribute the level of my Master's degree to his encouragement and effort and without him, this thesis, too, would not have been completed. One simply could not wish for a better or friendlier supervisor. I would like to thank each and every member of Data and Design Lab, the University of Dhaka for assisting me throughout this research providing extra ordinary effort in data collection, digitisation, data organisation, statistical analysis, ensuring access to the state of the art tools etc.

I would like to thank my parents for inspiring me in every step of my life, supporting me spiritually and providing me with the best education, my whole family who gave me the strength to complete this degree and my wife for her unconditional support and trust on me. This thesis is dedicated to them.

To my parents, sisters and wife

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

List of Abbreviations

MPEMR: Ministry of Power, Energy and Mineral Resources

PD: Power Division

DESCO: Dhaka Electric Supply Company Limited

DPDC: Dhaka Power Distribution Company LTD

PGCB: Power Grid Company of Bangladesh

BBS: Bangladesh Bureau of Statistics

SG: Smart Grid

NN: Neural Network

MAPE: Mean Absolute Percentage Error

MAE: Mean Absolute Error

List of Publications

1. **S. Islam**, A. Ahamed, A. Ali, M. Zaber. A smart grid prerequisite: survey on electricity demand forecasting models and scope analysis of demand forecasting in Bangladesh. 6th International Conference on Informatics, Electronics & Vision (ICIEV), & AMEC 7th International Symposium on Medical Engineering. 1-3 September 2017, Himeji, Huyogo, Japan. **(Accepted)**

2. N. Rakib, **S. Islam**, F. Alam, D. Chaki, M. Zaber, A. Ali. Can deeper analysis of electricity consumption be a proxy of household economic health?: A case study of Bangladesh. 6th International Conference on Informatics, Electronics & Vision (ICIEV), & AMEC 7th International Symposium on Medical Engineering. 1-3 September 2017, Himeji, Huyogo, Japan. **(Accepted)**

ABSTRACT

In this research, electricity data of Dhaka city, the capital city of Bangladesh has been analysed to use the insights for social good and betterment of electricity sectors. Bangladesh has a very complex electricity infrastructure for both generation and supply sector. According to Power system master plan, Bangladesh mainly produces electricity from gas mine and supply to the grid line. From the gridline, electricity supplies to households. Under the jurisdiction of the Ministry of Power, Energy and Mineral Resources (MPEMR), the Power Division (PD) oversees the whole electricity utility. There are two parts in Dhaka in the historical evolution: old Dhaka and new Dhaka. The responsible department for supplying electricity in these areas are DESCO and DPDC. DESCO is mostly responsible for new Dhaka and extended urban area of Dhaka. The data we have collected from DESCO consist of billing (monthly consumption) data, supply (hourly load) data and load shedding data. Monthly consumption data spans from 1995 to till date, hourly load data and load shedding data span from 2015 to 2016. The objectives of this research can be classified into two parts: one is to analyse monthly consumption/billing data and propose a consumption forecasting model which will predict the consumption in user level. Second, analyzing load/supply data (along with load-shedding data) to understand how legacy method works, addressing key points to ensure a better forecasting, how forecasting will help in future, a brief study of recent forecasting techniques, load shedding scenario, area specific impacts and proposing a forecasting technique which ensures granularity and relatively higher accuracy. It has been found that electricity consumption varies a lot for different tariff bracket consumers in a zone. Consumers from same tariff bracket act differently in different zones. Moreover, electricity usage is strongly correlated with temperature, seasonal change and an occasional change. If temperature increases, electricity usage also increases, usage changes a lot due to particular events. So, forecasting the demand (consumer and substation level) is a crucial part. A proper flow of information is a must from consumer level to the generation plants to predict future demand with minimal error.

CHAPTER 1

1.1 Introduction

Electricity is a prerequisite for economic prosperity. If the current global energy consumption pattern continues, by 2030 overall consumption will be increased by 50% [1]. These energy transitions are more visible in a country experiencing an economic shift. A country while in its developing phase, should ensure efficient use of its resources, proper management, prioritisation of needs, forecast the demand etc. In a developing country, the industrial sector generally consumes 45% to 50% of the total commercial energy [2]. Electricity as one of the fuels of the development thus needs to be managed in an organised fashion. Policy makers need to come up with a better plan than they currently own where statisticians, researchers, data scientists, practitioners can come into play.

Bangladesh, as one of the developing countries in the world, is blooming in economic development. The International Monetary Fund (IMF) has published a report that indicates, Bangladesh is now ranked among the top ten fastest growing economy nations in 2016 [3]. In correspondence with instantaneous economic development, the demand for energy and consumption of electricity increase at a faster rate. Electricity is a major contributor towards the improvement of the standard of living of any individual, family and society at large. Economic development of a country and usage of electricity are strongly correlated [4, 5]. Supply of electricity is vitally important to meet the growing demand for electricity and to improve the country's economic condition [6]. However, the impact may vary depending on the context of a certain country. Thus, consumption and supply pattern can indirectly indicate the country's economic health. Direct methods of understanding economic health of a country, such as a census entail colossal expenditure and utterly time-consuming. Hence cannot be repeated frequently. Thus, these methods can be the bench mark but fail to be a real time indicator of macro level economic health. With the advent of machine learning algorithms and high performing computing devices, use of indirect methods to understand macro level socioeconomic conditions are gaining popularity. Two most frequent approaches are using secondary data sources as proxy indicators [7-10] and processing satellite images to understand economic health of an area [11].

Significant elements of socioeconomic changes are taking place throughout the world where the dynamic development is powerful and pervasive information technology is becoming an important role to play [12]. Many companies or organisations are currently working to determine a way to predict the socioeconomic factors or related attributes based upon previous data which are already in storage. But in low and middle-income countries like Bangladesh, there are insufficient attempts which are being performed so far to make proper use of data. This paper is intended to show a pathway that elaborates the use of data for the public good in LMIC (Lower Middle-Income Countries) like Bangladesh.

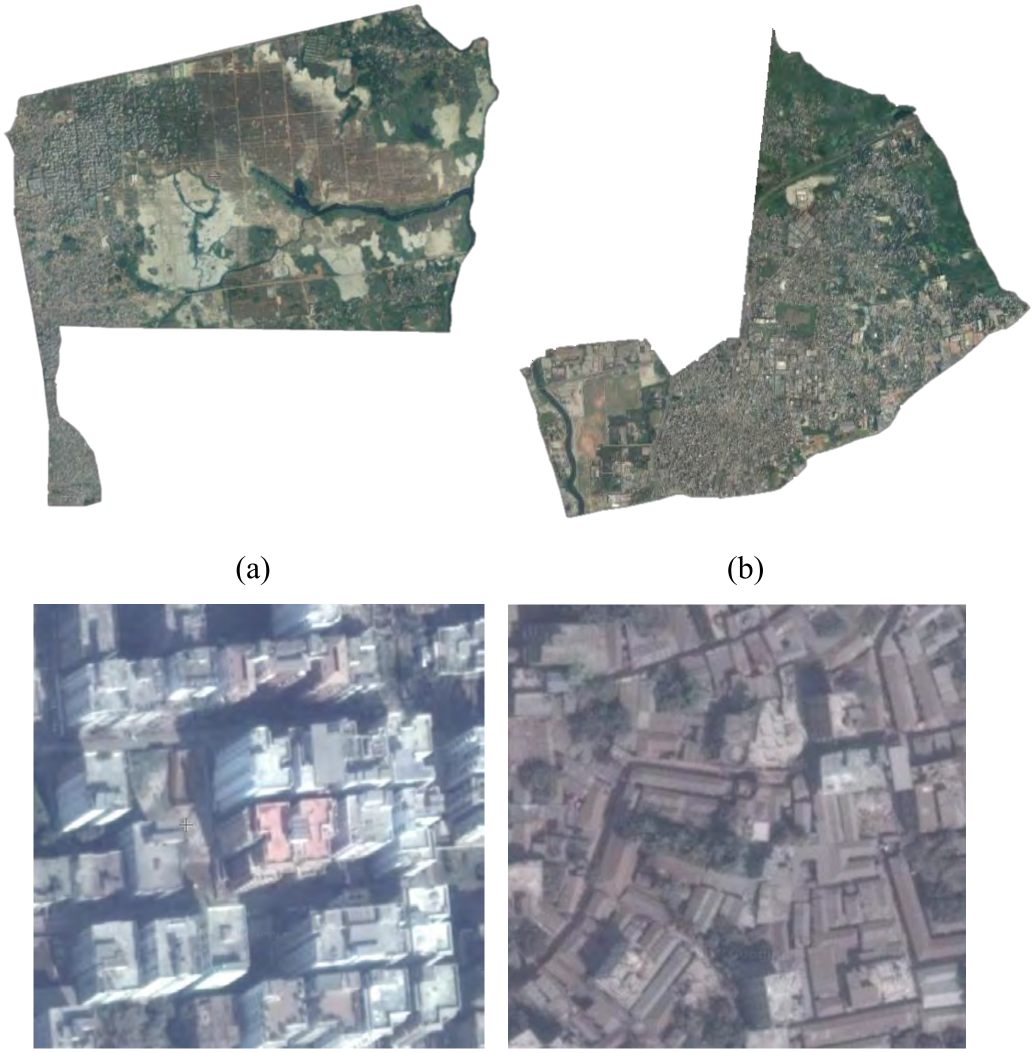
1.2 Problem Statement

A country conducts census or economic surveys periodically to understand the current socioeconomic status and how this changes over the year. These methods are highly expensive and time-consuming from conducting the fieldworks to impute, clean, analyze and publish the data. As this also requires extensive manpower to execute, most of the census have been carried out with gap of couple of years. However, the outcomes of these surveys e.g. group(s) representing the total population, change in economic status, how people handles the price hike, number of people and their actions related to socio-economy, how it differs etc. are extremely important to design policies; hence, a faster way of understating the socioeconomic status has become a crucial factor.

Apart from socioeconomic condition, as a usual person, one would like to know how s/he consumes electricity. This has been hardly understandable from our monthly electricity bill which is just a representation of consumed unit and relative price along with few credentials of the consumer. A person doesn't get to know how anything about his/her consumption habit, trend of previous consumption or how much he will be consuming in a coming month etc. To the best of my knowledge, no such approach has been taken to understand these patterns. Also, the legacy method of load forecasting works on an aggregate data. Both load and load shedding data are still maintained in hand written log books where a summarized data has been sent to upper level authorities on which distribution policies are prepared. This action results to a significant amount of uncertainty and ultimately responsible for the load mismatch in an area and causes load shedding.

1.3 Aim of Study

Dhaka is the capital city of Bangladesh and one of the heavily populated cities in the world [13]. While this Dhaka being a densely populated city of Bangladesh, many sources are still not being used to predict and address the socioeconomic health. Electricity is one of those 'not yet fully considered' sources those might have a huge impact to draw a statistics of socioeconomic status to the residents. As mentioned above, there are two ways of understating socioeconomic status, image processing and secondary data sources are being used as a proxy. Apart from the total area of different regions, number of households, population density differs a lot. Figure 1 demonstrates this statement while there is an obvious difference in satellite images of Tongi and Baridhara, two areas under Dhaka.



(a) (b)

(c) (d)

Fig. 1. (a,b,c,d) Satellite images of Baridhara (left; a and c) and Tongi(right; b and d): showing clear difference in household number and population density

Figure 1 represents the clear difference between two areas of Dhaka city. In the left images (a and c), household types, the density level of Baridhara is reflected while the right images (b and d) are representing Tongi's condition. About all the households of Tongi (in the above images) are tin-shed while Baridhara is more concrete settlement representing a socioeconomic difference. Many researchers are using these satellite image sources to analyse and derive the unseen.

However, we are following the other option, using a secondary data source, electricity consumption data as a proxy. Findings of this research enhance the understanding of economic activities and electricity demand. The aim of this paper is to analyze consumption data to find different types of electricity consumers, understand their consumption characteristics, impacts of tariff rate and tariff bracket change along with the area wise impact of electricity consumption, how a specific trend in consumption can derive an economic of shift of that area, how seasonal pattern exits and impacts on consumer usage etc. to establish electricity data as a proxy indicator for a better understanding of country's economy. With the available data sources, a user level consumption forecasting model has been proposed and accuracy evaluation has been performed too. Findings suggest that there are many low consumption users exist and use electricity to meet up their basic needs. However, there are many users exist in certain zones where their consumption is relatively high in terms of other zones. Also, findings suggest that there are many low-income dwellers live in all areas and their distribution is relatively almost same.

From load data to perform a load forecasting scope analysis on Bangladesh, necessity of studying load forecasting techniques, recent load forecasting models which have been conducted on different countries' data, variation of load in different areas, load shedding statistics and pattern in a day and in a week has been performed. A granular level short-term load prediction has been performed later and accuracy and prospects have been discussed.

1.4 Thesis Outline

The rest of this report is organized as below:

- In chapter 2, brief introduction of electricity consumption and socioeconomic status, load forecasting and necessity of studying this before going to smart grid and previous works related to this research have been discussed.
- In chapter 3, electricity overview of Dhaka, data description along with data acquisition, imputation, cleaning etc. have been explained briefly.
- Chapter 4 represents statistical analysis, forecasting models, results and comparative analysis.
- Chapter 5 concludes the paper with general remarks and future works.

CHAPTER 2

2.1 Electricity Consumption and Socioeconomic Status

A portion of this research is a case study of electricity consumption pattern of Dhaka city dwellers of Bangladesh. By analysing the consumption pattern we try to indirectly understand the economic health of the households/consumers. The approach can be an effective measure of household level economic condition especially because various other direct methods such as survey which may be expensive and time-consuming. Electricity is a vital resource for country's economic development. Bangladesh is one of the developing countries in the world where electricity plays a vital role in development. Electricity price has been increased multiple times in Bangladesh and this paper is intended to find out the impact of the price hike on the residential consumer. Also this paper shows, it is possible to use electricity consumption data as a heterogeneous data source for better understanding of country's economy. It has been found that the number of electricity consumers is rapidly increasing over the year. But most of them are belong to the specific group. High consumption users belong to mainly in two specific zones but the total number of consumers are very low in terms of other zones.

2.2 Brief Introduction on Load Forecasting

Forecasting is evident to ensure better management of generation plants, supply grids and the transmission system. Short-term, medium-term and long-term forecasting is helpful for generation-transmission scheduling, fuel purchase planning, management of generation units, transmission-distribution channel expansion respectively. As a candidate, scopes have been discussed on the basis of electricity load data of Dhaka, the capital city of Bangladesh; one of the world's most highly populated city. This chapter discusses on three average sized DESCO zones of its total 16 while Shah Ali substation has been chosen for further analysis to draw a few statistical key points, analyze load-shedding scenario, day wise impact, impact in distributed load and all of these used to perform load forecasting. Substations of zones record this hourly load data in a handwritten log book.

Load forecasting can be classified into 3 parts: up to 1 day/week ahead for short-term, 1 day/week to 1 year ahead for medium-term, and more than 1 year ahead for long-term models [14, 15]. Short-term forecasts are useful in the scheduling of generation and transmission of electricity. Medium-term forecasts are necessary for fuel purchase planning while long-term forecasts are used to develop the strategic models which will be executed in long run; like, management of generation units, expansion of transmission and distribution system, if necessary. These statements are also supported by Weron [16] and Pedregal and Trapero [15]. Load prediction may depend on several factors including time, social, economic, and environmental variables by which the pattern will form various complex variations [17, 18]. Social and environmental factors are sources of randomness found on the load pattern [19]. There are some prediction models based on architectonic features such as heat loss surface, building shape factor, building heated volume and so on [20, 21], or housing type and socioeconomic features such as age of the dwelling, size of the dwelling, monthly household income, number of household members, etc. [22] along with historical load data. As in Bangladesh, such specific data is not available or not easily accessible, one can conduct forecasting based on historical load data only.

2.3 Smart grid and necessity of a Study on Demand Forecasting

Electricity supply via smart grid mechanism is gaining importance in many country's priority lists. A detailed study on electricity forecasting is required to ensure a smooth transition to the smart grid. Smart grid (SG), the fast development of new infrastructure in the electricity sector is opening a wide range of opportunities. An SG is an advanced electricity transmission and distribution network that uses information, communication, and control technologies to improve the economy, efficiency, reliability, and security of the grid [23]. Many countries worldwide now either have switched or are planning to replace their old electric grid system with the smart grid. For example, in 2010, the US government spent $7.02B on its SG initiative, while the Chinese government used $7.32B for its SG program [20]. In Bangladesh, The distribution system loss is high and the customers face daily planned load shedding. To address the power crisis and other problems, the conventional distribution system should be restructured to the smart distribution system, a part of Smart Grid. Though it is a very new and expensive concept, yet Bangladesh Government has shown positive approach [21]. Shifting to SG is not a straightforward concept to acquire all of a sudden. Moving from old conventional grid systems to SG, there are several prerequisites to ensure. One of these is to study and understand load forecasting techniques and perform scope analysis while judging present condition. This article can bring a positive impact towards it.

Smart grid, designed with digital technologies channels a two-way communication between the customers and the utility. It includes different operational and energy measures through smart meters, smart appliances and energy efficient resources [24]. It enhances the efficacy in electricity usage for supply planning and understanding demand and can quickly change electricity supply by controlling the production and distribution of energy. The data downpours the individual smart meters and distribution channel can be used for analytics to derive information allowing demand forecasting, planning of the distribution and the production of the electricity. A smart grid accompanies better control and efficient balance of electric energy demand and supply through copious invigilation over the consumption pattern and energy generation. Demand forecasting is one of the crucial steps to the success of the smart grid as the error in forecasting might bring considerable economic damage to a country. The production and proper distribution are greatly depended on the metric from the energy forecasting, moreover, the knowledge extracted from the grid is also critical to the estimation of the investment of the infrastructure to meet the demand. Through the smart grid, demand forecasting even in end-users level can be achieved which provides the necessary information for having a better awareness of individual usage pattern and efficient pricing strategy, in addition with, planning for growth [25]. As with smart grid, this is ensured that more granular level forecasting can be performed, before switching to SG we need to understand the different forecasting techniques and their applicability. While the transition is being planned, currently available data, patterns in it, techniques those can be used should be observed thoroughly as a preparation step for the final goal.

2.4 Previous Works Related to Proxy Indicators

A number of studies have shown the potential of various indirect data sources such as satellite images, cellular telephony calls records, bank transaction records etc. to understand temporal changes in the economic activities, social-behavioral pattern, the standard of living and even growth of cities [7-11].

A possible way to estimate and monitor economic growth rate at a high spatial resolution which can support countries, those have limited provision for conducting a survey for data collection has been proposed in a research conducted by Sundsøy, R. and co-authors. Authors have developed a model to bring out insight of spatial distribution of poverty from aggregated data of mobile operator and geospatial data which is widely available. Three geographically referencing poverty measure (2011 Bangladesh Demographic and Health Survey, Financial Inclusion Insight Survey-2014, and a national household survey conducted by Telenor Group from November 2013 to March 2014 collecting household income data) have been used in this study [7].

A machine learning approach for mining socioeconomic indicators from raw satellite imagery have been projected in the research of M. Xie, N. Jean, M. Burke, D. Lobell and S. Ermon. Nighttime light intensities were used as a proxy for economic activity. To ascertain complex features such as roads, urban areas and various terrains from satellite images, the convolutional neural network was used. This research has exposed that transfer learning succeeds in learning features relevant not only for nighttime light prediction but also for poverty mapping. Another research group of J. Blumenstock, G. Cadamuro and R. On have used past events of mobile phone metadata. Authors have constructed the asset distribution for small areas consisting fewer household and extend the model to reconstruct the country's wealth distribution. They have used Call Details Record (CDR) data from Rwanda's largest mobile network and have merged it with several survey responses related to wealth. They have found that a mobile phone subscriber's wealth can be predicted from his or her historical pattern of phone use and then they have used the same model for calculating district's average wealth index [8, 9].

J. Falkingham and C. Namazie have shown that asset based indicators define household socio economic or wealth status based on dwelling characteristics, access to basic services like electricity and clean water, and asset ownership such as refrigerator, air-conditioner, television, etc. Authors have claimed that asset based measures have been considered as a better proxy for the long-term status of households as they are thought to be more representative of permanent income or long-term control of resources [10].

Night-time-light (NTL) data can be used as a proxy for urbanisation, density and economic growth has been stated in a research conducted by C. Mellander, J. Lobo, K. Stolarick and Z. Matheson. However, the authors have warned that, even though NTL correlates positively with economic activity, it may over or under estimate the relationship. The household level electricity consumption data may give a slightly better estimate as it is not constrained by night time emission of lights and a direct proxy of household appliance usage pattern [11].

2.5 Previous Works Related to Load Forecasting

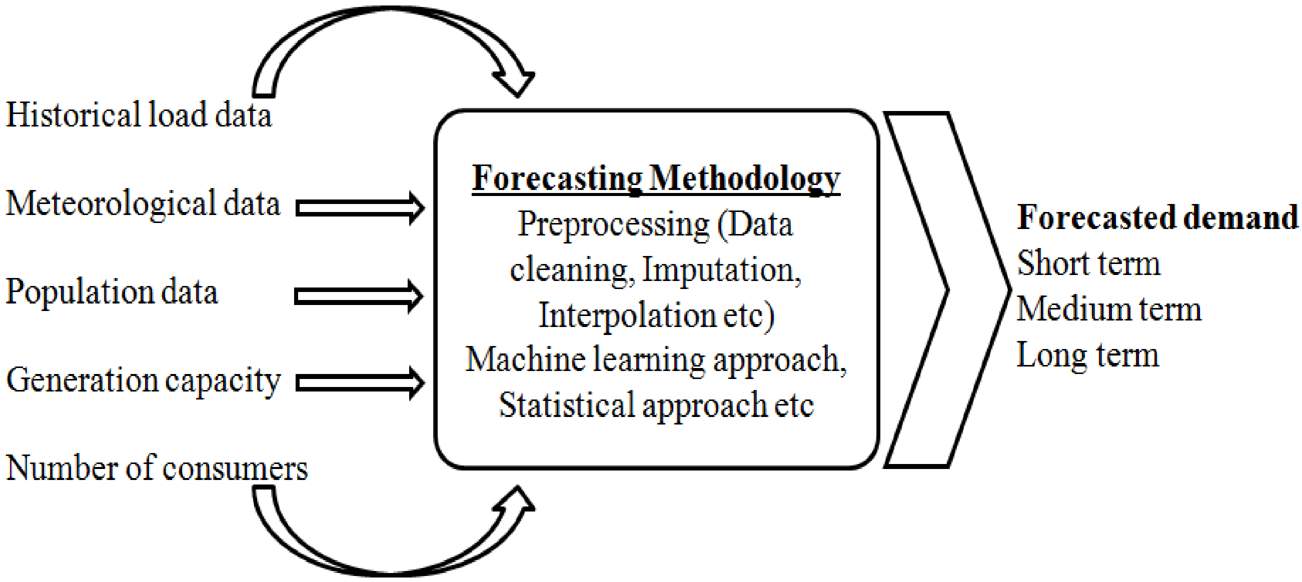


Fig. 2. General electricity load forecasting model

Figure 2 represents a basic structure of an electricity load/demand forecasting model. On the basis of input data, expected outcome forecasting methodology differs, so is a model. Hence, to understand how forecasting works, a study on forecasting in inevitable. This section contains a summarised technical review of a few electricity load forecasting related research. This summary focuses on a particular research at a time. It will summarise the data related information followed by the models/algorithms used and performance. Research chosen here are published on or after 2012. Primarily around 10 such forecasting papers were chosen and from there 4 papers have been selected and discussed here as because of page limitation. The reason for this selection is these differ in the long and short term, differ in space and context. Among these, one emphasises on building level analysis and others are country level. This differences in context help us understand the applicability of different algorithm in various context. Goal is to discuss and analyse recent approaches taken by the researchers related to developing countries (those who haven't yet switched to SG with one exception where forecasting technique on smart grid is discussed to point out how SG helps in more granular level prediction) and also extend the view of a previously published review paper of Suganthi L, Samuel A.A. back in 2012 [2]. Their research was not specifically on electricity load forecasting rather that was on overall energy demand prediction. Before going into the reviews of selected papers, a brief summary from the view of electricity load prediction on Suganthi L, Samuel A.A.'s review paper is included here. From there, the discussion will be taken further by classifying those recent researches based on different datasets used and also on different conclusions were drawn.

Suganthi L and Samuel A.A's review paper discusses several models of load prediction highlighting the uses in previous research. As demand is changing over time, so as data and over models. Moreover, performance is varying for previously introduced models against the same data sources. So, with the change in trend, an updated review is also needed. In that energy forecasting review paper [2], models like time-series, regression, econometric, decomposition, genetic algorithms etc. have been discussed. Here, we are promoting similar sort of models those have been used recently. On the basis of the articles selected here, techniques can be classified as time series models, regression analysis, group method of data handling, classifiers, fuzzy inductive reasoning and hybrid techniques.

Here, each model and its performance have been discussed article-wise with a data description of each research on which that model was executed.

Long term forecasting: Kaytez F. et al. conducted comparative analysis among regression models, neural networks and least squares support vector machines [26]. This research has been carried out on Turkish electricity transmission data which spans over the time span of 1970~2009. The indicators that were used as an independent class variable to predict the net electricity consumption were the number of population, installed capacity of the power generation, gross electricity generation and total subscription. The Turkish Electricity Transmission Company (TEIAS) statistical database was used for Turkey's total installed capacity and gross electricity generation [22, 27, 28], and the Turkish Statistical Institute (TIE) database was used for population data [29]. Both the number of subscribers and the net electricity consumption values for Turkey were taken from Turkish Electricity Distribution Company (TEDAS) and other private electricity distribution companies [30]. However, a total number of subscribers and population data are not much reliable for Turkey as the census was performed every 5 years during that time span, there were some missing values as well in the TIE records which has been handled through interpolation. Data were split into 2:1 ratio for training and testing purposes respectively while the process was validated using 2010~11 data. A multilayer feedforward backpropagation neural network was one of three models used in prediction of the net electricity consumption. Multiple linear regression models was another method used in addition to the least square support vector machine (LS-SVM) to estimate the target class. The predictive accuracy of all the model was impressive but the LS-SVM outperforms other models in various performance indicators. ANN's lack of convergence to the actual value in some of the test data reduces the success rate in comparison to LS-SVM and the fast convergence and high sensitivity to almost all of the test data helped LS-SVM to achieve the best result compared to other models [31].

Short term forecasting: A short-term forecasting research work has been performed on Korean hourly electricity load data. This one day ahead of load forecasting technique was performed on a single Korean electric load data set sourced by Korea Electric Power Corporation. This has been carried out on a sample set of two weeks of data where weekdays' data was used for forecasting load for an another immediate weekday and likewise, weekend data was used to predict the load of another subsequent weekend.

The data was classified before making a forecasting model using K- means and k- NN to eliminate error from calendar based classification. Then ANN, Group Method of Data Handling (GMDH) in addition with Simple Exponential Smoothing (SSE) were used to design the predictive model. They also gave a comparative study among the models. GMDH which had the least Mean Absolute Percentage Error (MAPE) provided the best predictive accuracy in compared to other models. GMDH uses Kolmogrov-Gabor polynomial which is inductive self-organizing data driven approach which provides scintillating performance in modelling small data set. In this study, the data set was very small as mentioned earlier and no data other than the previous load data to predict the load caused the GMDH to be the best performer because of having strength in providing predictive accuracy in small data-driven predictive design [31].

Another short-term load forecasting research was done by Lee C.M, and Ko C.N. Electric load data set of the year 2007 of Taiwan has been used here to perform 24 hours ahead of load forecasting. Three different patterns of load data are utilised to evaluate the effectiveness of the proposed algorithm. Working days were considered from Monday to Friday, Saturday as weekends and Sunday and other national holidays as holidays. Whole year data was divided into these 3 groups and different training and testing data have been selected from each group to perform and evaluate forecasting efficiency [32].

Previous 41 days of weekdays data used as an indicator to predict the load of the 42nd date. Likewise, few weekends and holidays data was used to predict a load of the subsequent weekend and holiday. Here, a model was proposed to improve the accuracy of the short-term load forecasting which integrated radial basis function neural network (RBFNN), support vector regression (SVR), and adaptive annealing learning algorithm (AALA). Determining the initial structure of the RBFNN using an SVR was the first step of designing the proposed methodology. In the intention of optimising the initial parameters of SVR-RBFNN (AALA-SVR-RBFNN), an AALA along with the time-varying learning rates was applied. In order to overcome the stagnation for searching optimal

RBFNN, a particle swarm optimisation (PSO) was applied to simultaneously find promising learning rates in AALA. Finally, the short-term load demands were predicted by using the optimal RBFNN. Comparing DEKF-RBFNN, GRD-RBFNN, SVR-DEKF-RBFNN, and ARLA-SVR-RBFNN with the proposed AALA-SVR-RBFNN, the error of MAPE is reduced by 44.94%, 68.79%, 12.50%, and 20.97%, respectively. Moreover, the SDAPE value of the proposed algorithm is 0.38%, smaller than those obtained by using the four approaches. These results verify the superiority of the proposed AALA-SVR-RBFNN over other prediction methods.

Short term forecasting (consumer specific): Another research has been carried out in small scale with some new fields as predictors. Here, data of three buildings of the Universitat Politecnica de Catalunya (UPC-Technical University of Catalonia) was obtained; 1) the Library of the ETSEIAT2 faculty in Terrassa of more than 500 m2; 2) the bar of the ETSAV3 faculty in Sant Cugat of 150 m2; 3) a building with different classrooms at FIB4 faculty in Barcelona of 630 m2. For determining the scalability of the models in any type of building, three zones with different profiles of consumption and locations are used here affecting different climatology (temperature, humidity, solar radiation, etc.), consumption patterns, schedules and working days. There was a one-hour data acquisition frequency. Therefore there are 24 recordings per day per location. For all three zones, the data set comprise a whole year of electricity consumption, from 13/11/2011 to 12/11/2012. Training and test data were split respectively as 91% and 9%. The testing data consists of 35 different days distributed equally through the whole year; meaning around 9 days per season. And taking into account the seven days of the week (from Monday to Sunday) will evaluate the models against the changes caused by a seasonal period(s) and day of the week [33].

For understanding the seasonal effect, the whole dataset was split into four chunks; autumn, winter, spring and summer. So, they proposed four models where for each model, both training and test data was comprised of that particular chunk. Each model has been trained using 7896 data points that correspond to the period 13/11/2011 to 12/11/2012 excluding all the test days.

In this research, two Soft Computing techniques and one traditional statistical method were used for the hourly energy forecasting in buildings and the techniques were RF (Random Forest), Artificial NN, FIR (Fuzzy Inductive Reasoning) and ARIMA (Auto Regressive Integrated Moving Average), respectively. At first, a feature selection methodology was applied on previous consumption data to dig out features having the highest impact on prediction. The design of the predictive model was divided into two stages: 1) Feature selection applied before training to all the methods 2) All the models used without any feature selection. In this study, FIR outperformed, most of the cases, all the methodologies followed by the RF and NN. Regarding ARIMA which had the least errors but when it came to predictive accuracy, ARIMA was far from most of other methodologies [33].

TABLE 1. SUMMARY OF FORECASTING TECHNIQUES

Ref	Data used	Time span and Country	Subgrouping	Methodology
26	Load data Number of population Installed capacity Gross electricity generation Total subscription	40 years Turkey		Multilayer feedforward backpropagation neural network Multiple linear regression models **Least square support vector machine**[a]
31	Load data	2 weeks Korea	Weekday and Weekend	Artificial neural network **Group Method of Data Handling**[a] Exponential Smoothing
32	Load data	1 year Taiwan	Weekday, Weekend and Holiday	**Radial basis function neural network**[a] Support vector regression Adaptive annealing learning algorithm
33	Load data (Data of 3 buildings of Universitat Politecnica de Catalunya were used)	1 year Spain	Autumn, Winter, Spring and Summer	Random forest Artificial neural network **Fuzzy Inductive Reasoning**[a] Auto Regressive Integrated Moving Avg.

a. Best performing methodology in each technique

In light of these research and from the summary of Table 1, it has been seen that to perform load forecasting, along with historical load/transmission data, the number of population, the number of subscription, power generation statistics, meteorological data etc. are required. For more granular level forecasting, data is needed from the smart meters (smart grid is required to perform this level of forecasting) about particular household load requirements at different hours of the day, randomness in requirements, appliances used, shape and size of the household etc. (last technique of Table 1 is such approach).

In many research, data has been further subdivided into some groups to ensure better forecasting and to accommodate seasonal impacts and occasional impacts (subgrouping column of Table 1). There could be grouping in terms of working day/weekend. Incorporating holidays as another group might help. Days of Religious events, national events etc. could be another subgroup. There could be seasonal breakdowns like summer vs winter or four seasonal division like the reference paper 33.

For long term forecasting, a longer period of historical load data along with other heterogeneous sources are required while for short-term forecasting, a couple of years/months data are required.

A detailed study on the available data, finding patterns, understanding the relevance of heterogeneous sources etc. are a huge concern before starting forecasting. Hence, such survey on forecasting techniques and scope analysis is a must.

CHAPTER 3

3.1 Electricity overview of Dhaka

Bangladesh is one of the developing countries in the world with an area of around fifty-seven thousand square miles. Country's demand in the energy sector is increasing highly. But the scenario of electricity utility structure (generation, supply, distribution) is very complex (see Figure 3). As a sample of it, electricity data of Dhaka north has been used for this research. This section includes details of data description of three formats; consumption data, load data and load shedding data collected from DESCO, distributor of Dhaka north.

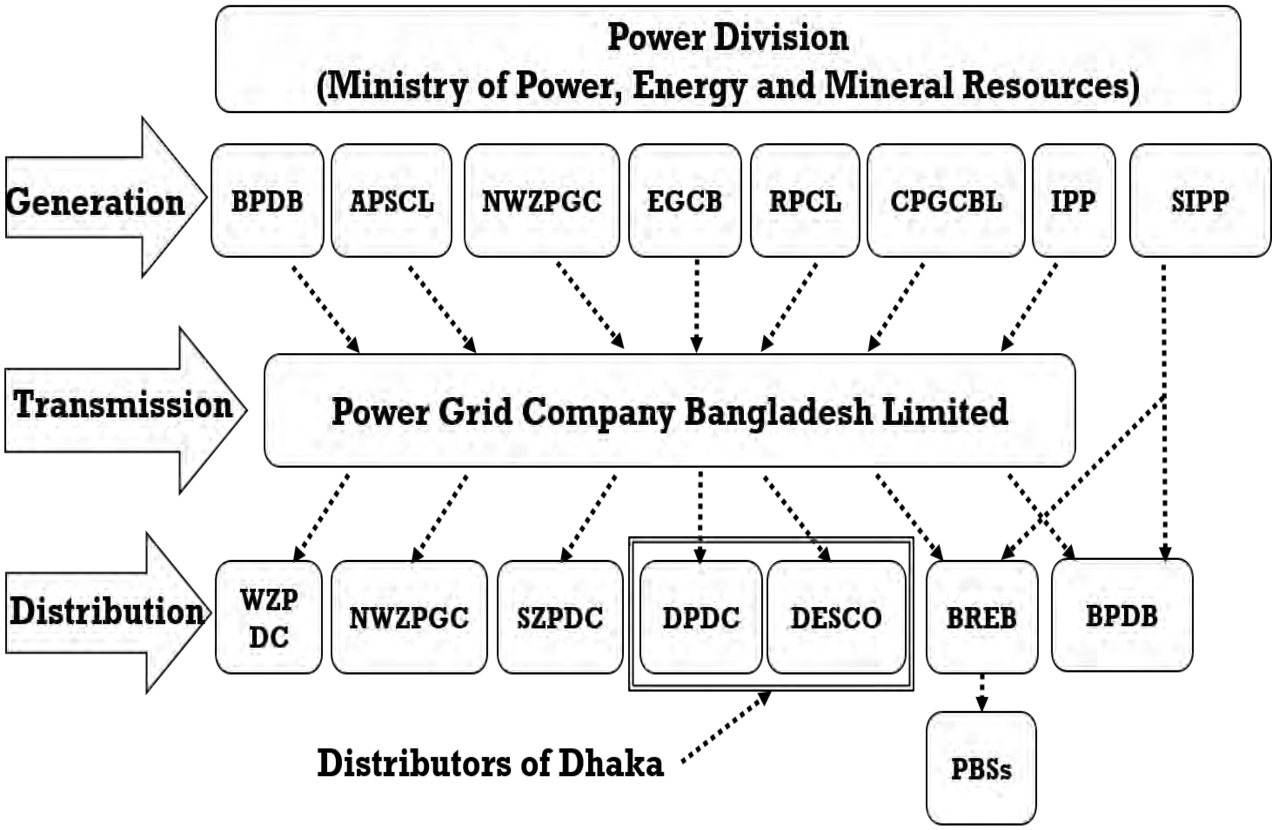


Fig. 3. Electricity sector structure in Bangladesh: Distribution companies of Dhaka (marked with a box)

3.2 Data Description

3.2.1 Consumption Data Description

Electricity distribution to consumer level in Dhaka is maintained by Dhaka Electric Supply Company Limited (DESCO) and Dhaka Power Distribution Company LTD (DPDC). Dhaka is divided into two City Corporations; one is Dhaka North City Corporation (DNCC) and another is Dhaka South City Corporation (DSCC). A portion of this analysis has been conducted on monthly electricity consumption data from the year 2005 to 2015, collected from DESCO, distributor of Dhaka north. DESCO is responsible for sixteen zones in Dhaka for electricity distribution. Eight zones have been used for the analysis. Among the eight zones, Tongi-West is outside of Dhaka City Corporation but still maintained by DESCO. Consumption data is consists of several fields like, account number of a consumer, meter number, address, zone, block, route, load (maximum allowance), tariff (category), monthly consumed unit, month and year.

Table 2 represents the summary of the monthly electricity consumption data of eight zones from DESCO.

TABLE 2. NUMBER OF UNIQUE CONSUMERS, RESIDENTIAL CONSUMERS, NUMBER OF FEEDERS, NUMBER OF TRANSFORMER FOR EIGHT ZONES

Zones	**Unique consumers**	**Residential consumers**	**Number of feeders**	**Number of transformers**
Agargaon	47198	44863 (95%)	39	601
Kafrul	62240	57918 (93%)	29	1465
Shahali	40037	36897 (92%)	N/A	N/A
Tongi-West	57850	52695 (91%)	44	2314
Pollobi	42338	38467 (91%)	N/A	N/A
Badda	52118	46211 (89%)	14	504
Baridhara	16876	15079 (89%)	N/A	N/A
Gulshan	31030	25611 (83%)	52	388
Total	349687	317741 (91%)	178	5272

From the Table 2, it can be seen that the number of consumers is relatively high in Kafrul and Tongi-West while Baridhara and Gulshan have a significantly low number of consumers. Five out of eight zones (Agargaon, Kafrul, Shah ali, Tongi-West, Pollobi) have more than ninety percent, residential consumers. Percentage of residential consumers from the rest of zones are near ninety percent except Gulshan which contains eighty-three percent. Ninety-one percent residential consumers exist within these eight zones. Number of the transformer is high in Kafrul and Tongi-West zone which might indicate denser population.

TABLE 3. COUNT, MEAN, 25TH PERCENTILE, 50TH PERCENTILE AND 75TH PERCENTILE VALUE OF CONSUMPTION

Zones	Count	Mean	25%	50%	75%
Agargaon	4956709	384.452	117	187	296
Kafrul	6778862	374.454	121	191	303
Shahali	4084847	241.889	117	189	298
Tongi-West	3639404	316.343	106	182	311
Pollobi	3839337	227.806	111	181	282
Badda	882272	280.784	93	182	328
Baridhara	1380825	8428.219	129	231	412
Gulshan	2993273	1018.962	160	333	736

Table 3 represents the count, mean, 25^{th} percentile, 50^{th} percentile and 75^{th} percentile value of unit column in the data set. The unit of the unit column is kilowatt per hour and the column represents the monthly consumption of specific users. Badda zone is declared on 2012. Even though number of users is low in Baridhara and Gulshan zone, the mean consumption is significantly high compared to other zones. The values of the 25^{th} percentile are almost same for all zones, which indicates the low consumption users are distributed evenly over all eight zones. The value of 75^{th} percentile represents the usage of high consumption users where that value in Baridhara and Gulshan is significantly high in terms of other zones while Gulshan being the highest. There were several negative values in the unit column, which has been replaced with imputed values. For a specific user, if a negative value found in the unit column, that value is replaced by the mean value of one-year user's consumption.

3.2.1.1 Tariff Bracket

In this article, as mentioned earlier, tariff analysis and key findings have been discussed on the monthly consumption/billing data of residential consumers for the time span of 2005 to 2015. There are 8 different tariff categories in Bangladesh. From these 8 different types of connections, residential consumers have been chosen as approximately 90% of overall consumers are from residential category (Table 2). During this time period, tariff category hasn't been changed, though there has been a change in the internal structure of the residential category. This category has been subdivided into a few blocks based on the unit consumed. We are denoting this as residential tariff brackets. Tariff rate differs among these brackets while low consumption bracket will incur less per-unit price and relatively upper brackets will cost more. From 2005 to 2015, in these sub categories of the residential category, there has been one change in September 2012 [34]. The Figure 4 below is a detailed presentation of when and how that restructuring was formed.

According to the Figure 4, on September 01, 2012, there was a major change took place in residential tariff bracket. Before, there were three brackets in terms of unit consumed in a month; 1) 0~100 unit, 2) 101~400 unit and 3) 400 and above unit. After that alteration, first two brackets have been broken down as 1) 0~75 unit, 2) 76~200 unit, 3) 201~300 unit and 4) 301~400 unit. 400+ bracket was divided into another two blocks; 5) 401~600 unit and 6) 600 and above unit.

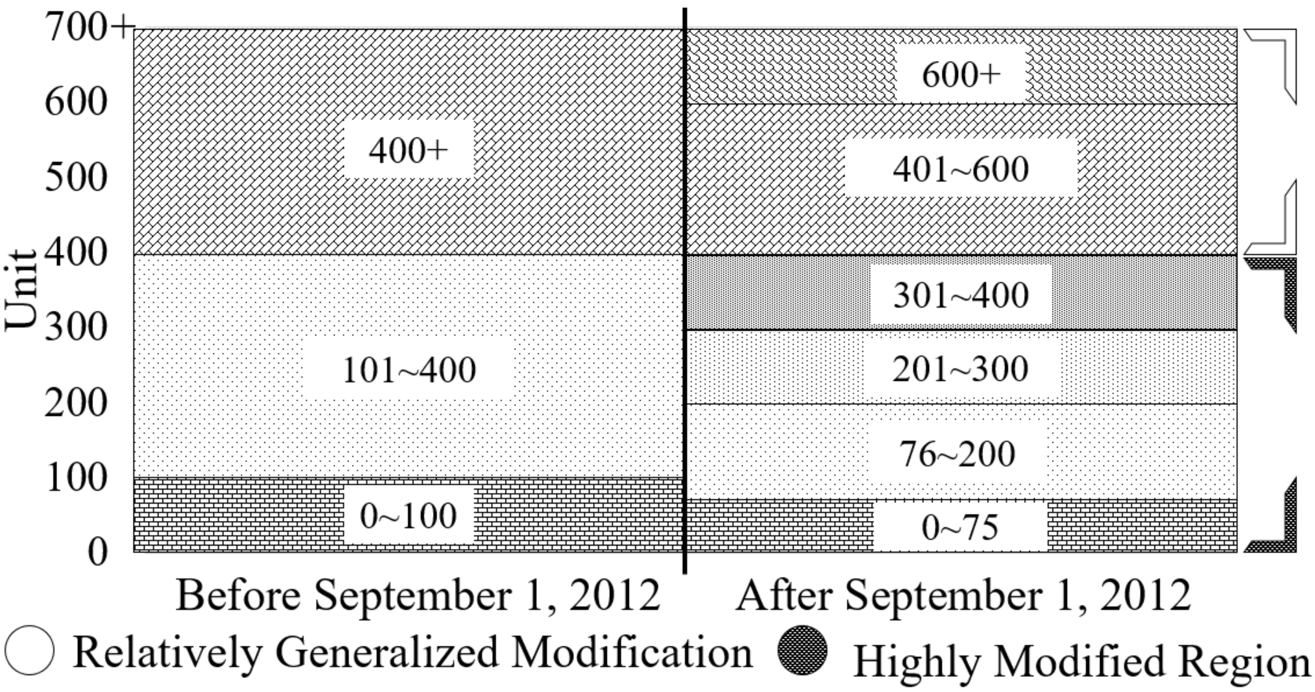


Fig. 4. Change in Residential Tariff Brackets on September 01, 2012

It has been clearly seen that specific attention was given to 0 to 400 units range (low to mid consumer range) which has been completely redesigned whereas 400+ bracket was just divided into two parts. We are highlighting 0~400 range as a highly modified region and the rest as a general modification. When we were trying to trace back to the reasons of that 0~400 units' modification, it seems to vary price more on the consumption is the base reason. The range of the new brackets has been finalised brilliantly to accommodate different consumers in terms of their consumption behaviour. To understand, follow the Tables 4 and 5.

TABLE 4. APPLIANCES, REQUIRED WATT AND RUNTIME

No.	Appliances	Watt	Runtime (hour/day)
1	CFL bulb	15	8
2	Ceiling fan	60	20
3	Refrigerator (165 litres)	180	20
4	Refrigerator (210 litres)	250	20
5	Television (small)	110	4
6	Television (big)	220	4
7	Oven	1000	0.5
8	AC (1.5 ton)	2250	3

TABLE 5. MONTHLY CONSUMED UNIT AND TARIFF BRACKET IN TERMS OF DIFFERENT APPLIANCES USED

No.	Appliances	Unit/month (kw/h)*
1	1 Bulb, 1 Ceiling fan	39.5 [0-75]
2	2 Bulbs, 1 Ceiling fan	43 [0-75]
3	2 Bulbs, 2 Ceiling fans	80 [76-200]
4	2 Bulbs, 2 Ceiling fans, 1 Refrigerator (165 l), 1 TV (s)	200 [76-200]
5	2 Bulbs, 2 Ceiling fans, 1 Refrigerator (210 l), 1 TV (b)	255 [201-300]
6	2 Bulbs, 2 Ceiling fans, 1 Refrigerator (210 l), 1 TV (b), Oven, AC (1.5 ton)	470 [401-600]

* Assuming that all appliances run for 30 days per month

3.2.1.2 Tariff Rate

Tariff rate is electricity prices based on tariff brackets. Though there was just a single change in tariff brackets, tariff rate has been changed 6 times from 2005 to 2015; 3 of these were before September 1, 2012, and 3 were after [34]. Although it seems before and after period of bracket changes have a similar impact, it is not, and rather there is a significant difference and interesting findings. Table 6 and 7 show the changes in tariff rates.

TABLE 6. CHANGE IN TARIFF RATE BEFORE SEPTEMBER 01, 2017

Unit	0 to 100	101 to 400	Above 400
March 1, 2007	2.5	3.15	5.25
March 1, 2008*	2.5	3.15	5.25
February 1,2011	2.6	3.46	5.93

* On March 1, 2008, overall tariff rate was changed but not in residential brackets

TABLE 7. CHANGE IN TARIFF RATE AFTER SEPTEMBER 01, 2017

Unit	0~75	76~200	201~300	301~400	401~600	600 +
Sept 01, 2012	3.33	4.73	4.83	4.93	7.98	9.38
Mar 13, 2014	3.53	5.01	5.19	5.42	8.51	9.93
Mar 13, 2015	3.8	5.14	5.36	5.63	8.7	9.98

These tariff rates are in nominal values; inflation adjustment is not reflected here. From the tabular view of these changes, significances are hard to be seen. To understand, please follow the line graphs of Figure 5 and 6.

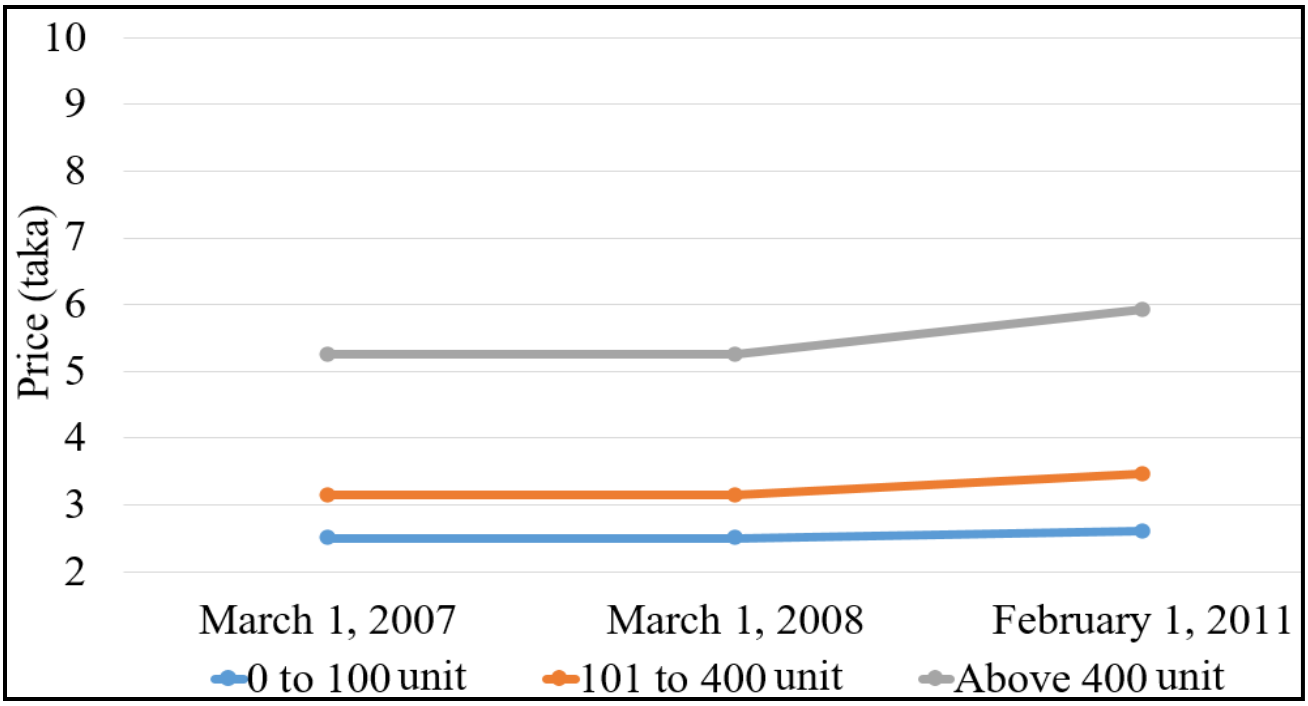


Fig. 5. Price change from March 1, 2007 to September 1, 2012: More impact on mid to high consumption users

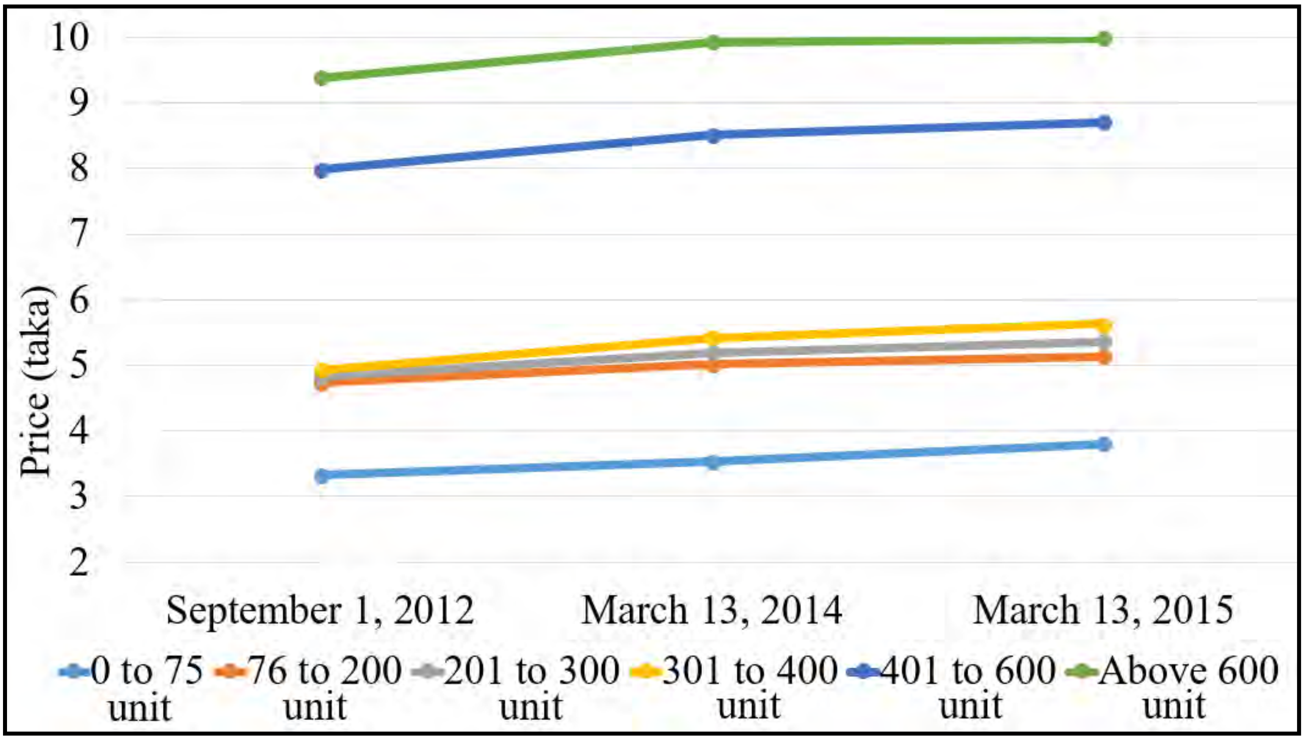


Fig. 6. Price change from September 1, 2012 to March 13, 2015: More impact on low to mid consumption users

Figure 5 and 6 represents tariff rate change over the years. Before tariff bracket change, in Figure 5, is has been seen that price change had relatively more impact on high consumption users while in Figure 6, we can follow that, low to mid consumption brackets are affected more by the change. And, Pricing almost gets doubled from 2007 to 2015. Though we haven't considered inflation adjustment, such increment is hardly the reason of inflation.

As after September 01, 2012, there are 5 tariff brackets replacing previous 3, tariff rate change impact variability was more from consumer to consumer. As we are claiming that most of the consumers fall under 'highly modified region' and as these brackets have more impact of tariff change after September 01, changes in usage pattern can be a good indicator for community detection and how different people handle these incremental changes. Figure 7 and 8 is the bar graph representation of these rate changes. Both of these graphs are drawn to the same scale to show the visual differences in changes and it has been seen that the price has been almost doubled in 2015 compared to the tariff pricing of 2007.

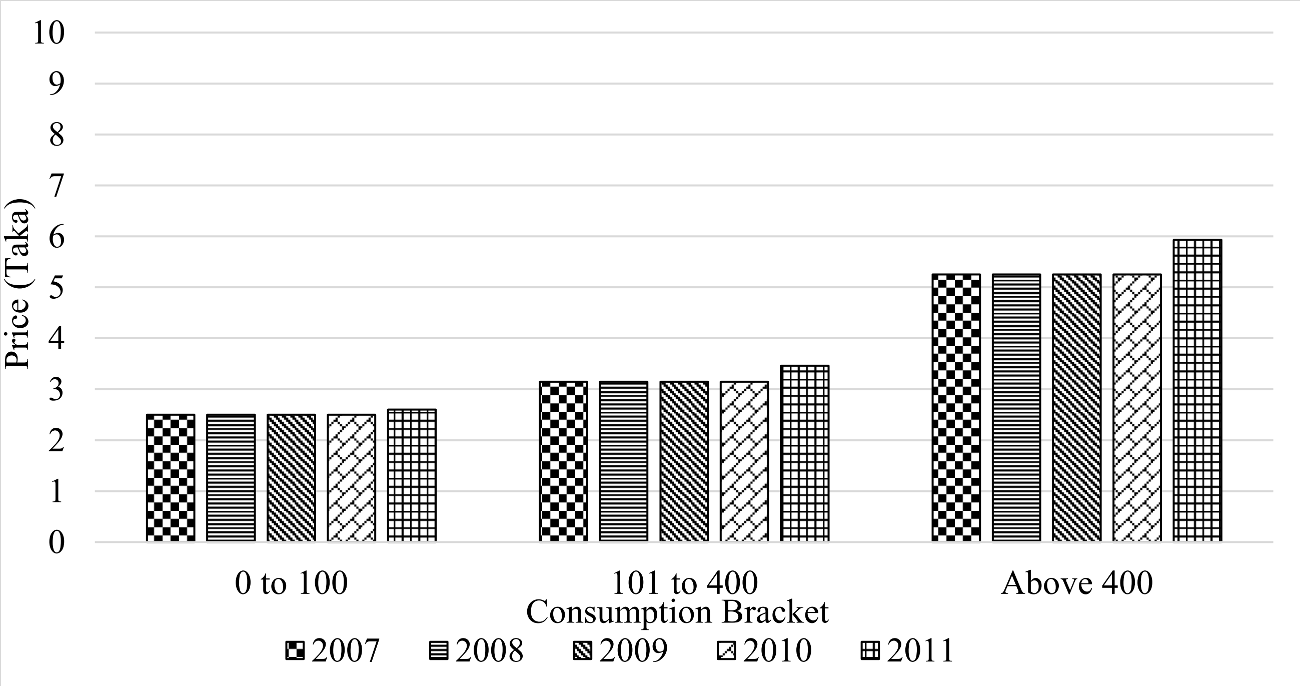


Fig. 7. From 2007 to 2011, in residential tariff brackets price changes once, on 2011

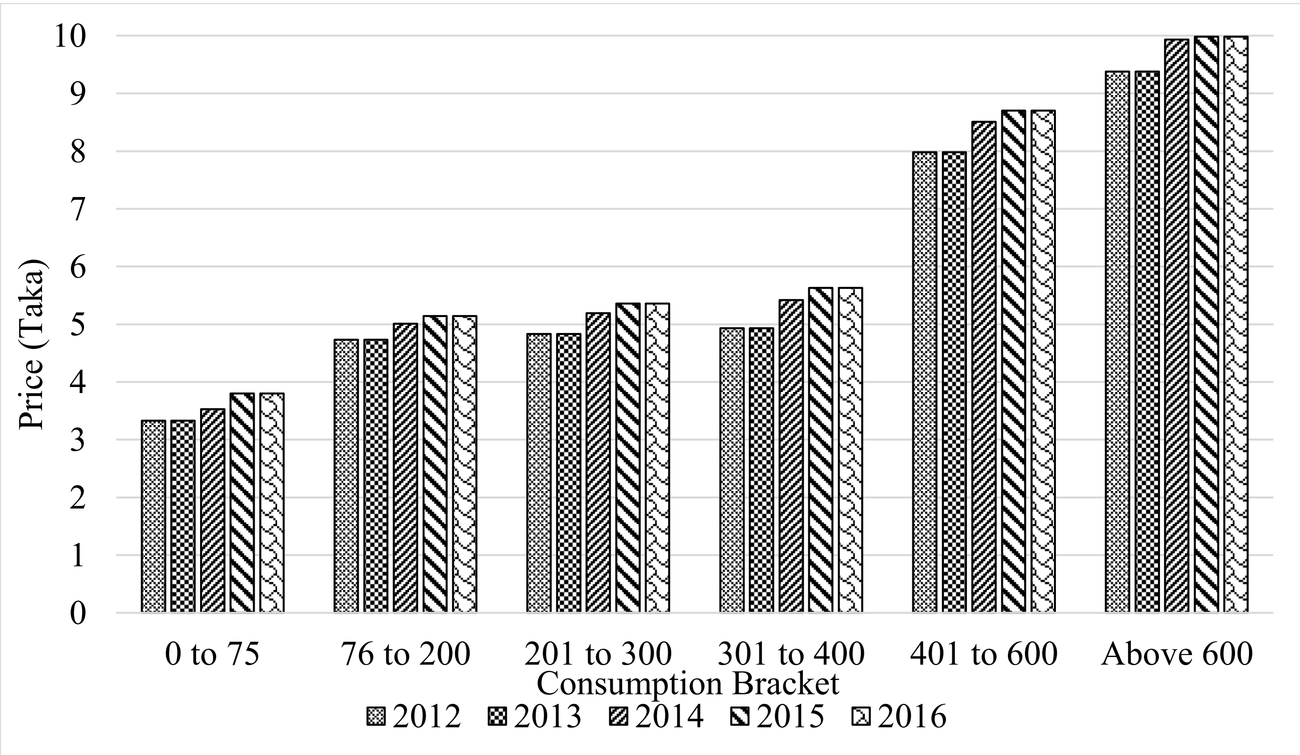


Fig. 8. From 2012 to 2016, in residential tariff brackets price change occurs in 2012, 2014 and 2015

3.2.1.3 Consistent Consumers

It has been mentioned before that for this analysis, consumption data has been selected from year 2005 to 2015 of 8 zones of DESCO. To maintain balance and perform a comparative analysis between consumers, all the new connections or newly connected consumers and those who got disconnected or left his/her electricity connection between this ranges have been kept aside. So, discussion has been carried out on those consumers only who were continuously in consumption from the year 2005 to 2015. We highlighting this group of customers as consistent consumers.

3.2.1.4 Budget Consumers

Statistical analysis of the next section specifically highlights that specific consumption ranges users are very high in numbers; in another way, it can be said that, their impact can be used to judge the overall scenario of an area. All of these groups have one thing in common. They all use electricity to meet up basic needs and to run appliances those are mostly not used to ensure comfort but a necessity. Table 4 and 5 can be used altogether to point out these consumers while in Table 5, first 5 consumers type can be the examples of them. Here, we placing consumers whose consumption is between 1 and 300 units are using electricity for basic needs and hence denoted as budget consumers.

3.2.2 Distribution Data Description: Load and Load Shedding Data

Under the jurisdiction of the Ministry of Power, Energy and Mineral Resources (MPEMR), the Power Division (PD) oversees the whole electricity utility. Electricity is generated by the Bangladesh Power Development Board (BPDB), a company spun off from BPDB, Independent Power Producers (IPPs) and private power producers. Generated electricity is supplied via the Power Grid Company of Bangladesh's (PGCB) power grid to the capital area and then distributed by Dhaka Power Distribution Company (DPDC) and Dhaka Electricity Supply Company (DESCO); local areas are supplied by BPDB and West Zone Power Distribution Company Limited (WZPDCL); and farming areas are supplied by Palli Bidyuit Samity (PBS). Figure 3 represents the overall (generation to distribution) architecture of electricity network [35]. In the Figure 3, companies above the power grid are involved with generations while distribution companies are located in the row below of the power grid. From generation to distribution, it seems quite a complex architecture; a few generation plants (SIPPs) are directly distributing to the distributors while others are delivering through the national power grid. Under these distributors, there are several sub stations from where power goes area specific feeders. Feeders are directly responsible for consumer-end distribution. So, from the top level, we can name this tier flow as: Generation > Transmission (PGCB) > Distribution (Zonal Substation) > Feeder > Consumer end. Data is being recorded at every substation in a handwritten log book; load data is being recorded in every hour for every substation and load shedding data is being recorded for every occurrences.

Load data is consists of feeder name, hours of a day, load of that hour in each feeder, power factor etc. So, for every feeder, 24 load data points are being collected for a particular day. Load shedding data is being recorded whenever a load shedding/blackout occurs by enlisting the specific day, calendar date of that day, name of the feeder that is affected, a starting and ending time with the total duration along with cause of load shedding. There are three reasons of load shedding. 1. Preplanned load shedding scheduled by NLDC, 2. Sudden rise in demand and 3. System or machine failure. NLDC is the National Load Dispatch Center who are responsible to oversee the distribution network to come up a policy and load shedding schedules to keep a harmony between generation, supply and demand. As both load and load shedding data are being maintained in paper, an aggregate information goes to upper level resulting a 'lost in translation' effect. NLDC works on the aggregate data, workout to point deficiency in supply capability (which could be avoided with more granular data) and design a load shedding planning area wise. Still, due to demand uncertainty, there are often sudden demand rise has been seen which is another reason of load shedding. Last one mostly hardware components issue which may result a blackout as well.

CHAPTER 4

Chapter 4 comprises of statistical analysis, proposed forecasting models, experimental setup and result evaluation. Statistical analysis have been conducted on both consumption and distribution (load and load shedding data), key findings have been examined and these were used to develop forecasting models. Each of these models has been have been trained and tested with related data sources and experimental results have been showcased.

4.1 Statistical Analysis

4.1.1 Consumption Data Analysis

Figure 9 represents the user community of eight zones from the selected month when a price hike (change in tariff rate) had been conducted or major change in consumption bracket happened from 2005 to 2015. In Bangladesh, electricity price increase five times for residential users. Bill calculation is different for different consumption bracket. Before 2012, 3 consumption groups were existed to calculate the bill for the residential user, but after 2012, six consumption groups had been introduced.

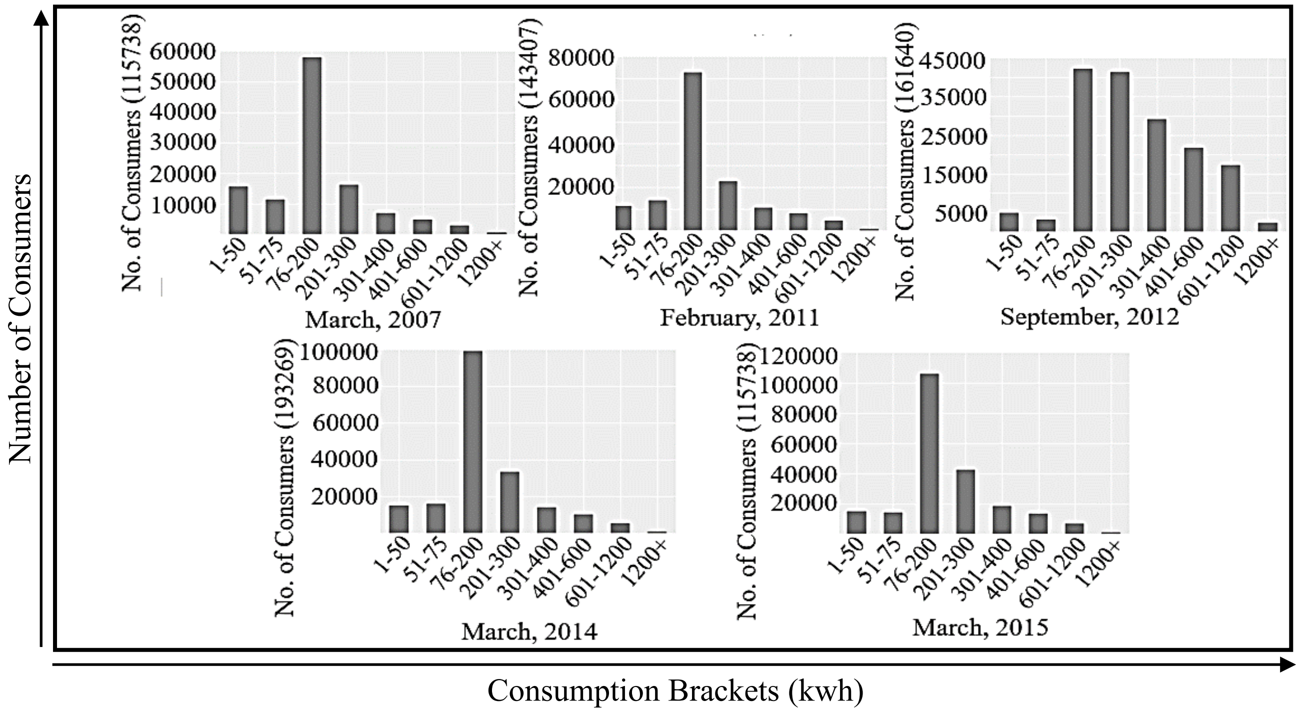


Fig. 9. Most users belong to 76-200 and 201-300 tariff bracket for 8 zones (tariff rate has been changed in the following months on the graph)

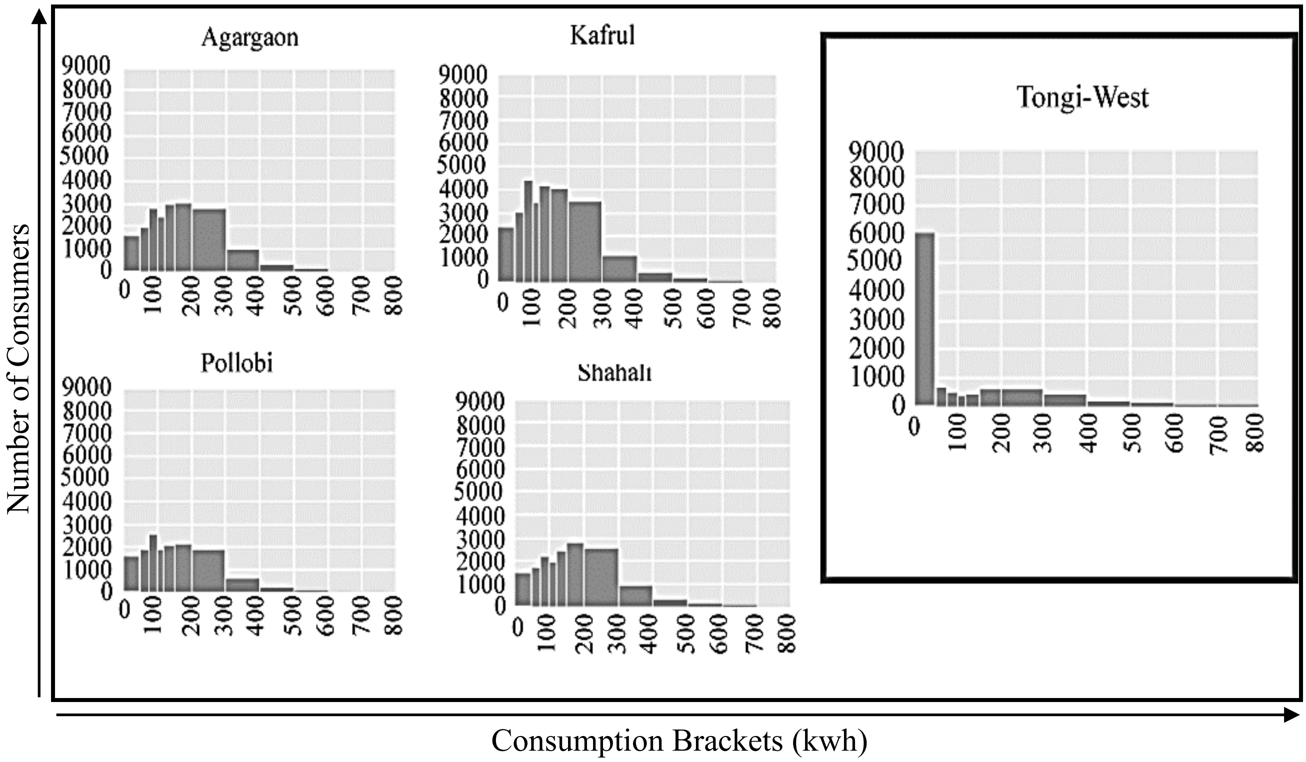


Fig. 10. Agargaon, Kafrul, Pollobi, Shah Ali had many low and mid consumption users on March 2007 (Skewed Right) whereas Tongi had many users within 0-50 range (official tariff range was 0-100, 101-400 and 400+)

From Figure 9, it is obvious to deduce that user count is significantly high for 76-200 range and gradually increasing (please, follow the vertical axis labels) over the year. But there are a sufficient number of users who are in 1-50 range and 51-75 range. User count in the first bracket (1-50) is gradually decreasing. The most unsteady group is 51-75 range. There is a possibility that from 76-200 bracket, many users exist who are near to 76 unit consumption and shifted mostly within 51-75 and 76-200 bracket. Users who live under 1-50 bracket, it is possible that their ability to expense behind electricity is relatively low. In 2012, when tariff bracket shifted from three groups to six, users suddenly increased their consumption. Which indicates that users could not control their consumption pattern due to sudden change on tariff bracket. But over the year, count of high consumption users get reduced.

More granular level analysis has been conducted on these eight zones for a better understanding of user communities. Figure 10 represents the division of each user monthly consumption (using 0-50, 51-75, 76-100, 101-120, 121-150, 151-200, 201-300, 301-400,401-500,501-600,601-700,701-800 bracket). The scenario reveals that plenty of consumers are there who are near to lower bracket. Official tariff bracket was 0-100, 101-400 and 400+ for the residential user. A number of users are found who consumed near to 100 unit (actual consumption was 110, 120 unit) for Agargaon, Kafrul, Pollobi and Shah Ali (the graph skewed right). Interestingly, a lot of consumers found within 1-50 bracket in Tongi-West zone in 2007. Consumers who consumed more than 300 hundred unit are relatively low in these five zones. These findings indicate that most of the electricity users are consuming to meet up their basic needs.

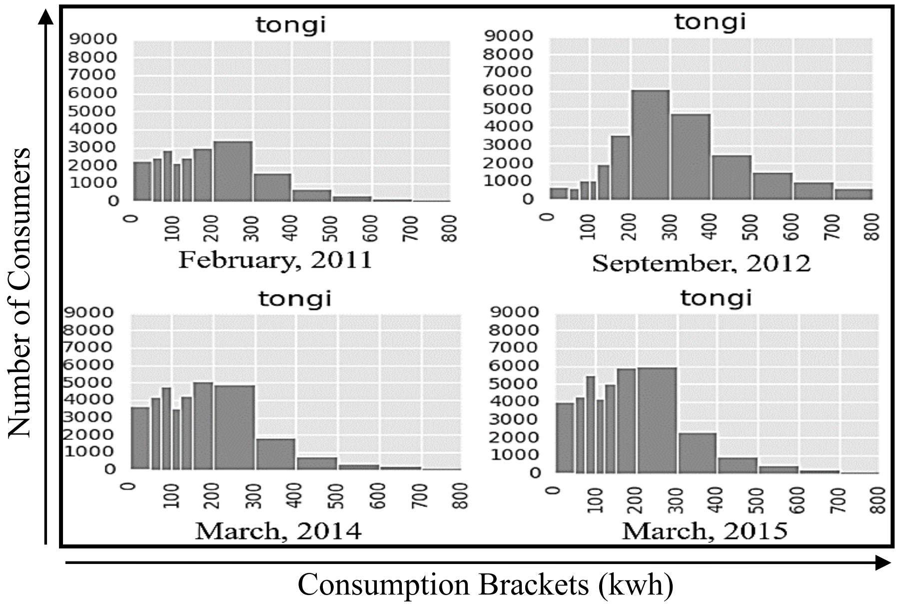


Fig. 11. Gradually Tongi-West users got symmetric over the years and again returns to the previous pattern

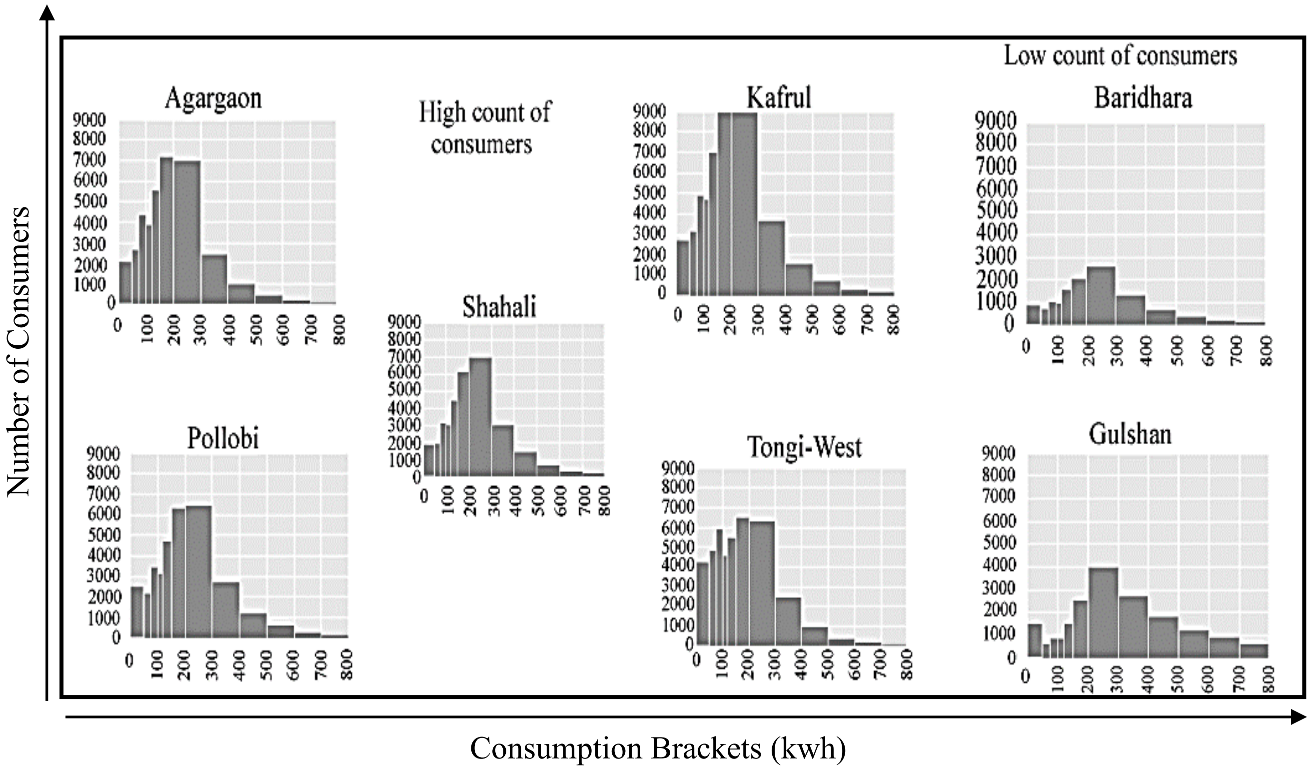


Fig. 12. Agargaon, Kafrul, Pollobi, Shah Ali and Tongi have large number of low and lower-mid consumption users rather than Baridhara and Gulshan (December 2015)

Figure 11 shows that users from Tongi-West zone get symmetric over the year from 2011 to 2015. Tongi users have been mostly increased on 2014 and 2015 and they distributed mostly on the first few groups. Earlier we have found that number of the consumer is relatively low in Baridhara and Gulshan which is represented by Figure 12. However, compare two other zones, consumers from Baridhara and Gulshan use more electricity than other five zones (Agargaon, Kafrul, Pollobi, Shah Ali, Tongi-West) in December 2015. A number of high consumption users are found mostly in Gulshan but low consumption user also can be found in this zone.

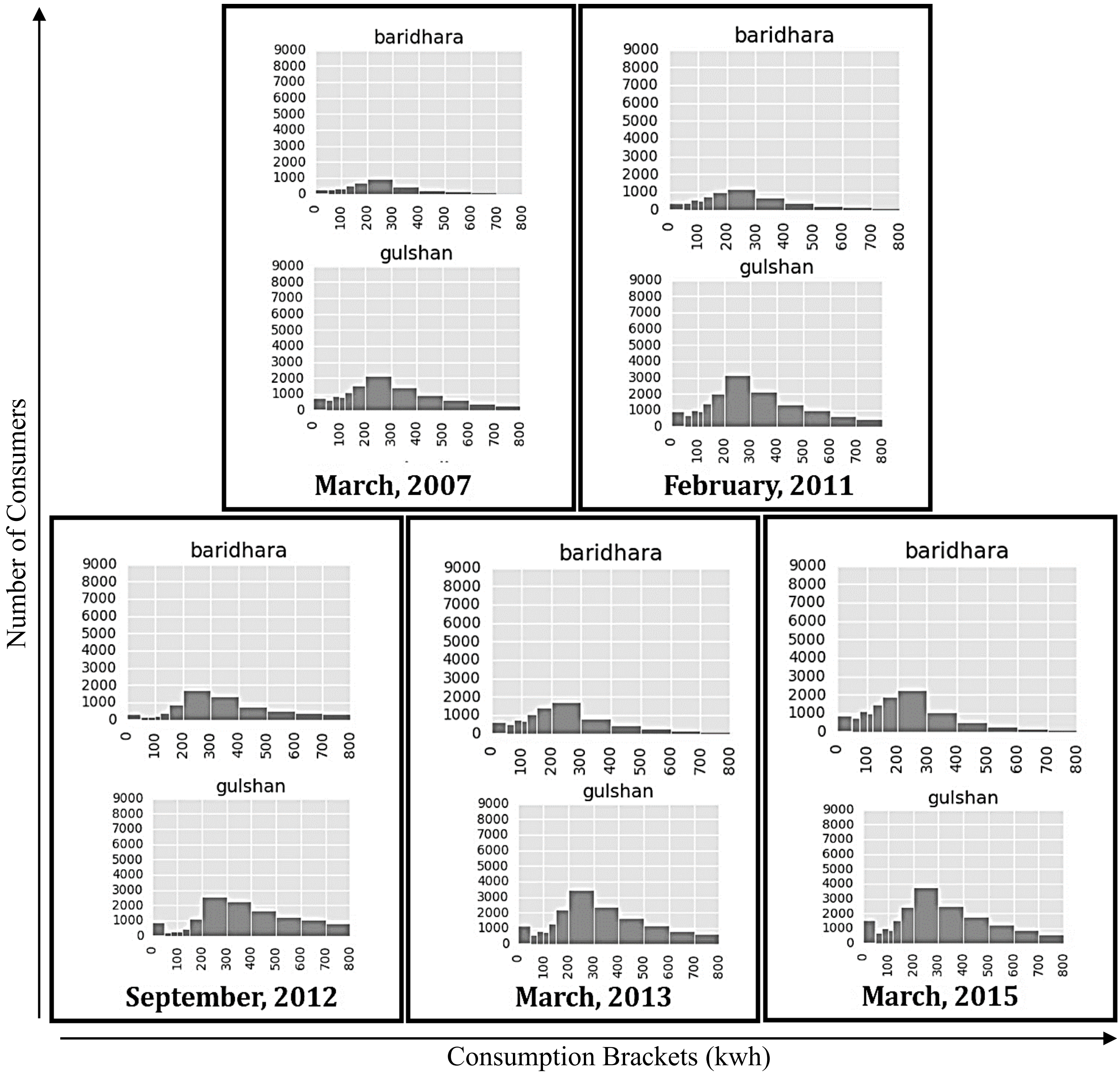


Fig. 13. Consumers count is low in Gulshan and Baridhara but Baridhara and Gulshan has many high consumption users

Figure 13 represents the scenario of Gulshan and Baridhara from 2007 to 2015. Consumers' counts are relatively low in Baridhara rather than Gulshan. But from the beginning, many high consumption users (consumed more than 300 units per month) can be found in Gulshan. Mid-level users (consumed 100-299) are relatively high in Baridhara in terms of Gulshan. But over the year, a number of consumers who consumed more 400 units in a month also got increased in Gulshan.

There are ample reasons to believe that, overall budget electricity consumers are significantly higher in numbers. These statistics are helpful to identify the groups of consumers with relatively higher impacts and requires a detailed study. Although over the year, the scenario is changing, e.g. from 2007, tongi's border line consumers' levelled up and distributed over several brackets ranging from 1 to 300 units. However, the change rate of budget consumers who fall under 300 units is not gaining speed while reaching a bracket above 300 unit. This deduction may misdirect as the situation is mostly stable which is not. Figure 14 and 15 can be a very good representation of seasonal impacts and how consumers are switching tariff brackets for this reason.

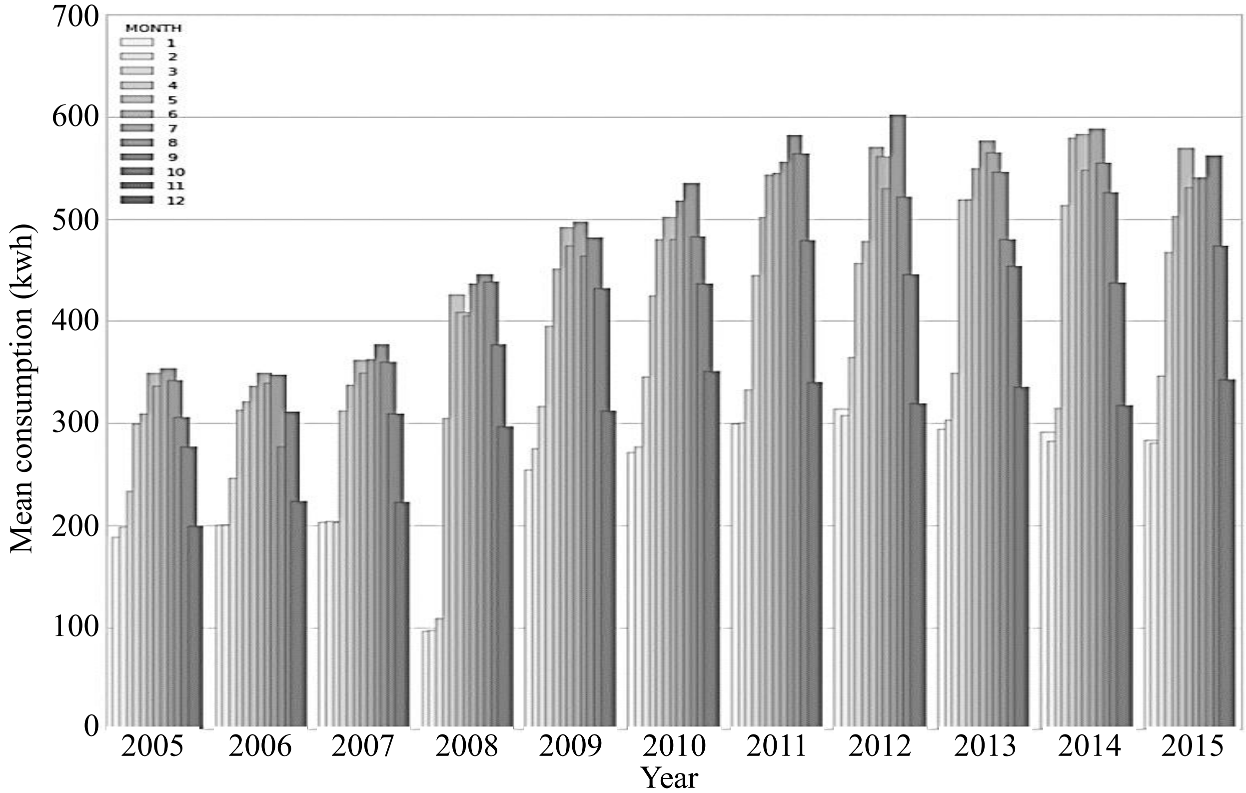


Fig. 14. Seasonal pattern in electricity consumption

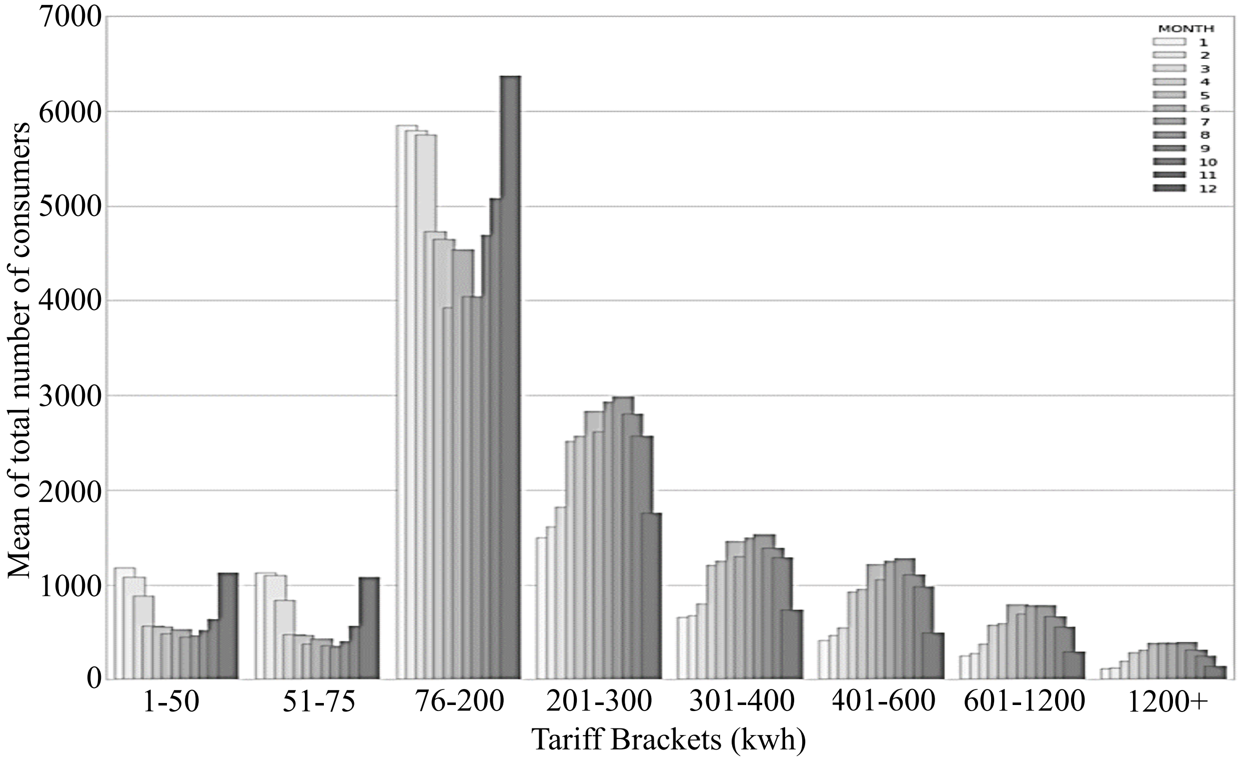


Fig. 15. How inter-tariff bracket consumer movement happens due to seasonal changes

Figure 14 denotes mean consumption in terms of different months of different years. It is clearly visible that overall consumption is rising over the year where in each year, during winter, less power has been consumed than summer. This seems a trivial conclusion. However, this practice forces consumers to shift down to a lower bracket during winter. Figure 15 shows the number of consumers in those months. During winter, many users switched down to a lower tariff bracket; thus, less power is required. During summer, a huge number of consumers started using cooling related appliances more which requires a lot of power. As a healthy portion of consumers is involved in this chain, usually there is a major demand supply gap occurs. To address, generation and distribution channel must be well aware of this situation and for this a prediction on the demand side or consumer side is inevitable. Predicting consumers' consumption may lead us to a better management in two ways; one, this information could be used to design a more effective generation/distribution network and two, prepare a consumer level feedback mechanism through which a consumer can be updated about his/her consumption habit, future probable consumption and ways to reduce his/her consumption optimally.

4.1.2 Distribution Data Analysis: Load and Load Shedding Data

As already mentioned in data description, in every zone, there are a few substations from where the electricity is distributed to the feeders. Feeders are responsible for ensuring direct supply to the households or consumer ends. In every substation, under each feeder, the load has been recorded hourly which means 24 entries each day for each substation. In Bangladesh, these entries are not maintained in a digitised format rather it is being handwritten in a logbook. As the actual load data is not digitised, a summarised report is being sent to the upper level on which legacy methods of load prediction works.

Because of this aggregation (summarized report), we lose some key factors like, feeder wise load, the difference in peak hours in different feeders etc. which could lead us to a better management.

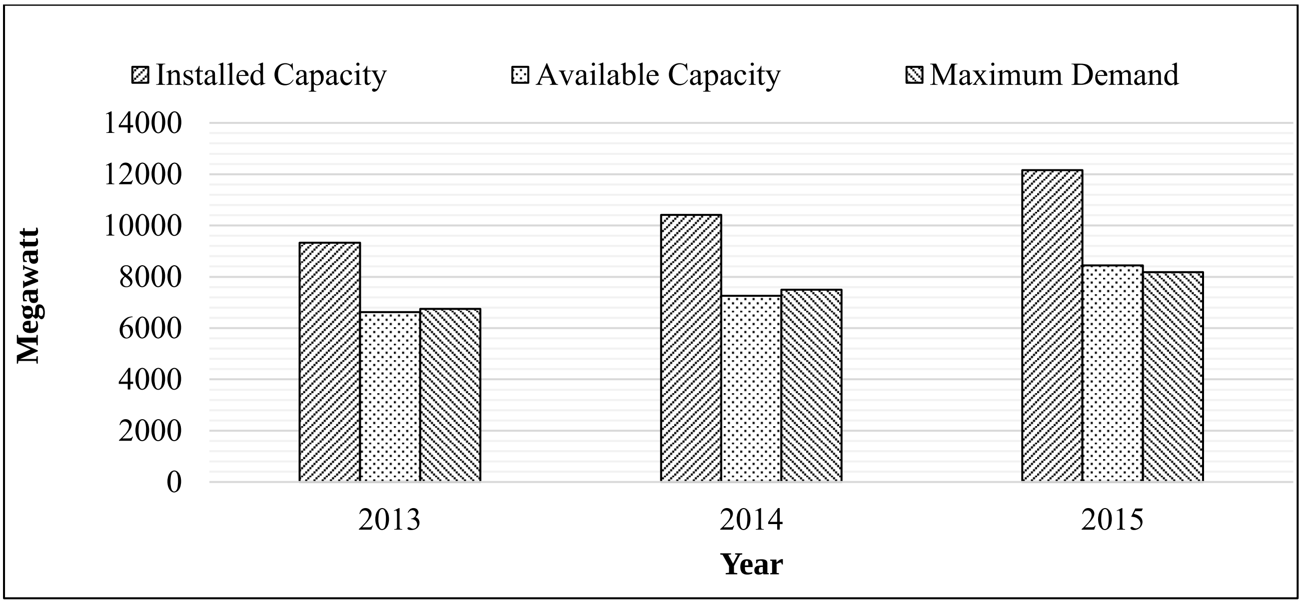


Fig. 16. Supply/Demand (Max) Balance in a Day: in 2015, available capacity was enough to satisfy maximum demand

Figure 16 represents a comparison between installed capacity, available capacity and maximum demand for the year 2013 to 2015 [35]. Approximately 30% of installed capacity was not available due to decreases in the output and thermal efficiency and failures of power generators mainly due to the insufficient periodic maintenance etc. However, it is clear that in the year 2015, available capacity was sufficient to satisfy maximum demand. However, load shedding is still a huge concern here in Bangladesh. Without knowing area specific demand, even if with available resources, demand cannot be satisfied. Thus, aggregated data and legacy method needs to be changed.

Seasonal pattern, variations of the load in different zones in different seasons, patterns and impact variance of load shedding in different zones etc. have been pointed out below to support the claim that better forecasting can be achieved with the data in hand and data from meteorological department. After a successful transition to the smart grid, more granular level data will be available; so, another level of betterment in load forecasting could be achieved. Pallabi, Shah Ali and Baridhara, 3 zones of DESCO out of its 16 have been chosen to perform basic analyses to support this claim.

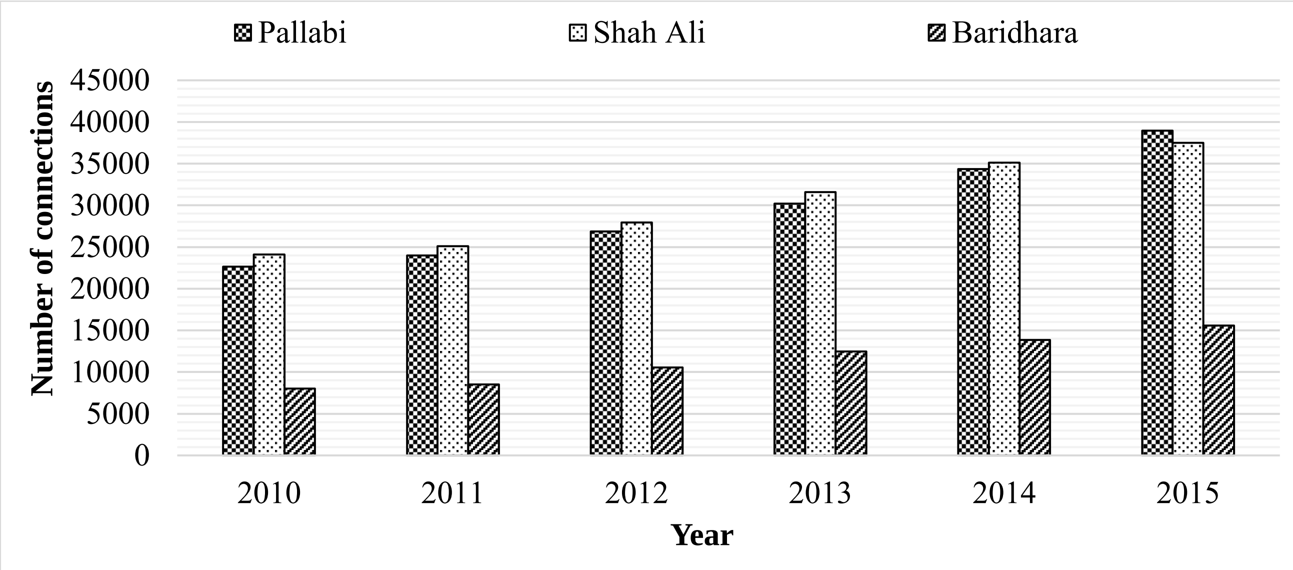


Fig. 17. Incremental Pattern in Number of Connections (year 2010 to 2015)

In Dhaka, electricity consumers are increasing day by day. If Pallabi, Shah Ali and Baridhara have been considered, number of connections has been almost doubled from 2010 to 2015. Figure 17 illustrates year wise consumer connections number in a bar chart. An increment in number of consumers is equivalent to the complexity in management which demands a better forecasting. There are an occasional impact, seasonal impact in load variability, the effect of industrial working hours or feeder-wise pattern difference in same substation etc.

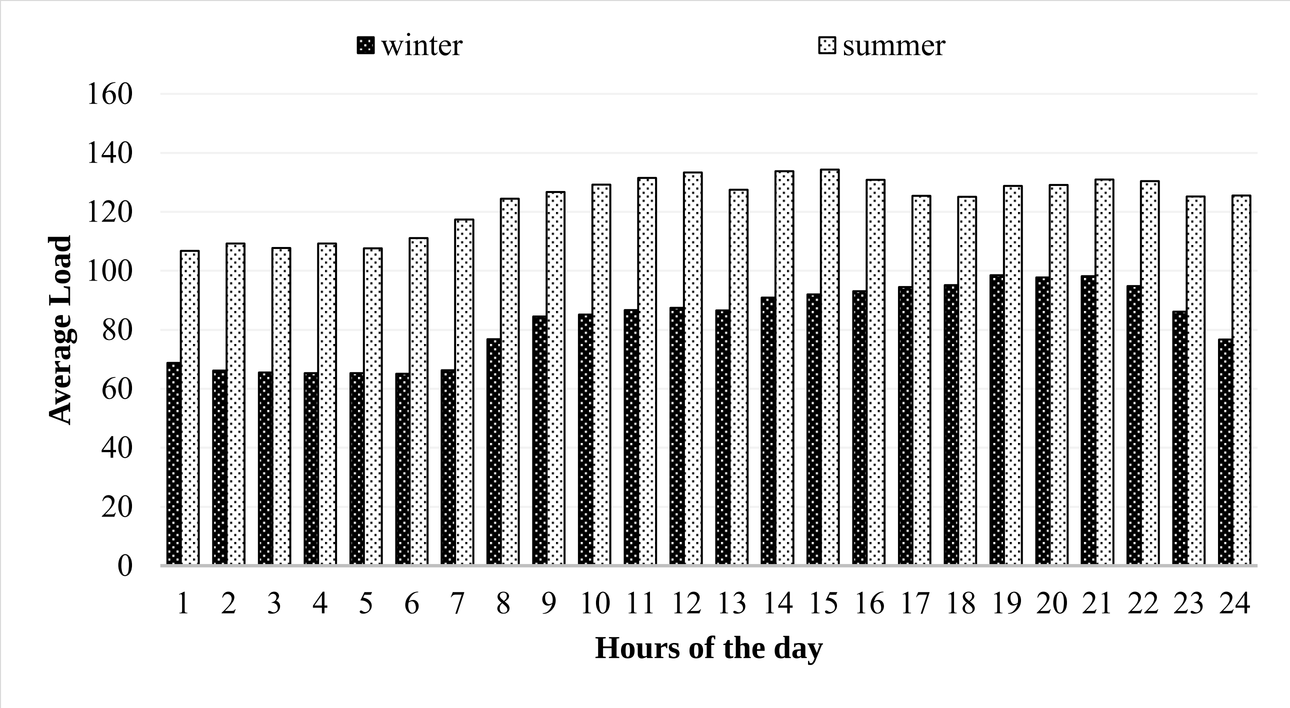


Fig. 18. Hourly Average Load of Kalyanpur: loads are almost stationary throughout the day where loads in winter are around 40 kWh less than the loads in summer

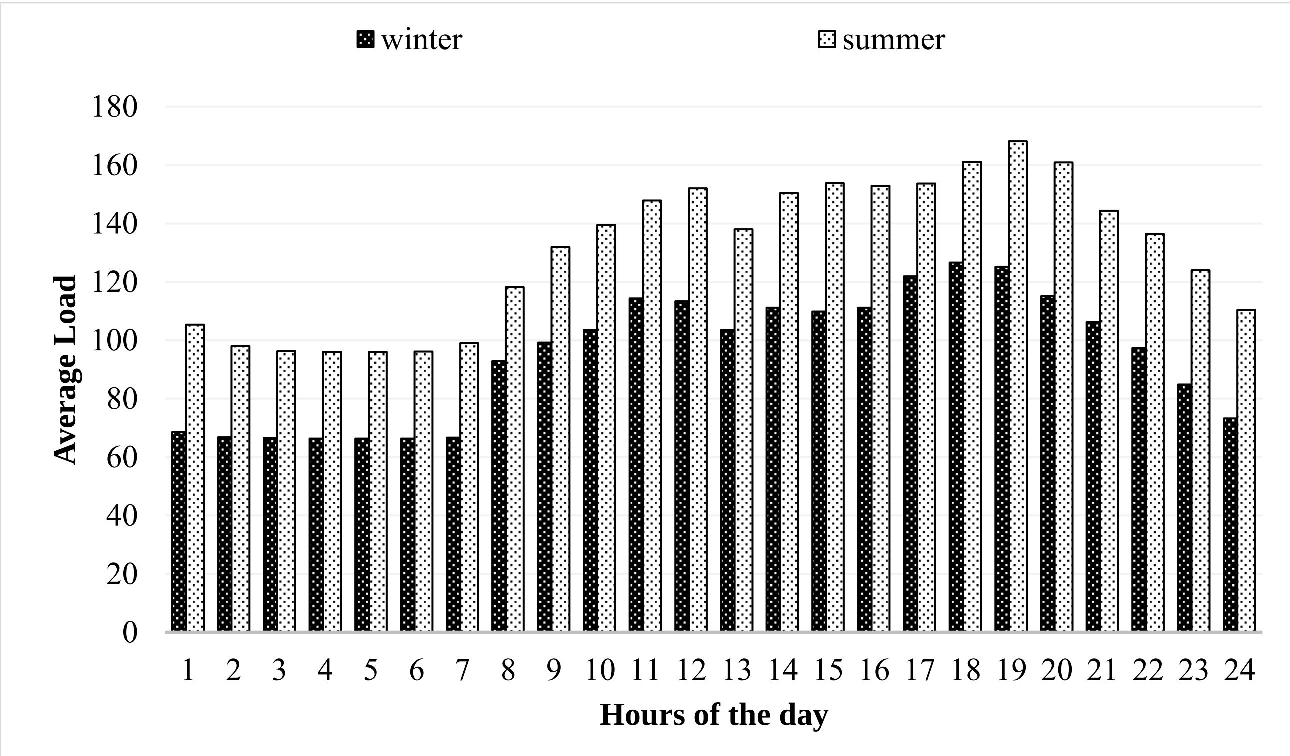


Fig. 19. Hourly Average Load of Shah Ali: the graph is following a hockey stick pattern

Figure 18 and 19 represents a comparison between winter and summer season in Kalyanpur and Shah Ali sub station of Shah Ali zone respectively. As, during winter, people most likely do not use room heaters and during summer many people use ACs, it is clearly visible that in summer, consumption is quite high. However, in a particular season (summer or winter), in Kalyanpur substation, the load varies a little throughout the day while in Shah Ali, from midnight to 7 AM in the morning it remains almost same but then gradually rising till 12 PM, mostly remain constant till 8 PM and then again gradually drops till midnight. The main point of bringing these statistics is to clear the point that even in the same zone, different sub stations act differently. As the load data are available from the logbook of previous years, substation-wise or feeder-wise forecasting could lead to a better management and distribution in the electricity sector.

Along with the load data, load-shedding data can further help to understand consumer characteristics of each feeder. While analysing load data, few key points are found which need to be incorporated in forecasting and defining distribution related policies. In legacy methods, these data were left behind and an aggregate summary was rather used and it is impossible to understand these key components from this aggregated data. From data description section of chapter 3, we already know that there are three reasons of load shedding. These are:

- ✓ Pre-Scheduled load-shedding proposed by NLDC
- ✓ Sudden rise in demand
- ✓ System failure

Frist two reasons can be addressed with a better planning. Because of aggregation, the decision is being taken on an average scenario where granular level analysis could be the deal breaker. Even two areas with an almost same number of consumers and area can act differently when it comes to electricity consumption. Required load varies from sub station to the substation, feeder to feeder. Number of load shedding occurrences and actual impact is not analogous as the duration of load shedding varies. Demand also varies in different days of a week and at different times of a day.

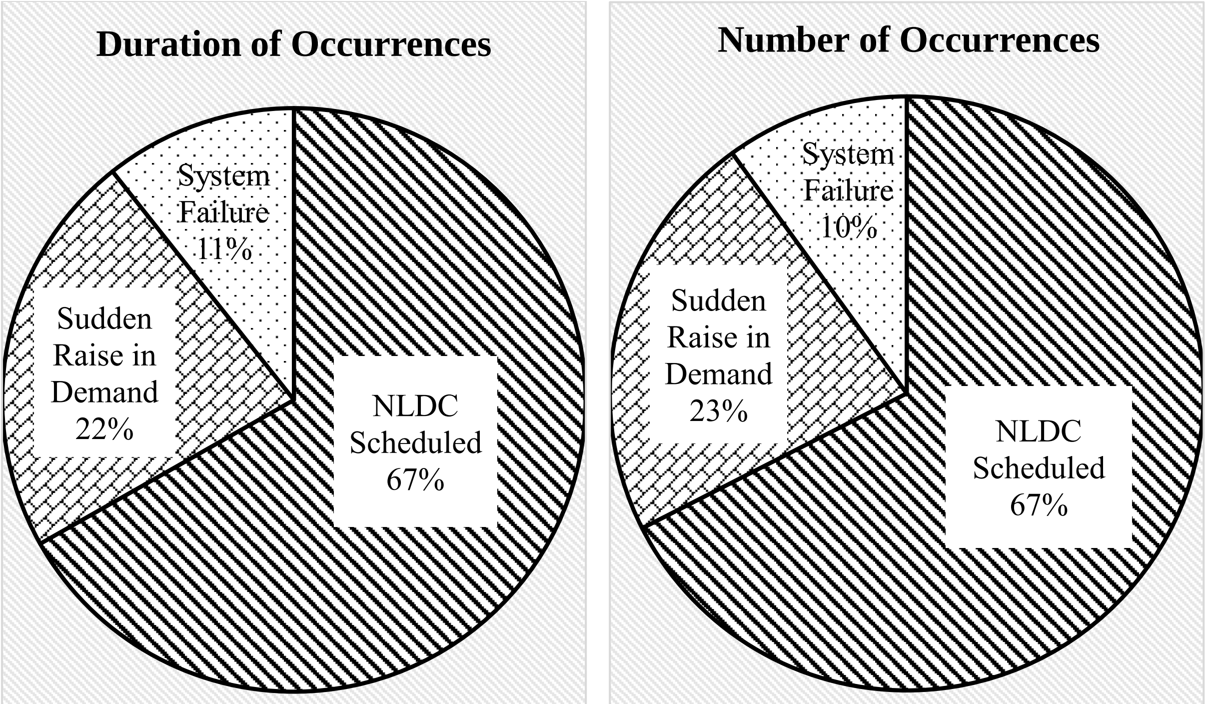


Fig. 20. Reason wise load shedding scenario of Shah Ali substation

Figure 20 shows reason wise load shedding statistics. Circle on the left represents duration wise distribution where circle on right signifies number of occurrences. Both of these distributions are very similar from where it can be concluded that the first two reasons of load shedding cover 90% of all the load-shedding scenario. If we consider each feeder, the scenario mostly remains same for many feeders but changes in a few. Feeder wise scenario can be seen from Figure 34 to Figure 44, in the appendix. A mentioned above, these are the reasons can be addressed with better management policies. So, by incorporating granular level data insights, up to 90% of load shedding can be avoided. Figure 21 could be another way of better representation showing this distribution in a line graph where lines marked with box and circle (representing first and second reasons of load shedding relatively) are with more impact than the line marked by a triangle. Figure 45 in the appendix represents number of load shedding in each feeder in a bar graph (aggregation of all reasons).

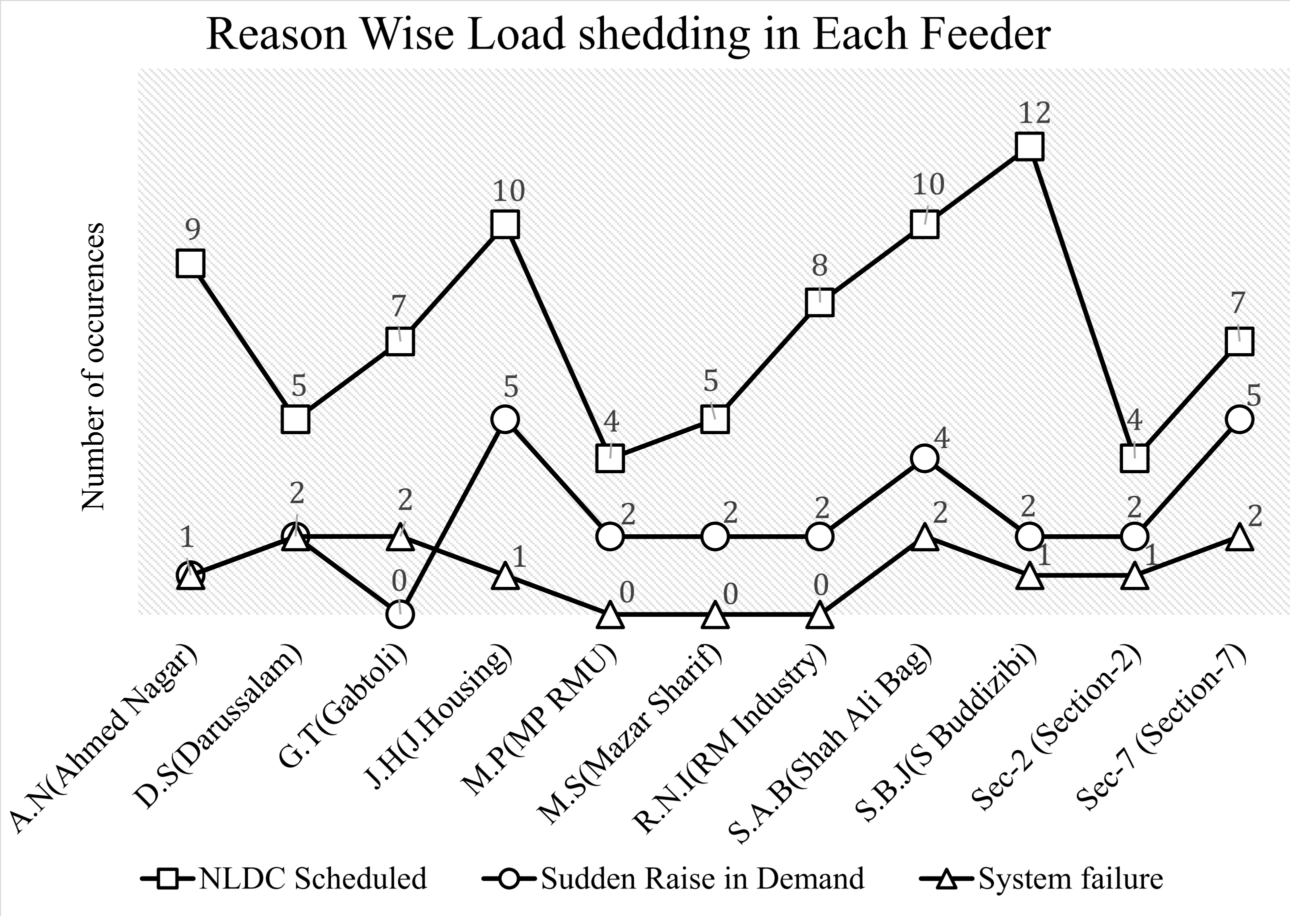


Fig. 21. Reason wise load shedding impact in every feeder of Shah Ali sub station

Though it has been seen that duration and number of load shedding are quite similar in terms of overall comparison (figure 20), it has different impacts in different feeders in different days. Figure 22 represents such comparison between occurrences of load shedding against duration impact for each day of a week. To calculate the impact of load shedding for a particular day on the basis of the duration, total duration of load shedding for that day has been divided by the average load shedding duration of all days. In most of the cases, number of load shedding over judge the actual impact while during Monday and Wednesday, though the number of load shedding is relatively lower, its' impact is relatively higher because of the duration. In both cases, Sunday and Monday are more prone to load shedding. Follow Figure 46 and 47 in the appendix which also supports this by representing a stacked and layered area chart relatively. Feeder wise daily load shedding statistics have been represented in Figure 48 in the appendix. To scale the duration to make it comparable with occurrences, the total duration of load shedding for each day was divided by the average load shedding duration of the entire dataset. As the duration varies, for further analysis, duration of load shedding has been considered as relative impact.

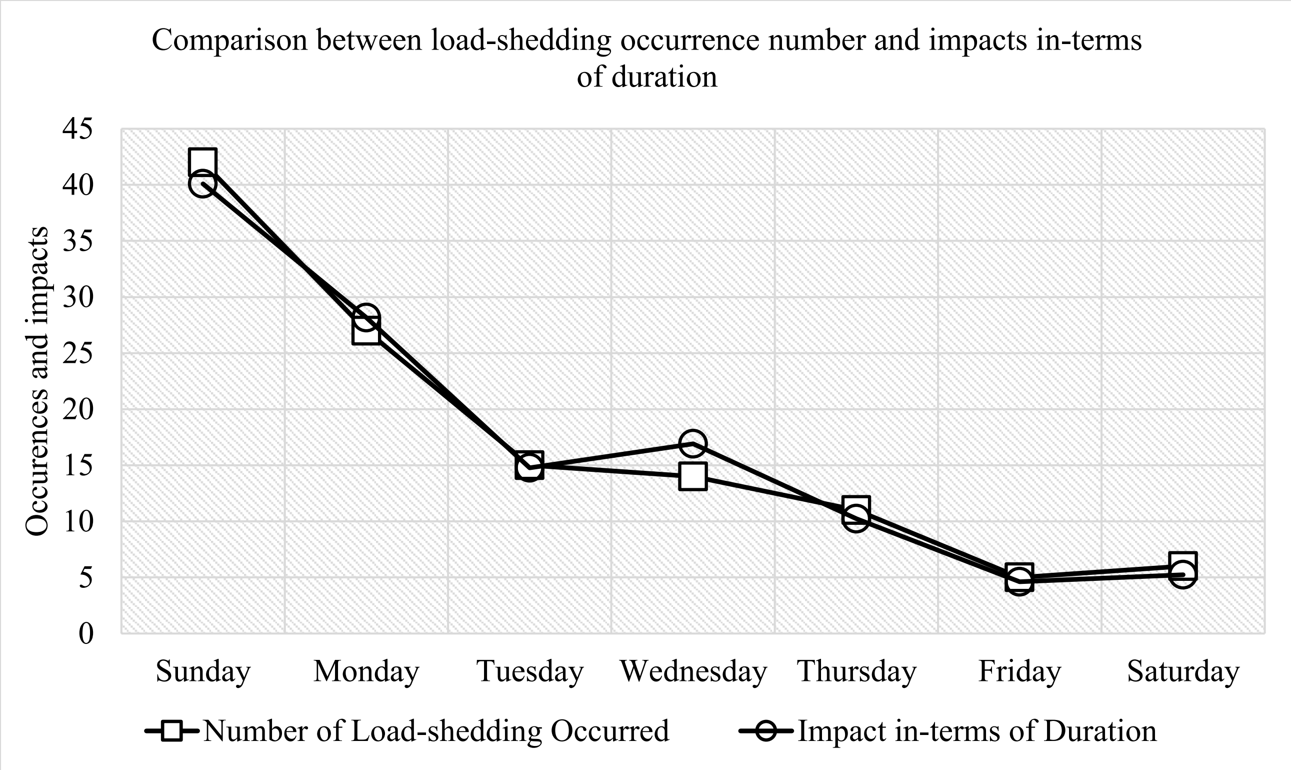


Fig. 22. Comparison between the day wise occurrences of load shedding against actual impact in terms of duration

It has been shown already that load shedding varies in different days of a week. However, in a particular day, this is not evenly distributed. If we divide the 24 hours into four equal chunks, one of this chunks is very prone to load shedding which has been rated as peak hours. Each day from 5 PM to 11 PM is considered as peak hours when the number and duration of load shedding is significantly higher than rest three chunks. Figure 23 shows that overall in Shah Ali, more than 60% load shedding takes place during peak hours of a day. This overall scenario can be better understood if feeder wise situation can be taken into consideration. Figure 49 in the appendix represents a feeder wise breakdown of this scenario.

Peak Quartile (5PM to 11PM) vs. Rest (11PM to 5PM) in Shah Ali Substation

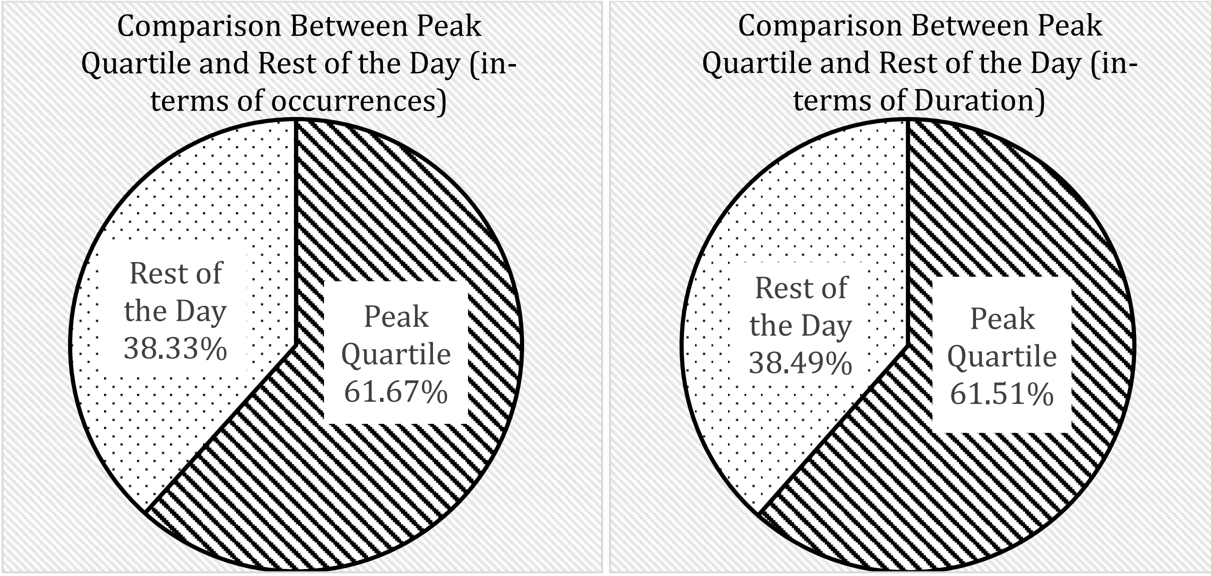


Fig. 23. Number of occurrences and duration of load shedding in peak quartile (5 PM to 11 PM) versus rest of the day (11 PM to 5 PM)

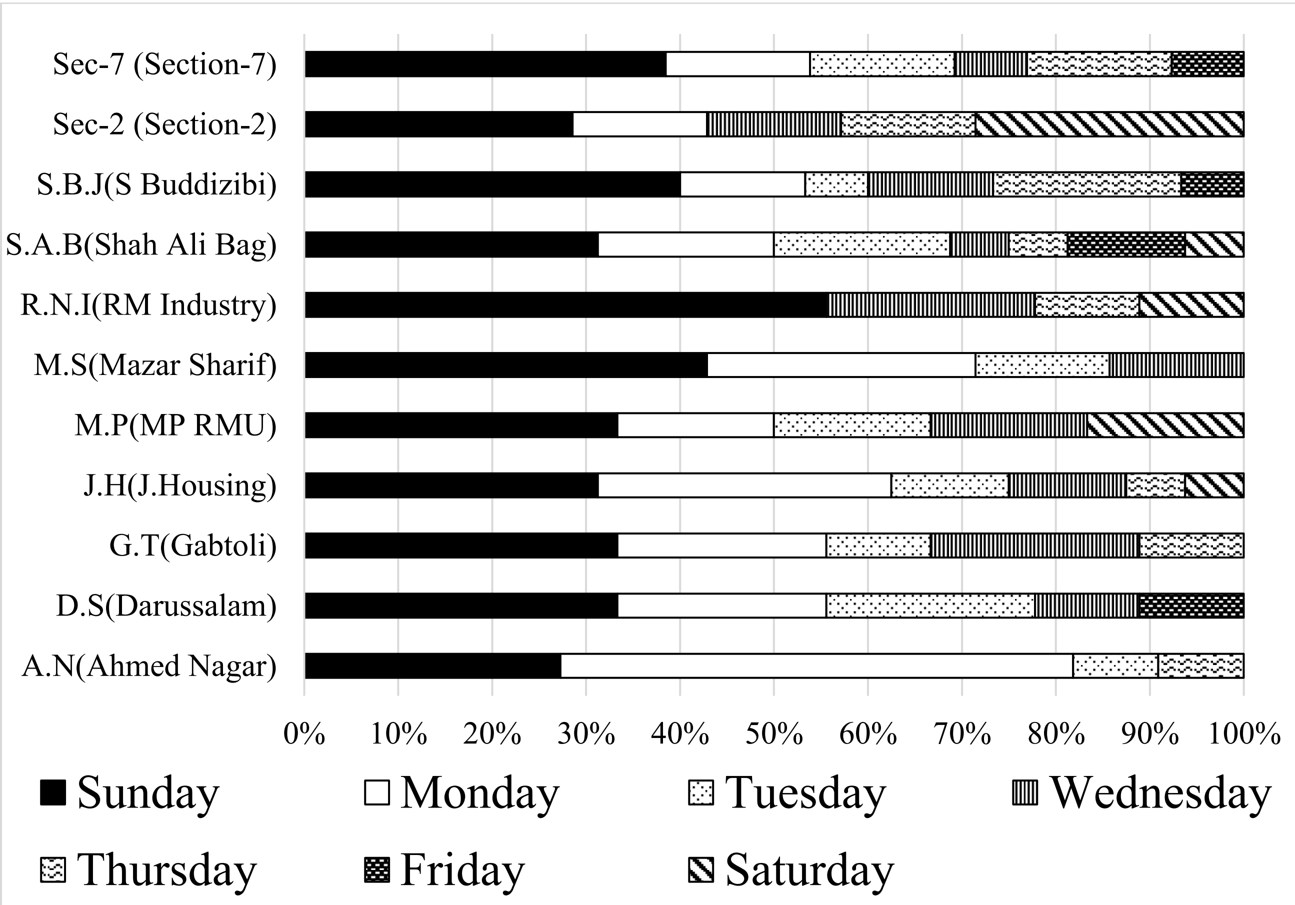


Fig. 24. Percentage of Load Shedding Occurrence on a Day at Shah Ali: More than 50% load shedding occurs during first two days of a week

Figure 24 represents a comparison of the percentage of load shedding occurrence on a day of a week between different feeders of Shah Ali substation. Around 30% of load shedding occurs on Sunday while 50% occurs in first two weekdays (Sunday and Monday). For some feeders, i.e. Mazar Sharif and Ahmed Nagar this crosses even 70%. So, satisfying the need of only two days of a week can reduce the load shedding for more than 50%. Here, the only snapshot of entire data has been used to show the statistics and showcased a glimpse of what can be achieved. Forecasting at this granular level will ensure better management, efficient generation and distribution which might have a positive impact on reducing load shedding.

4.2 Forecasting models and experimental results

Forecasting always will have uncertainty as we are predicting a future state which depends on several 'frequently changing' data like, meteorological data, economic status, consumption behaviour, historical data etc. We all that can ensure is a good forecasting method with a minimum possible error. A proper management and good forecasting require detailed and necessary data flow from consumer end and feeders to distributors/substations (for understanding area specific need and ensuring timely distribution) and to generation plants (to design an optimum generation schedule for satisfying the need). This section further focuses on consumption and load forecasting.

4.2.1 Consumption Forecasting and Performance Analysis

Consumer level prediction depends on several data points. Along with historical data, temperature data for understanding seasonal impacts, number and types of electrical appliances to understand possible consumption habit, household types, socioeconomic information of households etc. are required for an optimal forecasting model. The frequency of the data points is another factor while hourly consumption data and temperature data will outperform the accuracy of daily or weekly data. However, in our case, we have only month wise historical consumption data. To predict the consumption, we run a web scraping to collect month wise temperature data from World Weather Online website [36]. We run forecasting techniques on the data of Shah Ali zone from September 2014 to September 2015. A total of 13 months' data were selected where first 12 months data was for training purposes and September 2015 was for evaluating the performance.

TABLE 8. DIVISION OF CONSUMPTION DATA FOR TRAINING AND PERFORMANCE EVALUATION

Data Period: September 2014 to September 2015 (13 months)	
Training	Testing
September 2014 to August 2015 (12 months)	September 2015

Table 8 represents the division of data for training and testing. As input of the forecasting model though several key points are required, to reach an optimal performance, we could not manage all the sources. Household types and socioeconomic condition of households are available through Bangladesh Bureau of Statistics and we are still trying to get that the data from them. This data exchange requires permission and clearance and the overall process is very time-consuming. Another key factor, types and number of appliances data are not available, to the best of our knowledge. This would require an extensive survey which is also time-consuming and very much expensive. Having this challenges in mind, we proceed with the historical consumption data and web scrapped temperature data to forecast. As forecasting features, we have used previous month's consumption, monthly maximum, minimum and average temperature. We first tried with linear regression which obviously didn't perform well as variation in the data doesn't follow a linear pattern. We then run the neural network on the training data of Table 8 with 70% data in training, 15% in validation and 15% in testing to create a neural net. Later, testing data of September 2015 which was not a part of training process was sent to judge. Figure 25 represents the flow of our suggested model. Figure 26 and 27 represent actual and predictive output comparison. Green line denotes the actual output while blue line signifies the predictive outcome.

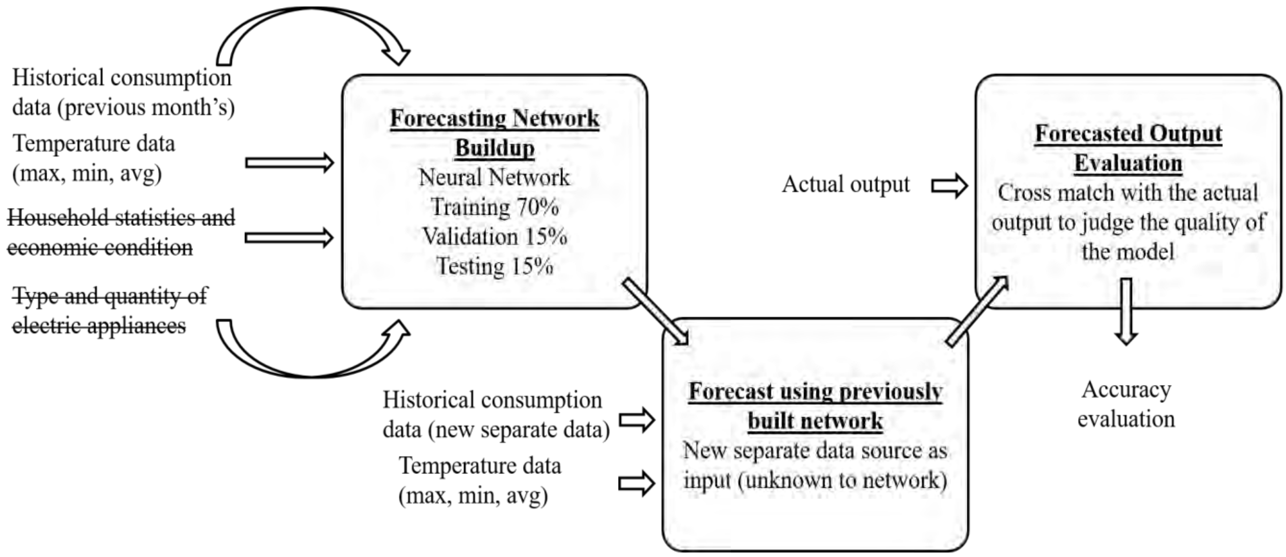


Fig. 25. Workflow of the forecasting network buildup and performance evaluation

Using the proposed model of Figure 25, we performed forecasting. In the Figure 25, household statistics and economic condition and type and number of electrical appliances as input sources are strikethroughs. These are very much important data sources and strongly suggested to be used as forecasting input. However, due unavailability of these sources, we omitted these from the input. Figure 26 and 27 represents actual and predictive output comparison. Green line denotes the actual output while blue line signifies the predictive outcome.

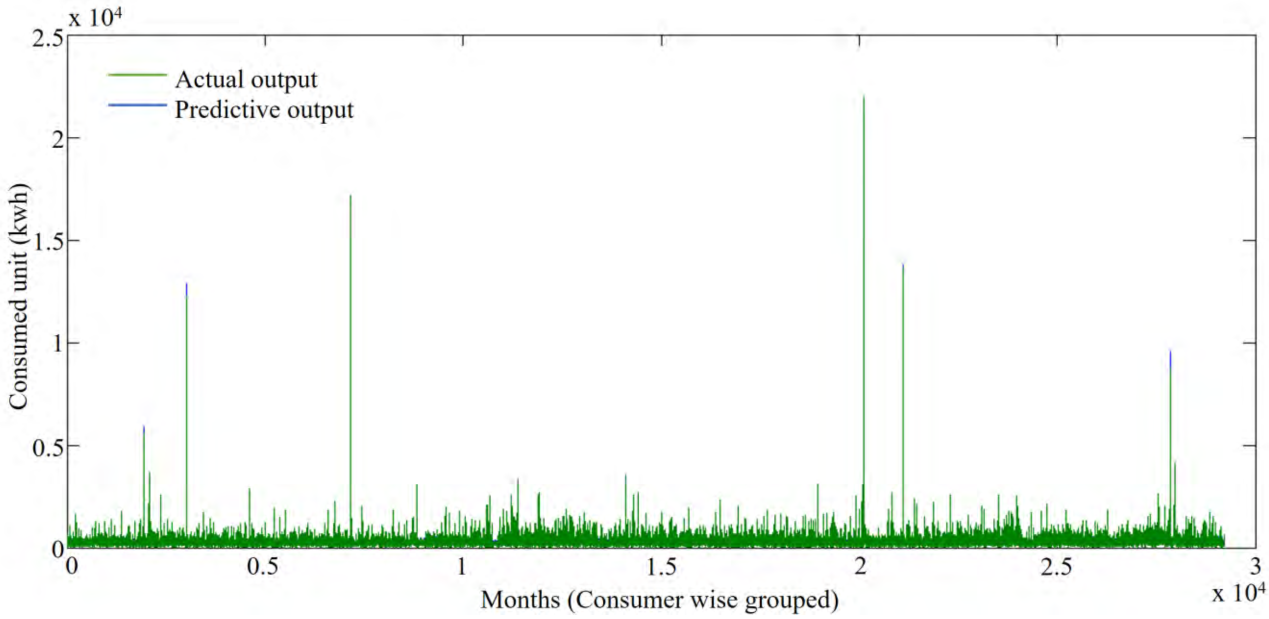


Fig. 26. Actual and predictive consumption of all consumers of Shah Ali for September 2015

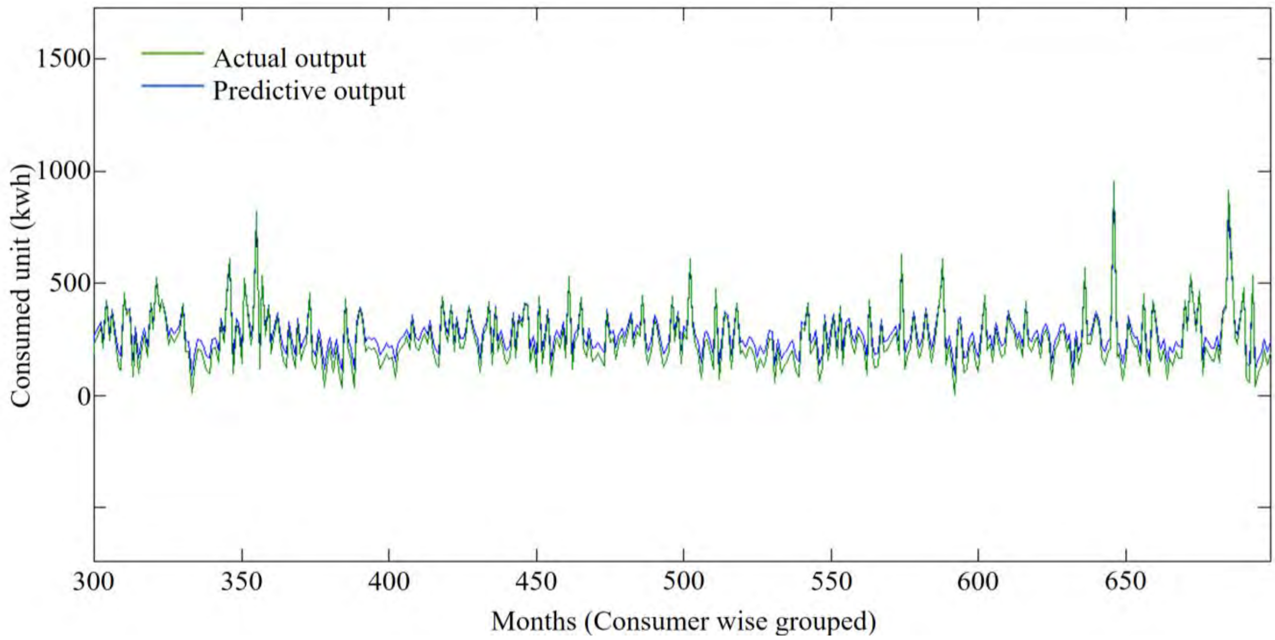


Fig. 27. Selected portion of actual and predictive consumption of all consumers of Shah Ali for September 2015

We have computed accuracy using MAPE (mean absolute percentage error) and MAE (mean absolute error) formulae and it turns out our predictive model is not performing well. We are achieving roughly a little above of 50% accuracy. The reason of such outcome is we are predicting consumer end consumption with a very little knowledge about that particular consumer. Everyone doesn't act in the same way for weather change; action depends on their consumption behaviour, economic status, appliances they are using and many more factors. The meter of uncertainty is very high here and this leads to such poor performance. As we do not have access to those data sources, we started to see our outputs from a different perspective. A prediction can go wrong in two ways, by overestimating and by underestimating. For example, the actual consumption of Mr X is 150 units and there are two predictive models, A and B. These two models predict X's consumption as 145 units and 155 units relatively. So, both of these models outcome are of 5 units variance from actual consumption. Though numerically they are wrong by the same absolute distance, these two models' impact is quite different on both consumers and policy makers. For model B, consumers will be planning for 155 units and the actual cost will be lower which will be comfortable to them than incurring a 5 units cost more by using model A. Also for policy makers, if they design the network by following model A, there will be shortage while B will ensure a surplus. Though surplus is not good for long-term planning as well; but, this will at least allow avoiding power shortage. I am bringing this debate because although our model's performance is not good, this can be followed as this is a B type model which will mostly show overestimation for inaccurate prediction and thus will assist to diminish power shortage due to demand side uncertainty.

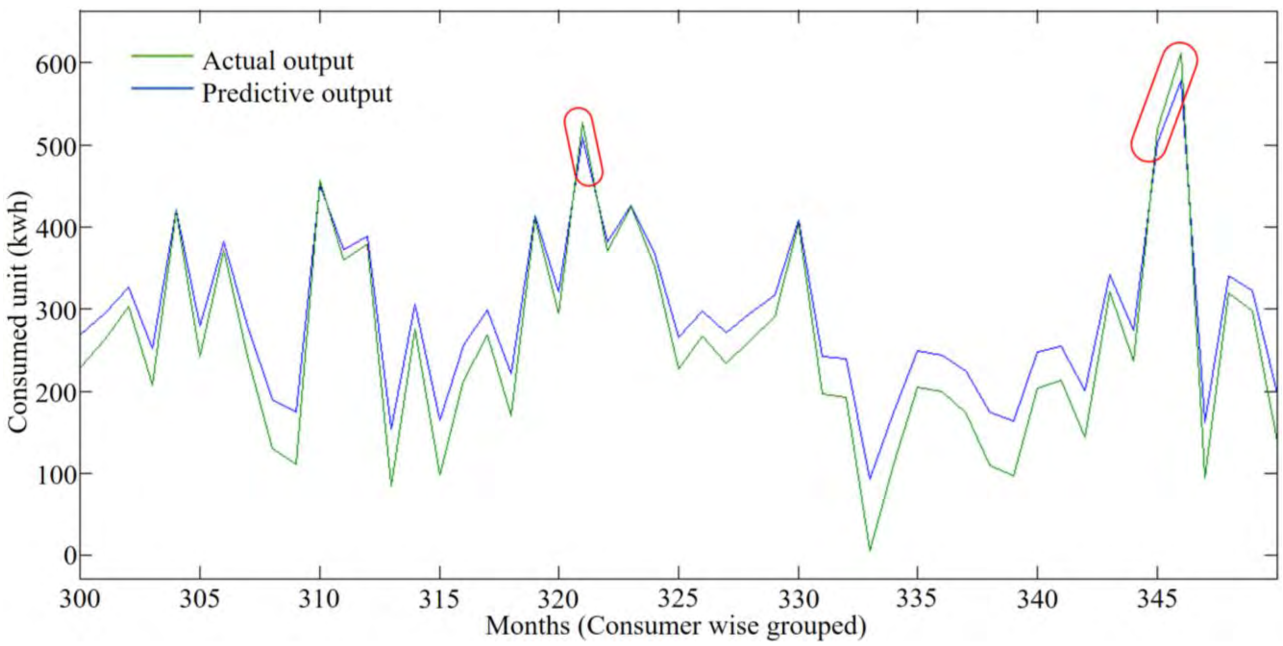


Fig. 28. Red marked points where proposed model underestimates the consumption

Figure 28 represents the red-marked portion where actual demand was more than the predictive demand. However, these underestimation is very rare in our model and if we consider the only underestimation as the wrong prediction, our accuracy bumps up to more than 95%. So, we can say less than 5% consumers in total will suffer a negative impact of forecasting while in other cases consumers will be satisfied with the forecasted consumption and this will also help policy makers to design an optimal network. Table 9 represents formulae of MAPE and MAE, modified formulae to judge only underestimation cases and accuracy of our neural network.

TABLE 9. PERFORMANCE ANALYSIS OF CONSUMPTION FORECASTING MODEL

	MAPE	**MAE**
Stock	$MAPE = \frac{100}{n}\sum_{j=1}^{n} \lvert\frac{\hat{y}_j - y_j}{\hat{y}_j}\rvert$	$MAE = \frac{1}{n}\sum_{j=1}^{n} \lvert y_j - \hat{y}_j\rvert$
Modified	$MAPE = \frac{100}{k}\sum_{j=1,k=0}^{n} \lvert x_j\rvert$ x_j = 0 if $(y_j > \hat{y}_j)$ Otherwise, x_j = $(\frac{\hat{y}_j - y_j}{\hat{y}_j})$ and $k = k + 1$	$MAE = \frac{1}{n}\sum_{j=1}^{n} \lvert x_j\rvert$ x_j = 0 if $(y_j > \hat{y}_j)$ Otherwise, x_j = $(y_j - \hat{y}_j)$
Accuracy	Stock: (100-47.5182) = 52.4818% Modified: (100-4.0168) = **95.9832%**	Stock: (100-46.8421) = 53.1579% Modified: (100-4.6083) = **95.3917%**

*y_j and $\hat{y}_j$are predictive and actual consumption of consumer j relatively

Section 01 of this chapter introduced the data points and discussed few key findings to support the claim that most of the electricity consumers are budget consumers, consumes only to run necessary appliances. Although, there number is significantly high, they are not equally distributed. For example, tongi had more borderline consumers (1 to 50 units) than any other zones in 2007 which has been distributed among other budget brackets over the year. Though, high consumption users are relatively low, Baridhara and Gulshan have more high consumption users than other zones. While consumption is rising over the year, a same seasonal impact has been seen in each year's consumption. Consumers from every bracket has a tendency of consuming low power during winter than summer. To understand a future demand, this variability needs to be studied and a better forecasting should be proposed. However, to develop a better forecasting model it requires several data sources as inputs in which we currently do not have access. A forecasting model has been proposed in this chapter to predict future consumption of a consumer based on the available sources. As few sources couldn't be managed, accuracy has been found around 55% which is not satisfactory; however, if only underestimation is considered, this model ensures a 95% of accuracy overall.

4.2.2 Load Forecasting and Performance Analysis

To forecast load, granular level data is a prerequisite such as half hourly or hourly historical load data, generation capacity, meteorological data (like hourly temperature, rain prediction index, humidity etc.), wealth index of an area, population of an area, number of active connection, frequency and impact of load shedding etc. More in depth data ensures more accurate forecasting generally. Unfortunately, all the data sources are not available to design an optimum model. As the load data and load shedding data is being recorded in a handwritten log book, this data need to be digitised before stepping into forecasting. Digitising such data is a time-consuming process. Because of such limitation, four parts of data has been digitised from two substations (two parts from each substation) in a one week span. One week of summer and one week of winter data has been digitised to compare seasonal load variability. On the basis of above scope analysis, with the currently available sources, a model has been proposed to perform a short term load prediction. For performing forecasting, one-week summer data of Shah Ali substation has been taken from 15th May 2017 to 21st May 2017; a total of 7 days. Daily temperature has been web scrapped from Weather Underground website [37] and based on the maximum and minimum temperature, a uniform distribution has been run to mimic the hourly temperature.

TABLE 10. DIVISION OF LOAD DATA FOR TRAINING AND PERFORMANCE EVALUATION

Data Period: 15th May 2015 to 21st May 2015 (1 week)	
Training	Testing
15th May 2015 to 20th May 2015 (6 days)	21st May 2015

Table 10 represents the division of data for training and testing. Two different forecasting model has been proposed varying the inputs and has shown that better forecasting can be achieved with more granular level data. These two models have been denoted as '5 input model' and '8 input model'. In 5 input model, an hour of the day, whether this is a peak hour or not, due point, temperature (hourly) and previous hour's load have been used as features to predict future demand. In 8 input model, along with the features of the first model, load of 24hours earlier, average load of the previous day and peak hours' average load of previous day have been used as predictors. Just like chapter 1, we first tried with linear regression which obviously didn't perform well as variation in the load data doesn't follow a linear pattern. We then run a neural network with 70% data in training, 15% in validation and 15% in testing to create a neural net. Later, testing data of 21st may 2015 which was not a part of training process was sent to judge.

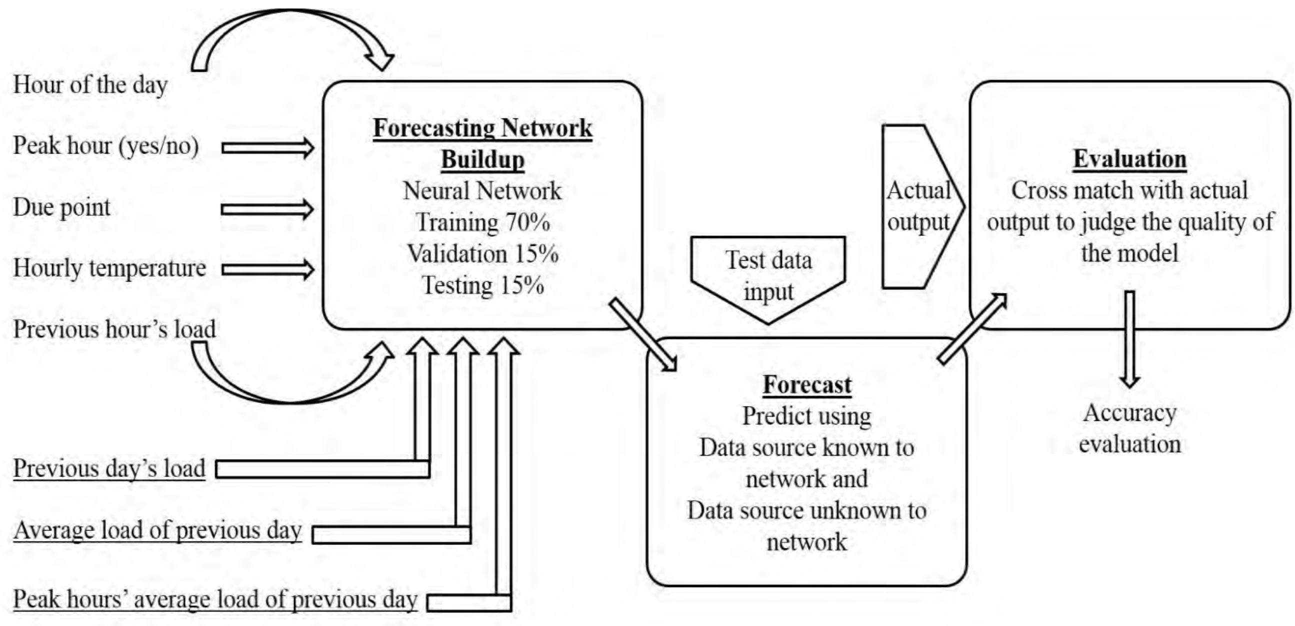


Fig. 29. Workflow of the load forecasting network buildup and performance evaluation

Figure 29 represents the flow of our proposed model. Underlined input sources are only for 8 input model. Both the train data (that has been used to build the model) and a separate test data has been run through the model. Figure 30-33 represent actual and predictive output comparison. Green line denotes the actual output while blue line signifies the predictive outcome.

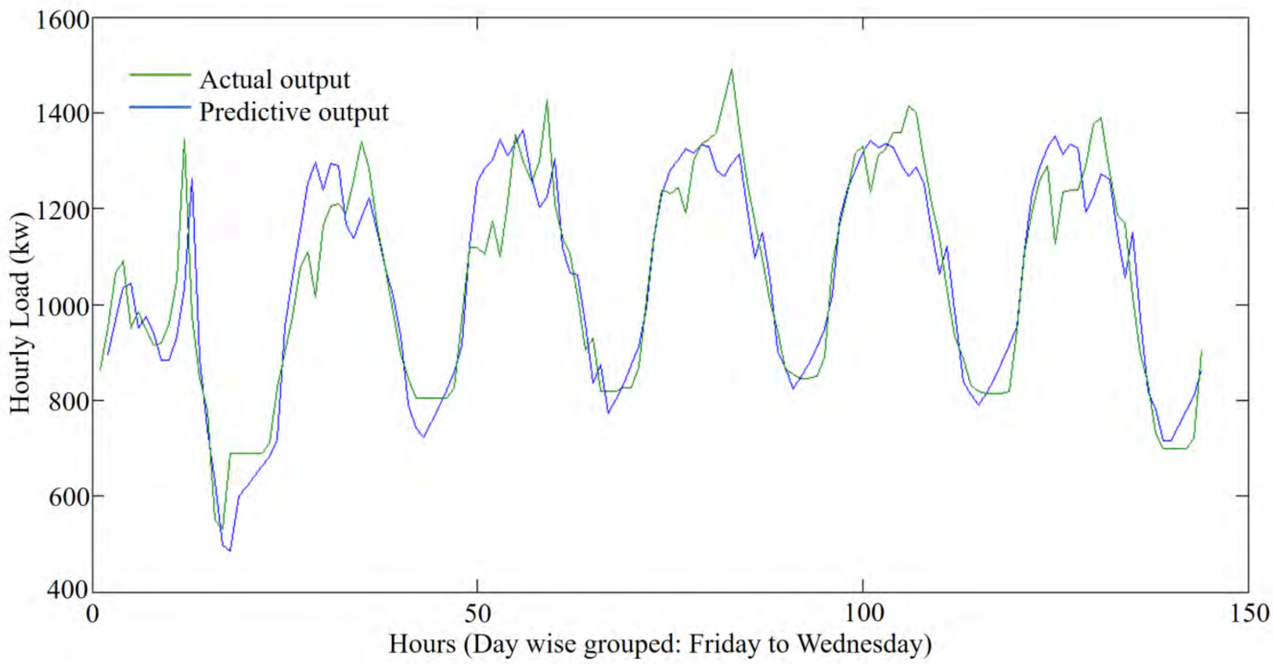


Fig. 30. Performance of the training set that has been used to train the network in 5 input model

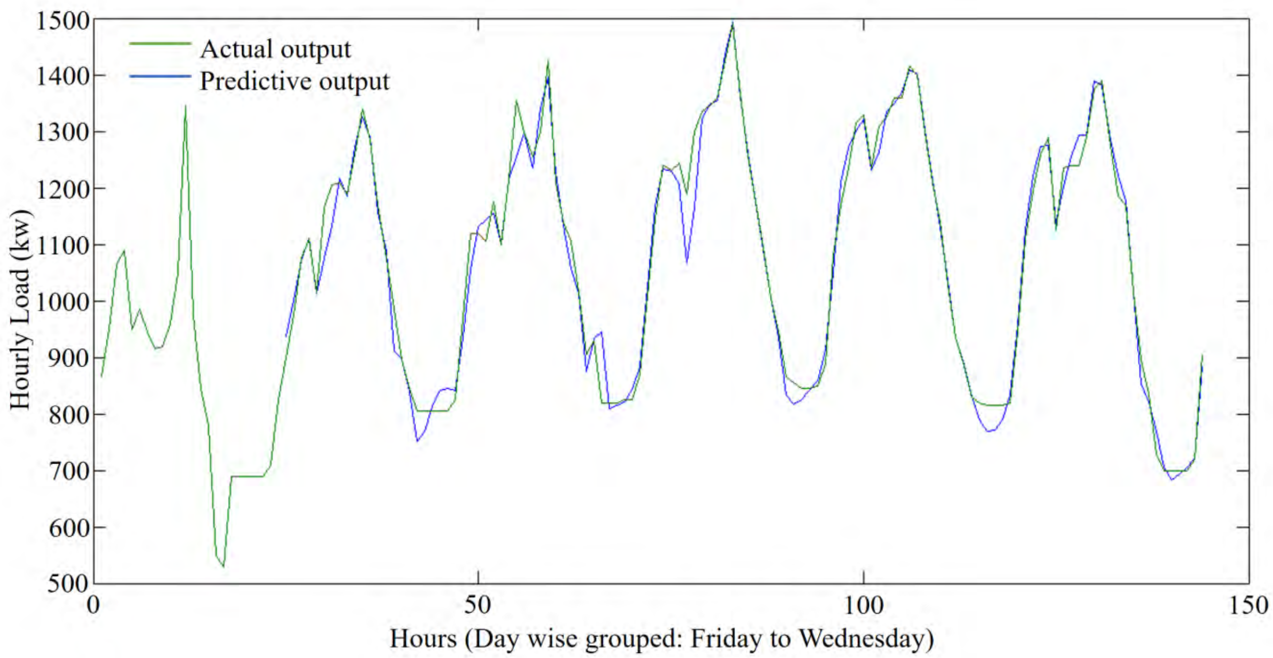


Fig. 31. Performance of the training set that has been used to train the network in 8 input model

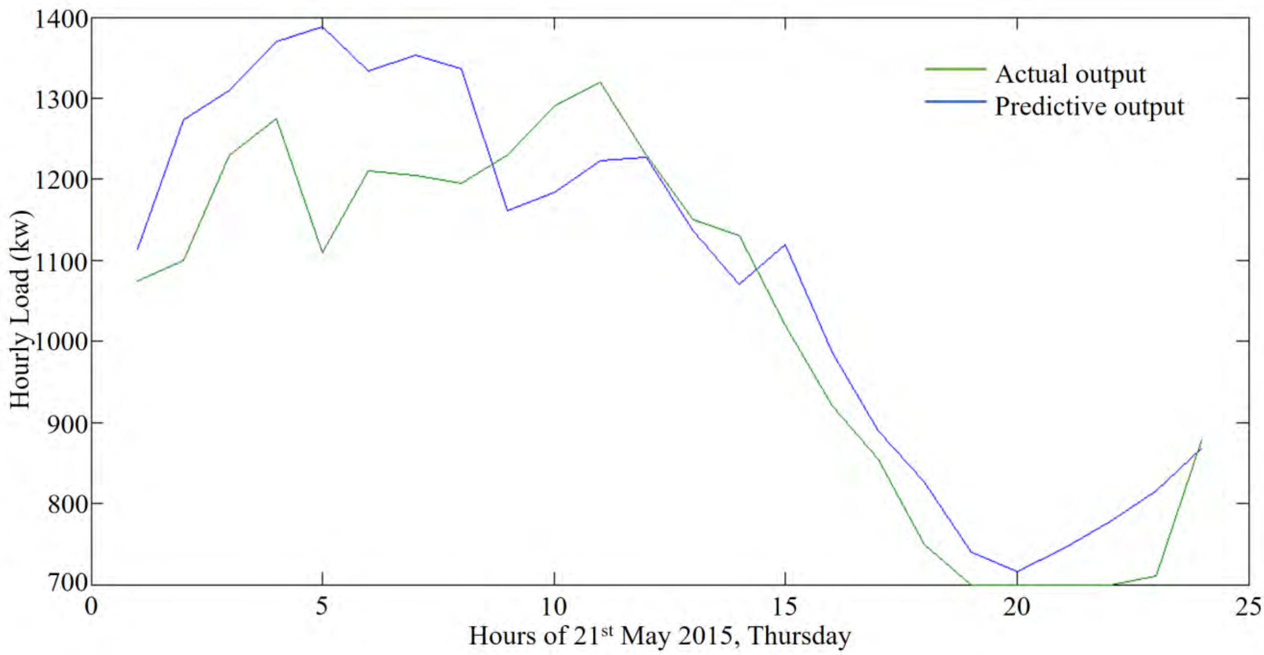


Fig. 32. Performance of the test set that has not been used to train the network in 5 input model

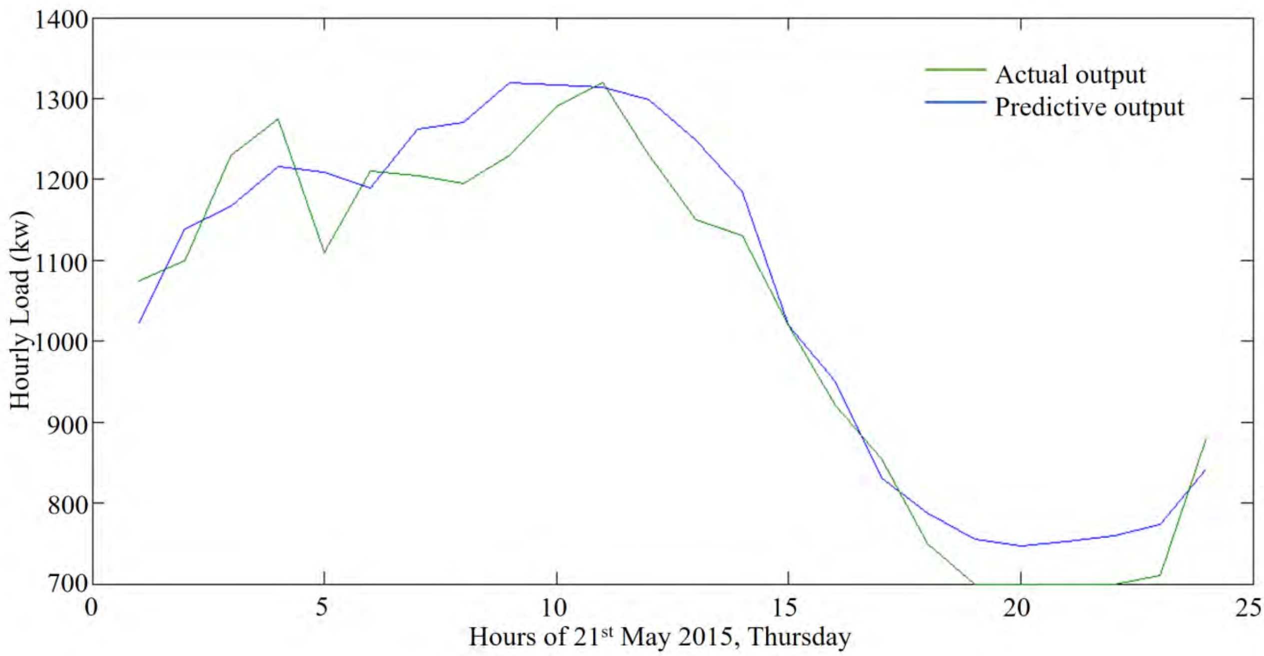


Fig. 33. Performance of the test set that has not been used to train the network in 8 input model

We have computed accuracy using MAPE formula (Table 9) and it turns out our 8 input predictive model is performing better. We have achieved around 95% of accuracy while in 5 input model, this is around 90% for unknown datasets (those not have been used to train the network).

Few more sources have been used in 8 input model which results in a better forecasting. Table 11 represents the accuracy of the models and comparison shows that 8 input model performs better.

TABLE 11. PERFORMANCE ANALYSIS LOAD FORECASTING MODELS

	5 input model	**8 input model**
Accuracy	Dataset known to model: (100-6.4817) = 93.5183% Dataset unknown to model: (100-8.0615) = **91.9385%**	Dataset known to model: (100-2.0672) = 97.9328% Dataset unknown to model: (100-5.1488) = **94.8512%**

Section 02 of chapter 04 covered a discussion on load data, its' variability, why load forecasting is needed, a brief statement of smart grid and necessity of studying state of the art forecasting techniques. A short review of recent forecasting techniques has been also showcased. Scope analysis of demand forecasting in Dhaka and two proposed short term forecasting models are the core of this chapter. Among those two, 8 input model performed better with an accuracy of approximately 95%.

CHAPTER 5

5.1 Conclusion

Proper management of electric energy is crucial for the development of any country. An accurate and adaptive electricity demand prediction can efficiently address the ever-asking question, how much electricity needs to be generated. The precise accuracy of prediction of electric demand is an essence of economic attainments; overly estimated demand can lead to paramount wastage of resources, conversely, lower estimation is imperial for the industrial growth having an adverse impact on the economy of a country. This paper provides both a digest and an analysis of the electricity consumers in eight zones of Dhaka. A series of statistical analysis has been presented to prove that electricity data can be a good heterogeneous data source to understand country's economy better. While this paper has not related to any other data source yet e.g. dwelling number of bedrooms, dwelling size, asset index of the resident, monthly income with electricity data, but the results are promising that it could be a good source to understand granular level socioeconomic status. Also, there is proof that urbanisation and development are ongoing in several zones e.g. Tongi-West zone has many low consumption users in 2007. But in recent year, evidence suggests those users got distributed within later tariff brackets and their number is increasing. It indicates that it is possible that rate of urbanisation and development is happening in that area rapidly. Later in this chapter, a consumption level forecasting technique has been proposed. In chapter 02, we have analysed a variety of aspect and approaches to predict electricity demand including studying different predictors which are extremely correlated to the target class, in addition to that, exhaustive research on machine learning algorithms and how those methods behave and perform diversely with the changing nature of the dataset have been studied. The study we have conducted provides us with indispensable instruments to predict electricity consumption and demand for a developing nation.

5.2 Future Work

Further work can be done by conducting several surveys to specific households to find out what kind of electrical appliance is used by residents and cross match with consumption data. Along with the data collected from survey, data of Bangladesh Bureau of Statistics (BBS) can be used as input to engage a much more granular level consumption forecasting. By ensuring a better consumption prediction, we can develop a system e.g. web or mobile app to let consumers know about their consumption habit, month wise forecasting and to route them to their own optimum consumption bracket. Results of these research can be used to the betterment of societies where these evidences can be linked to the policy level implications. From the perspective of load data findings, peak usage hours can be used to open another research track where scopes of semi-dynamic tariff planning concept can be analyzed. Another future implication can be combining load data with other heterogeneous sources e.g. CDR data, transportation data etc. to understand day to day urban movement of city people and their area wise consumption impact.

REFERENCES

[1] R. Pachauri, "Energy transformed: sustainable energy solutions for climate change mitigation," CSIRO & The Natural Edge Project, Melbourne, 2007.

[2] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting—A review," *Renewable and Sustainable Energy Reviews,* vol. 16, pp. 1223-40, 2012.

[3] "Which are the world's fastest-growing economies?", World Economic Forum, 2017. [Online]. Available: https://www.weforum.org/agenda/2016/04/worlds-fastest-growing-economies. [Accessed: 20- Jun- 2017].

[4] R. Ferguson, W. Wilkinson and R. Hill, "Electricity use and economic development", Energy Policy, vol. 28, no. 13, pp. 923-934, 2000.

[5] M. Golam Ahamad and A. Nazrul Islam, "Electricity consumption and economic growth nexus in Bangladesh: Revisited evidences", Energy Policy, vol. 39, no. 10, pp. 6145-6150, 2011.

[6] G. Altinay and E. Karagol, "Electricity consumption and economic growth: Evidence from Turkey", Energy Economics, vol. 27, no. 6, pp. 849-856, 2005.

[7] J. Steele, P. Sundsøy, C. Pezzulo, V. Alegana, T. Bird, J. Blumenstock, J. Bjelland, K. Engø-Monsen, Y. de Montjoye, A. Iqbal, K. Hadiuzzaman, X. Lu, E. Wetter, A. Tatem and L. Bengtsson, "Mapping poverty using mobile phone and satellite data", Journal of The Royal Society Interface, vol. 14, no. 127, p. 20160690, 2017.

[8] M. Xie, N. Jean, M. Burke, D. Lobell and S. Ermon, "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping", Arxiv.org, 2017. [Online]. Available: http://arxiv.org/abs/1510.00098. [Accessed: 10- Jun- 2017].

[9] J. Blumenstock, G. Cadamuro and R. On, "Predicting poverty and wealth from mobile phone metadata", Science, vol. 350, no. 6264, pp. 1073-1076, 2015.

[10] J. Falkingham and C. Namazie, "Measuring health and poverty: a review of approaches to identifying the poor", http://www.eldis.org, 2017. [Online]. Available: http://www.heart-resources.org/wp-content/uploads/2012/10/Measuring-health-and-poverty.pdf. [Accessed: 15- Jun- 2017].

[11] C. Mellander, J. Lobo, K. Stolarick and Z. Matheson, "Night-Time Light Data: A Good Proxy Measure for Economic Activity?", PLOS ONE, vol. 10, no. 10, p. e0139779, 2015.

[12] A. Katal, M. Wazid and R. Goudar, "Big data: Issues, challenges, tools and Good practices", 2013 Sixth International Conference on Contemporary Computing (IC3), 2013.

[13] "These are the world's most crowded cities", World Economic Forum, 2017. [Online]. Available: https://www.weforum.org/agenda/2017/05/these-are-the-world-s-most-crowded-cities/. [Accessed: 10- Jul- 2017].

[14] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: literature survey and classification of methods," *International Journal of Systems Science,* vol. 33, pp. 23-34, 2002.

[15] D. J. Pedregal and J. R. Trapero, "Mid-term hourly electricity forecasting based on a multi-rate approach," *Energy Conversion and Management,* vol. 51, pp. 105-11, 2010.

[16] R. Weron, Modeling and forecasting electricity loads and prices: a statistical approach, England: John Wiley & Sons Ltd., 2006.

[17] M. G. Al-Shakarchi and M. M. Ghulaim, "Short-term load forecasting for baghdad electricity region," *Electric Machines and Power Systems,* vol. 28, pp. 355-71, 2000.

[18] H. S. Keyno, F. Ghaderi, A. Azade and J. Razmi, "Forecasting electricity consumption by clustering data in order to decrease the periodic variable's effects and by simplifying the pattern," *Energy Conversion and Management,* vol. 50, pp. 829-36, 2009.

[19] E. Almeshaiei and H. Soltan, "A methodology for electric power load forecasting," *Alexandria Engineering Journal,* vol. 50, pp. 137-44, 2011.

[20] "Green Energy Reporter, LLC," [Online]. Available: http://greenenergyreporter.com/renewables/cleantech/china-to-invest-7-3-bln-in-smart-grid-projects-in-2010.

[21] T. Jamal and W. Ongsakul, "Smart Grid in Bangladesh power distribution system: Progress & prospects," in *Students Conference on Engineering and Systems*, 2012.

[22] "Turkey Electricity Statistic 2010," Turkish Electricity Transmission Company (TETC), Ankara, Turkey, 2012.

[23] "National Institute of Standards and Technology," [Online]. Available: https://www.nist.gov/engineering-laboratory/smart-grid.

[24] "Federal Energy Regulatory Commission Assessment of Demand Response & Advanced Metering," United States Federal Energy Regulatory Commission, 2008.

[25] H. Hahn, S. Meyer-Nieberg and S. Pickl, "Electric Load Forecasting Methods: Tools for Decision Making," *European Journal of Operational Research,* vol. 199, pp. 902-7, 2009.

[26] F. Kaytez, M. C. Taplamacioglu, E. Cam and F. Hardalac, "Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines," *Electrical Power and Energy Systems,* vol. 67, pp. 431-8, 2015.

[27] "Turkish Electrical Energy 10-Year Generation Capacity Projection," Turkish Electricity Transmission Company (TETC), Ankara, Turkey, 2009.

[28] "Annual development of Turkey’s gross electricity generation–imports–exports and demand," Turkish Electricity Transmission Company (TETC), Ankara, Turkey, 2012.

[29] "Population, annual population growth rate and projections," Turkish Statistical Institute (TURKSTAT), Ankara, Turkey, 2009.

[30] "Turkey Distribution Statistics," Turkish Electricity Distribution Company (TEDC), Ankara, Turkey, 2009.

[31] B. Koo, S. Lee, W. Kim and J. Park, "Comparative Study of Short-term Electric Load Forecasting," in *Fifth International Conference on Intelligent Systems, Modelling and Simulation*, 2014.

[32] C. M. Lee and C. N. Ko, "Short-Term Load Forecasting Using Adaptive Annealing Learning Algorithm Based Reinforcement Neural Network," *Energies,* vol. 9, p. 987, 2016.

[33] S. Jurado, À. Nebot, F. Mugica and N. Avellana, "Hybrid methodologies for electricity load forecasting: Entropy-based feature selection with machine learning and soft computing techniques," *Energy,* vol. 86, pp. 276-91, 2015.

[34] "Annual Reports", Dhaka Electric Supply Co Limited [BD], 2017. [Online]. Available: https://www.desco.org.bd/?page=annual-reports. [Accessed: 18- Jun- 2017].

[35] "Survey on Power System Master Plan 2015: Draft Final Report [PSMP2015]," Ministry of Power, Energy and Mineral Resources; Bangladesh Power Development Board, Bangladesh, 2015.

[36] "Word Weather Online website". [Online]. Available: https://www.worldweatheronline.com/dhaka-weather-averages/bd.aspx. [Accessed: 18-Jun-2017].

[37] "Weather Underground website". [Online]. Available: https://www.wunderground.com/history. [Accessed 17 01 2017].

APPENDIX

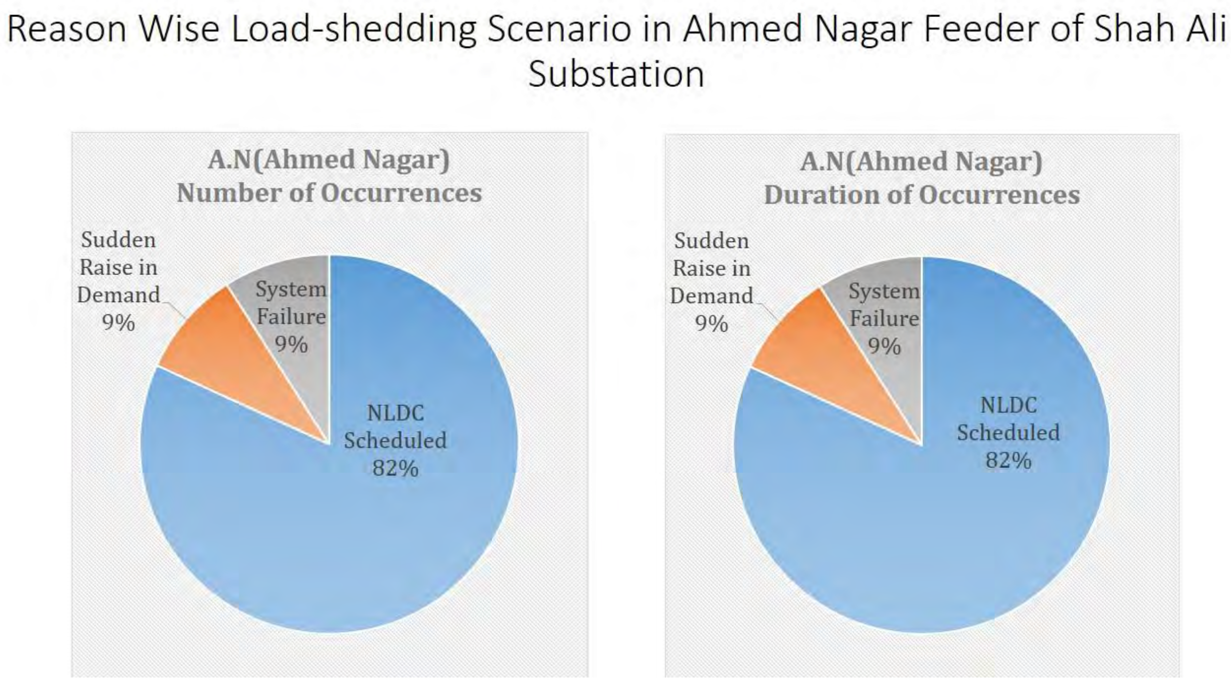


Fig. 34. Reason wise load shedding scenario of Ahmed Nagar

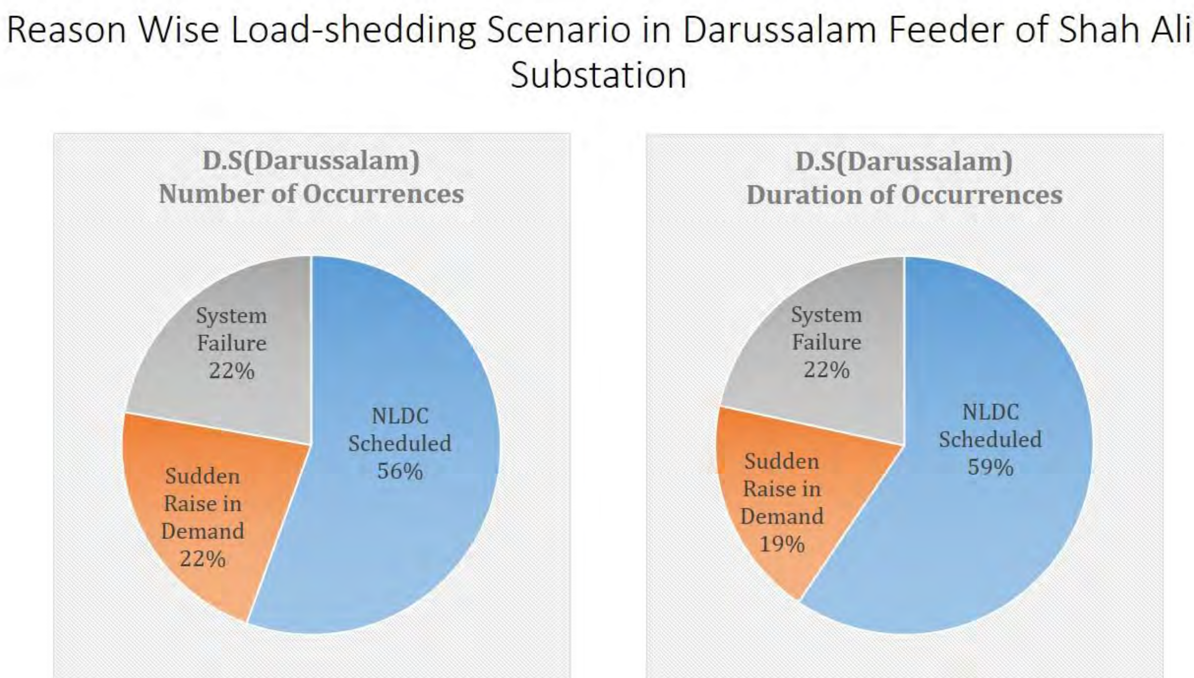


Fig. 35. Reason wise load shedding scenario of Darussalam

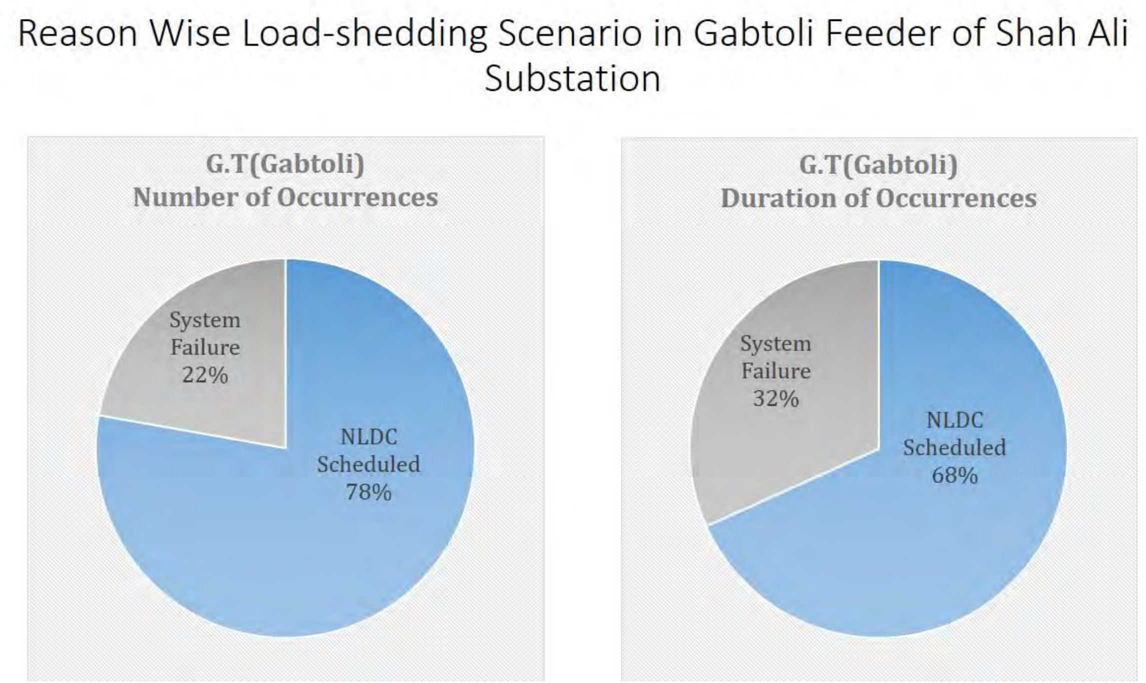


Fig. 36. Reason wise load shedding scenario of Gabtoli

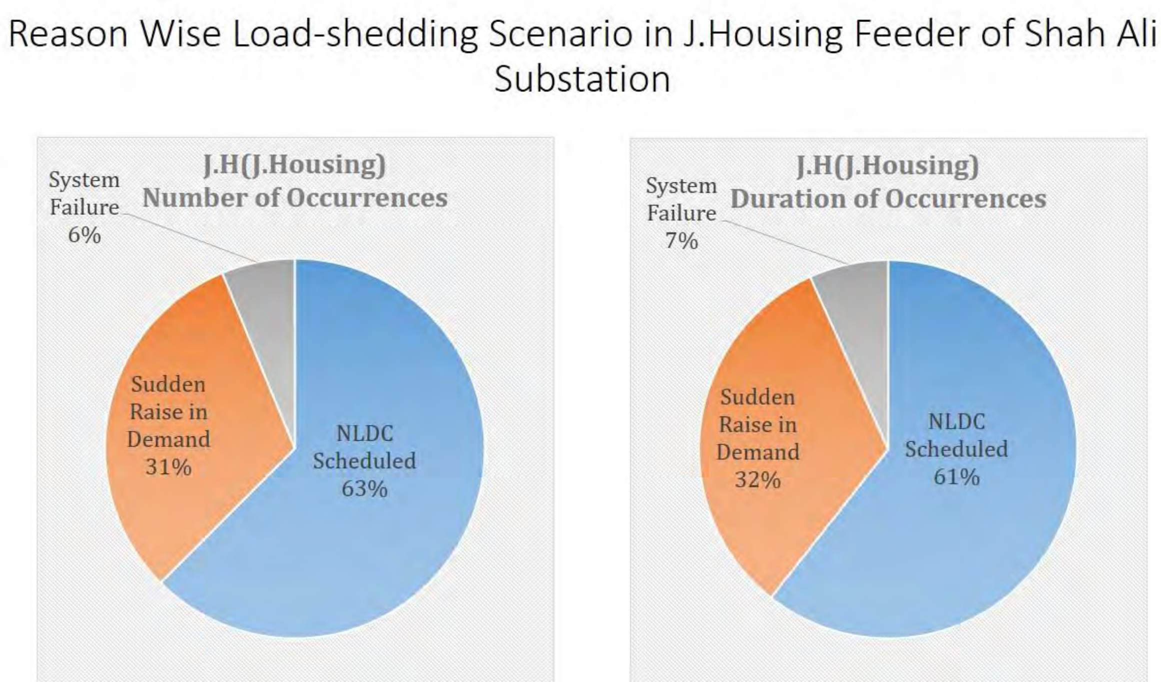


Fig. 37. Reason wise load shedding scenario of J. Housing

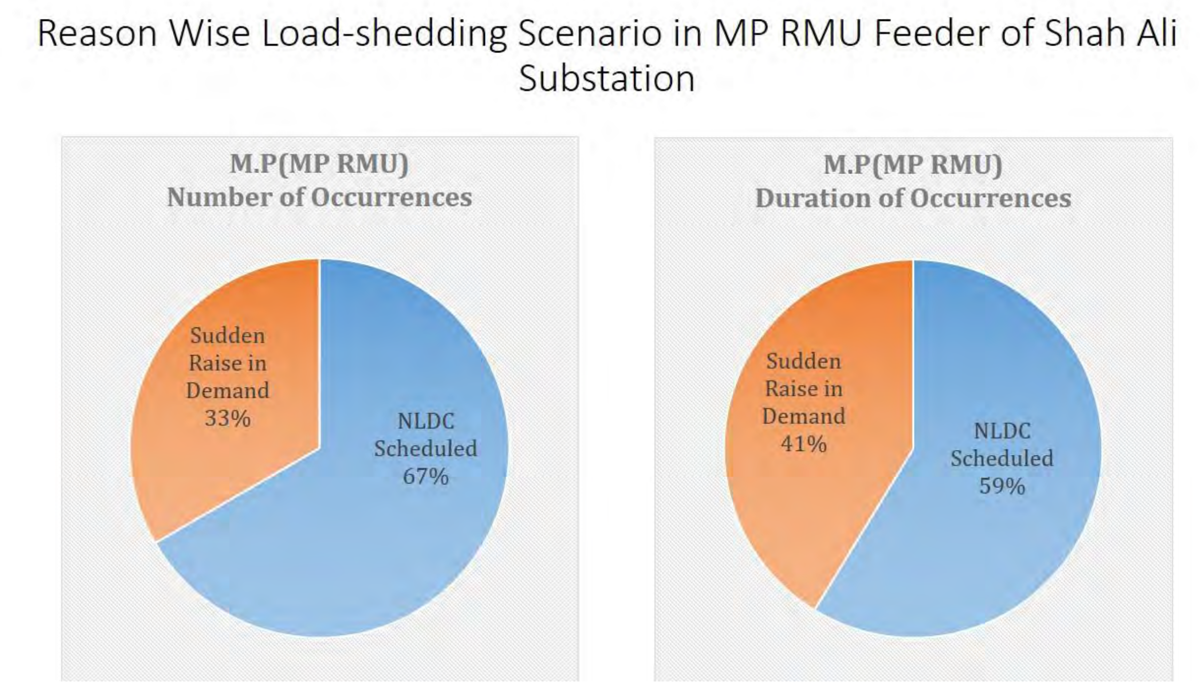


Fig. 38. Reason wise load shedding scenario of MP RMU

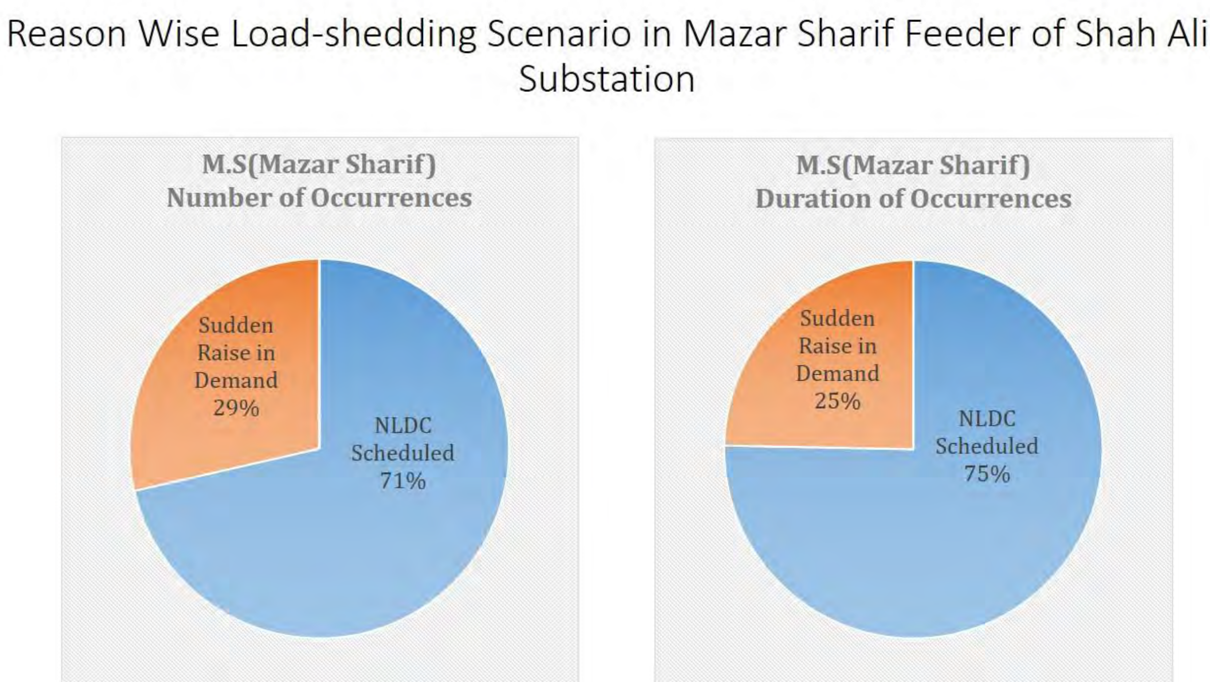


Fig. 39. Reason wise load shedding scenario of Mazar Sharif

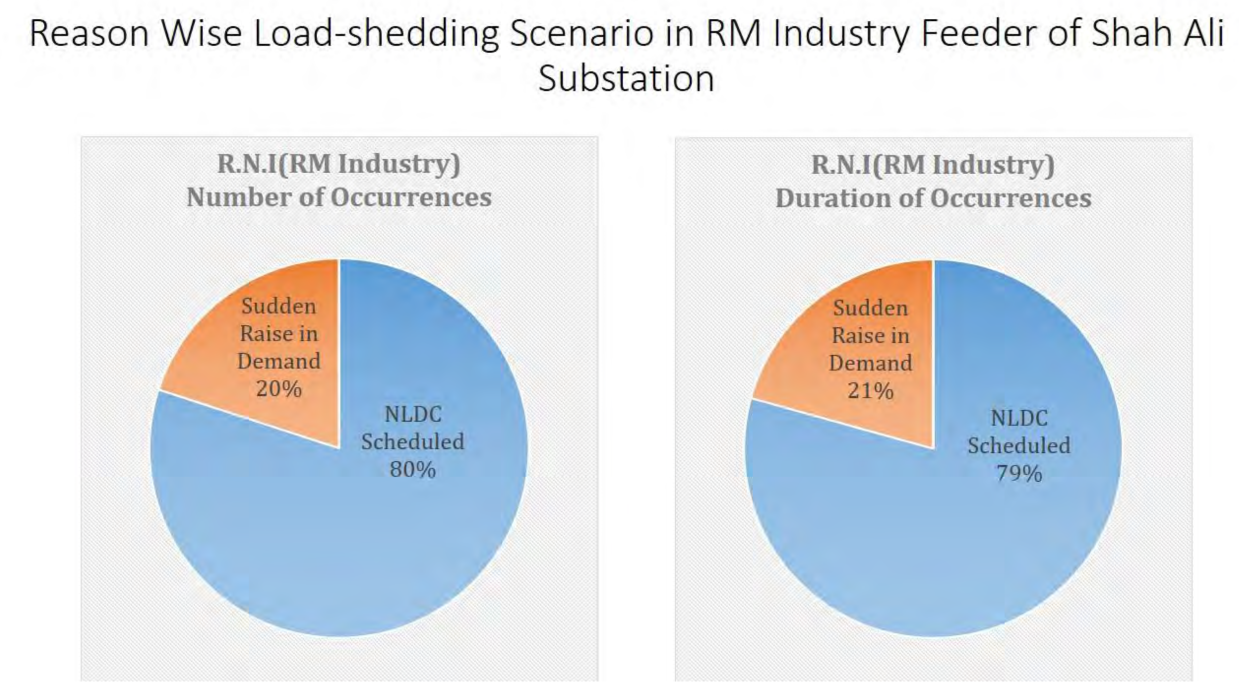


Fig. 40. Reason wise load shedding scenario of RM Industry

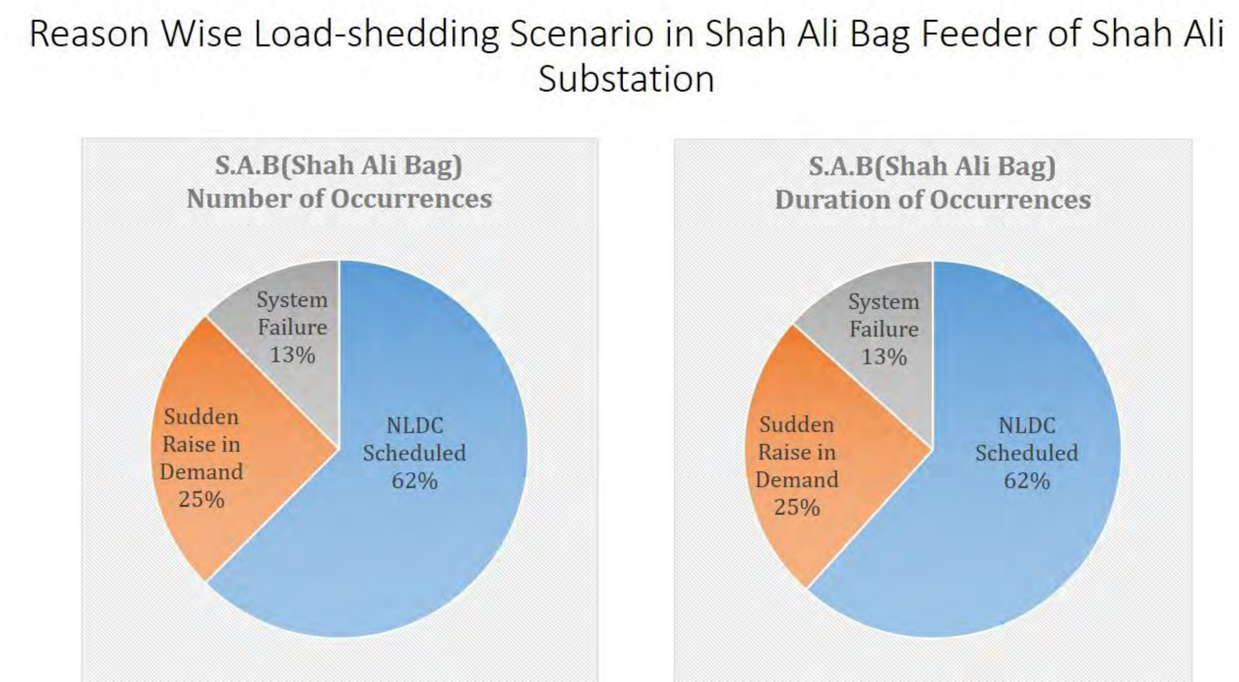


Fig. 41. Reason wise load shedding scenario of Shah Ali Bag

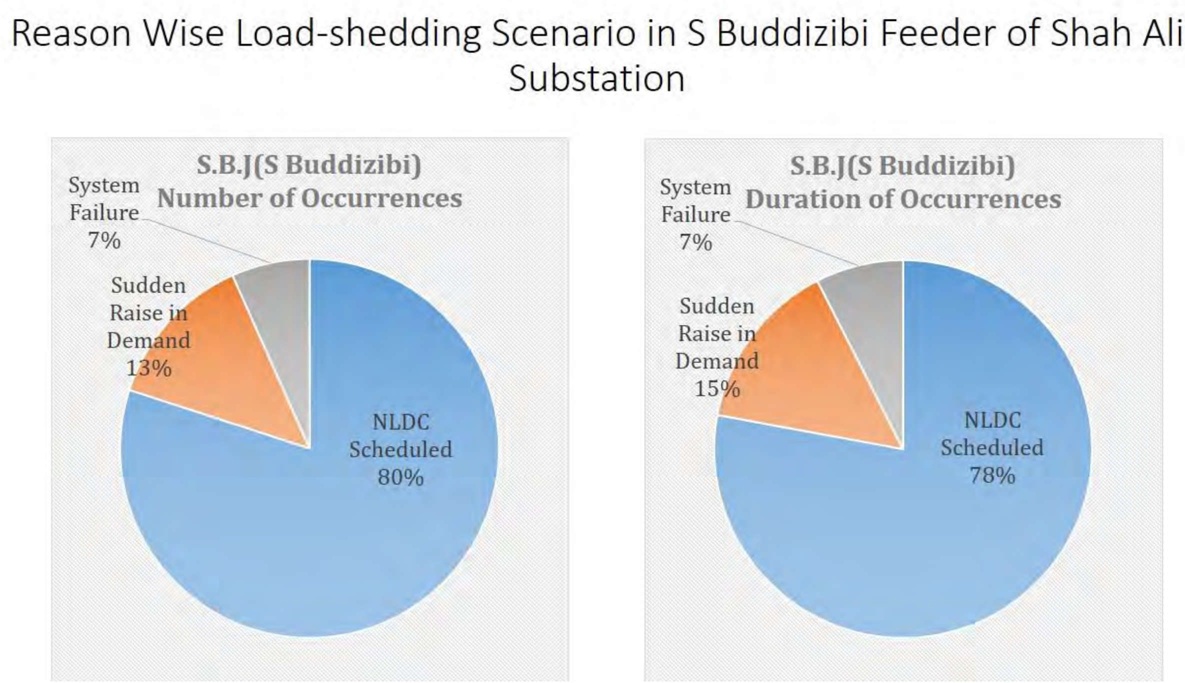


Fig. 42. Reason wise load shedding scenario of S. Buddizibi

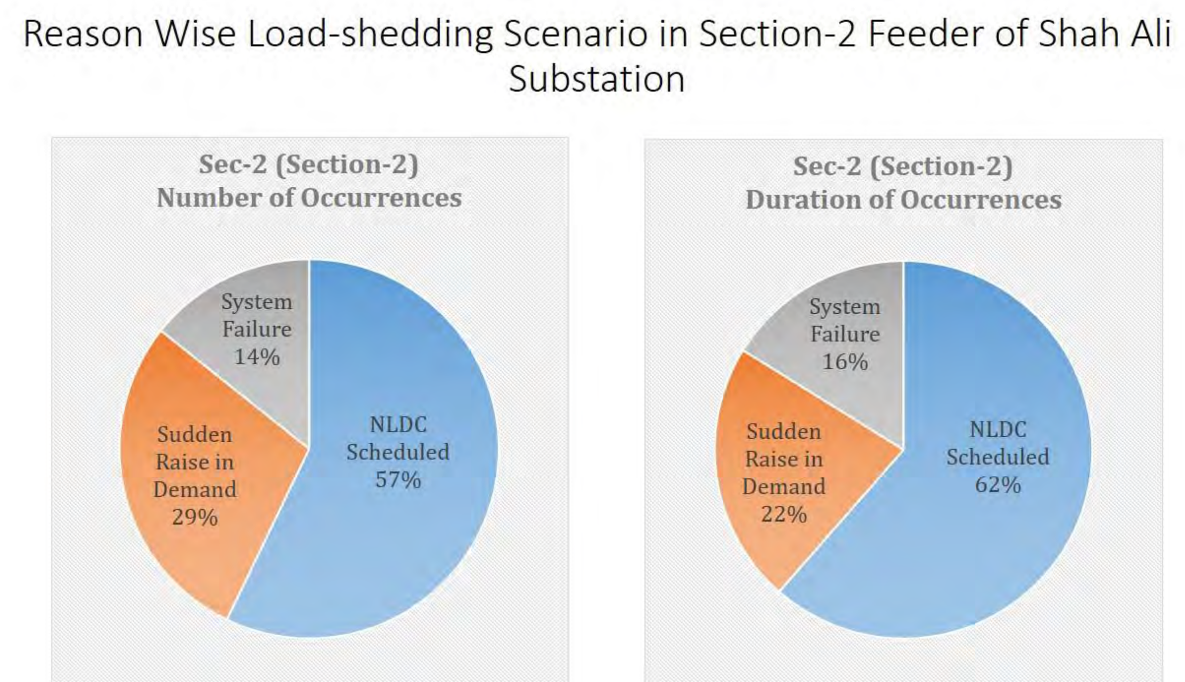


Fig. 43. Reason wise load shedding scenario of Section-2

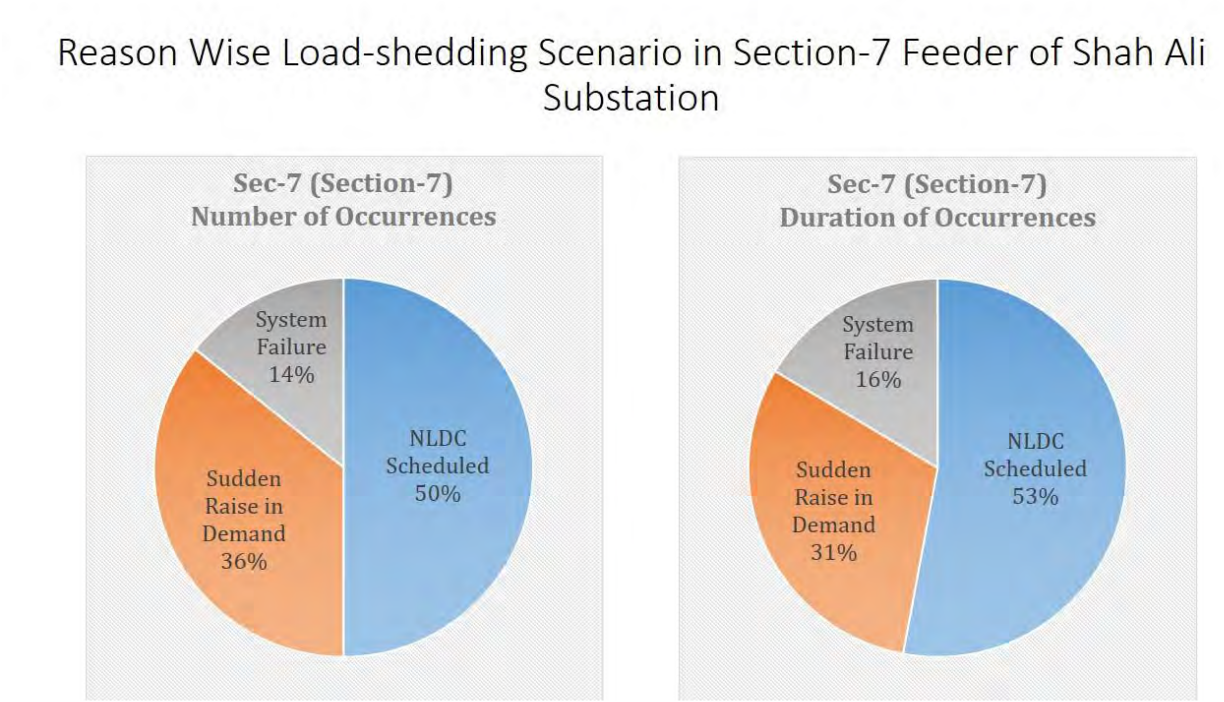


Fig. 44. Reason wise load shedding scenario of Section-7

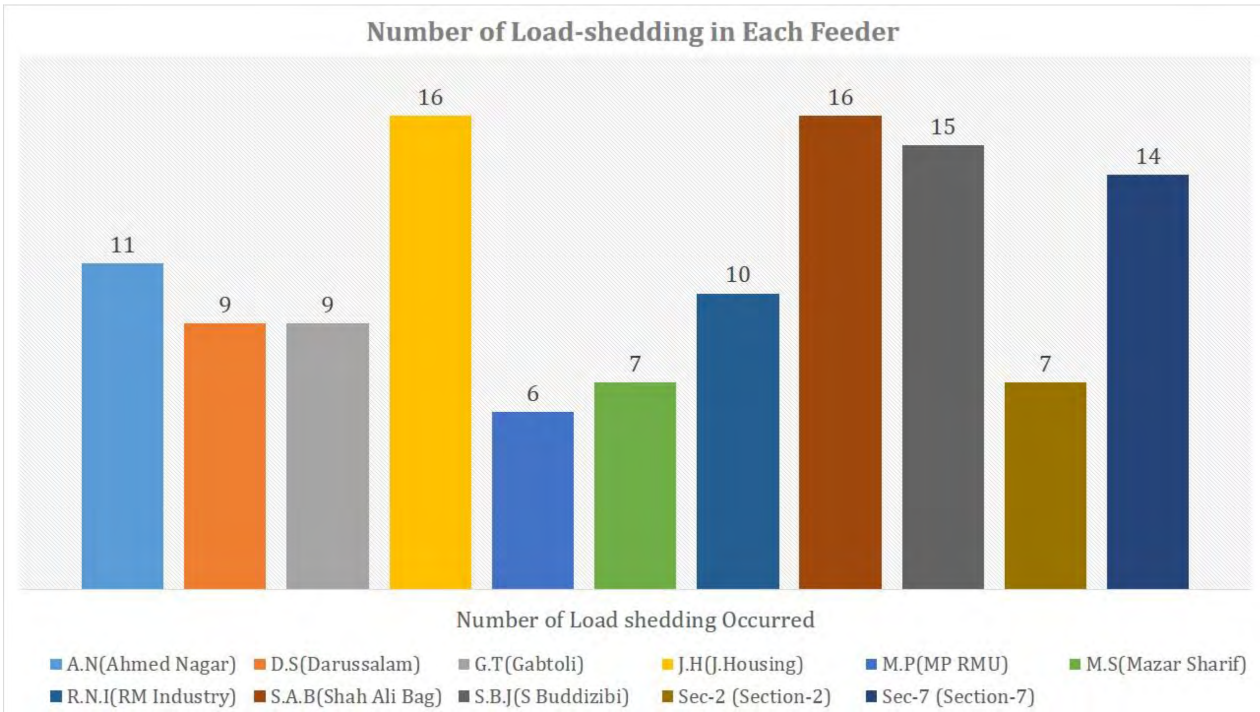


Fig. 45. Differences in number of load shedding occurrences in different feeders of Shah Ali

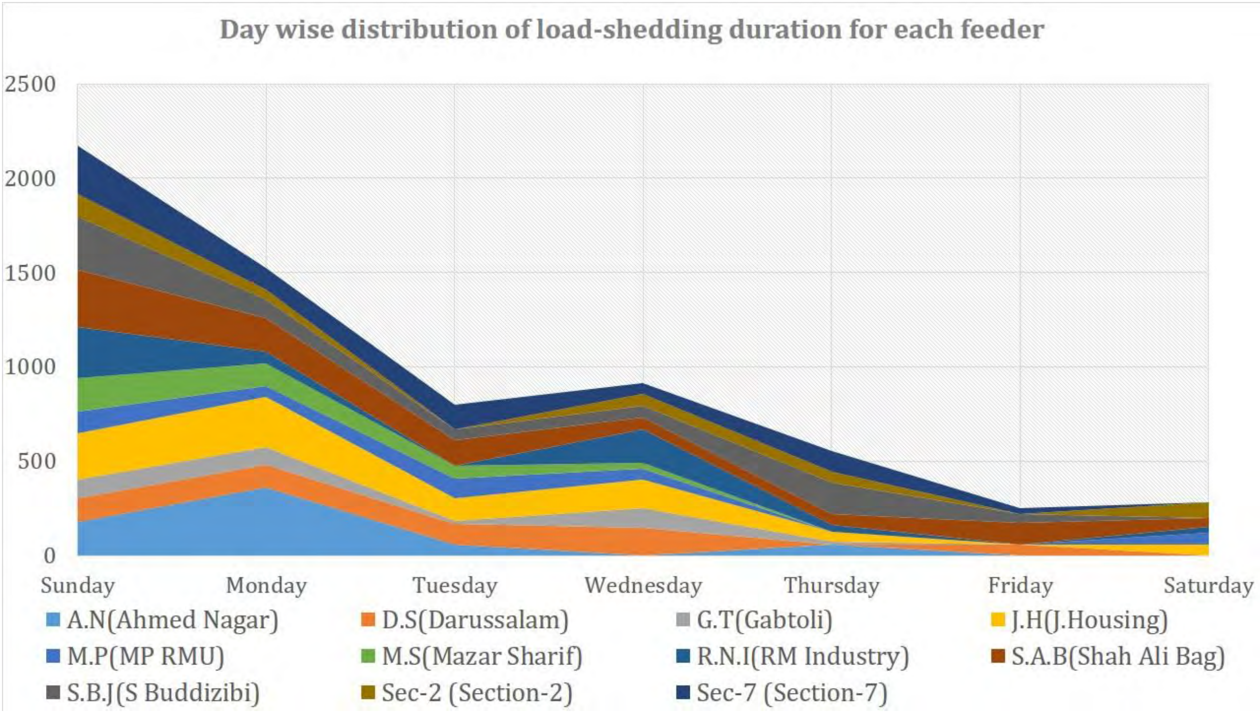


Fig. 46. Day wise stacked area chart of impact of load shedding duration

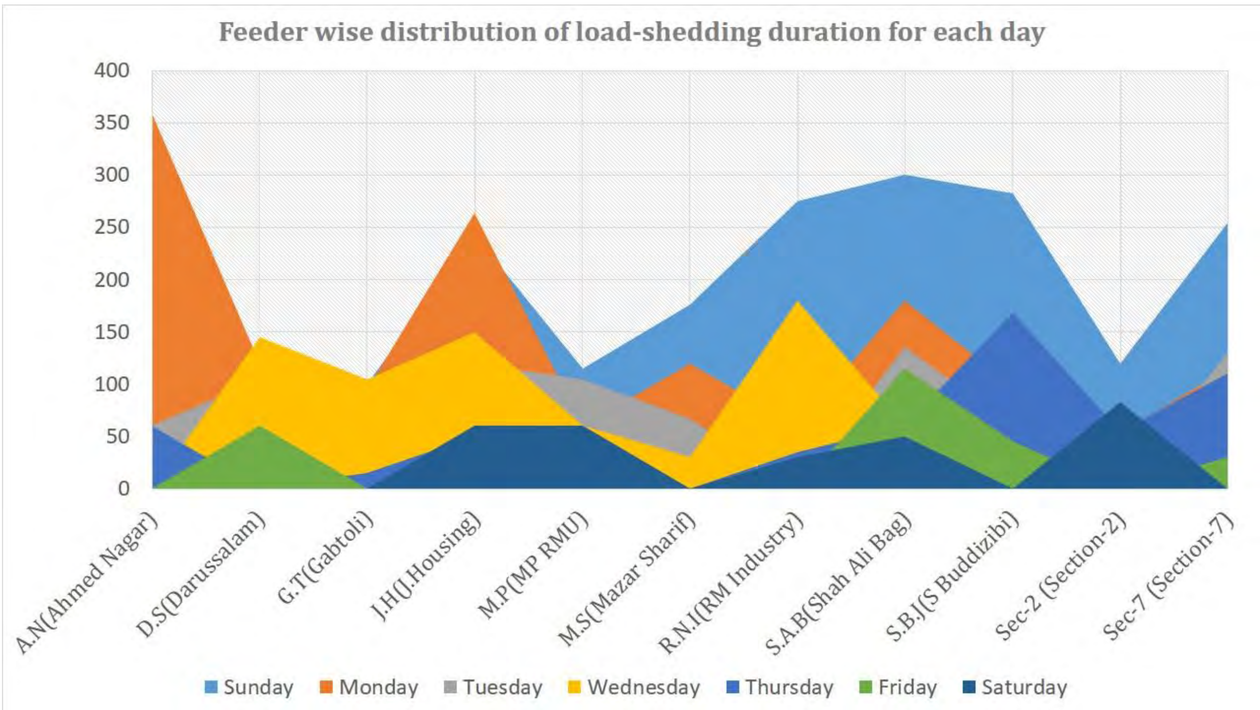


Fig. 47. Feeder wise layered area chart of impact of load shedding duration

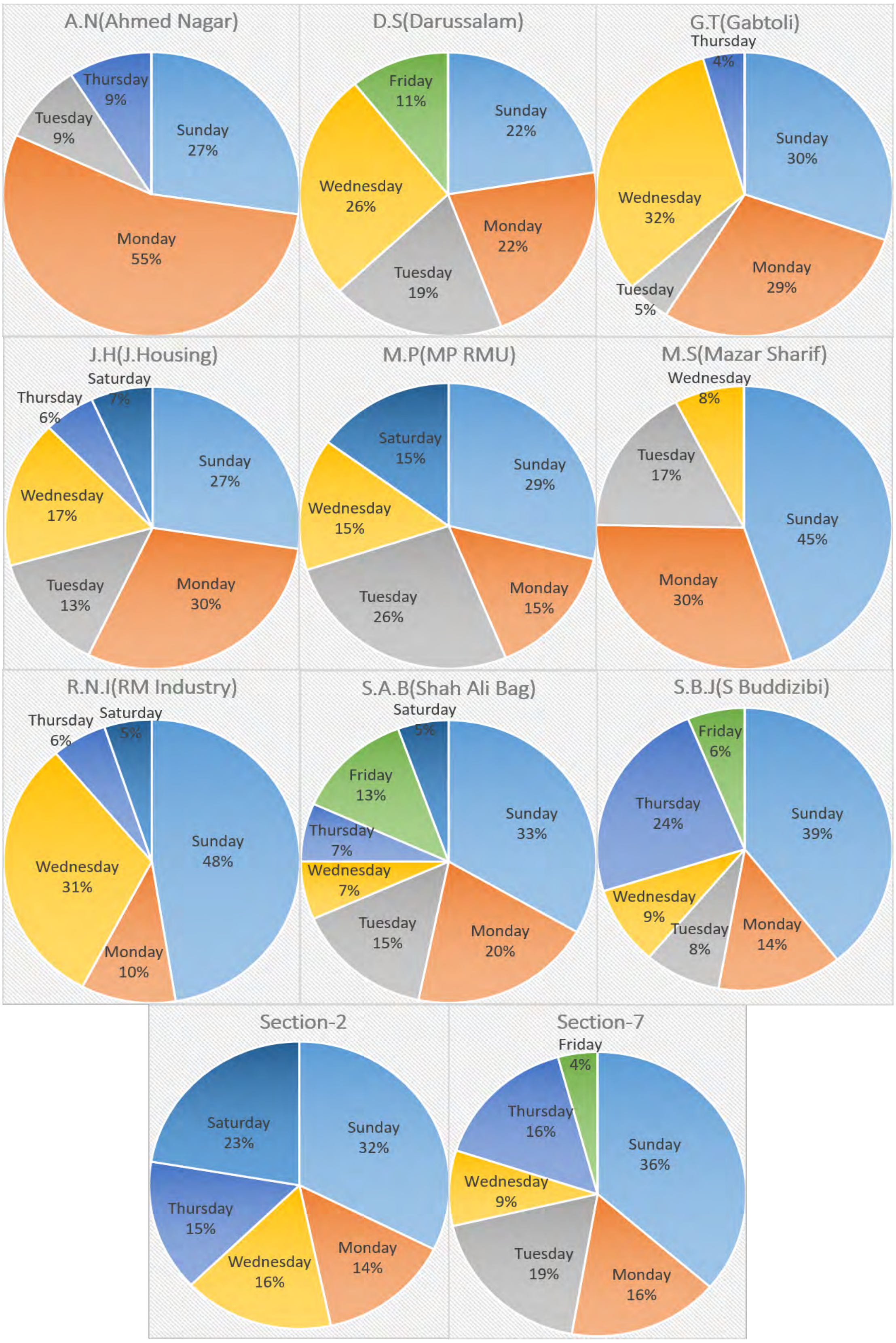


Fig. 48. Day wise load shedding impact on each feeder of Shah Ali sub station

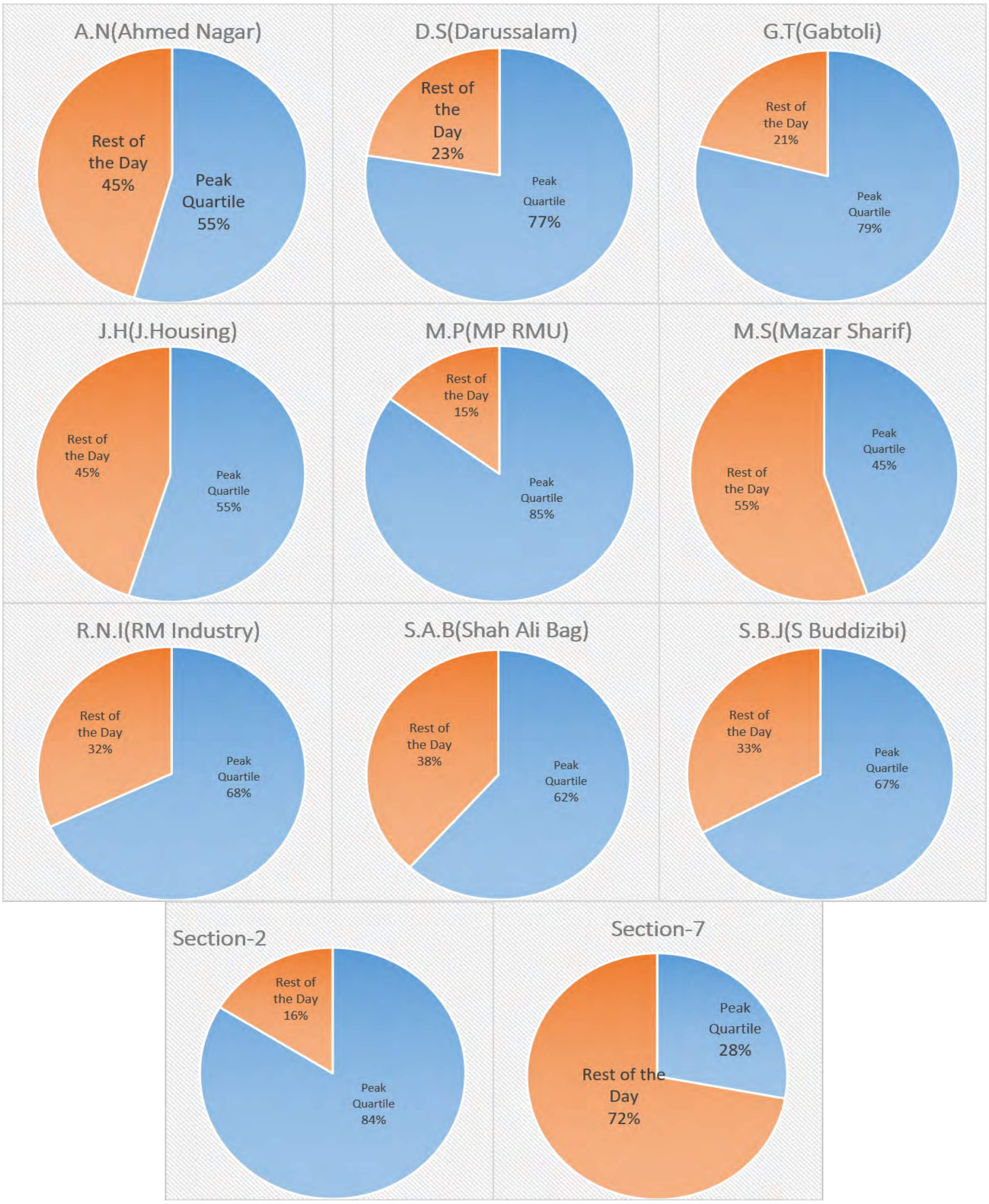


Fig. 49. Peak hours versus rest of the day's load shedding impact on each feeder of Shah Ali sub station

---

THE END