CHAPTER 1

INTRODUCTION

The purpose of this chapter is to provide the demonstration of the research. The sections are divided into the research background, which illustrates the goal of our study and the sources of the findings. Afterwards the objective, thesis queries and the scope of the research are conferred. At last the approach and the rundown of the thesis are described.

1.1 Research Background

Worldwide energy demand is increasing day by day[1]. This demand has reached a new peak level and created special requirements on balancing the demand and generation. Like other form of energy, from 1974 to 2000, electricity production has increased at an average annual rate of 4.6% in non-OECD countries, with OECD countries at 3.0%. This trend however changed from 2000 to 2015, with the average annual growth fell to only 0.9% in OECD countries while it grew by 5.9% in non-OECD countries, and consequently, in 2011, non-OECD electricity production exceeded OECD production. 1974 energy demand was 500 TWh but 2015 it demand was 25000 TWh, Electrical energy demand increasing every year in every sector.

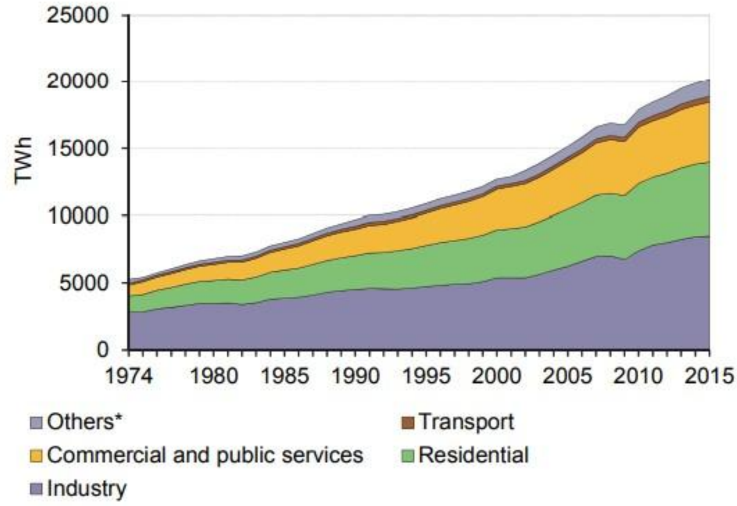
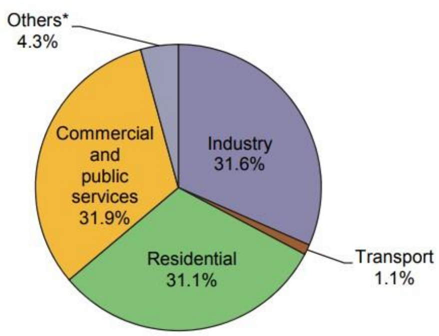


Figure 1.1: World-wide electricity consumption by sector[2].

As seen in figure 1.1, residential sectors account for almost one third of total energy demand in recent years. Power consumption in residential sector is also becoming unpredictable and fluctuating due to introduction of different distributed energy resources (DER) and renewable energy resources (RES) such as rooftop solar photovoltaic (PV) system, battery storage system, etc. The combined share of total consumption of the residential and commercial/public service sectors increased from 48.4% in 1974 to 63.0% in 2015. Although the amount of electricity consumed in industry increased from 1874 TWh in 1974 to 2970 TWh in 2015, its share of total electricity consumption in the OECD fell from 48.7% in 1974 to 31.6% in 2015. From the figure 1.2 we can see that the commercial and public services consume electricity 31.9%, others consumption is only 4.3% ,the residential sectors use 31.1% and the industry sectors use 31.6%. Commercial and public , residential and industry sector is also same .

In 2015, final electricity consumption in non-OECD countries was 10803 TWh, an increase of 2.7% from 2014. Between 1974 and 2015, final electricity consumption increased at an average annual rate of 5.1%. Non-OECD countries share of world electricity final consumption has been experiencing sustained growth, increasing from 27.1% in 1973 to 53.5% in 2015. From the figure 1.2 we can see that the commercial and public services consume electricity 14%, others consumption is 9%, the residential sectors use 23% and the huge amount of electricity consume by industry which is 51%.



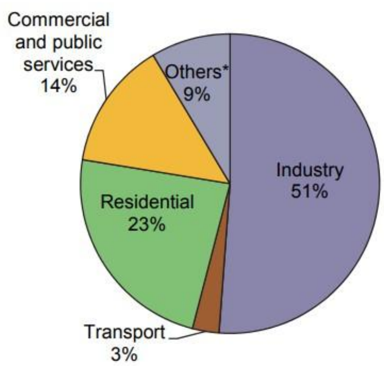


Figure 1.2: OECD & Non-OECD country’s electricity consumption by sector, 2015[3].

Energy consumption in the residential sector includes energy used for heating, cooling, lighting, water heating, and consumer products. Energy consumption in the residential sector is affected by income levels, energy prices, location, building and household characteristics, weather, efficiency and type of equipment, energy access, availability of energy sources, and energy-related policies, among other factors. As a result, the type and amount of energy consumed by households can vary significantly within and across regions and countries.

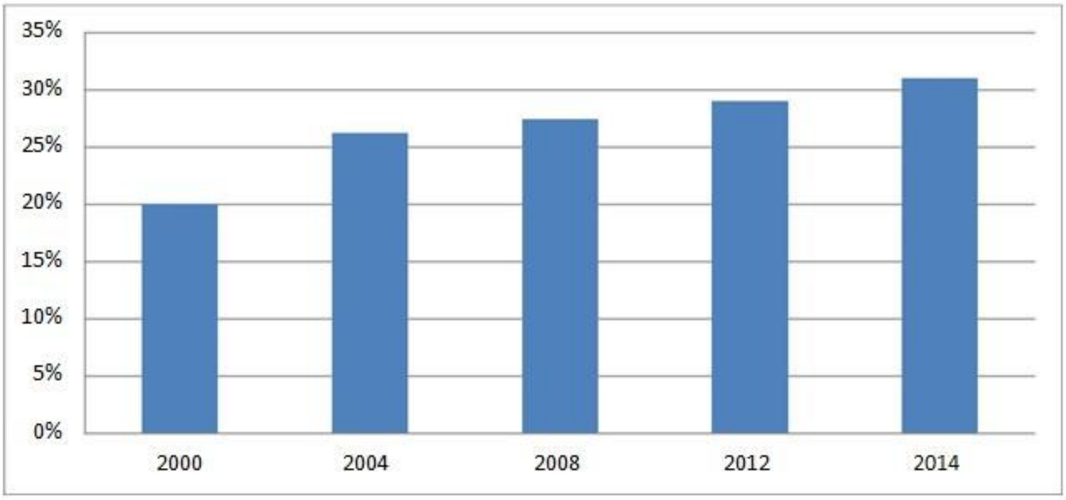


Figure 1.3: Growth rate of electrical energy consumption in residential sector[4].

From figure 1.3 we can see increasing energy consumption 2000-2014 in residential sector. If compare with 2000 to 2004 residential sector energy increasing 6% it’s increasing every year. From 2004 to 2014 energy demand percentage was 12%.

To cope up with the recent trend, every sector of the energy needs to be self-dependent. As a part of that this research work is mainly focused on making an individual, in this case a BTS, self-dependent in the sense of energy consumption. Again, it is obvious to not using the sources, which is available to them, in an efficient way. For this purpose it would be very helpful to predict the pattern of the energy consumption for each individual a day ahead. Meaning, a model that can forecast the next day energy consuming pattern will be feasible. Every BTS has almost common equipment, which is used in a daily basis. In fact the machines are almost operated 24/7. So, the best way to reduce the energy loss and produce for itself is to know the future consuming model. The main idea is to build a forecasting model, which can forecast the day-ahead consumption profile and schedule a storage device, in this case battery, to become grid dependent.

For making the consumer grid free, PV panels will be used. And for storing the energy battery will be a better option. So, the consumer can produce electricity by its own as per his necessity and thus can use it later on. Now the main challenge is to build an efficient forecasting model, which will be used for the main purpose of this work.

A vast number of forecasting models are available to build the model. But among them, artificial neural network is a kind which can be trained based on the previous days consumption and forecast the energy profile for the next day. The model needs the previous day's energy consumption data of the specific BTS for forecasting the model on the next day. Similar training based on the previous day's consumption was given to support vector machine and auto regressive integrated moving average and the forecasted data of these three popular models being compared to find the most efficient one among this three.

As every BTS needs different amounts of power at the different part of the day, the main focus would be given on the higher consuming times of the day. Though the forecasted model would not be properly accurate with the actual data, an idea about the energy consuming pattern can be obtained. Which can be further used to build a suitable solar system or any other renewable sources to produce energy for its own. Then scheduling has to be modeled to store the right amount of energy at the right time.

1.2 Literature Review

Short term load forecasting plays a vital role in the daily generation, efficient power system planning, unit maintenance, determining unit commitment and secured power system operation. There are number of approaches for short term load forecasting but it is observed that time series approach is most feasible and provides more reasonable accurate forecast for linear type of data set [5]. The time series Autoregressive gives better forecasting results for 4 to 6 Hours ahead. The load demand forecasting is very important for the management of power system and is a vital and fundamental factor for a successful operation of a power system [5]. In order to operate the electrical system efficiently and optimally the future load data has to be precisely predicted. Recently, Short Term Load Forecast (STLF) has become more important for two main reasons: the deregulation of power systems and the fact that no two utilities are the same, which necessitates a detailed analysis of the different geographical, meteorological, load type, and social factors that affect the load demand [6]. On the other hand if the system load forecasting is understated, the reliability and security of the system may be compromised which may result in power interruptions [5]. The load forecasting is basically classified into two categories viz. short-term and long-term load forecasting. Long term load forecasting usually deals with longer lead time, whose results may be used for long-term maintenance, economic and capacity expansion of power stations [5]. Whereas, the short-term load forecasting predicts information about the load ranging from one hour to a couple of hours. It is necessary for many functions such as managing and controlling of power stations, to provide real-time electricity tariff to consumers and planning load handling capacity of individual power station [7]. In present power scenario, due to increase in the number of customers, the load demand is increasing tremendously. The load requirement of different industries is influenced by different factors, so accurate forecasting becomes a tedious task. So to overcome these problems the short-term load forecasting is used to estimate load flows and to take appropriate decisions which can prevent the system from overloading. Timely implementations of such decisions lead to reliable, secure and healthy operation of power system leading to minimal equipment failures, brownouts and blackouts. In order to achieve higher prediction accuracy of the system, the different methods and techniques have been developed like time series analysis, metro numerical predictions and hybrid fuzzy neural methods [8-9]. Amongst which the time series analysis is found to be the main area with rich research effort, with specially formulated methods for data in various contexts [5]. After conducting a literature survey it has been found that the time series model is a powerful and effective tool amongst the other tools [10].

Short-term power load forecasting means load forecasting where time units are the hour, day or month. Because it has strong randomness, formulation of mathematical models is difficult. Improvement of forecasting accuracy is also difficult. Traditional load forecasting methods, including time series, regressive analysis and so on, can’t meet the accuracy standards demanded in practical situations. On the contrary, modern artificial intelligence methods, represented by artificial neural network (ANN), support vector machines (SVM), etc., have a certain ability of self-learning, self-adapting and powerful continuous function approximation. As their outstanding performance in the field of non-linear applications indicates, modern methods are more feasible than traditional ones. However no single model can perform well enough because each one just takes several or only one relevant factor into consideration. The optimization combination forecasting method, which makes good use of valuable information from several models, has become one of the most popular subjects in the field [11].

Support vector machines (SVM) has been used in load forecasting field. The noise and redundancy of sample data are important factors to the generalized performance of SVM. They can cause some disadvantages of slow convergence speed and low forecasting accuracy. Short-term load forecasting has always been an important issue in power system operation, and also an important basis for making plans for energy transactions and dispatching scheduling of generating capacity. A number of operating and management decisions are based on load forecasting, such as dispatch scheduling of generating capacity, reliability analysis. In order to improve safety and economy of grid operation, improve the quality of power supply and realize automation of power system, higher accuracy is necessary in short-term load forecasting for the modern power system [12]. Load forecasting can be influenced by a wide variety of facts, such as weather conditions and previous load data. We cannot get higher forecasting accuracy only by advanced forecasting models. During the past years, most of techniques based on time-series analysis have been applied in load forecasting [12]. The time-series model mainly includes approaches based on statistical methods and artificial neural networks (ANNs). ANNs model has shown to have the ability to learn nonlinear relationship however, ANNs do not usually report good performances in forecasting new samples. For example, the problem of over fitting usually exists. Because this algorithm mainly minimizes the objective function by using gradient descent method to train network weights, it results in bad generalized performance. In particular, forecasting accuracy will be affected to a large extent with minor training samples. Recently, a new method based on machine learning techniques and support vector machines (SVMs) has been used for load forecasting or classification and can get good performances [12]. SVM is a new and powerful machine learning technique for data classification and regression based on recent advances in statistical learning theory. But classical SVM has low convergence speed and low forecasting accuracy under the condition of much redundancy and noise in training data [12]. Various fundamental short term load forecasting (STLF) techniques such as dynamic linear and non-linear models [13], Kalman filtering [14] and some optimization methods for load forecasting [15] are found in literature. Machine learning techniques have also seized the attention of researchers, such as fuzzy-neural model, support vector machines (SVMs) [16], artificial neural network (ANN) [17] and genetic algorithms to optimize ANN structure [18]. Apart from machine learning techniques, some researchers focused on time series forecasting models such as auto regressive integrated moving average (ARIMA) [19] [20], auto regressive moving average (ARMA) [20], seasonal ARIMA (SARIMA) or autoregressive fractionally Integrated Moving Average (ARFIMA). The selection of these models depends on variation of input data and forecasted time period. The models compared and evaluated are based on ANN and AIRMA, very popular choices for STLF [21] [22] [23]. The forecasted consumption profile maybe used for scheduling storage devices. The initial step for any kind of forecasting approach is to collect historical data of the system to be studied. Depending on the models input could be of the same type (only electrical energy or heat production) or a combination of different types of data (energy consumption with temperature or seasonality etc.). Using the forecasted electricity consumption, it is possible to have an idea about the amount of stored energy. After mitigating self-consumption from PV generation, surplus energy maybe used to charge the storage devices.

The main objective of the power system is to balance the load and generation. The electricity should be generated as soon as demanded as electricity energy on a large scale cannot be stored and moreover it is presently uneconomical. Therefore the amount of load required by the consumer has to be forecasted in prior, for the reliable operation of the power system. In the deregulated power industry the accurate STLF is most important as it’s more useful in purchasing the power from the generating companies at real time, switching of the loads and most important factor is to avoid congestion in the transmission line network by dynamic control. High forecasting exactness and speed are the two most critical necessities of here and now stack anticipating and it is imperative to dissect the load attributes and distinguish the fundamental variables influencing the load [24]. In power network, the conventional load influencing components, for example, season, day and climate, power value that have deliberate and may have a convoluted association with framework stack [25], [26]. The time series method is advantageous based on the linear analysis and it is unable to forecast the nonlinear sensitive part of the load. But the nonlinear external factors such as the weather, days of the week and so on are also important for the accurate load forecasting. The Artificial Intelligent techniques (ANN and SVM) can accurately forecast the non-linear sensitive part of the load. But ANN method lacks in providing the perfect result and quantitative analysis since it take on empirical risk minimization principle according to the Statistical Learning Theory (SLT) which only tries to minimize the experience risk whereas SVM is based on the Structural Risk Minimization (SRM) which minimize the generalization error rather than the empirical error [22], [27].

Since the forecasting methods consider the historical loads, it is important that the historical data is precise, unfortunately the data recorded may have some missing values, error in the value or there will be some abnormal values due to outages, load shedding [28], [29], [30], sudden rise or dip in the demand. These data's are termed as the Outliers. So it is important to incorporate an Outlier detection method for load forecasting.

1.3 Research Motivation

The main objective of this research is to identify the appropriate model of forecasting. Moreover, to facilitate the investors with some aspects which might be very important, mentioned as follows -

- Finding the appropriate forecasting model among the popular short term load forecasting models.
- Importance of storage device scheduling to facilitate local market.

At a deregulated energy market where customers are encouraged for distributed generation and demand response activities, short term load forecasting at individual level has a great impact for scheduling the storage system at end user level. Moreover, the basic problem for the electric utilities, and more recently for the system operators, is to maximize operation system performance, given the operating and economic characteristics of the generating units, the transmission line constraints and the limited amounts of capital available for new units and equipment. For electricity suppliers it is important to schedule the energy transaction at different stage of market with small error margin and for the establishment of bidding strategy. In the transition to a smarter electricity grid, the need for information is increasing. However, the precise idea about the electricity consumption can handle all of these circumstances. As the main focused area of this research is base station storage devices scheduling, though the load prediction gives an insight of the available capacity for smart charging/discharging plan of storage devices. From application point of view, these storage devices of a base station have an influential impact at demand response activity of consumer. This is a scope to mitigate peak load demand, load shifting which also improve the utilization of distribution grid and enhance the quality and reliability of grid. However, controller of these storage devices as an agent should have prior knowledge about upcoming electrical loads to minimize losses of the storage devices and thereby possibility to have all the benefits. So to propose an appropriate electricity consumption forecasting model to schedule base station storage devices as a use-case is the core objective of this research.

1.4 Research Goal

The research goal of this thesis is:

Establish a proper electricity consumption forecasting technique for optimization of storage devices of Base Station.

The few main objective of this research can be formulated as follow:

1) To investigate different types and categories of short term load forecasting techniques.
2) To investigate different types and categories of battery storage.
3) Define the best forecasting techniques to optimize battery storage scheduling.
4) To build up the most appropriate forecasting model.
5) To design a scheduling model for battery storage.

1.4.1 Research questions

The research questions that belong to the research goal are:

I. What will be a suitable approach for scheduling the storage devices at base station level to make it energy neutral or grid independent?

II. Model development for forecasting the energy consumption of a base station and evaluate the result with real consumption data.

III. Which forecasting technique will be taken and why?

1.5 Scope of the Research

Before constructing a forecast model it is important to identify the scope of the research. As explained in the research motivation section, the day-ahead electrical load forecasting of a particular base station will be used to schedule storage devices based on the available capacity of PV as local generation. To fulfill this goal the electricity consumption need to be predicted per phase. To identify an appropriate day-ahead load forecasting model is the main focusing point of this research. However, it is very important to identify the proper time horizon and frequency/sampling rate of the historical data. All the historical data used to train the models are real consumption profile of the year 2013 [31] of different base stations and in 15 minute sampled. After evaluating the models for STLF, the well-performed model will forecast the load profile of a consumer. In reality the forecasted load profile and PV generation will be used to scheduling the storage devices but to predict the consumption profile is the scope of this research. Prediction of PV generation is not the ultimate target of this work.

1.6 Research Approach

The process of achieving the research goal is mainly divided into two main parts; Assessing the forecasting model and scheduling of storage device. Some other steps are also involved within this methodology. The modeling and simulation is performed in MATLAB®. All the steps are discussed below.

Stage 1: Review and study several kind of theory and literature. Study works covered:

a) Study about forecasting, process of forecasting in different condition.

b) Study and investigate different types and categories of battery storage.

c) Study about battery scheduling, how to schedule them.

d) Study about different type of optimize technique, calculation of optimize technique; best optimize technique for a situation.

Stage 2: In this period several of work will be done. Like

a) Select an area for applying the research.

b) Collect data for the specific area.

c) Forecast data for that specific area.

d) Investigate the forecast data.

e) Correction the false data.

Stage 3: Confirm a fixed forecasting data to apply. Study about various types of forecasting technique. Select an appropriate forecasting technique for applying in the research. March-2013 to May-2013 [32]: In this period the data had been applied.

a) Develop a short term load forecasting model.

b) Simulate the data.

c) Study and modify the simulation output data.

d) Find a optimize output.

Stage 4: In this period activity was

a) Analysis the performance of the forecasted data.

b) Find the appropriate forecasting model depending upon the forecasted data and the actual one. Though the forecasted data was not accurate yet these models can perfectly identify the shape of load pattern.

Stage 5: Write and rearrange the research report and make a presentation to present the research.

1.6.1 Historical Data

As mentioned earlier section, energy consumption (in Whr) of base stations of an area is used to train and evaluate the performance of the forecasting modes. Moreover, PV generation profile of the same base stations is also provided to scheduling the storage devices. However, all of these data is collected for some measuring devices, thus some analytical pre-processing is essential to ensure accuracy of the data.

Data pre-processing:

Acquire representative data, remove unusual consumption hike and other inconsistencies, define proper format for time stamping (day, month, hour, minute and weekend days' identification) and split up the data in identification and validation sets.

Data-analysis:

Analyze the data to find underlining mechanisms, trend and variations in the data, use of clustering to get more insight in intraday correlation.

1.6.2 Forecasting of Consumption:

To achieve ultimate goal to schedule the storage devices it is very crucial to have prior knowledge about electricity consumption on day-ahead. Therefore, short-term forecasting is an important step to ensure better scheduling of the storage devices. Pre-processed data is used to evaluate the performance of the forecasting models. Initially three forecasting models, ANN, ARIMA and SVM were chosen to forecast the consumption profile. This model is evaluated with respect to some evaluation criterion.

1.6.3 Storage Device Scheduling:

The forecasted consumption profile of a base station will be used to achieve optimum scheduling of storage devices. The main idea is to utilize the surplus of PV generated energy after mitigating self-consumption. If a particular consumer can have an idea about the level of stored energy, it is possible to utilize the energy in various way like, load shifting, include some flexible loads to consume the extra energy, valley filling etc. However, for this research our target is to feed the extra stored energy in local market, so that the consumer can have some financial benefits from the trading.

1.7 Outline of the Report

The previous data of 92 days has been collected from 32 different BTS for training the models. The assumption was that 92 days is enough as an adaptive period for the agent to capture changes in electricity consumption, while the 15 minutes resolution is a common practice from a data acquisition point of view [32].The consumption data was in Whr. The accuracy criteria was the regression factor, MAPE and MSE.

In chapter 1, introduction describes the background, literature review and motivation of the research, research scope, objective of the research, some question which will answer throw the research and stages of work to fulfill the research.

In chapter 2, the background part describes definition of load forecasting, classification load forecasting, different short term load forecasting models, different type storage technology, working principal, advantage, disadvantage etc.

In chapter 3, methodology part describes the selection and description of the electricity consumption data and the different variables are discussed. The processing and the detection of missing values being discussed. The methods are discussed to discover the cohesion and pattern in the selected data. After which the procedure of evaluation is appointed.

In chapter 4, model building covers the details classification of neural network. And after a discussion a suitable form of neural network which is FNN is finalized to compare with the others (SVM and ARIMA).

In chapter 5, parameters of evaluation are described in brief.

In chapter 6, simulation and results develops a MATLAB code for the simulation, show the simulation set up, optimization of this thesis and result of this thesis.

In chapter 7, showing the conclusion and future works of the research. References are added at the end of this report.

CHAPTER 2

BACKGROUND

The Background section presents various methods of energy consumption forecasting that have been developed through the years to determine base station electricity demand forecasting and analysis for electrical energy storage device optimization by using forecasted energy consumption. An immense amount of literature is available on the topic of load forecasting. An attempt is made to show the most in influential work by a review of the literature on different topics of the aspect of forecasting.

2.1 Base station Electricity Usage and storage

In the characteristics of electricity consumption at commercial sector and the way to control it by scheduling energy shortages are presented. Proper scheduling of storage devices can also provide benefits for the commercial customer and the supplying utility by more optimally managing the consumption of electricity. However, the electrical energy plan for base station consumption is a key element for the utilities' for integrating demand-side and supply-side resource options into plans that provides reliable service at the lowest reasonable cost. Moreover, base station electrical energy consumption also depends on socio-economic structure of a society as a hole, which also has influential impact on base station consumption. In Falong Yan has tried to illustrate this impact from Chinese society, more or less it has the same effect on every society. Since 1980, with reform of economic structure and opening to the outside world, urban base station real income has increased rapidly. Therefore, less electricity-intensive appliances entered urban base stations rapidly. First, electric fans, tape recorders and black-and-white TV sets, then color-TV sets, washing machines and refrigerators represented the primary electric load. Since 1990, electricity use has grown continuously with resident income rise. At the sometime, income disparity has gradually widened. Electricity-intensive appliances such as air conditioners and electric heaters have begun to enter primary the high- and middle-income base stations. In 1993, urban base stations consumed 57.14% of the electricity used by all base stations, although the urban population was only 28.14% of the total. However, at liberalized market situation, rapidly increasing electricity consumption make the situation more and more complicated from reliability and market point of view. So predict the energy consumption for proper planning for future network extension and also to ensure the market stability consumption forecasting is very important. Though, one of the major parts of total consumption is from residential sector, so proper prediction of base station consumption can be beneficiary for all market parties.

2.2 Benefits of forecasting

Load forecast has been a central and integral process in the planning and operation of power systems to ensure reliability and stability of the network [33]. A. Ghanbari mentions that “The ability to forecast electricity load has always been one of the most important targets for power systems. This advantage enables to manage generation with a much higher performance”. The use of load forecasting is widely accepted as operation aid for the control of the electric power system as well as to enhance consumer participation in local energy market through providing financial benefits. The use of load forecasting is divided into several applications which will be described here.

2.2.1 Applications in deregulated electricity system

In a deregulated electricity market load forecasting is becoming increasingly important. For suppliers it is important so that accurate energy transactions can be scheduled with small error margin and bidding strategies can be established. In the transition to a smarter electricity grid the need for information is increasing. In smart grid context, due to instigation of huge amount of local or distributed generation bi-directional power flow can create unwanted congestion in distribution grid. Moreover, the ultimate aim of this research is to scheduling of storage devices by means of consumer participation in local market and provides more flexibility in electricity consumption. Load forecast can enhance the possibilities of consumer being financially facilitate from local market participation. The load forecast gives insight in the available capacity of the BTS storage devices which can be used to mitigate self-demand. The use of a battery energy storage unit can be enhanced if the controller has knowledge of the upcoming load. The controller can act upon coming electricity demands and hereby minimize losses by optimizing charge and discharge schedule and thereby improving the use of the battery storage [34].

2.2.2 Generation scheduling

Generation capacity scheduling in one of the main purpose of generation scheduling in aggregated level [35]. If it is possible to have prior knowledge about local production and amount of network imbalance power, generators can be scheduled a day ahead to satisfy the predicted load demand after taking into account the physical constraints arising from transmission capacity and physical ramp up and down time. The switching on or off generation units at what time is called unit commitment and the scheduling problem is referred to unit commitment problem [35].

2.2.3 Control of the distribution network

Other use of load forecasting are more focused on the control of parts of the system, for example to determine the capacity of spinning reserve [35] (which is closely related to generation scheduling), the effective use of demand response mechanisms or the proper control of a smart storage unit [35].

All these applications of the forecasted load are subject to the accuracy of this forecast. Therefore one can imagine the importance of a high accuracy of the forecast and the error rate to take into account for operation. An improvement in prediction error will results in a strong influence of the aim of this research by improving the storage device scheduling performance. Therefore the area of load forecasting methods is of constant interest to scientist to minimize the prediction error actions which will be described here.

2.3 Historical view of load forecasting

The first paper describing the topic of load forecasting is published in 1944, where three distinct factors are identified which determine the total system load of a utility. In this paper, authors tied to explain the temperature effect on base load. Temperature, cloudiness and wind are the described weather factors in influencing the load and given weights based on trial and error [36]. Since then numerous papers have been published on this topic focusing on different aspects. Forecasting electricity demand is vital for planning and investment purposes. In estimating future electricity demand, it is important to assure high responsibility that there will be no supply shortages. Moreover, form policymaking point of view energy demand forecasts have traditionally played a key role; like in Chinese development planning process the government uses aggregate approaches to determine future energy consumption [37]. Apart from total national demand, electricity consumption form particular sector is also important for planning and operation of the network. As an example, proper prediction of base station electricity consumption a functional form of measuring willingness to pay or increasing cost for electricity demand for base stations and firms, proposed in [38]. The measuring willingness to pay for electricity relies critically on a reliable estimate of the demand for electricity function. The usual approach is to calculate consumer surplus (CS) on the basis of a linear electricity demand function.

Moreover, the purpose or utilization of forecasted result is also important to choose the appropriate model for forecasting. In this research, base station electricity consumption is focused and the target is to utilize this forecasted consumption for scheduling of storage devices as a demand side management tool. C.W. Gellingset presented the choice among aggregate and individual methods or models are made in the context of a definite objective. If the objective is to forecast quickly and have a model that is easily and transparently described in terms of the key sensitivities (e.g., elasticity's), then an aggregate econometric model may be appropriate. However, if it is focused to develop a detailed demand-side plan in which various DSM strategies are evaluated and incorporated into the forecast, then an individual forecasting model at end-use model is required.

On the other hand, forecasting or target period so called lead time is also depended on the purpose of the forecasting. Abu-El-Magd & Sinha distinguish forecasting for different lead times (minutes to weeks to longer times). Multiple regressions, spectral decomposition, exponential smoothing stochastic time series approach are reviewed as part of time series solutions. As it is seen, forecasts are made for various purposes depending on the time span: the day-to-day operation of the power system [39] requires the prediction of the load for a day ahead whereas the decision whether to undertake major structural investments require a far longer prediction horizon.

Forecasts can be distinguished therefore firstly by the time-horizon or the lead time: Short-term load forecasting (STLF) usually aim to predict the load up to day ahead to one-week ahead [39].

STLF is a subject that has been studied extensively and several different methods have been developed during the last few decades. A comparative review of five different approach of STLF where each method is applied for same database is done by Moghram and Rahman. These are multiple linear regression, stochastic time series, general exponential smoothing, state space method and knowledge-based approach.

2.4 Different approach of STLF

In general, STLF techniques can be classified as either traditional or modern models [34]. Traditional statistical load forecasting techniques, such as regression, time series, Kalman filters etc., have been used in practice for a long time, showing that the forecasting accuracy is system dependent [34]. This part of report is mostly focused on machine learning techniques for short term electric load forecasting. The history of machine learning in STLF is outlined, leading to a discussion of the various approaches as well as the current research directions and bears the witness to the enthusiastic application of AI technologies (mainly of expert systems (ESs) and artificial neural networks (ANNs)) in the field of STLF. During the study, it is noticed, both ESs and ANNs methods have strong justification and different methods to solve any problems and researchers are still nourishing these methods to find so novel approach. A humorous but fully justified approach of this discussion has already been given by Professor Jay Liebowitz[35] in the Editorial of an issue of “Expert Systems with Applications’’

‘‘If you are a dog lover, build expert systems; if you are a cat lover, build neural networks’’.

ESs is one of the most commercially successful branches of AI. In other words, an ES is a computer system containing a well-organized body of knowledge, which emulates expert problem solving skills in a bounded domain of expertise. The system is able to achieve expert levels of problem solving performance, which would normally be achieved by skilled human resources, when confronted with significant problems in the domain of academic research groups.

2.4.1 Time Series Forecasting Model

Time series can be defined as a sequential set of data measured over the time, such as hourly, daily or weekly peak load. The basic idea of forecasting is to first build a pattern matching available data as accurate as possible and then it obtains the forecasted value with respect to time using established model [41]. The figure 2.1 shows steps for selection of the model by using the historical data.

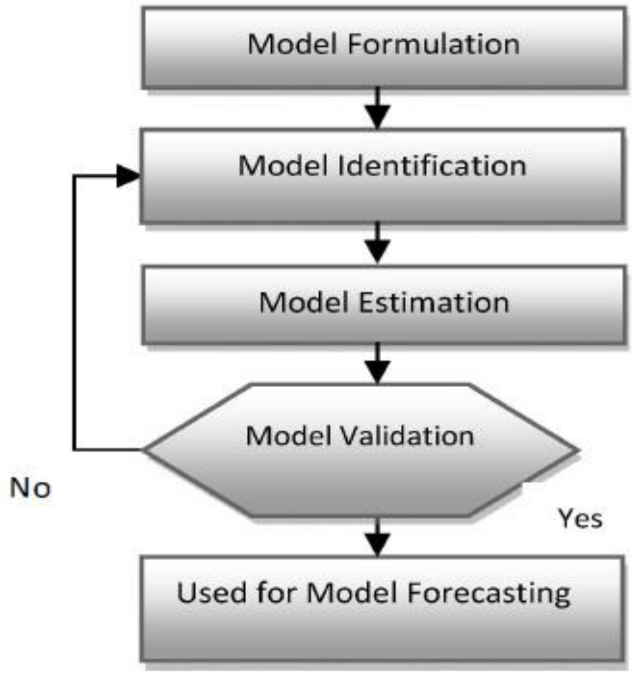


Figure 2.1: Stages in the time series model building [41]

STLF is a necessary element of Energy Management System (EMS). It has an important role in purchasing the power from the generation companies (GENCOS) in real time scenario. Moreover, avoiding the congestion needs to be ensured in the transmission line network during the switching of the loads. The energy consumption profile is a function of time which can be expressed as a time series Y_t. The consumption profile Y_t may consist of some linear and non-linear part. The factors responsible for the high accuracy rate are the parallel processing of the input data and the model building process. Besides, the model is independent of any previous presumptions. The difference in accuracy of different forecasting model depends on the proper training algorithm and structure of input- hidden-output layers.

2.4.1.1 Auto Regressive Integrated Moving Average

Data can be represented with the help of different stochastic processes in models involving general time series. There are two linear time series models which are used vastly: Autoregressive (AR) and Moving Average (MA) models. The ARIMA model is the combination of these two models [19], [20]. It is extensively popular among all other time series approaches because of its adaptive nature to handle linear patterns [22]. In 1970, Box and Jenkins first proposed this model, so it is also familiar as Box-Jenkins methodology [20]. Besides, the ARIMA model is composed of a simplified algorithm which operates with advanced technology. The mathematical model of ARMA (p,q) is illustrated in [19], [20] by combining AR (p) and MA (q). But, ARIMA (p,q) lacks the ability to utilize a non-stationary time series properly, which is the key point for this paper.

Apparently it contains non-stationary type data. In the ARIMA (p,d,q) model [36], finite differentiation of the data points is adopted to transform the non-stationary time series to a stationary one. Equation (2.1) and (2.2) represent the mathematical expression of the ARIMA (p,d,q) model,

$$\varphi(L)(1-L)^d y_t = \theta(L)\varepsilon_t \tag{2.1}$$

$$(1-\sum_{i=1}^{p}\varphi_i L^i)(1-L)^d y_t = (1+\sum_{j=1}^{q}\theta_j L^j)\ \varepsilon_t \tag{2.2}$$

where, p, d, and q are positive integers, which indicate to the order of the autoregressive, integrated, and moving average parts of the model, respectively. Obviously, the accuracy level is highly dependent on the appropriate selection of p, d, and q. A general method to find out p and q is to study the Autocorrelation function (ACF) and Partial Autocorrelation Functions (PACF) of the training data. The PACF plot helps to figure out the maximum order of AR (p), while ACF plot can identify the non-stationary time series.

2.4.2 Machine Learning Model

The ARIMA model does not consider the non – linear sensitive factors of the load such as weather, temperature, day of the week etc. In order to improve the forecasting accuracy by considering these factors we introduce artificial intelligent methods basically ANN and SVM. ANN is based on ERM (Empirical Risk Minimization), this will lead to the local optimum value and hence increases the generalization error whereas SVM is based on SRM (Structural Risk Minimization) reduces the generalization error.

Machine learning, a branch of artificial intelligence, concerns the construction and study of systems that can learn from data. In [40] Tom M. Mitchell provided a widely quoted, more formal definition: "*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E*". This definition is notable for its defining machine learning in fundamentally operational rather than intellectual terms, thus following Alan Turing's proposal in Turing's paper "Computing Machinery and Intelligence" that the question "Can machines think?" be replaced with the question "Can machines do what we (as thinking entities) can do?"[1]

So as a branch of AI, machine learning has an enormous scope in the field of load forecasting and previous literature researches have shown attentiveness for machine learning as forecasting model. A wide variety of models, varying in the complexity of functional and estimation procedures, has been proposed for the improvement of load forecasting accuracy. In [41] Matthewman and Nicholson (1968) conducted an early survey of electric load forecasting techniques. Then in [42] and [43] also reviewed load demand modeling and electric forecasting techniques using machine learning. Recently, in [44] made a survey of previous proposed forecasting models and includes new categories that reflect recent research trends which are mainly focused on machine learning. They proposed nine categories; fuzzy logic, neural network and knowledge based export systems are related to machine learning because these models are used with knowledge based algorithms to perform measurements. So [41]–[44] gives a justification to use ANN for STLF of an individual customer.

Several techniques have been developed to represent load models by fuzzy conditional statements. In [45] an expert system using fuzzy set theory for STLF is proposed which is used to do the updating fuzzyfication and defuzzyfication function. This Short-term forecasting was performed and evaluated on Taiwan power system. A fuzzy linear programming model of the electric generation scheduling problem, representing uncertainties in forecast and input data using fuzzy set notation is described in [46]. S. Bataineh et al. discussed the implementation of a fuzzy-logic approach to provide a structural framework for the representation, manipulation and utilization of data and information concerning the prediction of power commitments in [47]. But fuzzy based models have some limitation of huge amount of data handling time compared to Neural Network. Neural networks are used to accommodate and manipulate this large amount of sensor data in several cases like in [48]. It proposed a hybrid fuzzy-neural technique to forecast load to combine the neural network modeling and techniques from fuzzy logic and fuzzy set theory.

An approach for short-term load forecasting by combining self-organizing map (SOM) and support vector machine (SVM) is proposed in [49] by clustering historical load data by SOM and daily 48-point load values are vertically predicted respectively based on SVM. Date type is classified according to profile of load curve, magnitude of load, weather information and date type where load profile and magnitude of load was unknown for that particular case study. Moreover SVM gives considerably high error rate for day ahead individual consumption side forecasting as described in [50].

Expert systems are new techniques that have emerged as a result of advances in the field of artificial intelligence. An expert system has the ability to reason, explain and have its knowledge base expanded as new information becomes available to it. To build the model, the 'knowledge engineer' extracts load forecasting knowledge from an expert in the field by what is called the knowledge base component of the expert system. This knowledge is represented as facts and IF-THEN rules, and consists of the set of relationships between the changes in the system load and changes in natural and forced condition factors that affect the use of electricity [44]. The typical variables in the process are the season under consideration, day of the week, the temperature and the change in this temperature. Algorithms of this model used in [51] and [52] combine features from knowledge-based and statistical techniques, using the pair wise comparison technique to prioritize categorical variables. Model developed in [52] was tested using data from several sites around the USA, and the errors were negligible so it can be called a site-independent expert system for STLF. Several hybrid methods combine expert systems with other load-forecasting approaches. In [53] Dash et al. combined fuzzy logic with expert systems. In [54] a two-step approach in forecasting load for Korea Electric Power Corporation is proposed where at first an ANN is trained to obtain an initial load prediction, and then a fuzzy expert system modifies the forecast to accommodate temperature changes and holidays. Though this is a new approach of load forecasting, it has some difficulty to forecast base station consumption because of complicacy of its computer programming and nonlinearity of the historical data.

All models and algorithms described above is found with strong evidence in literature, to be a successful STLF model but this paper is focused on individual consumption side load forecasting where previous load profile will use as historical data. And historical data should have both nonlinear and linear part because of different electrical appliance and characteristics of the consumer. So machine learning approach is preferable than the time series forecasting model because of its capability to deal with nonlinearity of the data in following section.

2.4.2.1 Support Vector Machine

SVM, the Support Vector Machine is a Machine learning algorithm used for predicting the nonlinear factors of the load in this paper. SVM is proposed by V. N. Vapnik and so forth in 1995 through statistical learning theory. SVM have more subtle tunable parameters than NNs and the decision of parameter values might be less important for good forecast comes about. The SVM is intended to efficiently advance its structure based on the input training data. The training of a SVM includes understanding a quadratic optimization, which has one interesting arrangement and does not include the random initialization of weights as preparing NN does [55]. So any SVM with similar parameter settings prepared on identical information give identical outcomes. This builds the repeatability of SVM forecast thus incredibly lessens the training runs required to discover the ideal SVM parameter settings when contrasted with NN training. Also it can be utilized to take care of the issues that have components of small sample size, high measurement and neighborhood minima with great generalization capacity so are worried about progressively by analysts in STLF field. The essential structure for SVM is appeared as Figure 2.2 [56].

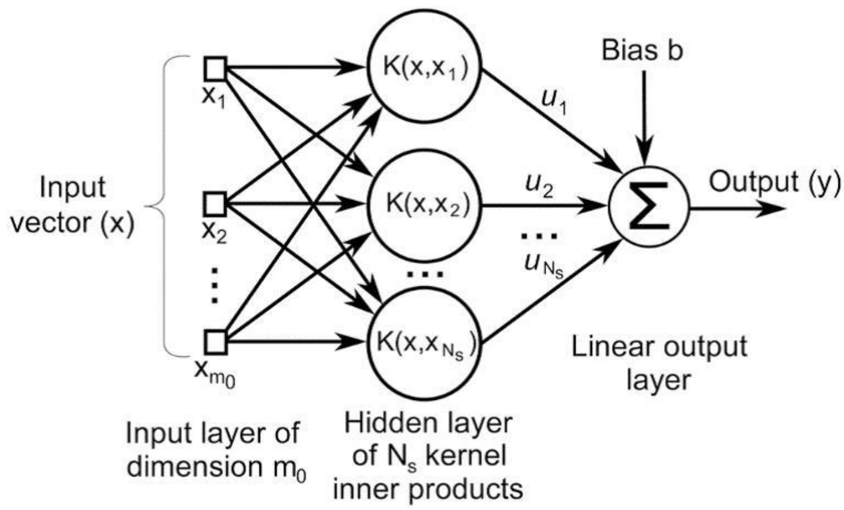


Figure 2.2: Basic Structure of SVM

SVM is frequently utilized to tackle two sorts of issues: classification problem and regression estimation one. The investigation of STLF has a place with the last that is called Support Vector machine Regression (SVR), whose propose is to discover the mapping relationship between the input vectors and output vectors [57], [58].

2.4.2.2 Artificial Neural Network

ANNs are an information processing technique based on the way biological nervous systems, such as the brain, process information. The fundamental concept of ANNs is the structure of the information processing system. Composed of a large number of highly interconnected processing units (‘‘neurons’’) connected into networks, a neural network system uses the human-like technique of learning by example to resolve problems.

One of the earliest research works to introduce ANN application in STLF is done by Lee et al., who proposed an innovative ANN methodology for the STLF problem. The multilayer network with three layers, i.e. one input, one hidden and one output was proposed. The training of the network was performed through a simple back-propagation algorithm. Using historical consumption profile and weather information, the system produced three different forecast variables (peak load, total daily load and hourly load). In [60] presented a search procedure for selecting the training cases for ANNs to recognize the relationship between weather changes and load shape, while [61] implemented a multilayer neural network with an adaptive learning algorithm. Lu et al. [62] used a feed-forward neural network incorporating the previous load demands, the day of the week, the hour of the day and temperature information for load forecasting.

A part from load forecasting with weather data or other variable, B. M. J. Vonk et al. proposed a suitable approach of STLF for energy storage system [63]. It considers the nonlinearity of load profile due to behavior of consumer which will change during the transition towards a more sustainable society. ANN is preferred as an adaptive or learning algorithm. Therefore, the forecaster will have to use historic results of the system, to train itself to the new situation. Proposed ANN model is able to map nonlinear input-output relations such as weather predictions on to power or energy consumption.

All of these literatures are focused to prove machine learning approach is an appropriate forecasting model for energy consumption. But in this case study it is mainly focused on base station consumption forecasting and individual customer should have some nonlinearity in their consumption profile. Machine learning approach can provide satisfactory result with nonlinear part of historical data. But to deal with linear part of consumption profile it needs a different forecasting model and that is the basic drive to choose mathematical/time series forecasting model like ARIMA or ARMA which is described in following section.

CHAPTER 3

Methodology

In this chapter the research approach is further elaborated. In Section 3.1 the selection and description of the electricity consumption data and the different variables are discussed. In Section 3.2 the processing and the detection of missing values. In Section 3.3, methods are discussed to discover the cohesion and pattern in the selected data. After which in Section 3.4 the procedure of evaluation is appointed.

3.1 Data

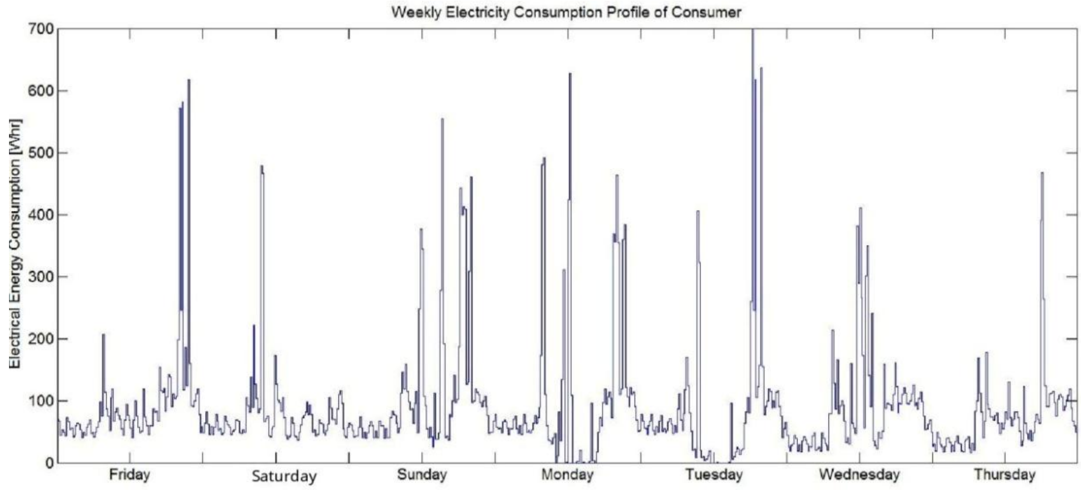
As discussed in Chapter 1, real data of electricity consumption is needed for estimation and validation purposes of the forecasted result. Moreover, to predict the amount of stored energy in the storage devices on day-ahead, the PV generated energy of the corresponding base stations are also needed. As the ultimate goal is to predict the base station forecasting, though additional Qualitative variables are introduced to recognize the precise pattern of the consumption behavior of each consumer.

3.1.1 Electricity use data

For the sake of forecasting, real energy consumption data of different BTS on a daily basis at 15 minutes of sampling is used. This data will be used to train the network and build the forecasted model. Then that forecasted model will be used to design the PV panel system and schedule the storage device. All the data are taken in Whr and thus will be forecasted in the same unit.

3.1.1.1 Normal Energy Consumption Data

On average the yearly consumption of an identical base station is 3500 kWh. So daily electricity consumption on average is 9.5 kWh [65]. Consumption profile shown in the figure 3.1 is the electricity consumption pattern of a particular base station for the first week of March, 2013 with average consumption around 9 kWh per day. However, the daily consumption pattern is different depending on week days and weekend, even similar week days does not have similar consumption patterns shown in figure 3.1.



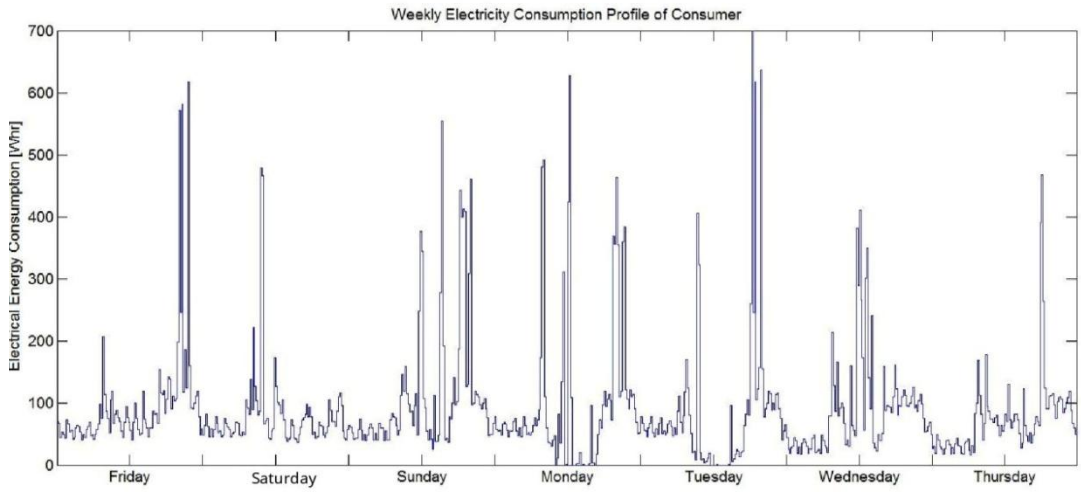


Figure 3.1: Weekly consumption pattern

It is very important to identify the non-linearity of the consumption pattern, because forecasting model selection depends in the degree of non-linearity of the input data.

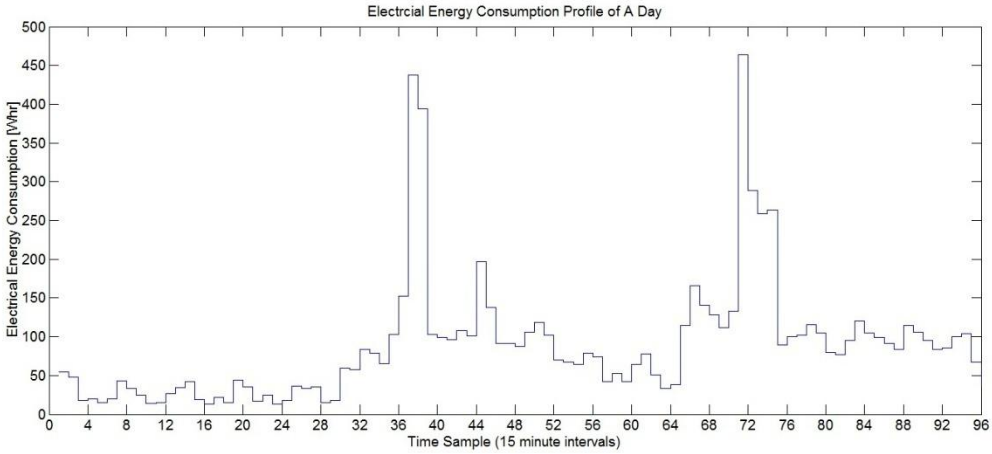


Figure 3.2: Daily load consumption.

3.1.1.2 Electrical Energy Consumption compromised with PV:

After careful analysis some data are found in the data set which follows a particular pattern but the average electricity consumption is less than the standard one. Figure 3.3shows a consumption profile of a day which goes close to zero at mid-day. Moreover this particular base station has installed PV as local generation source. So, it is very much understandable this consumption profile is the net amount of energy consumed from grid after mitigating some load with PV. Thus, the electricity energy consumption data of this type of base stations cannot be used to identify the proper forecasting model.

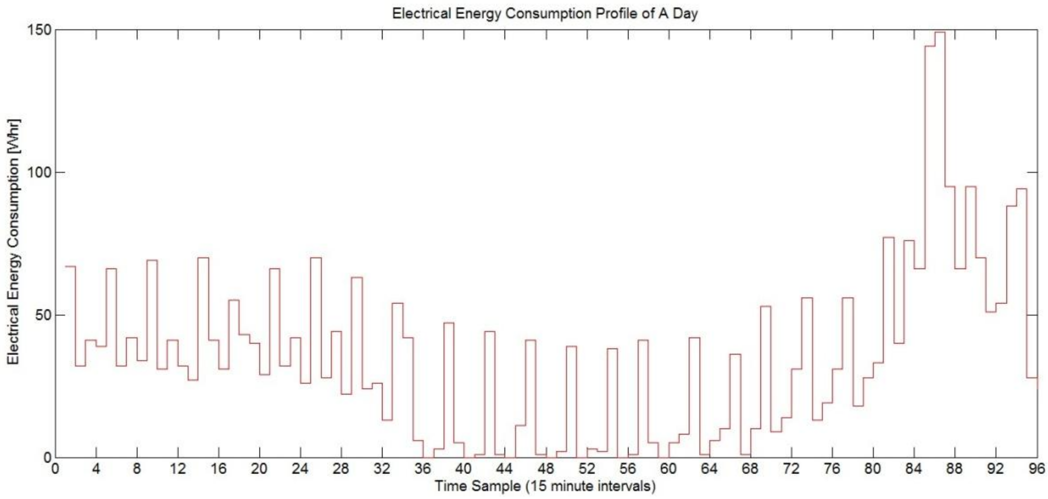


Figure 3.3: PV Compromised Electricity Consumption

Moreover, there are some consumption profile data sets with very low average daily consumption. A consumption profile of total 3.62 kWh is shown in figure 3.4. At figure 3.4

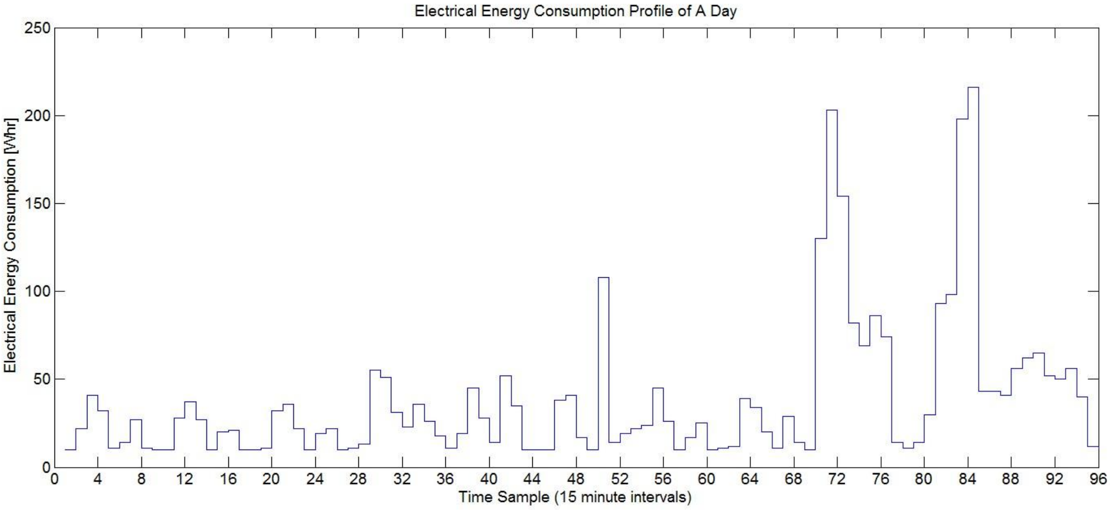


Figure 3.4: Electricity consumption of some particular equipment

During day time the consumption is very low but at night it goes high. So most probably it is the consumption profile of cooling system. Though this consumption profile cannot give the proper idea about the consumer load behavior.

3.1.2 Qualitative variables

Beside input variables (electrical energy consumption), also the qualitative variables are important for forecasting. Qualitative variables better known as dummy variables, do not have a natural ordering. These variables contain descriptive values, like the day of the week. Moreover, all the data are 15 minute sampled so hour information is also a qualitative input variable. Depending on the time electrical consumption varies like at working days from 9.00AM to 6.00 PM consumption should be low and at night when everybody is at home consumption goes high at base station. However, this hour based consumption pattern also depends on seasons. Thus, seasonal effect can be an input variable for forecasting. But this research is focused on STLF and to capture the consumer behavior, electricity consumption data of two or three months is used as training data. So most of the case seasonal identification remains same for all training data. Finally, day identification along with our identification is considered as qualitative input variables for the forecasting model.

3.2 Data Pre-processing

Real life data contains huge amount of noise and often has quality issues. Such volatility must be removed before simulation can be performed. If the input values to a forecasting model are poor, it will be hard to produce a good forecast, irrelevant of the quality of the forecast model. All the steps has to be taken into consideration before simulation as pre-processing is given below

- Duplicate data check
- Missing data check
- Filtering unusual and noise from PV generation data set

Duplicate and Missing Data check:

The electricity consumption data of different base station of a year is given as Wh/15 min. Smart meters are used as measuring device, thus it has high possibility of missing data and duplicate data. Initially the full data set is passed through some checking algorithm to identify duplicate data and missing data as Pre-processing step. However, data with same time stamp is treated as duplicate data. For missing data on weekdays, average value of the immediate 7 weekday's consumption data on the same time sample is taken. However missing data for weekend days, average consumption of the previous 4 same days on the same time sample is calculated. As an example, to find a missing data on weekdays at y_t it should take the average

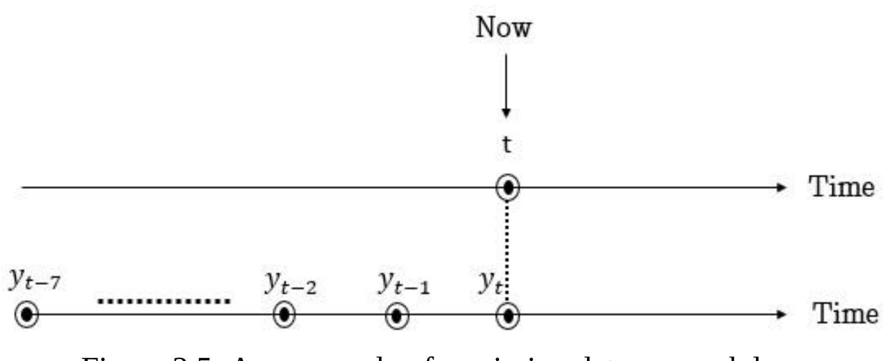


Figure 3.5: Average value for missing data on weekdays

Moreover, to find a missing data on weekend (as an example Saturday) it has to take the average of previous 4 Saturday of the same time sample shown in figure 3.6.

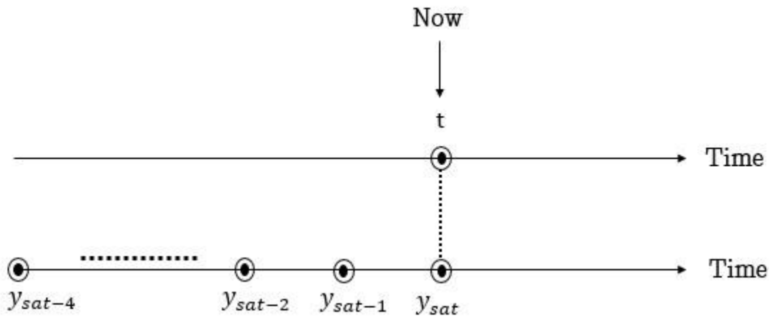


Figure 3.6: Average value for missing data on weekend

Filtering unusual and noise from PV generation data set:

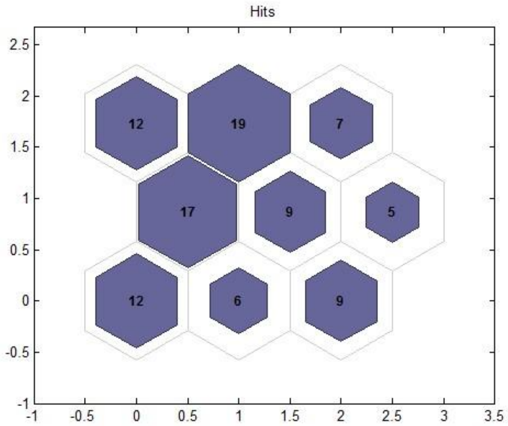
The PV generated data is also measured in 15 minute interval but it is very important to identify the noise or unusual production. Normally electronics based measurement devices are used to capture the data from controller [66]. So to have unusual production peak or noise (like production level 1 or 2Wh) is very common. Moreover, synchronize PV production and consumption data for the same consumer is also important for scheduling of stage device.

3.3 Data Analysis

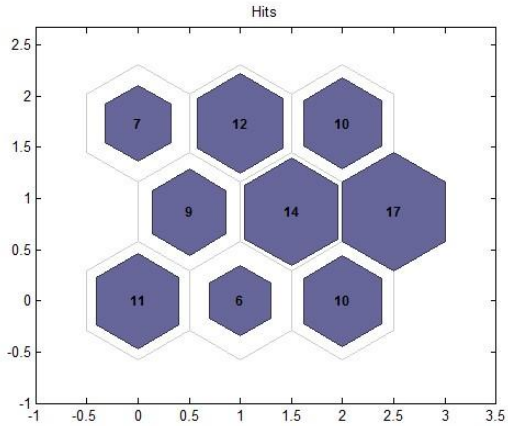
3.3.1 Day identification number as input of ANN

As it is described in section 3.1.2 it is very important to make some difference among the different days of the week so that the models can identify the target data set according to the train data set. Initially these identification variables are coded into integer value. The days of the week is represented by 1,2,3,4,5,6,7 respectively for Monday, Tuesday, Wednesday, Thursday, Friday, Saturday and Sunday.

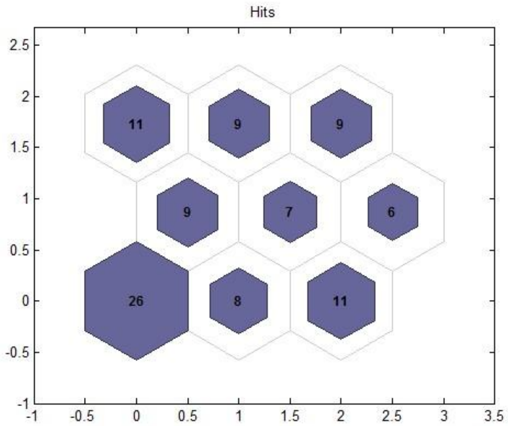
However, to construct a good forecasting model the simulation data need to be known in detail. Correlation or patterns within the electricity use can aid the forecast if well defined. To discover patterns in the electricity use, self-organizing map (SOM) is used. The motivation behind SOM analysis is to find the pattern of consumption of each day of a week. SOM clustering of seven days of a week is shown in Figure 3.7.



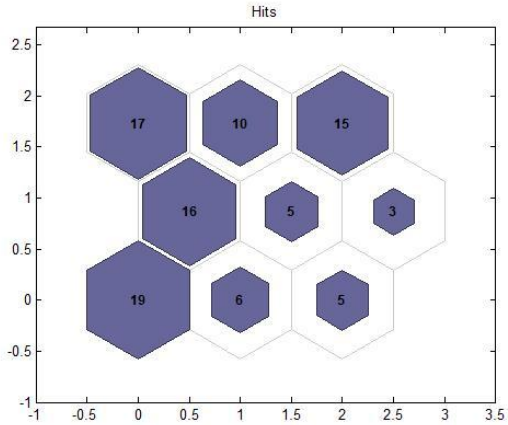
Monday



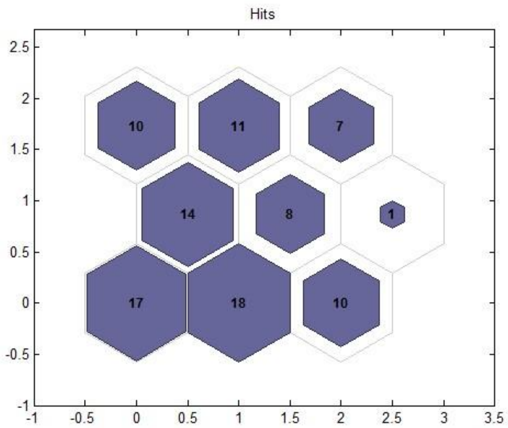
Tuesday



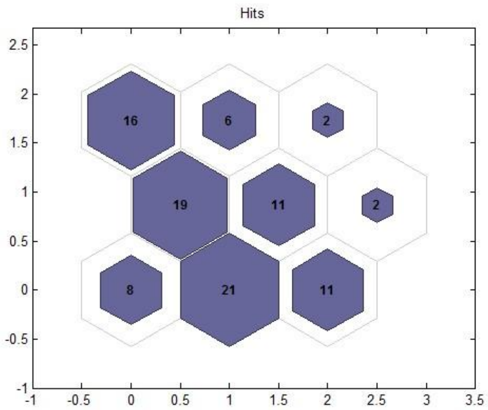
Wednesday



Thursday



Friday



Saturday

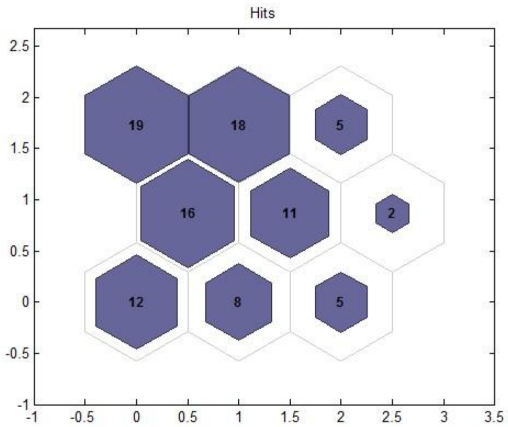


Figure 3.7: SOM clustering of electricity use per day

Sunday

However, Figure 3.7 illustrates different electricity consumption pattern of each day. So, as an input of ANN day identification number will be different for each day.

3.4 Evaluation method

In this section the evaluation method will be discussed. An evaluation method is necessary for three aspects;

- Model identification
- Performance comparison between models
- Insight in model performance for practical use

It will become clear in the next Chapter about model building, within each model different approaches can be used to achieve a forecast. For example, which set of input variables is taken? Or what should the period of the training set be? This is part of the model identification. To answer these questions an evaluation method is set up. The goal of the evaluation method is to select the best performing model. This immediately raises the question, how is 'best performing' defined? In order to answer that question, first the method of forecasting will be elaborated some more.

In order to evaluate a forecast made with a specific model, forecast-error metrics or so called performance indicators are defined. Three types of performance indicators are discussed. The first used performance indicators are the scale dependent metrics. These indicators are on the same scale as the data. They are easy to understand and to calculate. When comparing different data sets these performance indicators cannot be used. One of the most used scaled performance indicator is the mean are extremely large when the actual values reaches zero [67].When considering data which contain zeros, the metric is undefined. An example of percentage based metric is the MAPE, which is one of the most used indicator. Another well-known indicator is mean square error (MSE). The other category is percentage based performance indicators, which can be used to compare different data sets as they are scale independent. Although commonly used, the percentage based indicators are performance indicators in literature. To overcome division by zero, a scale free error is proposed by Hyndman and Koehler in 2006 [68]. The above mentioned performance indicators are given below.

In each of the forthcoming definitions y_t is actual value, f_t is the forecasted value, $e_t = y_t - f_t$ is the forecast error and n is the size of the test set. Also, $\overline{y} = \frac{1}{n}\sum_{t=1}^{n} y_t$ is the test mean and $\sigma^2 = \frac{1}{n-1}\sum_{t=1}^{n}(y_t - \overline{y})^2$ is the test variance.

The Mean Square Error (MSE)

The mean square error is defined as [69], [70]

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(e_t)^2 = \frac{1}{n}\sum_{t=1}^{n}(y_t - f_t)^2 \tag{3.1}$$

Its properties are –

- It measures the average squared deviation of forecasted values.
- It shows the magnitude of overall error, occurred due to forecasting.
- In MSE, the effects of positive and negative errors are canceled out.
- For a good forecast, the obtained MSE should be as small as possible.
- Extreme forecast errors are not panelized by MSE.

The Mean Absolute Percentage Error (MAPE)

This measure is given by [69], [71]

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} | \frac{e_t}{y_t} | \times 100\ \% = \frac{1}{n}\sum_{t=1}^{n} | \frac{(y_t - f_t)}{y_t} | \times 100\ \% \tag{3.2}$$

It's important features are:

- This measure represents the percentage of average absolute error occurred.
- It is independent of the scale of measurement, but affected by data transformation.
- It does not show the direction of error.
- MAPE does not panelize extreme deviations.
- In this measure, opposite signed errors do not offset each other.

Summary:

Here it is discussed some of the error measurement model for judging forecast accuracy of a fitted model. Each of these measures has some unique properties, different from others. In experiments, it is better to consider more than one performance criteria. This will help to obtain a reasonable knowledge about the amount, magnitude and direction of overall forecasted error. Though at individual level expected error is higher than aggregated one, it is preferred to use atleast two of them to evaluate the developed model.

CHAPTER 4

MODEL BUILDING

4.1 ANN based forecasting model

Artificial neural networks (ANNs) constitute a class of flexible nonlinear models designed to mimic biological neural systems of brain. Typically, a biological neural system consists of several layers, each with a large number of neural units (neurons) that can process the information in a parallel manner. Like biological neuron, ANN has also multi-layer structure such that the middle layer is built upon many simple nonlinear functions and able to receive multiple input signals from other neurons. In this chapter the procedure of selecting optimal network architectures and their learning approaches are described for forecasting the wind speed in two different time horizon.

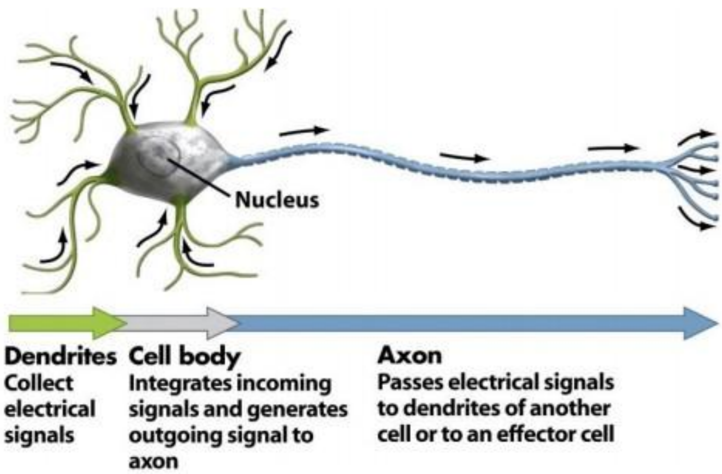


Figure 4.1: Flow of information in biological neuron

4.1.1 Historical Development of ANN

The research on using ANNs for forecasting is vast and growing. The first application dates back to 1964 when M.J.C Hu used Widrow's adaptive model in his thesis for weather forecasting. In the early stage the research was quite limited due to the lack of training algorithm. In the beginning of the 1980's, self-organizing map theory was developed by Teuvo Kohonen, which could represent highly dimensional input data by a low dimensional representation according to an unsupervised learning algorithm. In 1986 the neural network based forecasting revived when the back-propagation model was introduced by Rumelhart et al. The back-propagation algorithm gives a clear indication to adjust weights of neurons for better performance of the network. Lapedes and Farber (1987) conclude that ANNs can be used for modeling and forecasting nonlinear time series. Multi-layer perceptron (MLP), Hopfield's recurrent network and Kohonen's self-organizing network are considered as the most influential models based on ANN.

4.1.2 Working Principle of Forecasting Model

ANNs are networks typically composed of several layers with interconnected elements called neurons. The first or lowest layer is the input layer which gathers external information. The middle or hidden layers process this information in a fairly elementary way to produce signals for the connected neurons at output layer. The neurons or nodes at the adjacent layers are usually fully connected by acyclic arcs from a lower layer (input) to a higher layer (output). Input values are processed by this network to form output values as depicted in figure 4.2.

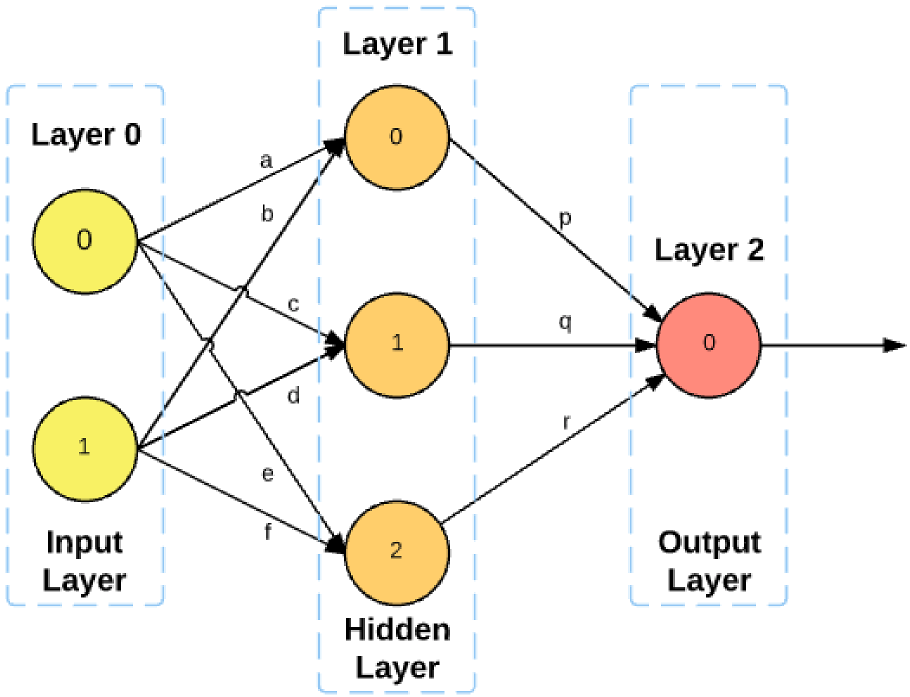


Figure 4.2: A general ANN structure with one hidden layer

Usually in a simple non-linear model, the inputs are independent variables. The functionality estimated by the ANN can be written as:

$$y_1, y_2, \dots \dots, y_n = f\ (x_1, x_2, \dots \dots, x_n) \quad (4.1)$$

Where, $x_1, x_2, \dots \dots, x_n$ are independent input variables and $y_1, y_2, \dots \dots, y_n$ are dependent output variables. Before using ANN to perform desire task, it must be trained. The knowledge learned by the network during training period is stored in the arc and nodes in the form of weight or bias. Weights are the key factors of network performance. Basically training is the process to adjust arc weights which is illustrated in the Figure 4.3 . These weights are continuously updated during the training period to carry out complex non-linear mapping. Training of network can be done in both supervised and unsupervised way. In this research the developed forecasting model is trained in supervised way. In the supervised training procedure the desired response (target) of the network for input pattern is always available. Once the network is trained with appropriate training data set, it is ready to perform desire task.

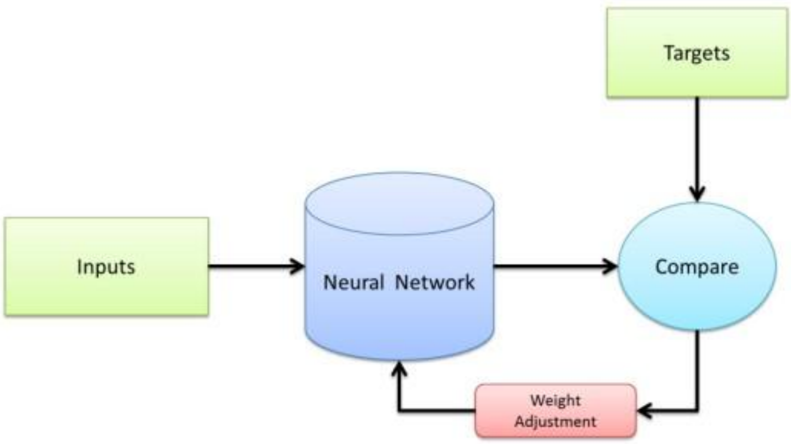


Figure 4.3: Illustration of networks weights in supervised learning process

It is recommended to retrain the network with updated data set especially for forecasting applications. An ANN with supervised learning algorithm is widely used in various fields like pattern classification, within-class categorization, waveform approximation, prediction, control, data analysis, data compression and associative memory. But one cannot generalize an ANN for multipurpose tasks. For different applications, the network architecture should be different.

4.1.3 Network Architecture

Complicacy of ANN mainly arises for the hidden layers which are used to build the connection between inputs and outputs. Thus, it creates an indirect relationship between inputs and outputs. Information received from the input layer is first processed in the hidden layer, and then transmitted to the output layer. So the learning capability due to nonlinearity of the input data is mostly depend on number of hidden layers. Depending on the requirement of the problem number of hidden layers may vary, for simple problem it may need just one layer whether some needs several to train up the model. But usage of unnecessary hidden layers may create problem during simulation by taking more training time. The optimum generalization performance can be obtained by trading the training error against network complexity.

In literature it has been found that a multilayer forecasting model with one hidden layer can perform an arbitrary convex approximation to any continuous non-linear mapping. According to the universal approximation theorem for neural networks, the standard multilayer feed forward network with a single hidden layer and finite number of hidden neurons is sufficient for any complex simulate nonlinear function with any desired accuracy. However it concludes with the view that number of hidden layers has an influence according to the complicacy of the problem otherwise system will be more complicated, even single hidden layer requires large number of nodes. So it is always a tradeoff of number of hidden layers depending on complicacy of the problem.

4.1.4 Selection of the Network

Depending on the structure of the network ANN can be classified in several model, among them Feed-forward and Recurrent network model are the simplest one. Figure 4.4 shows different types of neural networks used for forecasting applications.

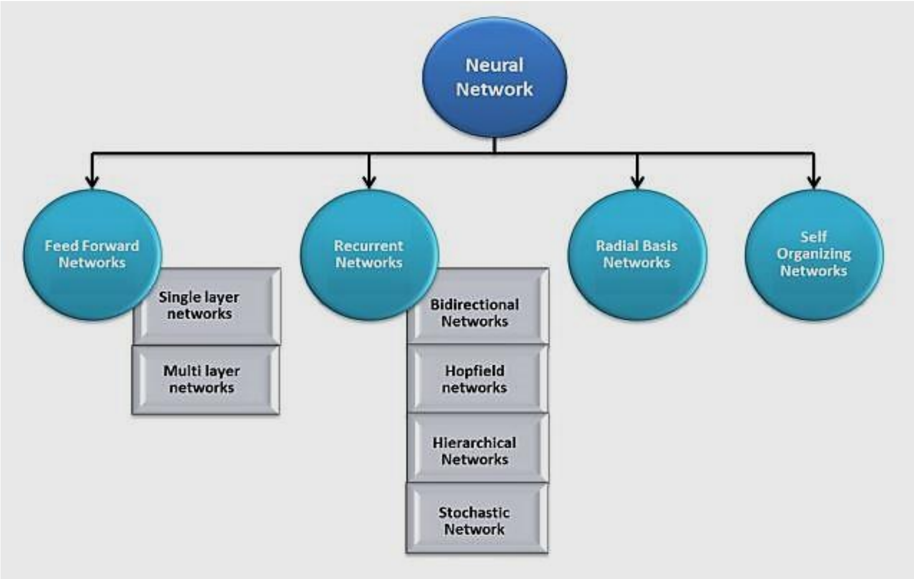


Figure 4.4: The taxonomy of ANN architecture

Feed-forward neural network (FNN) is fast and simplest ANN because in the case of Multilayer network it transmits information form input layer to output layer using some simple structured hidden layer. FNN is a biologically inspired algorithm so it is very easy to model the network. Normally it maps the static relationship between input and outputs through a unidirectional information flow (only form input towards output).

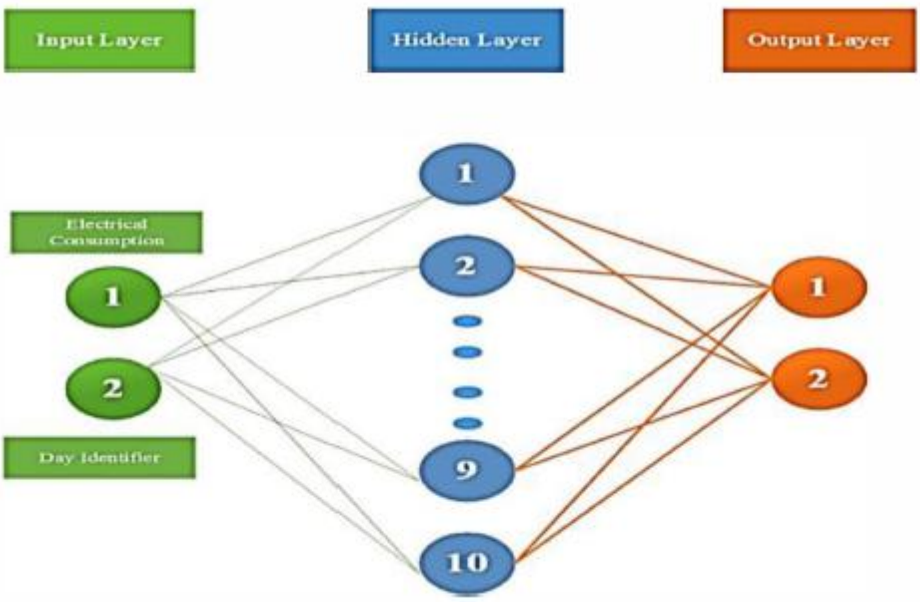


Figure 4.5: Structure of FNN

The inputs are fed directly to the outputs via a series of weights. These connections are not all equal; each connection may have different weights because these weights encrypt the knowledge of a network. A general structure of FNN is shown in figure 4.5. During normal operation there is no feedback between layers as refer to the name of the model, but if it is needed depending on the application FNN allows internal feedback. In that case they are computationally more powerful and biologically more acceptable than other adaptive approaches.

Due to lack of any feedback, FNN must use a large number of input neurons for learning the historical data and consequently achieving good results. However, they need large historical data and have a limited capability to predict loads of holidays and fast load changes. But recent studies reveal that randomly initialized Recurrent Neural Network (RNN) has interesting and useful features even without training. When the RNN is initialized with small random weights, the network like a memory is able to reflect the history of inputs successively fed to it. Thus, instead of learning the RNN which is a complex task, a large RNN as a hidden or internal layer is used actually as a rich reservoir of complex dynamics. The presence of internal dynamics increases its richness among other neural network models but the random weight choosing is much more related to expertise level of developer.

From the literature review it has been found that both ANN networks have better performance in nonlinear forecasting applications than other traditional methods. Despite of being simple, feed-forward networks are more efficient in complex nonlinear forecasting problem with bigger number of input variables.

So this is the motivation behind choosing FNN for its simplicity and scope of use large individual consumption data to train the model. On the other hand, FNN is fast and effective to use for both offline/online operations (as future scope of this research) especially in forecasting problem with 15 minutes time horizon.

4.1.5 Mathematical Model

In [71] a mathematical model has been developed referring to the figure 4.5

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i}\right) + \varepsilon_t , \forall t \tag{4.2}$$

Here y_{t-i} $(i = 1, 2, 3, \dots\dots p)$ are the input and y_t is the output. The integers p and q are the number of input and hidden nodes respectively. α_j $(j = 1, 2, 3, \dots\dots q)and\beta_{ij} = (i = 1, 2, 3, \dots\dots p; j = 1, 2, 3, \dots\dots q)$are the connection weights and ε_t is the random variable, α_0 , β_0 are the bias term. Usually, the logistic sigmoid function $g(x) = \frac{1}{1+e^{-x}}$ is applied as the nonlinear activation function. Other activation functions, such as linear, hyperbolic tangent, Gaussian, etc. can also be used depending upon the use case.

The feed forward ANN model in equation 4.2 in fact performs a non-linear functional mapping from the past observations of the time series to the future value, i.e$y_t = f\left(y_{t-1}, y_{t-2}, y_{t-3}, \dots\dots y_{t-p} , \boldsymbol{W}\right) + \varepsilon_t$ where $\boldsymbol{W}$ is a vector of all parameters and f is a function determined by the network structure and connection weights. To estimate the connection weights, non-linear least square procedures are used, which are based on the minimization of the error function

$$F(\varphi) = \sum_t e_t^2 = \sum_t (y_t - \hat{y}_t)^2 \tag{4.3}$$

Here φ is the space of all connection weights. The optimization techniques used for minimizing the error function equation 4.3 are referred as Learning Rules. The best-known learning rule in literature is the back propagation or Generalized Delta Rule [72].

4.1.6 Network Training

The training of the network is divided in three states: training, testing and validation. So the total dataset is divided into three samples. The training sample is used for development of the model so that network can be adjusted according to the error. The testing sample is used to evaluate the forecasting ability or accuracy of the network. It has no effect on training so provides an independent measure of network performance during and after training. The validation sample is used to measure network generalization. When the improvement of the generalization stops, the training process is halted. Thus the over-fitting of the network is avoided. For small dataset testing and validation are combined in one set. The division of the data into training, testing and validation is an important issue. In previous literature, it is common practice, so call the standard dataset partition i.e. 70% for training, 15% for testing and 15% for Validation. Once the network architecture and training dataset is ready, the network is trained and saved for the desire forecasting task.

Summary:

A neural network is basically combinations of many interconnected neurons which are categorized into different layers. Input and output neurons are specific groups only for input data and output of results but the hidden neurons can be distributed in one or multiple hidden layers. During the design steps of ANN one must determine the number of neurons in input and output layer and number of hidden layers. The selection of these parameters are mainly problem dependent. There are many different approaches exist in the literature to select these parameters. But there is no simple clear-cut method which can guarantee the optimal solution of all real forecasting problems so complete ANN model development is dependent to the problem and application.

CHAPTER 5

SIMULATION RESULTS AND ANALYSIS

The optimal structure of ANN, ARIMA and SVM used to forecast the BTS electricity consumption on day-ahead. The imbalance power is calculated by the power deviation between day-ahead forecaster and 15 minutes ahead forecaster. The flexible entities are activated in way to reduce this imbalance power. The model and simulation of the software for the use case is discussed in section 4.1. In section 4.2, a simple MATLAB-based GUI has been implemented to demonstrate how this tool can be user-friendly for a customer. Finally the summary of the chapter has been presented in section 4.3.

5.1 Assumptions

To maintain the time and simplify the method, some assumptions are made as following in this thesis work:

- The historical data are in Wh and the time interval is 15 minutes.
- As input, total 92 days of previous data is taken for preprocessing. The time range is March 1st, 2013 to May 31st, 2013 the forecasted data represents the consumption for the next day, June 1st 2013.
- The input data is divided in three data sets, for training, testing and validation respectively. For this study, 70% data is used for training, and the rest 30% for testing.
- The error calculated in the testing phase.

5.2 Electricity Consumption Forecasting

5.2.1 Model Evaluation

Model of Feed Forward Neural Network (FNN):

Designing a three-layer FNN model for STLF involves several major steps:

1. Determine the number of outputs: An FNN model may have only one output, which can be corresponding to the electrical load of a consumer, or several outputs, which can represent a 24-hour load profile of several BTS's. As some drawbacks of the multiple outputs FNN were discussed, in this chapter, a single-output FNN models, of which the output is individual BTS energy consumption.
2. Determine the number of inputs: The inputs may include weather variables, calendar variables, and the load of the preceding hours. For this study, previous consumption data of a particular consumer along with day identification number is used as target and input of FNN.
3. Determine the number of hidden neurons: There is no common rule for determining the number of hidden neurons. So it is followed a trial and error method to select an optimal hidden layer structure for used data set at FNN.

From the consumption profile, the amount of consumed power is highly nonlinear. So, an optimal hidden neural layer is required for best result. The number of hidden neuron layer is mostly depends on these three parameters.

1. MAPE – shows the level of accuracy of the forecaster
2. Elapsed Time – for this research MATLAB® has been chosen as simulation interface so the required time for simulation a FNN model is an important issue. If it is needed too much time, in practical use it may create problem for scheduling of storage devices when it has to handle too much data.
3. Epoch – it denotes the number of iteration to converge the model. More iteration gives better response and good predicted output.

Observing these parameters, the best output is obtained from single hidden layer amongst different options. Rest of the simulation is completed using 20 neurons. It is clear from the table that the lowest MAPE and lowest MSE are found for the chosen number of layer.

Model of Auto Regressive Integrated Moving Average (ARIMA):

To evaluate the performance ARIMA, forecasted electricity consumption profiles of 28 Base Station for the same day are compared. The comparison criteria are the MAPE and MSE. For the ARIMA, the previous 92 days (from March'13 to May'13) consumption data is used. But the model runs 96 times to predict the 96 points of the consumption profile of one day respectively. The consumption data is in Whr. The assumption of the authors is that 92 days is enough as an adaptive period for the agent to capture changes in electricity consumption, while the 15 minutes resolution is a common practice from a data acquisition point of view.

Model of Support Vector Machine (SVM):

Designing a three-layer SVM model for STLF involves several major steps:

1. The classifier was trained using SVM. For performing SVM operation the Radial Basis Function (RBF) Kernel Trick was used.
2. Determine the number of outputs: The SVM model may have only one output, which can be corresponding to the electrical load of a consumer, and represents a 24-hour load profile of several BTS's. A single-output SVM model, of which the output is individual BTS energy consumption.
3. Determine the number of inputs: The inputs may include weather variables, calendar variables, and the load of the preceding hours. For this study, previous consumption data of a particular consumer along with day identification number is used as target and input of SVM.
4. Determine the number of hidden Layer: There is no common rule for determining the number of hidden layer. For this study, a single hidden layer SVM model was used. From the consumption profile, the amount of consumed power is highly nonlinear. So, an optimal hidden neural layer is required for best result. The number of hidden neuron layer is mostly depends on these three parameters.
    - MAPE – shows the level of accuracy of the forecaster

Observing these parameters, the best output is obtained from single hidden layer amongst different options. It is clear from the table that the lowest MAPE, MSE is found for the chosen number of layer.

Process of fitting the data to a mathematical model:

- After the acquisition of the data to MATLAB, the data is fed to the 'Curve Fitting' application toolbox. The 'Curve Fitting' application works based on the principle of 'least square regression'.
- Different built-in and customized models are fitted for the data in that toolbox and the sum of square error (SSE) and the adjusted R-squared are observed for each. It is observed that, the exponential and double exponential models can closely follow the data pattern.
- The final model for each data is selected with the least SSE and adjusted R-squared close to 1.

5.2.2 Demonstration of day-ahead consumption forecasting

From the very beginning of the research work, the main goal is to forecast the next day power consumption of a BTS and compare the performance of the models. To achieve the goal, accuracy of the forecasted model is very much required. So, the selected optimal models ANN, ARIMA and SVM are used to forecast the energy consumption.

5.2.2.1 Consumption forecasting on mid-summer (June 01, 2013)

As it is mentioned earlier, practical consumption data from 28 BTS will be used to evaluate the forecasting performance of the models. In this context, at first it is tried to forecast the consumption profile of each consumer on the same day with training data set of previous 3 months (92 days). In the following sections the performance of FNN, ARIMA and SVM has been described.

To have the forecasted consumption profile at 01 June, 2013 which is a weekend (Saturday), training data has been chosen of previous three months. The noticeable aspect here, March is the end of winter and the target day is mid-summer. But here no seasonal identification number or variable is used. The external input variables are the day identification number and hour information as mention in Section 3.1.2. For all consumers the forecasting accuracy by means of different errors is shown in Table 5.1.

Table 5.1: Performance Evaluation

Consumer	ARIMA Evaluation		SVM Evaluation		FNN (10) Evaluation		FNN (20) Evaluation	
	MSE	MAPE	MSE	MAPE	MSE	MAPE	MSE	MAPE
1.	8584	320.53	4532	199	3665	33.9695	731.90	22.86
2.	13988	149.07	11017	141.06	1285	39.39	4693.80	29.02
3.	13589	163.85	10003	172.22	8725	28.5366	2410.6	23.02
4.	11721	183.41	9809	141.08	8496	52.8844	5925.5	31.10
5.	18198	108.15	15492	133.47	8735	15.9883	5463.9	10.59
6.	16749	83.59	14080	74.22	6854	103.0053	4398	32.77
7.	11217.4	105.02	9237	130.32	5879	47.334	109.45	28.48
8.	12667	165.06	8890	481.2	3492	25.3231	1245.6	33.84
9.	1481.9	55.52	9153	146.2	2691	26.5852	113.36	24.99
10.	8833	320.96	7638	186	8943	33.9	3007.1	22.38
11.	12743	181.79	8189	108.3	4791	52.6619	267.1	27.82
12.	9378.7	98.62	7892	84.46	3961	79.4627	706.92	13.65
13.	1468.5	85.97	7685	227.3	2819	227.0793	192.82	81.21
14.	8445	193.75	5892	101.1	8055	40.9806	8611.3	84.98
15.	8753	244.35	3889	145	1703	42.787	732.52	17.74
16.	7295.7	86.01	5610	146.9	1019	43.5497	254.87	21.21
17.	12156	200.95	9156	197	8053	19.4287	4580.1	26.85
18.	11631	263.32	8974	160.1	6872	40.17	3615.7	20.21
19.	9769	247.46	6978	171.9	5439	46.3432	2903.3	47.59
20.	7272	151.82	3941	153.03	3205	18.249	445.97	11.97
21.	9860.3	125.70	7928	180.53	5734	53.9584	1360.91	23.78
22.	4457.2	696.86	2644	150.78	1903	20.6975	181.29	22.25
23.	8294	237.52	6837	162.38	2938	40.1062	704.11	33.09
24.	6798.1	111.74	2789	186	1973	114.3078	98.39	25.09
25.	9620.1	205.43	7785	75.7	5493	72.9944	2438.01	11.13
26.	8414.9	311.23	7639	539.79	9835	68.4901	3489.15	20.48
27.	7658.1	487.03	4899	144.299	9345	28.0127	5450.49	24.49
28.	9806	598.88	8445	134.133	2109	20.6213	406.06	11.60

Form Table 5.1, it can be mentioned that FNN performed well. Some out of range data is found in MAPE which happened because those consumers had no usage of electricity in their consumer profile at some hours of the day. Now some best cases will be described. Considering the case of MAPE, consumer 5 has the best output as 15.9883% for FNN while both for ARIMA and SVM of the same consumer are 108.15 and 133.47 respectively. Though other error checking parameters have some issues, from the point of MAPE it shows the best outcome. Consumer 5 has a moderate level of error (MSE = 8735).

5.2.2.1.1 Forecasting with FNN

The forecasted pattern for consumer 5 is shown below in Figure 5.1. Almost at every point the forecasted data has matched the positions of the peaks. But at higher consumption areas some mismatch has found. But it’s not the error of the model. The consumer has changed the trend of electricity consumption. So, the model predicted as that trend, but it is clear that the model can identify the positions and the shape of the peaks. In Figure 5.2, the regression function and MAPE for the consumer is shown. The dotted line in the regression represents the exact result and the solid line represents the best fit regression line between the input and the output. The main reason for this deviation is for changing the behavior of the consumer at that particular time.

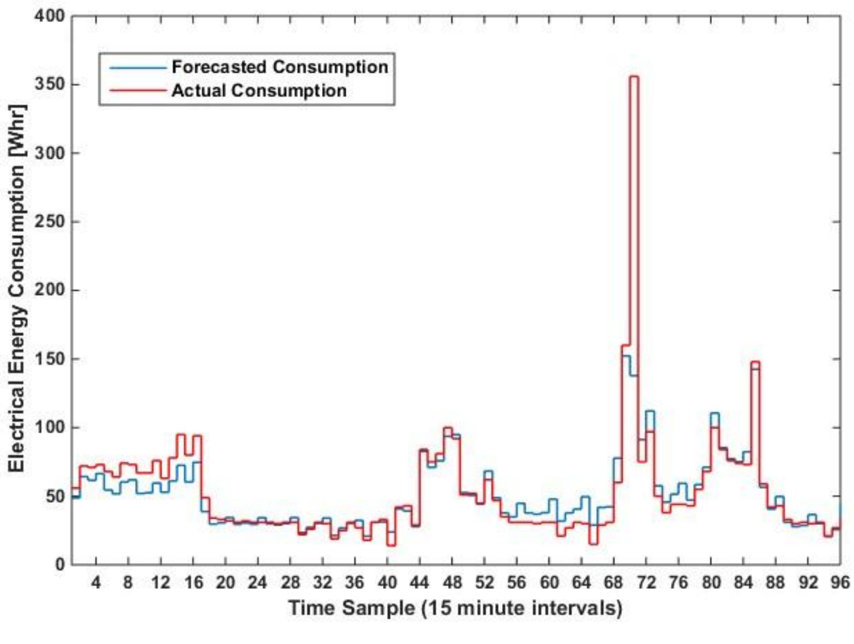


Figure 5.1: Forecasting performance of FNN for consumer 5.

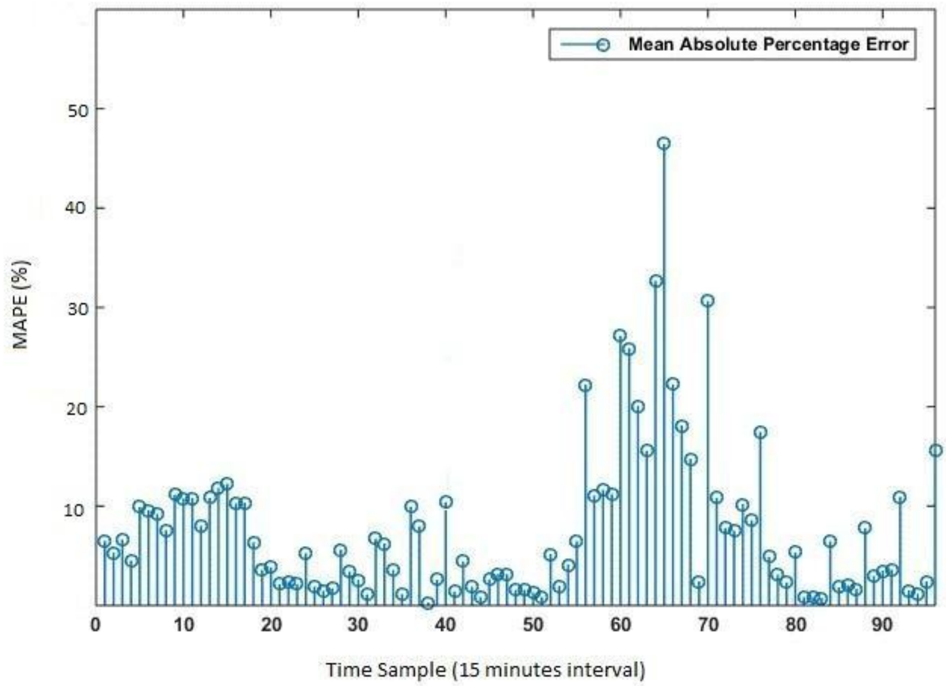


Figure 5.2: MAPE of forecasted output of consumer 5 by FNN.

For further evaluation of the performance the model, another output is exempted from the point of MSE. From Table 5.1, least MSE belongs to consumer 16 (MSE=1019). But the percentage error is as high as MAPE=43.55%.

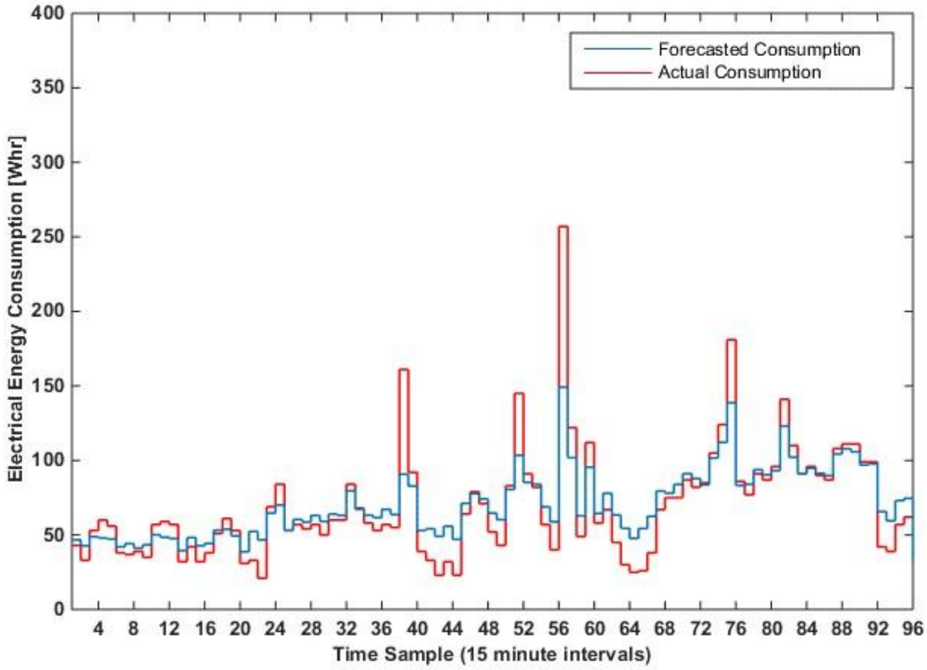


Figure 5.3: Forecasting performance of FNN for consumer 8.

In Figure 5.3 it is clearly visible that the output profile almost followed the trend. And most of the points are well predicted, though many of them are lower than the actual data. Other parameters are also shown in Figure 5.4, for demonstration.

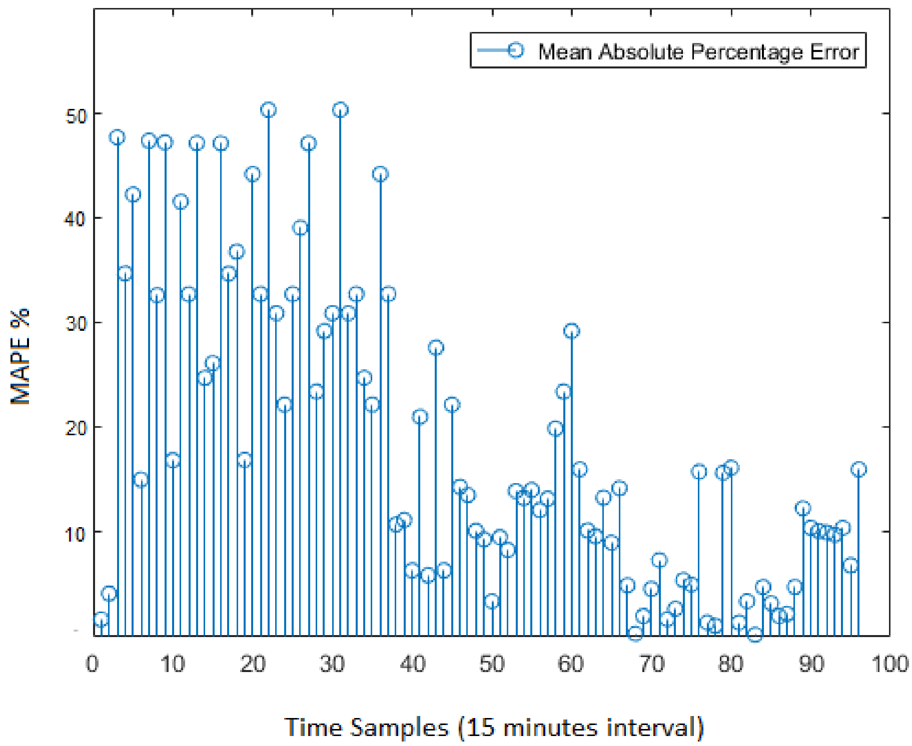


Figure 5.4: MAPE of forecasted output of consumer 8 by FNN.

From the above Figure 5.4, obtained regression function is comparatively good. The mismatched area lies in between 50 to 100Wh and the MAPE shows a scattered pattern of error. Moreover it is clear that the errors largely occurred at the peaks.

Finally, another consideration is done for other consumer for optimized evaluation. In this case the overall performance has been considered. The selected one is being consumer 2. For the selected consumer, obtained parameters are at moderate level as MSE=8285; MAPE=39.39. MAPE has been obtained in an acceptable range. But the MSE is so large that can affect the overall performance of the forecasting model.

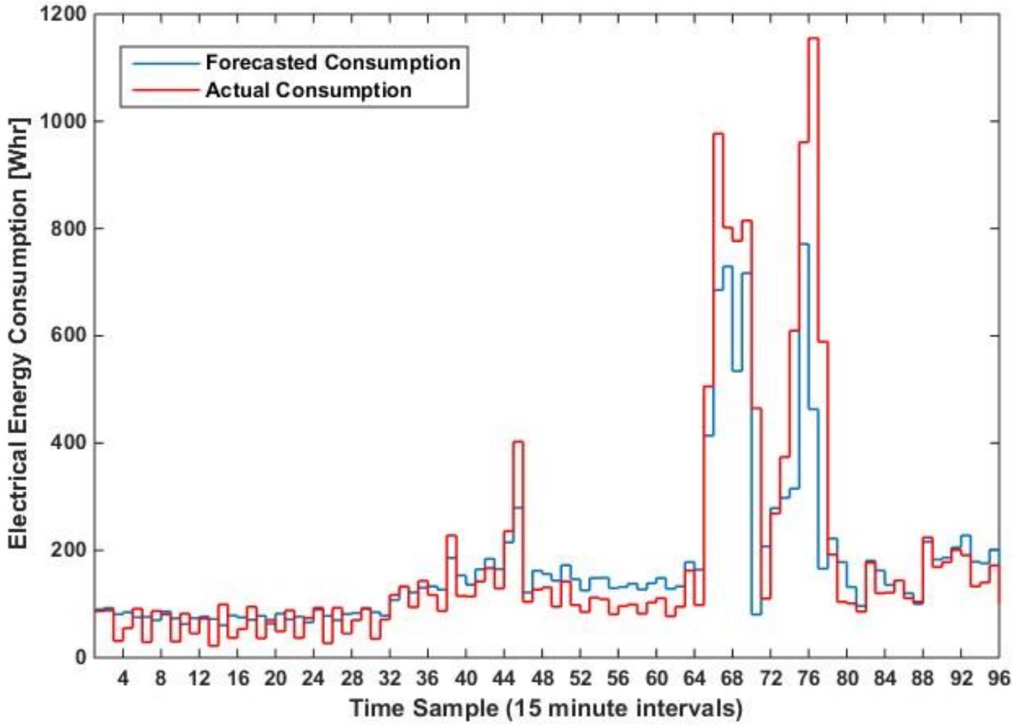


Figure 5.5: Forecasting performance of FNN for consumer 2.

From the Figure 5.5, the forecasted data is matched in almost every point with the actual consumption profile. The major mismatch is found at the higher peaks. Though the forecasted model has followed and detected the positions of the peaks well, but the values estimated are lower than the actual consumption. This mainly happens for the change in the consuming behavior of the consumer.

Detailed demonstrations about the regression function and percentage error is done is the next portion. Figure 5.6, illustrates the model of MAPE of consumer 2.

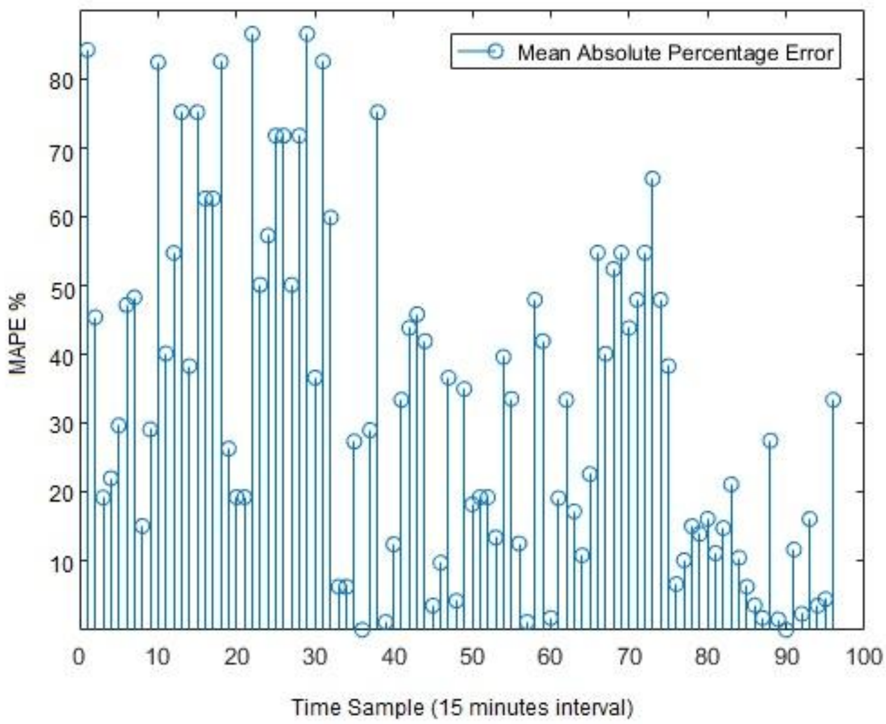


Figure 5.6: MAPE of forecasted output of consumer 2 by FNN

The regression function plotting is lineated based on the best points to be found. The most mismatching is found around 100 to 300Whr. Noticeable percentage errors are found at the peaks from the scattered plotting. And the reason behind this is discussed before. Hence the brief scenario for the forecasting model can be demonstrated.

5.2.2.1.2 Forecasting with ARIMA

The forecasted pattern for consumer 5 is shown below in Figure 5.7. Almost none of the points the forecasted data has matched the positions of the peaks. It is the limitation of this model. But it is clear that the model neither identify the positions nor the shape of the peaks. In Figure 5.8, the regression function and MAPE for the consumer is shown. The dotted line in the regression represents the exact result and the solid line represents the best fit regression line between the input and the output. The main reason for this deflection is ARIMA shows poor performance for nonlinear data.

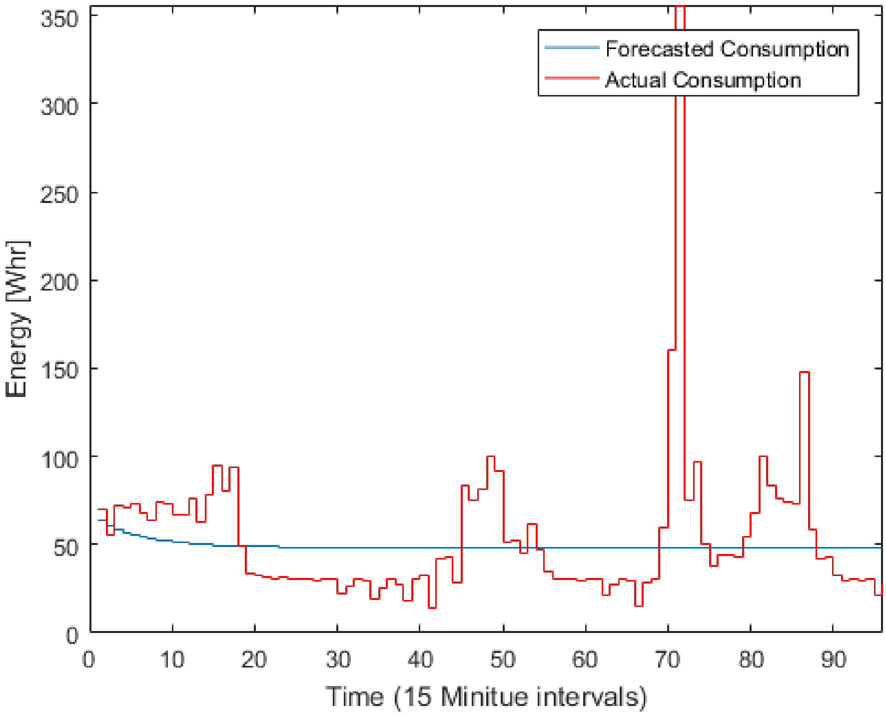


Figure 5.7: Forecasting performance of ARIMA for consumer 5.

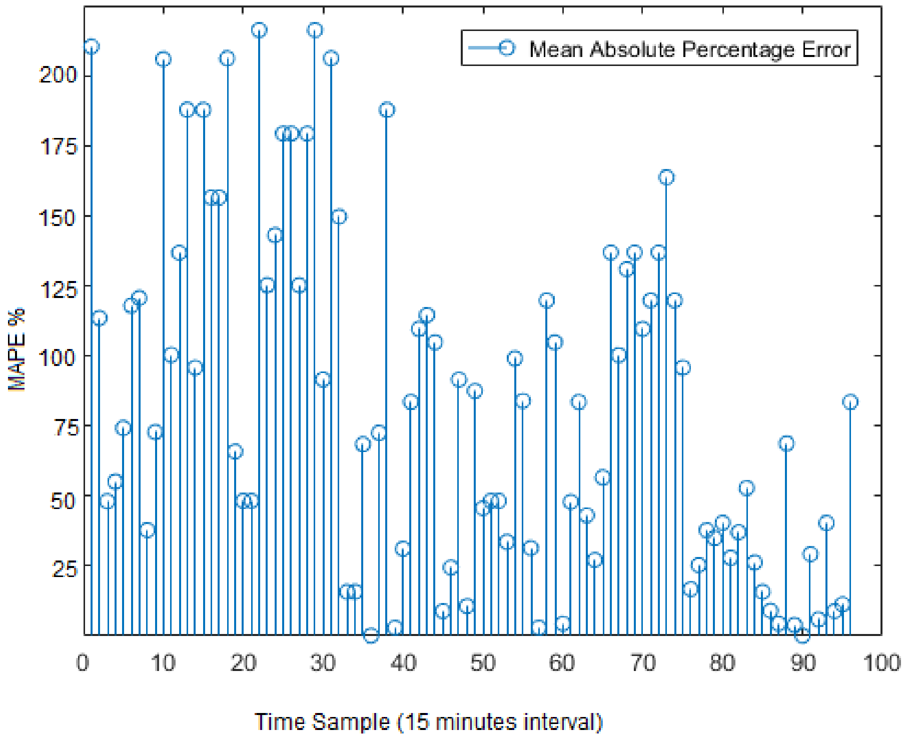


Figure 5.8: MAPE of forecasted output of consumer 5 by ARIMA.

For further evaluation of the performance the model, another output of consumer 8 is exempted as the same prediction has been done by FNN. To compare the forecasting ability of this model consumer 8 is being examined where from Table 5.1, MSE belongs to consumer 8 (MSE=12667). But the percentage error is as high as MAPE=165.06%.

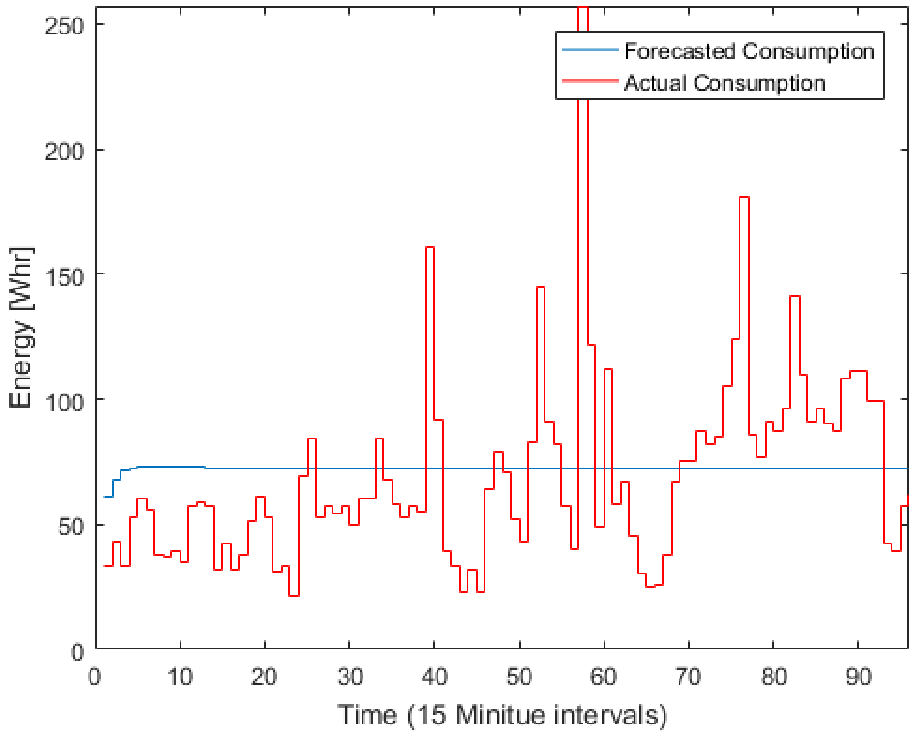


Figure 5.9: Forecasting performance of ARIMA for consumer 8.

In Figure 5.9 it is clearly visible that the output profile can't follow the trend. And most of the points are not well predicted, and many of them are lower than the actual data. Other parameters are also shown in Figure 5.10, for demonstration.

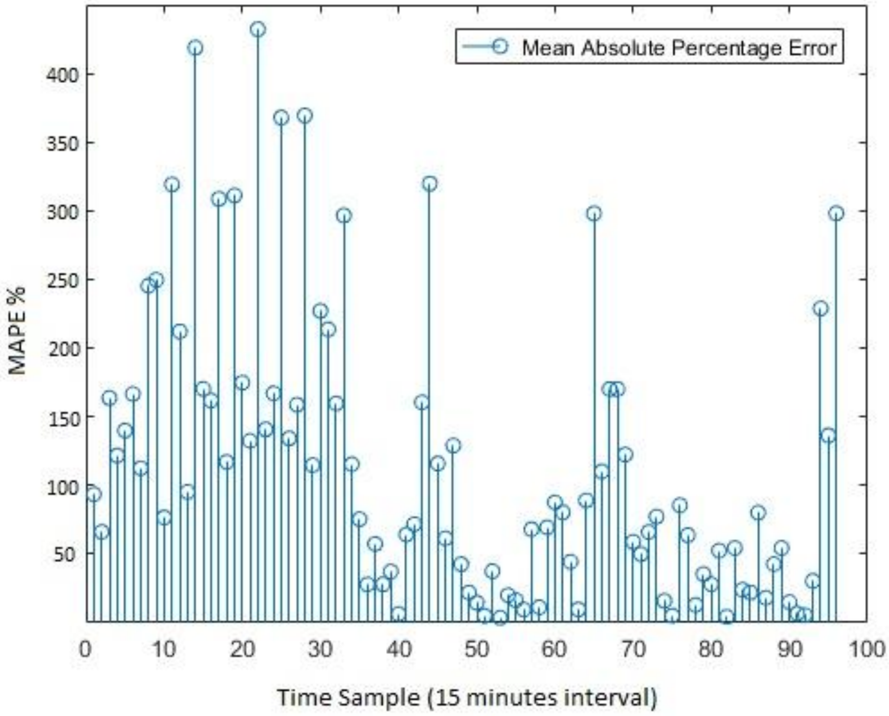


Figure 5.10: MAPE of forecasted output of consumer 8 by ARIMA.

From the above Figure 5.10, obtained regression function is not convincing at all. The regression curve shows a large area of mismatch and the MAPE shows a scattered pattern of error. Moreover it is clear that the errors largely occurred at the peaks.

Finally, as all the models are examined for a fixed consumer. The selected one is being consumer 2. For this selected consumer, obtained parameters are at moderate level as MSE=13988; MAPE=149.07. MAPE has been obtained in an acceptable range. But the MSE is so large that can affect the overall performance of the forecasting model.

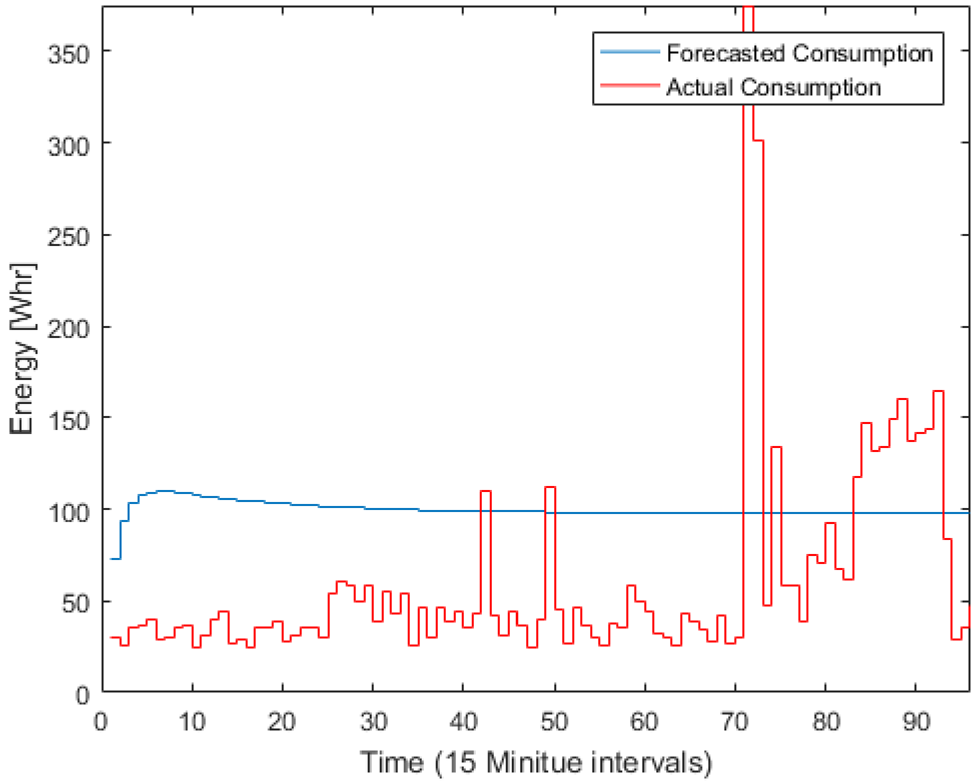


Figure 5.11: Forecasting performance of ARIMA for consumer 2.

From the Figure 5.11, the forecasted data is mismatched in almost every point with the actual consumption profile. The major mismatch is found at the higher peaks. Moreover the forecasted model has not followed and detected the positions of the peaks well, but the values estimated are higher while the actual consumption is low but when the actual consumption is higher than the estimated value is much lower than the actual consumption. This mainly happens for the limitations of handling nonlinear data and change in the consuming behavior of the consumer.

Detailed demonstrations about the regression function and percentage error is done is the next portion. Figure 5.12, illustrates the model of R and MAPE of consumer 2.

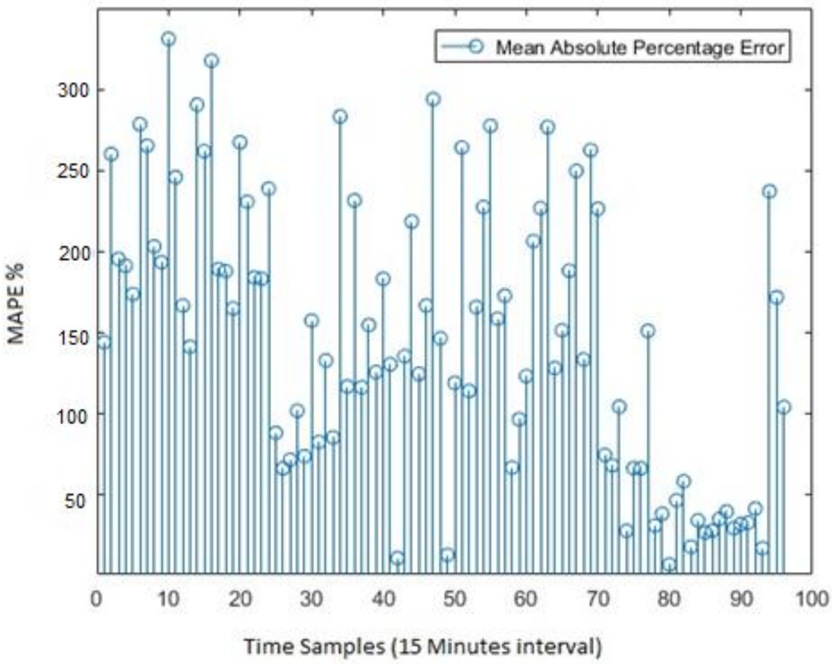


Figure 5.12: MAPE of forecasted output of consumer 2 by ARIMA

The regression function plotting is lineated based on the best points to be found. The curve mostly found mismatched. Noticeable percentage errors are found while the actual load consumption is very low. And the reason behind this is discussed before. Hence the brief scenario for the forecasting model can be demonstrated.

5.2.2.1.3 Forecasting with SVM

The forecasted pattern for consumer 5 is shown below in Figure 5.13. Almost at every point the forecasted data has matched the positions of the peaks. But at higher consumption areas some mismatch has found. But it's not the error of the model. The consumer has changed the trend of his using the electricity. So, the model predicted as that trend, but it is clear that the model can identify the positions and the shape of the peaks. The main reason for this deflection is for changing the behavior of the consumer at that particular time.

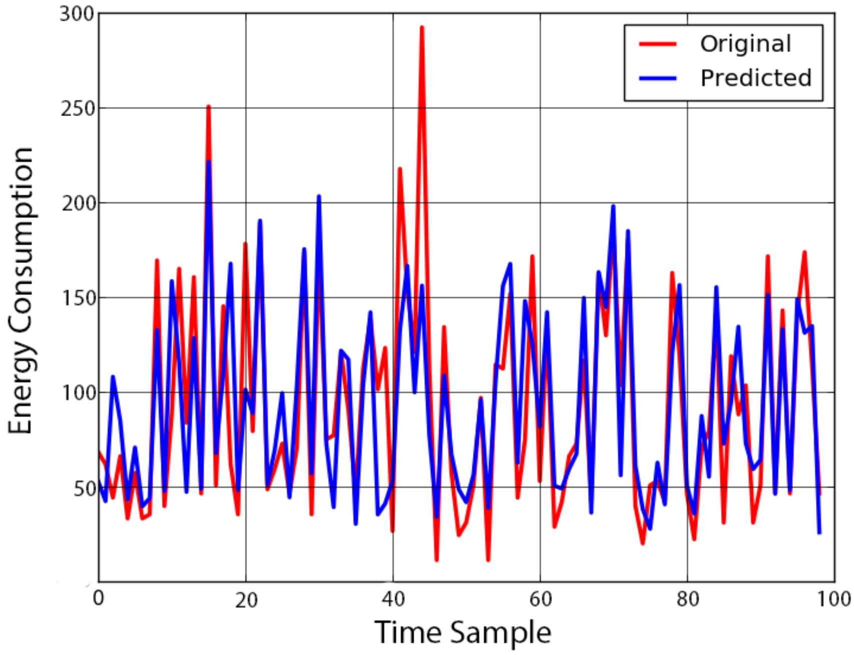


Figure 5.13: Forecasting performance of SVM for consumer 5.

For further evaluation of the performance the model, another output is exempted for consumer 8. From Table 5.1, MSE belongs to consumer 8 (MSE=8890). But the percentage error is as high as MAPE = 481.2%.

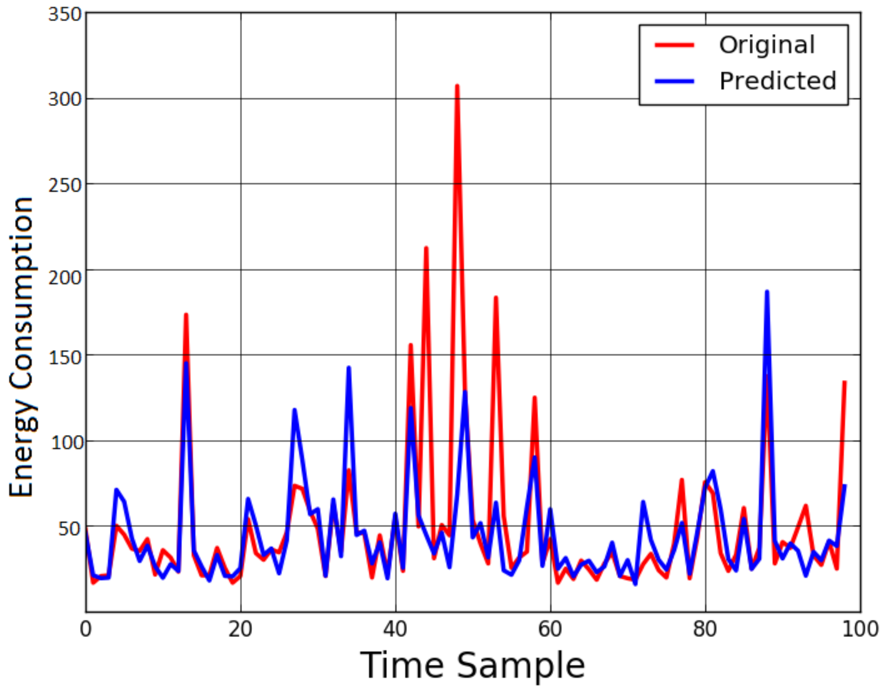


Figure 5.14: Forecasting performance of SVM for consumer 8.

In Figure 5.14 it is clearly visible that the output profile almost followed the trend. And most of the points are well predicted, though many of them are lower than the actual data.

Finally, another consideration is done for other consumer for optimized evaluation. In this case the overall performance has been considered. The selected one is being consumer 2. For the selected consumer, obtained parameters are at moderate level as MSE=11017; MAPE=141.06. In this case both MSE and MAPE has been obtained in an acceptable range.

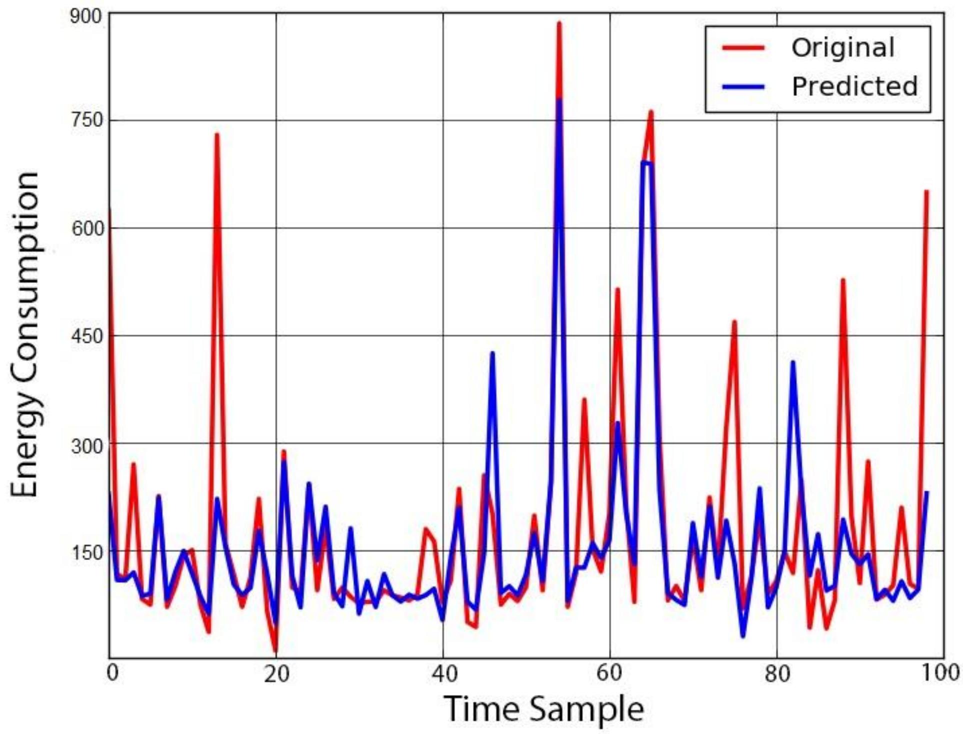


Figure 5.15: Forecasting performance of SVM for consumer 2.

From the Figure 5.15, the forecasted data is matched in almost every point with the actual consumption profile. The major mismatch is found at the higher peaks. Though the forecasted model has followed and detected the positions of the peaks well, but the values estimated are lower than the actual consumption. This mainly happens for the change in the consuming behavior of the consumer. Noticeable percentage errors are found at the peaks from the scattered plotting. And the reason behind this is discussed before. Hence the brief scenario for the forecasting model can be demonstrated.

From the above experiments, it is clear that FNN showed better performance than other short term load forecasting methods. Following these results a further simulation processed to find the effect of hidden layer or number of neurons. And it gives a obvious result shown in table 5.2 that with increasing number of neurons corresponding errors becomes reducing. But after the point corresponding 20 number of neurons gives no appreciable change in the errors. Table 5.2 shows that for given consumers the corresponding MSE is reducing with increasing the number of hidden neurons.

Table 5.2: MSE Comparison for different consumer varying number of neurons

Number of Neurons	Consumer No. 1	Consumer No. 5	Consumer No. 10	Consumer No. 15	Consumer No. 20	Consumer No. 25
5	5438	11392	10973	3182	5748	6019
10	3665	8735	8943	1703	3205	5493
15	1795	6785	5918	1071	1573	3316
20	731.90	5463.9	3007.1	732.52	445.97	2438.01
25	674	4973	2284	576	372	2073

Similar results shown in table 5.3 for MAPE. The obtained data is plotted in figure 5.16, which indicates that the mean absolute percentage error decreasing with increasing the number of neurons. Where from that graph it is clear that all consumers follow almost same shape of graph. So from that graph a pattern of the consumer can be identified.

Table 5.3: MAPE Comparison for different consumer varying number of neurons

Number of Neurons	Consumer No. 1	Consumer No. 5	Consumer No. 10	Consumer No. 15	Consumer No. 20	Consumer No. 25
5	47.87	28.87	36.75	53.54	28.82	96.57
10	33.9695	15.9883	33.9	42.787	18.249	72.99
15	29.36	13.87	28.79	35.56	13.36	57.63
20	22.86	10.59	22.38	17.74	11.97	11.13
25	19.78	9.87	16.69	10.85	9.78	9.93

And that pattern is an exponentially decreasing curve with increasing number of neurons. And there is an intersecting point at around 18 numbers of neurons. Also it is clear from that graph that after the points of 20 numbers of neurons the corresponding errors doesn’t improving significantly.

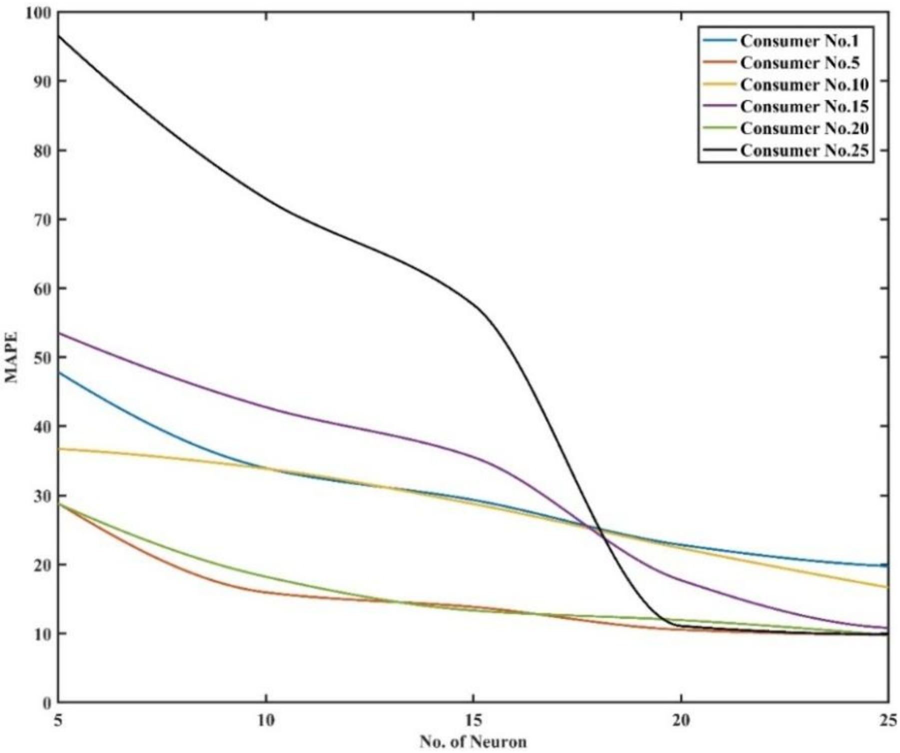


Figure 5.16: MAPE Vs Number of Neurons

Rather after this point the iteration or processing time increases prominently. Moreover from that graph it is seen that for consumer 20 the graph shows a stable pattern. So using the curve fitting technology an equation is drawn both for MSE and MAPE.

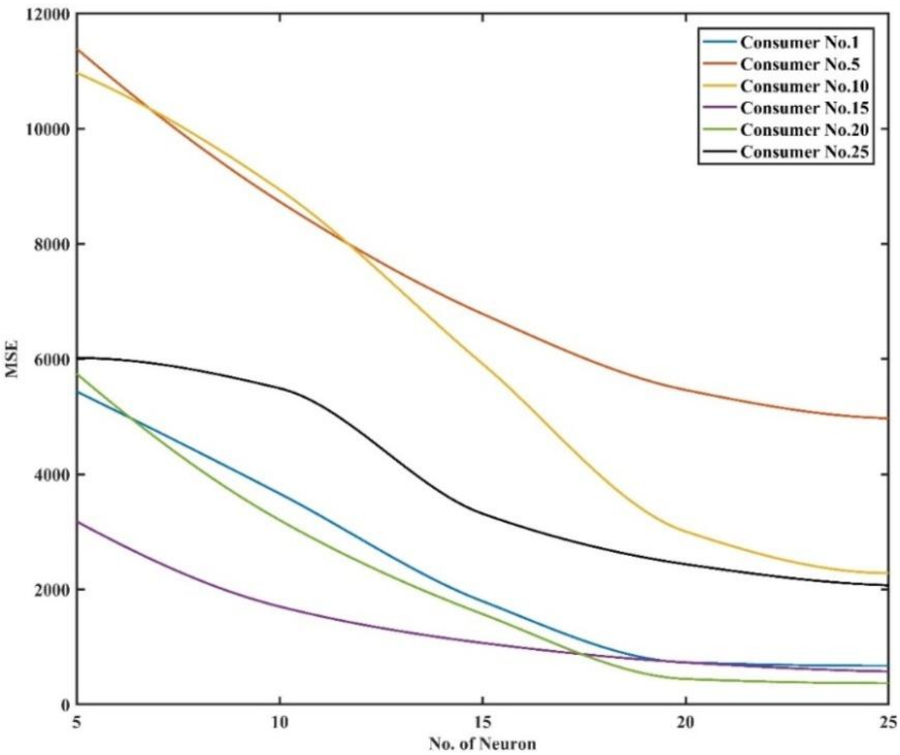


Figure 5.17: MSE Vs Number of Neurons

MAPE:

$$f(x) = ae^{bx} + ce^{dx}$$

Coefficient with 95% confidence bounds

a= 44.63

b= -0.2036

c=14.12

-0.01329

$$\text{MAPE (Number of Neurons)} = 44.63e^{-0.2036\ (number\ \ of\ neueons\ )} + 14.12e^{-0.01329\ (number\ \ of\ neurons\ )} \quad (5.1)$$

MSE:

$$f(x) = ae^{bx}$$

Coefficient with 95% confidence bounds

a= 12370

b= -0.1409

$$\text{MSE (Number of Neurons)} = 12370e^{-0.1409\ (number\ \ of\ neueons\ )} \quad (5.2)$$

From figure 5.16 and 5.17 using the curve fitting technique a mathematical model have been established from the data to independently measure the errors of the model.

Table 5.4: Comparison of the models for consumer 5

Forecasting Techniques	MSE	MAPE (%)
FNN (20)	5463.9	10.59
FNN (10)	8735	15.9883
SVM	15492	133.47
ARIMA	68198	108.15

Equation 5.1 can measure the MAPE of the forecasted energy for a given number of neurons, and the value of the coefficients are established with 95% confidence bound.

Table 5.5: Comparison of the models for consumer 2		
Forecasting Techniques	**MSE**	**MAPE (%)**
FNN(20)	4693.8	29.02
FNN(10)	1285	39.39
SVM	11017	141.06
ARIMA	13988	149.07

And equation 5.2 gives the value of MSE, which is an exponential equation. These two equations are not dependent to measure the errors. So these unique equations can be used to measure the errors of the neural network model.

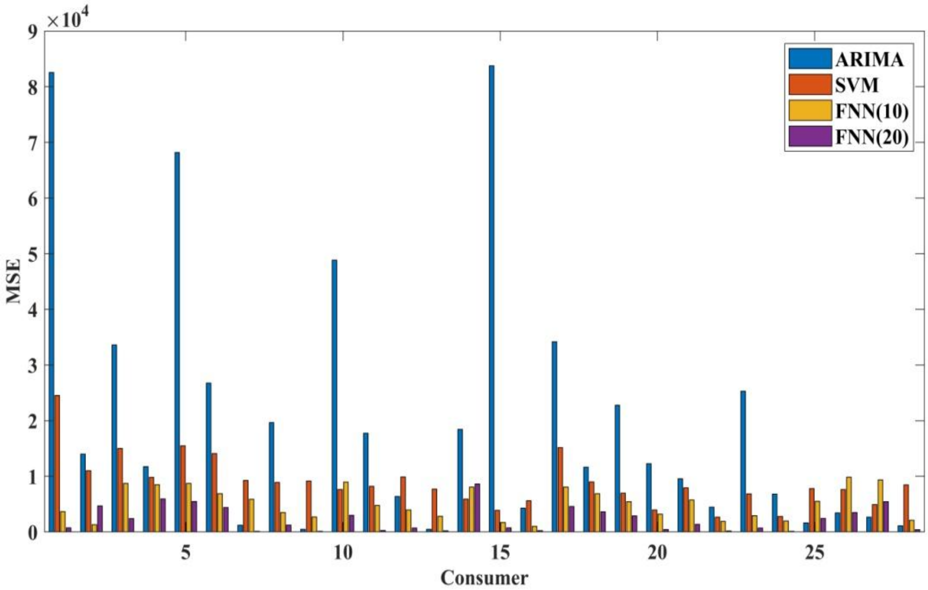


Figure 5.18: MSE comparison of consumers

For a single consumer, the forecasting ability of three popular STF methods, FNN, SVM and ARIMA have been compared to predict the consumption of electricity in table 5.4 and 5.5. It is evident from the result of figure 5.17 and 5.18 that FNN can predict the pattern of the electricity consumption profile with better accuracy than both the ARIMA and the SVM model.

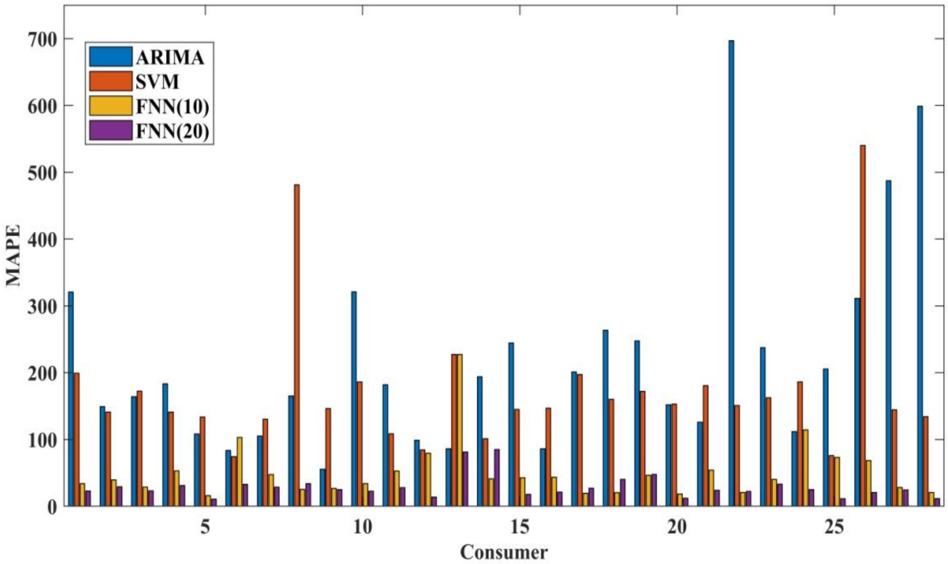


Figure 5.19: MAPE comparison of consumers

After performing the experiments, a manual process of calculating the Mean Squared Error (MSE) for consumer 8 for FNN method has been done. Here the MSE is being calculated from the graph drawn in figure 5.3. The results both calculated manually and from MATLAB Package shown in table 5.6. Here the result is being calculated 15 minutes interval. So that the error calculated is almost correct, and very close to our simulation error calculation. If we calculate our 96 data point they we can have the accurate result like MATLAB simulink output. So after performing this operation, we can say that our process of error calculation is the optimized model of estimation.

Table 5.6: Result comparison of Consumer 8 both by manual Vs MATLAB						
Observation, n	**Actual value, y_t**	**Forecasted Value, f_t**	**Error, $e_t = y_t - f_t$**	$(e_t)^2$	**MSE = $\frac{1}{n}\sum_{t=1}^{n}(e_t)^2$ From Graph**	**Using MATLAB**
1.	58	35	23	529	$\frac{39462}{25} = 1578.48$	3492
2.	58	40	18	324		
3.	46	33	13	169		
4.	68	42	26	676		
5.	48	32	16	256		
6.	52	28	24	576		
7.	52	78	-26	676		
8.	42	78	-36	1296		
9.	45	65	-20	400		
10.	44	75	-31	961		
11.	57	37	20	400		
12.	45	18	27	729		
13.	44	60	-16	256		
14.	82	110	-28	784		
15.	130	270	-140	19600		
16.	38	78	-40	1600		
17.	28	48	-20	400		
18.	52	87	-35	1225		
19.	58	98	-40	1600		
20.	97	73	24	576		
21.	78	108	-30	900		
22.	98	125	-27	729		
23.	98	138	-40	1600		
24.	27	67	-40	1600		
25.	38	78	-40	1600		

CHAPTER 6

CONCLUSION

The main aim of this research is to identify and justify an optimized forecasting model to forecast the electricity consumption of BTS. The forecasted data can be used for designing the scheduling and designing the independent power source for the BTS. This will help the consumer to use its own power to meet its requirement. Thus it can become grid free consumer.

To develop this model Feed-forward Artificial Neural Network being used along with auto regressive integrated moving average and support vector machine. Among these three FNN is the much more justified efficient model for forecasting as referred in previous chapters. Three major error measuring parameters are described and used to evaluate the result. The main focus of this research work was on optimizing the forecasting model and comparing the results on the basis of the major error measurement parameters and establishing a reliable and most accurate FNN forecasting model.

6.1 Research Questions

- What will be a suitable approach for scheduling the storage devices at base station level to make it energy neutral or grid independent?

  To evaluate the consuming data and thus scheduling the storage device, the suitable approach will be to forecast the day-ahead energy necessity with a suitable modeling. And based on that, renewable source i.e. Solar Panel System is the specified way for scheduling the storage device. Thus it can be made energy neutral or grid independent.

- Model development for forecasting the energy consumption of a base station and evaluate the result with real consumption data.

Modeling a feasible forecasting model to predict the energy consuming model and thus scheduling the storage device would be an option. ANN is more feasible and efficient in this purpose. To build the model and forecast the data MATLAB will be used as simulation reference.

For a single consumer, the forecasting ability of three popular STF methods, FNN, SVM and ARIMA have been compared to predict the consumption of electricity. It is evident from the result that FNN can predict the pattern of the electricity consumption profile with better accuracy than both the ARIMA and the SVM model.

However, FNN sometimes fails to predict the consumption mainly because of the abnormal changes of the consumer behaviors. But these short comings can potentially be avoided with higher number of training data. For the case of ARIMA, it does not provide the satisfactory results, although its performance can also be improved with more training data.

The errors tends to reduce with increasing number of neurons, the more the neuron the less the corresponding errors. But after the point corresponding twenty number of neurons the error reduction in compared to the processing time is insignificant. As twenty neuron gives a more stable error, using the curve fitting technique a mathematical model has been established from the data to independently measure the errors of the model.

In this work, we have predicted the consumption of several BTS; however this findings can also be utilized for predicting residential and industrial consumption. Besides, the better accuracy between prediction and result leads to increased monetary benefits for the consumer. Predicting the consumption for higher number of consumers with diverse energy usage pattern may be the possible drive for future research.

6.2 Future Works

Since the forecasting methods consider the historical loads, it is important that the historical data is precise, unfortunately the data recorded may have some missing values, error in the value or there will be some abnormal values due to outages, load shedding [28], [29], [30], sudden rise or dip in the demand. These data's are termed as the Outliers. So to make the forecasting error minimum the outlier detection technique will be used for a better accuracy.

REFERENCES

1. “Global Energy Demand to Increase by Thirty Percent.” [Online]. Available: https://www.maritime-executive.com/article/global-energy-demand-to-increase-by-thirty-percent#gs.NSleuMQ. [Accessed: 12-Mar-2018].
2. “Final energy consumption by sector and fuel — European Environment Agency.” [Online]. Available: https://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-9/assessment. [Accessed: 12-Mar-2018].
3. International Energy Agency(IEA), “World energy trends,” vol. 90, no. 2, pp. 23–26, 2015.
4. “About Electricity | Natural Resources Canada.” [Online]. Available: https://www.nrcan.gc.ca/energy/electricity-infrastructure/about-electricity/7359. [Accessed: 12-Mar-2018].
5. Jianguang Deng and Panida Jirutiti Jaroen, "Short-Term Load Forecasting Using Time Series Analysis: A Case Study for Singapore” in *IEEE conference on Cybernetics and Intelligent systems*, 2010.
6. Suci Dwijayanti and Martin Hagan "Short Term Load Forecasting Using A Neural Network Based Time Series Approach" in *First International Conference on Artificial Intelligence, Modelling & Simulation*. pp. 1217, 978-1-4799-3251-1113 2013 IEEE, DOl 10.1 109/AIMS.2013.11, 2013.
7. Tanel Kivipold and Juhan Val Tin "Regression Analysis of Time Series for Forecasting the Electricity Consumption of Small Consumers in Case of an Hourly Pricing System" in *Advances in Automatic Control, Modelling & Simulation*, ISBN: 978-1-61804-189-0, pp.127-132.
8. "Central Statistics Office National Statistical Organization Ministry of Statistics and Programme Implementation, Government of India www.mospi.gov.in.
9. Amjady, N., "Short-term hourly load forecasting using time-series modeling with peak load estimation capability" in *IEEE Transactions on Power Systems*, vo1.l6, no.4, pp.798-805, Nov 200l.

10. S. N. Dodamani, V. J. Shetty and R. B. Magadum “Short Term Load Forecast Based on Time Series Analysis: A Case Study” in *IEEE International Conference on Technological Advancements in Power & Energy*, 2015.
11. Qian Zhang, Kin Keung Lai and Dongxiao Niu “Optimization Combination Forecast method of SVM and WNN for Power Load Forecasting” in *Fourth International Joint Conference on Computational Sciences and Optimization*, 2011.
12. Zhang Jinhui and Dang Jiajia “Application of SVM Based on Rough Sets to Short-term Load Forecasting” in *Third International Symposium on Intelligent Information Technology Application*, 2009.
13. R. Sadownik and E. Barbosa, "Short-Term Forecasting of Industrial Electrical Consumption in Brazil" in *Journal of Forecasting*, vol. 18, no. 3, pp. 215-224, 1999.
14. S. Sargunaraj, D. Gupta and S. Devi, "Short-Term Load Forecasting for Demand Side Management" in *IEEE Proceedings-Generation, Transmission,* 1997.
15. Z. Yu, "A Temperature Match Based Optimization Method for Daily Load Prediction Considering DLC Effect" in *IEEE Transactions on Power Systems*, 1996.
16. Y. C. Li, T. J. Fang and E. K. Yu, "Study Of Support Vector Machines For Short-Term Load Forecasting" in *Proceedings of the CSEE*, vol. 26, no. 2, pp. 10-13, 2003.
17. F. Grimaccia, M. Mussetta and R. Zich, "Advanced Predictive Models towards PV Energy Integration In Smart Grid" in *IEEE International Conference on Fuzzy Systems*, pp. 1-6, 2012.
18. D. Bunn, "Forecasting Loads and Prices in Competitive Power Markets" in *Proceedings of the IEEE*, vol. 88, no. 2, pp. 163-169, 2000.
19. K. Hipel and A. McLeod, "Time Series Modelling of Water Resources and Environment System" in *Elsevier*, Amsterdam, 1994.
20. G. E. P. Box and G. Jenkins, "Time Series Analysis, Forecasting and Control" in San Francisco, CA: Holden-Day, 1970.
21. D. Park, M. El Sharkawi, I. Marks, L. Atlas and M. Damborg, "Electric Load Forecasting Using An Artificial Neural Network" in *IEEE Transaction on Power System*, vol. 6(2), pp. 442-449, 1991.

22. H. NIE, G. LIU, X. LIU and Y. WANG, "Hybrid of ARIMA and SVMs for Short-Term Load Forecasting" in *International Conference on Future Energy, Environment, and Materials*, vol. 16, pp. 1455-1460, 2012.
23. G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model" in *Neuro computing, Elsevier*, vol. 50, pp. 159-175, 2003.
24. E. E. El-Attar, J. Y. Goulermas and Q. H. Wu, "Forecasting electric daily peak load based on local prediction" in *IEEE Power & Energy Society General Meeting*, Calgary, AB, 2009, pp. 1-6.
25. D. Fay and J. V. Ringwood, "On the Influence of Weather Forecast Errors in Short-Term Load Forecasting Models" in *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1751-1758, Aug. 2010.
26. Engr. Badar Ul Islam, "Comparison of Conventional and Modern Load Forecasting Techniques based on Artificial Intelligence and Expert Systems" in *International Journal of Computer Science Issues (IJCSI)*, Vol. 8, Issue 5, No 3, September 2011.
27. Y. Wang, Q. Xia and C. Kang, "Secondary Forecasting Based on Deviation Analysis for Short-Term Load Forecasting" in *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 500-507, May 2011.
28. V. Margaret and K. U. Rao, "Partial load shedding using ant colony algorithm in smart grid environment" in *TENCON 2015 - 2015 IEEE Region 10 Conference*, Macao, 2015, pp. 1-5.
29. V. Margaret and K. Uma Rao, "Demand response for residential loads using artificial bee colony algorithm to minimize energy cost" in *TENCON 2015 - 2015 IEEE Region 10 Conference*, Macao, 2015.
30. Margaret V., Uma Rao K., and Ganeshprasad G.G., "Intelligent Load Shedding Using Ant Colony Algorithm in Smart Grid Environment" in *Artificial Intelligence and Evolutionary Algorithms in Engineering Systems (ICAEES 2014)*, Noorul Islam University, Kumaracoil, 22nd April 2014.
31. M. Abdullah Al-Amin and Md. Ashraful Hoque "Comparison of ARIMA and SVM for Short-term Load Forecasting" in *The 9th Annual Information Technology, Electromechanical and Microelectronics Conference*, March (13-15), 2019.

32. K. M. U. Ahmed, M. Abdullah Al-Amin and M. T. Rahman "Application of Short Term Energy Consumption Forecasting for Household Energy Management System" in *3rd International Conference on Green Energy and Technology (ICGET)*, September, 2015.
33. Kyriakides, E. and Polycarpou, M. s.l. "Short term electric load forecasting: A tutorial" Trends in Neural Computation, Studies in Computational Intelligence, 2007" in *Trends in Neural Computation, Studies in Computational Intelligence, Springer*, Vol. 35, pp. 391-418. (Chapter 16).
34. Sun, Wei, et al. and Guangzhou, "Application of Neural network Model Combining Information Entropy And Antcolonyclustering Theory for Short-Term Load Forecasting" in *Fourth International Conference on Machine Learning and Cybernetics*. pp. 18-21, 2005.
35. Liebowitz and Jay. "If you are a dog lover, buit expert system; if you are a cat lover, build neural networks" in *Expert Systems with Applications*, p. 30, 21, 2001.
36. Dryar and Henry A. "The effect of weather on the system load" in *Transactions of the American Institute of Electrical Engineers, IEEE Journals & Magazines*, Vol. 63, pp. 1006-1013, 12, 1944.
37. Hirschhausen, Christian von and Andres, Michael "Long-term electricity demand in China — From quantitative to qualitative growth" in *Energy Policy, Elsevier*, Vol. 28, pp. 231–241, 4, April 2000.
38. Choynowski and Peter. "Measuring Willingness to Pay for Electricity" in *Economics and Research Department, Asian Development Bank*, July 2002, ERD Technical Notes.
39. Kyriakides, E. and Polycarpou, M. "Short term electric load forecasting: A tutorial" in *Trends in Neural Computation, Studies in Computational Intelligence*, Springer, Vol. 35, pp. 391-418. (Chapter 16), 2007.
40. Mitchell and Tom M. Machine Learning, McGraw Hill, 1997. p. 02.
41. Matthewman, P. D. and Nicholson, H., "Techniques for load prediction in electricity supply industry" in *Proceedings of the IEEE*, Vol. 115, pp. 1451-1457, 1968.
42. Gross, G. and Galiana, F. D., "Short term load forecasting" in *Proceedings of the IEEE*, Vol. 75, pp. 1558-1573, 12, 1987.

43. Moghram, I. and Rahman, S., “Analysis and evaluation of five short term load forecasting techiniques” in *IEEE Transactions on Power*, Vol. 4, pp. 1484-1491, 1989.
44. Nazeeruddin, Hesham K and Mohammad, Alfares, “Electric load forecasting: Literature survey and classification of methods” in *International Journal of Systems Science*, Vol. 33, pp. 23-34, 2010.
45. Hsu, Y. Y and Ho, K. L., “Fuzzy expert systems : an application to short-term load Forecasting” in *IEEE Proceedings*, Vol. 139, pp. 471-477, 7992.
46. Liang, R. H and Hsu, Y. Y., “Fuzzy linear programming : an application to hydroelectric generation scheduling” in *IEEE Proceedings : Generation*, Vol. 141, pp. 568-574, 1994.
47. Bataineh, Sameer, Al-Anbuky, Adnan and Al-Aqtash, and Salem, “Expert system for unit commitment and power demand prediction using fuzzy logic and neural networks” in *Expert Systems*, Vol. 13, pp. 29-40, 1, 1996.
48. Srinivasan, D, Chang, C. S and Liew, A. C., “Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting” in *IEEE Transactions on Power System*, Vol. 8, pp. 343-348, 1992.
49. Bao, Zhejing, Pi, Daoying and Sun, Youxian, “Short-Term Load Forecasting Based on Self-organizing Map and Support Vector Machine Advances in Natural Computation” in *Springer Berlin Heidelberg*, 2005, pp. 688-691.
50. Li, Y., Fang, T. and Zheng, G., “Wavelet support vector machines for short-term load forecasting” in *Journal of University of Science and Technology of China*, Vol. 33, pp. 726-731, 6, 2003.
51. Rahman, S. and Shreshta, G., “A priority vector based technique for load forecasting” in *IEEE Transactions on Power Systems*, Vol. 6, pp. 1459-1465, 1991.
52. Rahman, S. and Hazim, O., “Load forecasting for multiple sites: development of an expert system-based technique” in *Electric Power System Research*, Vol. 39, pp. 161-169, 1996.
53. Dash, P. K., Liew, A. C. and Rahman, S., “Fuzzy neural network and fuzzy expert system for load forecasting” in *IEEE Proceedings: Generation*, Vol. 143, pp. 106-114, 1996.

54. Kim, K.H. and et al, “Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems” in *IEEE Transactions on Power Systems*, Vol. 10, pp. 1534-1539, 1995.
55. E. Ceperic, V. Ceperic and A. Baric, "A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines" in *IEEE Transactions on Power Systems*, Nov. 2013, vol. 28, no. 4, pp. 43564364.
56. Y. He, Y. Zhu and D. Duan, "Research on Hybrid ARIMA and Support Vector Machine Model in Short Term Load Forecasting" in *Sixth International Conference on Intelligent Systems Design and Applications*, pp. 804-809, Jinan, 2006.
57. E. E. Elattar, J. Goulermas and Q. H. Wu, "Electric Load Forecasting Based on Locally Weighted Support Vector Regression" in *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 4, pp. 438-447, July 2010.
58. Y. He, Y. Zhu and D. Duan, "Research on Hybrid ARIMA and Support Vector Machine Model in Short Term Load Forecasting" in *Sixth International Conference on Intelligent Systems Design and Applications*, pp. 804-809, Jinan, 2006.
59. Karthika S, Vijaya Margaret and Dr. K. Balaraman “Hybrid Short Term Load Forecasting using ARIMA-SVM” in *International Conference on Innovations in Power and Advanced Computing Technologies [i-PACT2017]*.
60. Peng, T, Hubele, N and G., Karady, “Advancement in the application of neural networks for short term load forecasting” in *IEEE Transaction Power System*, Vol. 7(1), pp. 250-258, 1992.
61. Ho, K, Hsu, Y and Yuan, C., “Short Term Load Forecasting Using A Multilayer Neural Network With An Adaptive Learning Algorithom” in *IEEE Transation on Power System*, Vol. 7(1), pp. 259-266, 1992.
62. Lu, C, Wu, H and Vemuri, S., “Neural Network Based Short Term Load Forecasting” in *IEEE Transaction on Power System*, Vol. 8(1), pp. 337-342, 2003.

63. Vonk, B M J., et al., et al. London : s.n., “Improving Short-term Load Forecasting for a Local Energy StorageSystem” in *47th International Universities Power Engineering Conference (UPEC)*. pp. 1-6, 2012.

64. G. E. P. Box, G. M. Jenkins and G. C. Reinsel, "Time series analysis: Forecasting and control” in *3rd ed. Englewood Cliffs*, NJ, Prentice Hall, 1994.

65. Verzijlbergh, R.A and et al., “Deriving electric vehicle charge profiles from driving statistics” in *IEEE Power and Energy Society General Meeting*, pp. 1-6, San Diego, CA, 2011.

66. Marion, B. and et al., “Performance parameters for grid-connected PV systems” in *Photovoltaic Specialists Conference. Conference Record of the Thirty-first IEEE*, pp. 1601 – 1606, 2005.

67. Hyndman, Rob J., “Another Look At Forecast-Accuracy Metrics For Intermittent Demand” in *International Journal of Forecasting*, 2006, Vol. 4, pp. 43-46.

68. Rob J., Hyndman and Anne B., Koehler, “Another look at measures of forecast accuracy” in *International Journal of Forecasting*, Vol. 4, pp. 679-688, 22, 2006.

69. Hamzacebi, C., “Improving artificial neural networks' performance in seasonal time”, in *Information Sciences, Elsevier*, Vol. 178, pp. 4550-4559, 2008.

70. Zhang, G Peter, “A Neural Network Ensemble Method With Jittered Training Data For Time Series Forecasting” in *Information Sciences, Elsevier*, Vol. 177, pp. 5329-5346, 23, 2007.

71. Adhikari, Ratnadip and Agrawal, R. K., “Time Series Forecasting Using Artificial Neural Networks. An Introductory Study on Time Series Modeling and Forecasting” in *LAP LAMBERT Academic Publishing*, 2013, pp. 25-30.

72. Kihoro, J.M., Otieno, R.O. and Wafula, C., “Seasonal Time Series Forecasting: A Comparative Study of ARIMA and ANN Models” in *African Journal of Science and Technology (AJST) Science and Engineering Series*, Vol. 5, pp. 41-49, 2, 2004.