Maninder Singh al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 1, Issue 2, Sept 2014, pp. 59-64

Design & Implementation of Energy Forecasting Using Regression Analysis & Curve Fitting

Maninder Singh[1], Dr. Kamal Sharma[2], Deepak Kumar[3]

[1]Student, M. Tech, ESEAR, Ambala

[2]Professor, Dept. of ECE, E-Max group of Institutions, Ambala

[3]Assistant Professor, Dept. of ECE, E-Max group of Institutions, Ambala

***Abstract*— Governments, global organizations, corporations and scientists publish every year 10-15 authoritative reports on energy scenarios and forecasts. Analysis of world consumption forecasts created by substantial volume of research, even when limited to an separate energy resource during next era, reveals a range of discrepancies in excess of 100 percent or over 3 billion tone of oil equivalent. These scenarios are used as decision support tools for major policies and investment projects. It is obvious that is not possible to forecast future. According to current predominant views, the field of energy will undergo significant structural changes in coming periods, making it radically different from what we know today. Thus, many researchers and organizations envision distinct paths of energy development. Managing electrical energy supply is a composite task. The most important part of electric utility resource planning is forecasting of future energy demand in regional or national service area. This is usually attained by constructing models on relative information, such as climate and previous load demand data. This research investigates forecasting of residential energy consumption by applying structural time series model to monthly data over the period January 2014. In this work, database contains day type information, monthly weather report and heating degree day data. This is used to analyse monthly report data and separately day data with the help of regression analysis and curve fitting method. The work is done with the help of MATLAB tool.**

***Keywords*— Data mining, energy forecasting, regression, curve fitting.**

I. Introduction

Energy is vitally important for modern economies. It enables use of daily appliances (such as computers, medical devices, telecommunication appliances, and transport vehicles) that upsurge people's quality of life. Most appliances used in daily life are powered by energy and it is generally regarded at least in developed world to be almost impossible to live without them. As a result, energy is seen as a necessity for social and economic welfare; it is essential to maintain economic activity in modern industrialized nations and social development. Moreover, one of main reasons for low social and economic progress in developing nations is limited access to modern energy services given appliances that need electricity (such as computers, televisions and radios) provide access to information that quickens social progress of societies [1].

Over centuries, humans have altered their lifestyles along with technological development and innovation. The exceptional economic growth and major improvements in standards of living over last two decades have mainly come about because of replacement of manpower with mechanical power throughtechnical progress. Energy consumption and technology have developed through history and modern societies' lifestyles became more energy dependent. These energy dependent lifestyles make energy crucial for life; societies want uninterrupted light, hot water, warm houses, to travel freely and to power industries. Humans have become accustomed to benefits that are provided by energy consuming appliances and arguably, itis impossible today to think about life without these appliances. [2].

The above highlight advantages of energy dependent lifestyles of modern societies but this also emphasizes the importance for need for modern societies to tackle energy security. However, this key energy policy objective is now coupled with need to tackle the problem of climate change. Since the beginning of industrial revolution, consumption of fossil fuels has substantially increased Green House Gas (GHG) emissions into atmosphere, which is generally regarded as cause of climate change. However, as discussed above, energy is important for social and economic progress and simply just reducing energy consumption in order to help solve climate problem is not an option since modern societies' given lifestyles are heavily dependent on energy. Moreover, it is commonly expected that this dependency will increase in foreseeable future. Furthermore, there are a significant number of studies that illustrate strong negative relationship between energy prices and macroeconomic performance, which is the main concern related to energy security. In order to sustain economic and social progress societies arguably need to secure access to energy resources at a reasonable price [3].

Given energy is generally accepted as being an important driver of economic growth, countries that focus on sustainable economic growth try to find ways to secure their future energy needs at a reasonable price.The rapid increase in demand from emerging economies, competition between nations to access energy resources, along with environmental problems, arouses another concern: whether or not there will be enough energy supply to meet future demand at reasonable cost. Arguably, this can be solved by long-term planning by developing scenarios for future evolution of energy demand and possibilities of meeting that demand in different ways. This can be achieved by a proper understanding of current and past energy demand and possible changes in terms of efficiency and structure, possible supply alternatives, possible technological change, etc.In order to develop successful policies to tackle issues of energy security and climate change it is important that energy demand is analysed and examined carefully [4].

The paper is organized as follows. In section II, we discuss related work with recognition of fingerprint images. In Section III, It describes basic motion detection system, its standards and block matching technique. In Section IV, it describes proposed technique of motion detection. Finally, conclusion is explained in Section V.

II. RELATED WORK

Authors proposed that handling electrical energy supply is a complex task. The most important part of electric utility resource planning is forecasting of the future load demand in the regional or national facility area. This is usually attained by constructing models on relative information, such as climate and previous load demand data. In this paper, a genetic programming method is proposed to forecast long term electrical power consumption in the area covered by a utility situated in southeast of Turkey. The empirical results demonstrate successful load forecast with a low error rate [5].

Some proposed that the information flow in pharmaceutical industry was relatively simple and application of technology was incomplete. However, as we progress into a more integrated world where technology has become an integral part of business processes, process of transfer of information has become more complicated. Today increasingly technology is being used to help pharmaceutical firms manage their inventories and to develop new product and services. The implications are such that by a simple process of merging drug usage and cost of medicines (after completing the legal requirements) with the patient care records of doctors and hospitals helping firms to conduct nationwide trials for its new drugs. Other possible uses of information technology in field of pharmaceuticals include pricing (two-tier pricing strategy) and exchange of information between vertically integrated drug companies for mutual benefit [6].

Some authors proposed a new approach for short-term load forecasting (STLF). Curve fitting prediction and time series models are used for hourly loads forecasting of week days. The curve fitting prediction (CFP) technique combined with genetic algorithms (GAs) is used for obtaining the optimum parameters of Gaussian model to obtain a minimum error between actual and forecasted load. A new technique for selecting training vectors is introduced. The proposed model is simple, fast, and accurate. It is shown that proposed approach provide very accurate hourly load forecast. Also it is shown that proposed method can provide more accurate results than conventional techniques [7].

Some proposed identification of model in functional form using curve fitting and genetic programming technique which can forecast present and future load requirement. Approximating an unknown function with sample data is an important practical problem. In order to forecast an unknown function using a finite set of sample data, a function is constructed to fit sample data points. This process is called curve fitting. There are several methods of curve fitting. Interpolation is a special case of curve fitting where an exact fit of existing data points is expected. Once a model is generated, acceptability of model must be tested. There are several measures to test goodness of a model [8].

III. DATA MINING

Data mining can be defined as “the process of discovering interesting knowledge from large amounts of data stored either in databases, data warehouses, and other information repositories” [10]. The process can be described in three steps—describe the data, build a predictive model, and verify the model. Data are described by its statistical attributes such as means and standard deviations, by visually reviewing charts and graphs, and looking for relationships among variables. Predictive models are built based on patterns found in known results.

Although there are number of approaches to modelling energy demand, the proposed modelling approach is thought to have a significant advantage in terms of identifying price responsiveness of energy demand and forecasting. Therefore, in this paper, a particular modelling approach is utilized to undertake energy demand modelling for a number of different sectors, energy types.

Different and unique data sets are used inthis work. They include the Weather report data, energy produced at the grid, day type and its temperature of complete month.Before any of these techniques were usedon the data sets, preliminary analyses were done on the data sets to gain at least a superficial knowledge of the data sets or to see the nature of the data sets. In real-life analysis, this preliminary data description is an important step in data-mining because in most cases, it helps the analyst in making a choice of the techniques to be used in the regression work [9].

The data mining process can be broken down into four distinct phases.

1. Decision, whether or not to carry out data mining on a given data set.
2. Data preparation, readying the data for analysis.
3. Model building, where the work of building a prediction model is carried out.
4. Interpretation, which is largely carried out by individuals but which can be greatly assisted using automated means, such as graphical representations of the results.

The first stage in the data mining process is that of deciding whether or not to go ahead with a given analysis. This is one of the most difficult and probably the most crucial of all the stages, as it is here that we decide whether or not we are going to spend our time and other resources investigating a given data set.

Once we have decided to go ahead with our investigation, it is vital that the data be in a format that can be easily interpreted by the model building tool. Data preparation is a vast topic. The scope of this study requires that we understand the importance of data preparation to the process as a whole and the need for such facilities in the tools which we investigate [11].

Examples of data preparation include:

- Search for outliers,
- Discretisation of continuous data
- Normalisation.

Model building is the core of the data mining process. This is where verifiable results are obtained. The scope of this study is limited to supervised learner models. What this essentially means is that the models we createwill have been trained using examples of known cases (from the data set) and then verified using further information from the data which has not yet been presented to the model. This stage is known as the training and testing or validation stage and once completed the model produced can be used to predict future outcomes, instances which have neither been seen by the model nor by individuals.The final stage of the data mining process is that of interpretation. This stage is vital to the process as, "it is the analysis of results provided by the human element that ultimately dictates the success or failure of a data mining work" [12].

IV. DESIGN AND IMPLEMENTATION

A time series is a set or sequence of observed data arranged in chronological order and in an equally spaced time intervals such as daily or hourly air temperature. Time series occurs in many fields and the analysis of time series has got a wide application in areas like process control, economic forecasting, marketing, population studies, biomedical science and many more areas. Time series analysis uses systematic approaches to extract information and understand the characteristics of a physical system that creates the time series. There are a number of different approaches to deal with time series analysis including dynamic model building and performing correlations [13].

A. Regression Analysis

Analysing a time series may arise from different objectives of the analyst. One may be interested in process or quality control; for this scenario a time series which measures the quality of a manufacturing process can be generated. It can also be used to obtain descriptive or statistical measures of a time series. In another scenario, if in case observations are taken on two or more variables, it may be possible touse the variation in one variable to explain the variation on the other; this can help to understand the nature of the relationship between the two. Finally, one may be interested to predict future values based on an observed time series [14].

Regression is a statistical technique to determine the linear relationship between two or more variables. Regression is primarily used for prediction and causal inference. In its simplest (bivariate) form, regression shows the relationship between one independent variable (X) and a dependent variable (Y). Regression thus shows us how variation in one variable co-occurs with variation in another. What regression cannot show is causation; causation is only demonstrated analytically, through substantive theory. For example, a regression with shoe size as an independent variable and foot size as a dependent variable would show a very high regression coefficient and highly significant parameter estimates, but it should not conclude that higher shoe size causes higher foot size. The simple (or bivariate) LRM model is designed to study the relationship between a pair of variables that appear in a data set. The multiple LRM is designed to study the relationship between one variable and several of other variables [15].

Many problems in engineering and science involve exploring the relationships between two or more variables. Regression analysis is a statistical technique that is very useful for these types of problems. Regression has got wide applications including prediction and process control. In regression analysis, the aim is to model the dependent variable in the regression equation as a function of the independent variables, constants and an error term. The performance of the model depends on the estimate of the constants and coefficients. A simple linear regression considers a single regressor or predictor x and a dependent or response variable Y. Assuming the relationship between Y and x is a straight line and that the observation Y at each level of x is a random variable.

B. Proposed Algorithm

The proposed steps for energy forecasting are:

1. Estimate the energy database which contains daily temperature, days, weather forecasting report and energy produced.
2. Forecast the energy with the help of presence of monthly database.
3. Use regression analysis and curve fitting method for forecasting data.
4. Estimate the power at any hour.
5. Estimate daily and monthly profiles w.r.t load.
6. Evaluate all profiles individually.
7. Evaluate the performance parameters like mean, S.D etc.

V. RESULTS

This showcases visualization and analysis (heavy statistics) for forecasting energy usage based on historical data. It has access to hour-by-hour utility usage for the month of January, including information on the day of the week and the Heating Degree Days (defined as 65 minus Average Temperature) of each day. Using this information, we will come up with an algorithm for forecasting future energy usage based on parameters such as day-type, forecasted temperature, and time of day [10].

A.Implementation

"MATLAB has excellent built-in support for many data analysis and visualization routines," in particular, one of its most useful facilities is that of efficient exploratory data analysis, which is a natural, fit in the context of data mining.MATLAB's portability comes from the fact that all MATLAB users will have the same range of basic functions at their disposal.

The representation which MATLAB implements, is dealing with all data in the form of matrices. Other advantages of MATLAB include its interactive interface, debugging facilities, object oriented nature and inparticular its high quality graphics and visualisation. MATLAB's add on feature, in the form of toolboxes, makes it possible to extend the existing capabilities of the language with ease.

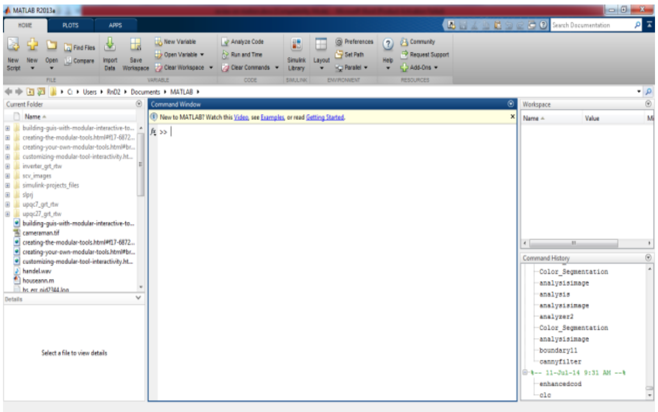

Figure 1: MATLAB Tool

B. Forecasting Results

1. Load data from worksheet (Refer to fig2)

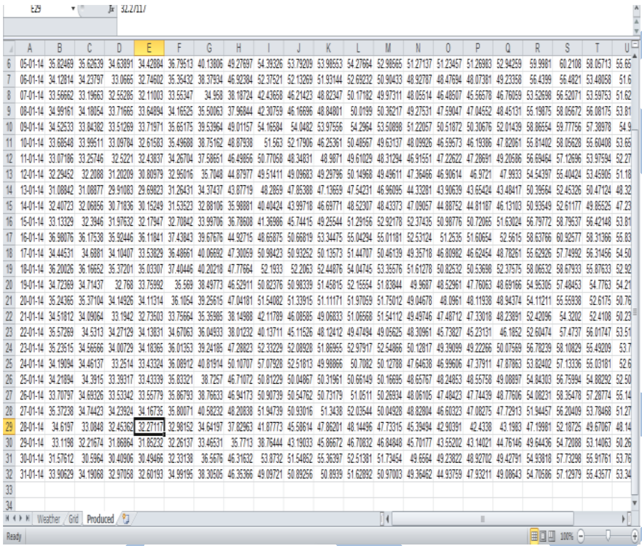

Figure 2: Excel Data

2. Input energy Data w.r.t System Load(shown in fig3)

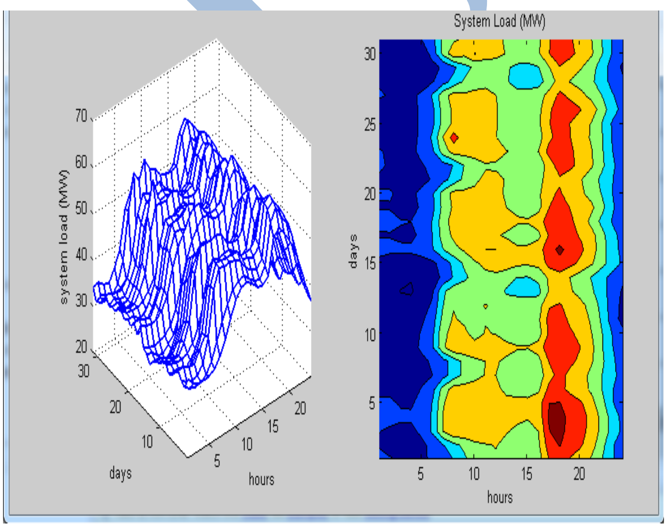

Figure 3: Input Data in 3-D Plot

3. Estimate Power at a particular hour of a Day(shown in fig4)

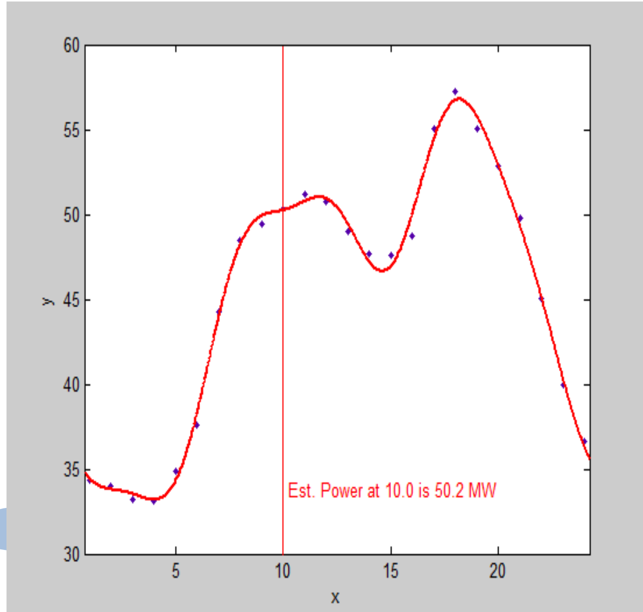

Figure 4: Estimate Power at a Particular Hour

4. Average Month Profile(shown in fig5)

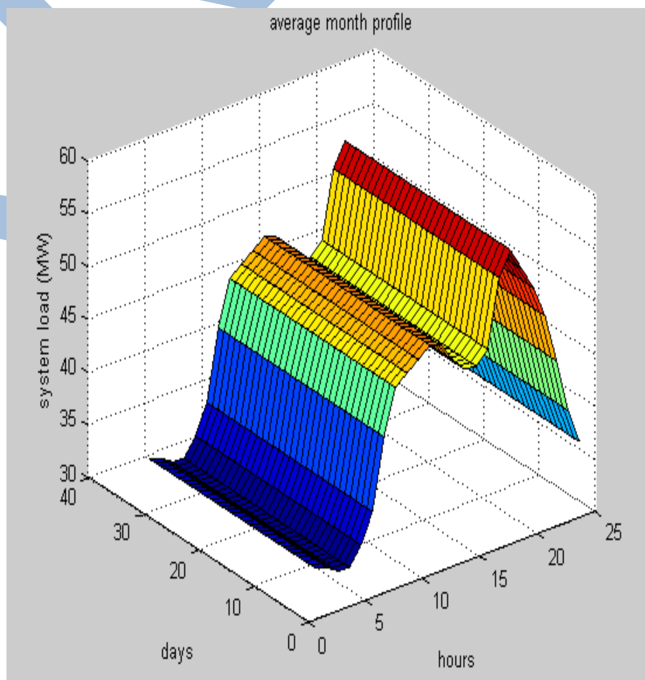

Figure 5: Average Month Profile

5. In order to do some statistics, it's going to determine the distribution type of this data set.From the distribution analysis, it will conclude that the data set has a normal distribution a shown in fig 6.

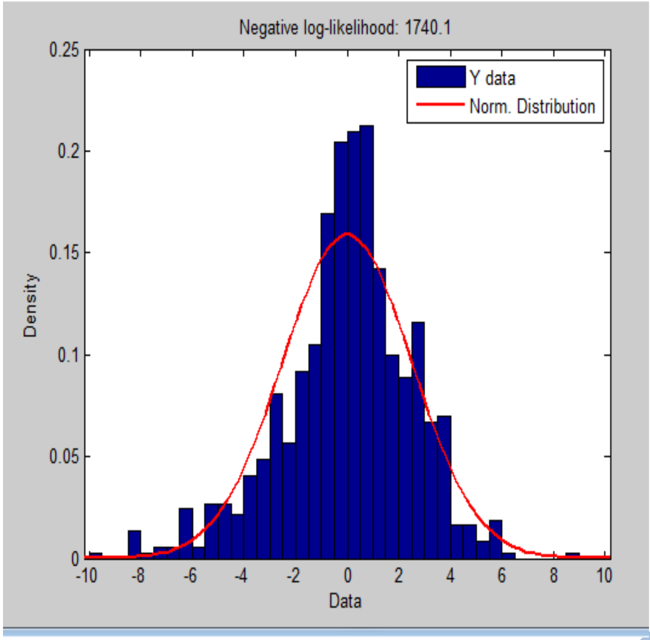

Figure 6: Normally distributed Type (with Log-Likelihood Value)

6. Daily and Monthly Profile w.r.t System Load(shown in fig7)

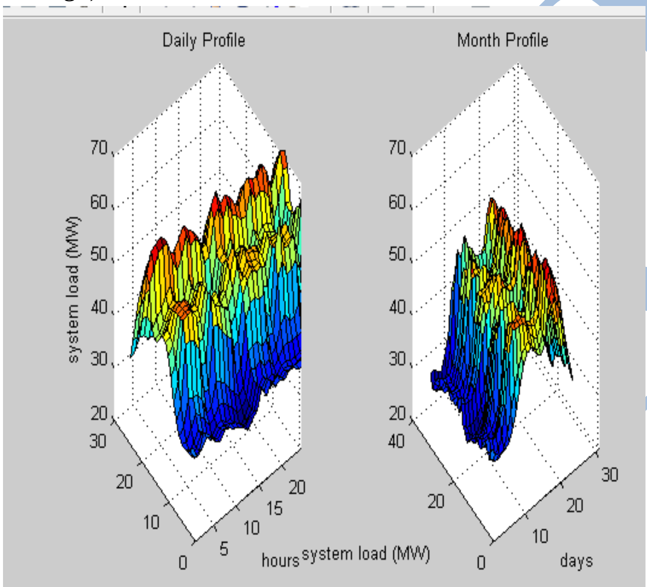

Figure 7: Daily and Monthly Profile

7. Mean, S.D and Confidence Interval of Daily and Monthly Profile.

In fig 8, the daily profile seems to have a very tight confidence interval, suggesting that the general trend throughout the day is similar from day to day.

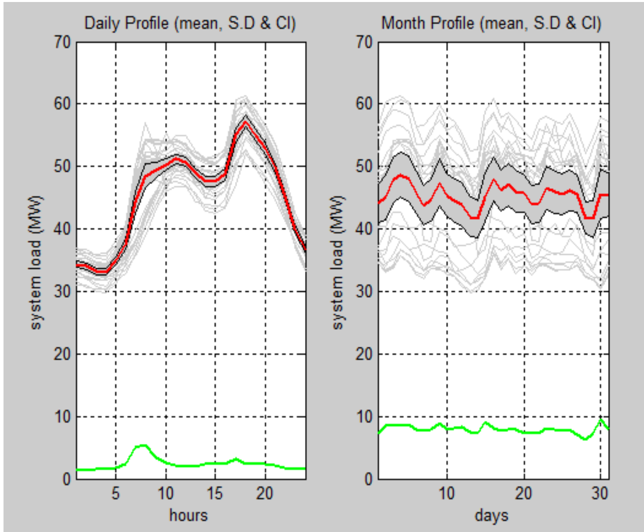

Figure 8: Performance Parameters Profile

8. Separate Each day for energy Usage(shown in fig9)

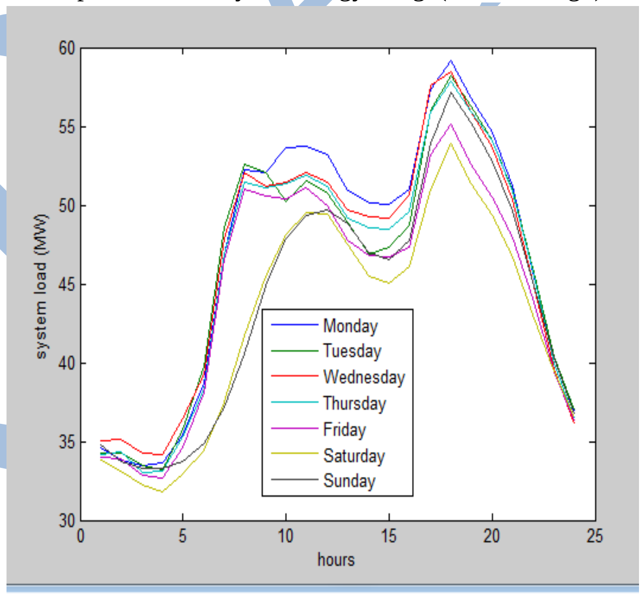

Figure 9: Each Day Profile

9. Estimate Energy at a Particular Day

In fig 10, we can see that the morning energy spike is not prominent on the weekends. Also, Mondays tend to have more usage throughout the day, and Saturdays have the lowest usage.

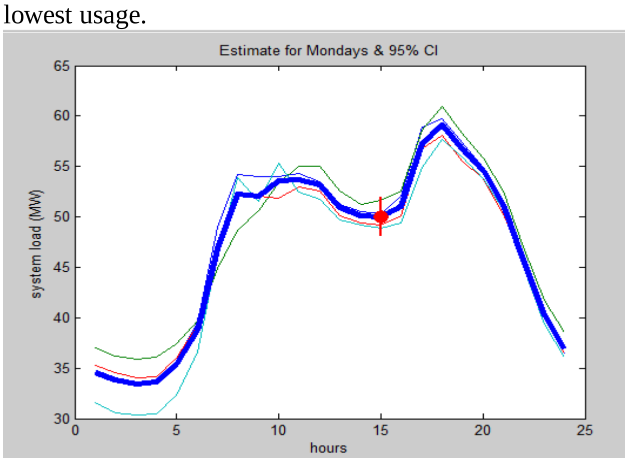

Figure 10: Energy at a Particular Day

VI. CONCLUSION

Long term energy consumption forecasting can provide important information for power distribution centres. In this thesis, it will investigate the energy forecasting using time series methods. The methods include regression analysis and curve fitting. In this, single day profile and average month profile will be evaluated. It also calculates power of any particular hour observed. It can say that the this approach can be used for electric utility resource planning and forecasting of the future load demand in the regional or national service area effectively. Daily profile seems to have a very tight confidence interval Mondays tend to have more usage throughout the day, and Saturdays have the lowest usage. The performance parameters are also evaluated.

REFFERENCES

[1] KorhanKarabulut, Ahmet Alkan, Ahmet S. Yilmaz, "Long Term Energy Consumption Forecasting Using Genetic Programming", Mathematical And Computational Applications, Vol. 13, No. 2, Pp. 71-80, 2008.

[2] JayanthiRanjan, *"Applications Of Data Mining Techniques In Pharmaceutical Industry*" Journal Of Theoretical And Applied Information Technology 2005 – 2007.

[3] M. A. Farahat, M. Talaat, "Short-Term Load Forecasting Using Curve Fitting Prediction Optimized By Genetic Algorithms",International Journal Of Energy Engineering 2012, 2(2): 23-28.

[4] R.Behera, B.B.Pati, B.P.Panigrahi, S. Misra, "An Application Of Genetic Programming For Power System Planning And Operation". Aceee Int. J. On Control System And Instrumentation, Vol. 03, No. 02, March 2012.

[5] Sanjeev Kumar Aggarwal, Lalit Mohan Saini, Ashwani Kumar, "Electricity Price Forecasting In Deregulated Markets: A Review And Evaluation" Electrical Power And Energy Systems 31 (2009) 13–22.

[6] Al-Hamadi, H.M., Soliman S.A.: Long-Term/Mid-Term Electric Load Forecasting Based On Short-Term Correlation And Annual Growth. Electric Power Systems Research 74 (2005) 353–361.

[7] Sisworahardjo, N.S., El-Keib, A.A., Choi, J., Valenzuela, J., Brooks, R., El-Agtal, I.: A Stochastic Load Model For An Electricity Market. Electric Power Systems Research 76 (2006) 500–508.

[8] CuaresmaJc, Hlouskova J, Kossmeier S, Obersteiner M. Forecasting Electricity Spot-Prices Using Linear Univariate Time-Series Models. Appl Energy 2004; 77:87–106.

[9] Zhou M, Yan Z, Ni Y, Li G. An Arima Approach To Forecasting Electricity Price With Accuracy Improvement By Predicted Errors. In: Proceedings OfIeee Power Engineering Society General Meeting, Vol. 1; 6–10 June 2004. P. 233–8.

[10] Zhou M, Yan Z, Ni Yx, Li G, Nie Y. Electricity Price Forecasting With Confidence interval Estimation Through An Extended Arima Approach. IeeProc Generation Trans Distribution 2006;153(2):187–95.

[11] ConejoAj, Plazas Ma, Espinola R, Molina Ab. Day-Ahead Electricity Price Forecasting Using The Wavelet Transform And Arima Models. Ieee Trans Power Syst 2005;20(2):1035–42.

[12] HaitengXh, Niimura T. Short-Term Electricity Price ModelingAnd Forecasting Using Wavelets And Multivariate Time Series. In: Ieee Power Systems Conference And Exposition Pes, No. 1; 10–13 October 2004. P. 208–12.

[13] Weron R, Misiorek A. Forecasting Spot Electricity Prices With Time Series Models. In: Proceedings Of International Conference On The European Electricity Market Eem. Lodz, Poland; 10–12 May 2005.

[14] Garcia Rc, Contreras J, Akkeren M, Van Garcia Jbc. A Garch Forecasting Model To Predict Day-Ahead Electricity Prices. Ieee Trans Power Syst 2005;20(2):867–74.

[15] Wu W, Zhou Jian-Zhong, Yu J, Zhu Cheng-Jun, Yang Jun-Jie. Prediction Of Spot Market Prices Of Electricity Using Chaotic Time Series. In: Proceedings Of The 3rd International Conference On Machine Learning And Cybernetics. Shanghai; 26– 29 August 2004. P. 888–93.

