Analysis of Household Electricity Consumption Using Smart Meter Data

Antara Mahanta Barua[1], Pradyut Kumar Goswami[2]

[1]Assistant Professor, Department of Electronics, K. K. Handiqui State Open University, Guwahati, Assam, India
[2]Ex-Vice Chancellor, Assam Science & Technology University, Guwahati, Assam, India

---

Abstract - *Deployment of smart meters in the household not only provides real time data consumption of electricity but also provides better services to the consumers. Consumers become responsive of their electricity usages as they can monitor the power consumption regularly in the smart metering system. The smart meter has the ability to communicate the power consumption data between the consumers and their suppliers. It also helps to track daily usage of power and to understand the consumption patterns to save excess consumption for the benefit of the consumer. In this paper, we present a thorough analysis of smart meter data of electricity consumption to study the behavior of the residential consumer's power usages as well as to predict their power consumption. It aims to help electricity suppliers and the consumers to realize their electricity consumption patterns. This paper provides day wise analysis for a month of power consumption data of randomly selected consumers which are divided into 6 consumer groups depending on their monthly load data. In the next step, simple linear regression method was applied to predict household electricity consumption.*

***Key Words*: Smart Meter, Power consumption profile, Smart meter data analysis, Simple Linear Regression.**

1. INTRODUCTION

Smart meters have been deployed around the world to measure energy consumption of a consumer and to provide additional information like the behavior and consumption profile of the customers to the electricity supplier and the consumer itself by using bidirectional communication process. Smart meters can read real time electricity consumption data including the values of voltage, current, phase angle and frequency and communicates that data [1]. These data aids the electricity suppliers to fulfill the demand of electricity and to manage the distribution of electricity.

The smart meter installation is implemented in the pilot projects basis in various places of our country, India. The pilot project aims in promotion of energy efficiency in residential, commercial and industrial consumers. In the North Eastern state Assam, it was implemented by the APDCL electricity supplier in 2018 by installing 11,523 nos. of smart meters divided in three sub-divisions namely Paltanbazar, Ulubari and Narengi of Guwahati city. Table 1 provides an overview of smart meter installation in the pilot project.

In this paper, we focused on the analysis of household smart meter data collected from 500 residential consumers which were randomly selected from three sub-divisions of the pilot project area. The Pilot Project involved installation of Smart Meters for the purpose of reduction in distribution losses, reliability improvement and power efficiency [2]. The Functionalities covered under this project are Advanced Metering Infrastructure (AMI), Power Quality Management (PQM), Outage Management System, Peak Load Managements (PLM), and Decentralize Generation (DG) [3].

Table - 1: Overview of Smart Meter installation in the Pilot Project

Sub-division	*Paltanbazar*	*Ulubari*	*Narengi*
Load Connected KW	5267	8526	1960
No. of Installed Smart Meter	5106	6350	67
Total No. of DCU	90	124	2
No. of meters per DCU	60 - 100	60 -100	50-120
Average distance from meter to DCU (Metre)	150	150	150

To implement the smart meter pilot project, the following infrastructure was developed (Fig.1) in which smart meter data are collected by Data Concentrator Unit (DCU) through a wireless RF mesh network. RF mesh systems are mainly used for remote reading, advanced metering and for some other applications, such as demand response or load management, that do not have strong requirements in terms of bandwidth and delay [4]. The communication channel used in the deployed mesh network is the unlicensed ISM band of 865-867 MHz. The advantages of this technology are low maintenance, low power consumption, wide range coverage, minimal infrastructure and high flexibility [5]. However, the individual connections are expensive and low power radio signals are liable to interference and link blockage that can affect the success and performance of the network [6].

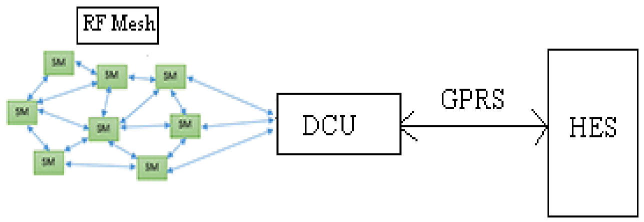


Fig-1: Smart Metering Infrastructure Implemented by APDCL Pilot Project (Source [2])

Meter data are communicated from DCU to the Head End System (HES) through General Packet Radio Service (GPRS) and Global System for Mobile Communications (GSM) technology. There are 216 nos. of Data Concentrator Units connected to collect the meter data.

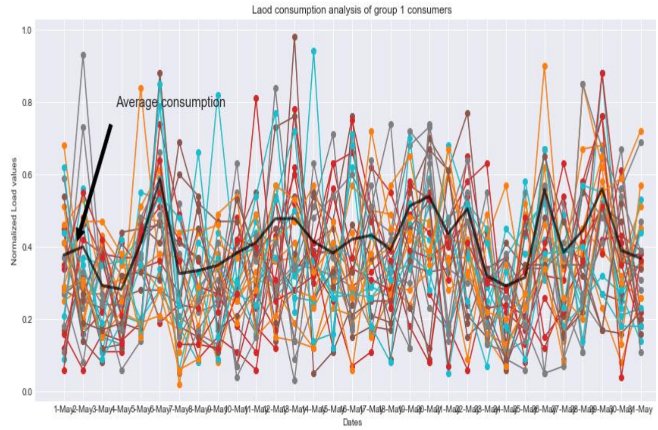
2. ALYSIS OF DAILY POWER CONSUMPTION DATA

During this research, we had collected load consumption data randomly from 500 smart meters for the period from 1st to 31st May, 2019 and segmented into 6 consumer groups by total monthly electrical power consumption (Table 2). For the considered month, the amount of collected data was quite significant: 500 consumers ×31 days = 15,500 data points.

For the evaluation of measured data statistical analysis methods were applied. Initially daily electricity consumption data (Fig.2) was analyzed using python program and select the average consumption for each of the consumer group.

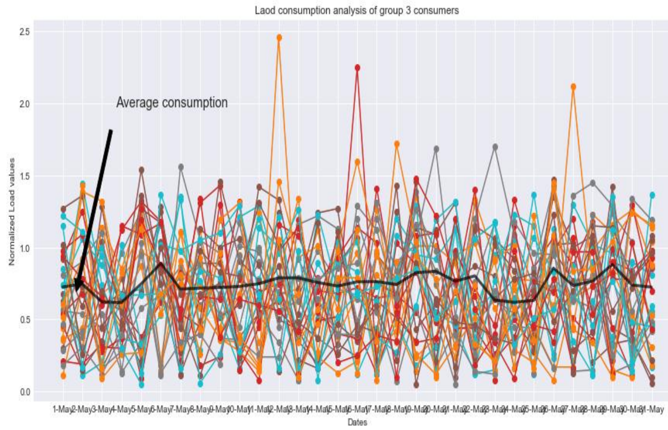
Table -2: Groups of Consumers involved in the pilot project

Groups	Electricity Consumption, kWh	No. of Consumers
Group 1	0 – 15	81
Group 2	16 – 20	57
Group 3	21 – 25	80
Group 4	26 – 30	72
Group 5	31 – 35	49
Group 6	Above 35	161
Total	500	

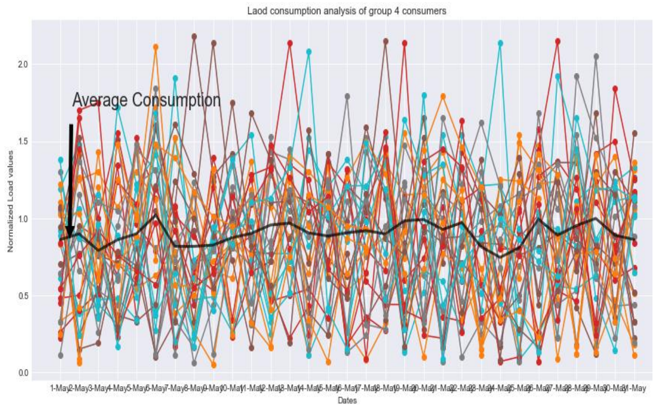


Group 1 (0-15 kWh)

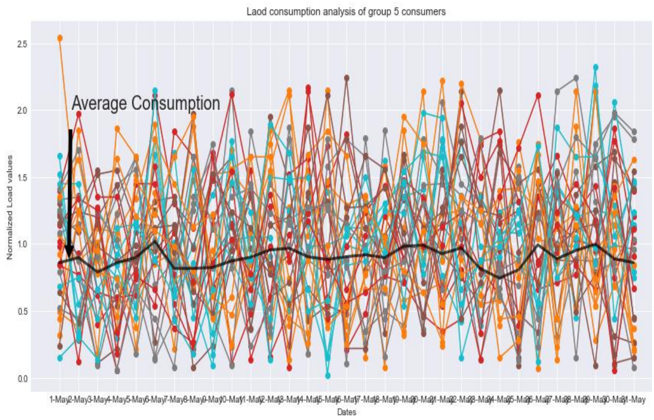
Group 2 (16-20 kWh)



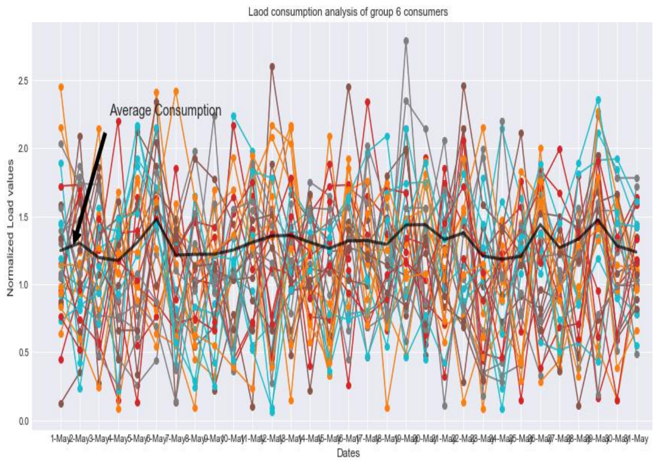
Group 3 (21-25 kWh)



Group 4 (26-30 kWh)



Group 5 (31-35 kWh)



Group 6 (Above 35kWh)

Fig - 2: Analysis of daily electricity consumption for different consumer groups

Evaluation the average energy consumption per day for each group and analyzed the consumption patterns as shown in figure 3.

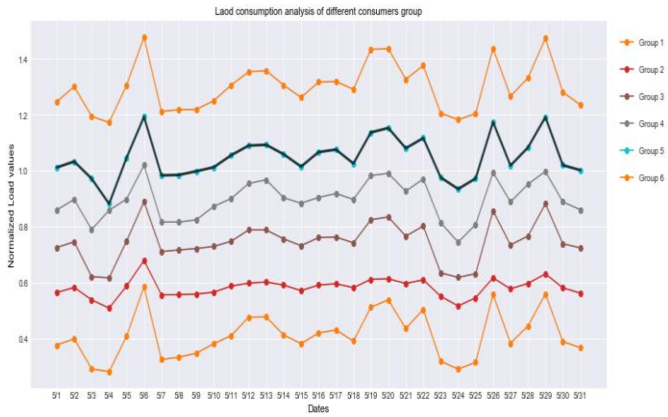


Fig -3: Analysis of average electricity consumption per day for six different consumer groups

3. STATISTICAL ANALYSIS OF CONSUMPTION DATA

The electricity consumption pattern of residential consumers depend on the weather. For describing energy consumption pattern in the pilot project area, the correlation between the energy consumption and outside maximum temperature which can influence the electricity consumption individually was analyzed through a linear regression model. Linear regression analysis is a methodology that allows finding of functional relationship among response or dependent variables and predictor or independent variables. Two types of linear regression techniques namely Simple Linear Regression (SLR) and Multiple Linear Regression (MLR) are used in finding the relationship among variables [7]. We have used SLR technique where total electricity consumption per day is the response variable and maximum temperature is the predictor variable. The dataset was prepared with smart meter consumption data collected from datacenter of APDCL and day wise temperature collected from the national weather database [8]. In this models, the relation between one response variable and one predictor variable is considered using the relation-

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (1)$$

where Y is the response variable, X is the predictor variable, β_0 and β_1 are the regression coefficients or regression parameters, and ε is the associated random error component to account for the discrepancy between predicted data from Eq. (1) and the observed data. The predicted value form of Eq. (1) is

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad (2)$$

where $\hat{Y}$ is the fitted or predicted value and $\hat{\beta}$ are estimates of the regression coefficients [9-10]. The difference between fitted and predicted values is that the fitted value refers to

the case where the values used for the predictor variable correspond to one on the n observations of the observed data used to find $\hat{\beta}$, but the predicted values are obtained for any set of values of the predictor variables different to the observed data [11].
For a set of *n* observed values of *x* and y, the simple linear regression equation represents in matrix form as-

$$\begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ . & . \\ . & . \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} \quad (3)$$

The quality of fit of the proposed model for the given dataset can be judged by using the co-efficient of determination (R^2) and the Root Mean Square Error (RMSE)

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}} \quad (5)$$

Where R is the correlation co-efficient and the term $\sum (y_i - \hat{y}_i)^2$ and $\sum (y_i - \bar{y})^2$ are called sum of squared errors (SSE) and total sum of squares (SST) respectively.
The value of R^2 varies between 0 to 1, when R^2=0.9, indicates that 90% of the total variability in the response variable is accounted by the predictor variable [12].

4. RESULTS AND DISCUSSION

For the given dataset, load consumption of Group 1, Group 2, Group 3, Group 4, Group 5 and Group 6 data represents in y_1, y_2, y_3, y_4, y_5 and y_6 respectively. And day wise maximum temperature data represents in x variable. The model of the form of equation (2), using the regression function on the Excel Data Analysis tool, the model for the daily consumption of six consumer groups is found to be-

Group 1 $y_1 = \beta_0 + \beta_1 x = 0.556+0.031x$
Group 2 $y_2 = \beta_0 + \beta_1 x = 0.187+0.013x$
Group 3 $y_3 = \beta_0 + \beta_1 x = -0.085+0.027x$
Group 4 $y_4 = \beta_0 + \beta_1 x = 0.108+0.026x$
Group 5 $y_5 = \beta_0 + \beta_1 x = 0.187+0.028x$
Group 6 $y_6 = \beta_0 + \beta_1 x = 0.298+0.032x$

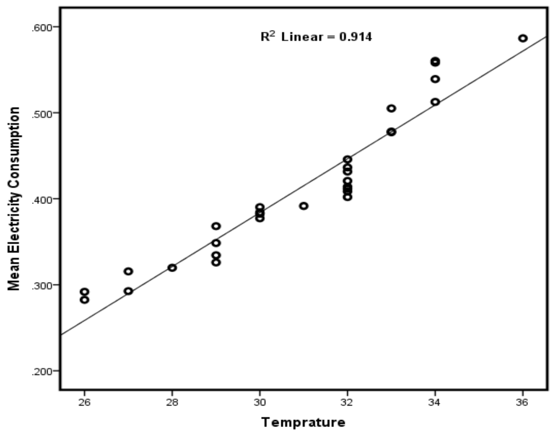
Values are taking up to three decimal places.

To validate the regression model and estimate its accuracy, different statistical tests were applied to evaluate how well the model explains the actual consumption data, an F -test to verify the presence of significant relation among the independent and dependent variables, a t -test for testing the strength of each of the individual coefficient of the model and R^2 [13].

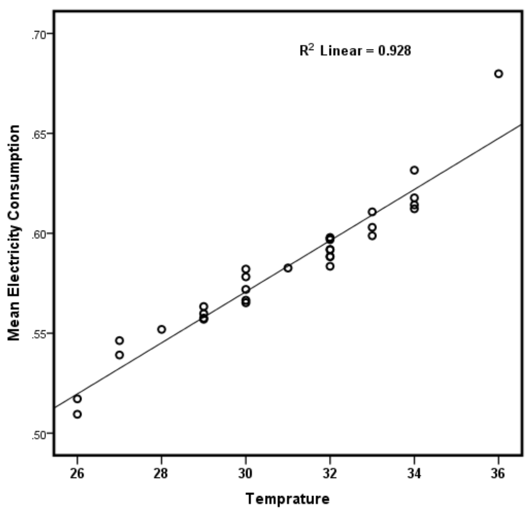
Table 3 represents the summary of regression results for each group. The quality parameters R^2 and RMSE obtained for the statistical model calculated as 0.914 and 0.024 for Group 1, 0.928 and 0.009 for Group 2, 0.921 and 0.019 for Group 3, 0.88 and 0.024 for Group 4, 0.91 and 0.022 for Group 5 and 0.906 and 0.026 for Group 6 consumers respectively. This indicates that predictor 'x' perfectly accounts for variation in 'y'. It means the high relationship between energy consumption and temperature per day. As temperature increases by one degree Celsius on an average a statistically significant increases in the mean electricity consumption by 0.031 kWh. Hence it can be noted that the daily energy usage of the smart meter installed residential consumers can be predicted by outside temperature data within acceptable errors.

Table -3: Regression result for different groups of consumers

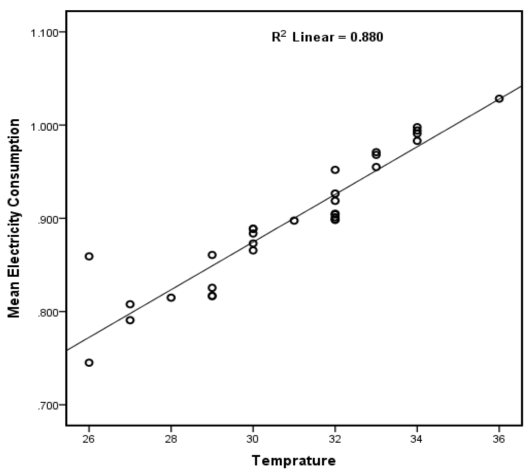
Consumer Groups	Intercept β_0	Regression Coefficient β_1	R^2	RMSE (kWh)
Group 1	0.556	0.031	0.914	0.024
Group 2	0.187	0.013	0.928	0.009
Group 3	- 0.085	0.027	0.921	0.019
Group 4	0.108	0.026	0.88	0.024
Group 5	0.187	0.028	0.91	0.022
Group 6	0.298	0.032	0.906	0.026



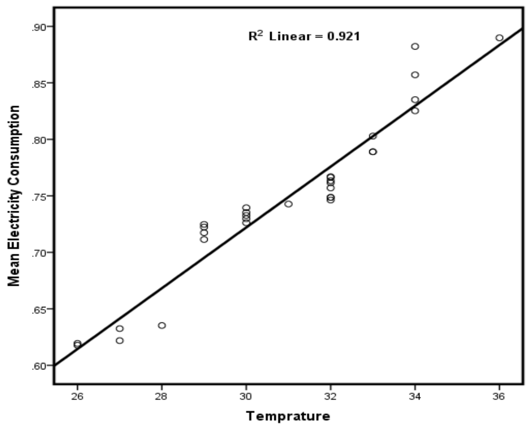
Group 1 Consumers



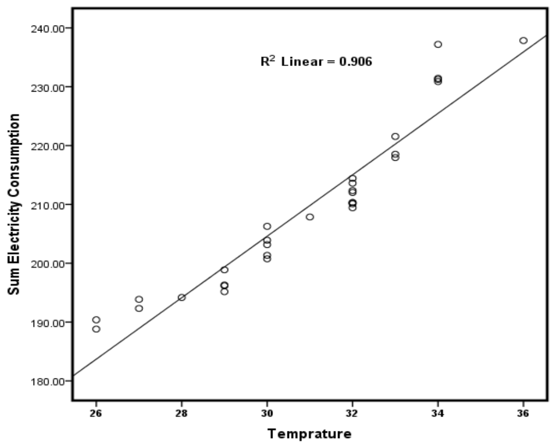
Group 2 Consumers



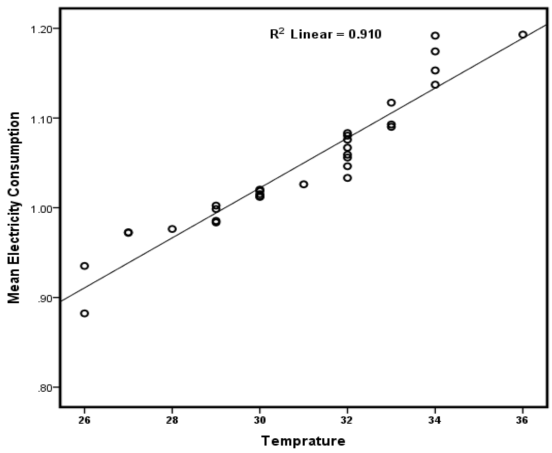
Group 3 Consumers



Group 4 Consumers



Group 5 Consumers



Group 6 Consumers

Fig - 4: Linear regression relation between temperature and mean consumption for six different consumer groups

5. CONCLUSIONS

This paper investigate the feasibility of usage of smart meter data for prediction of electricity consumption of household consumers in the pilot project area. An accurate prediction of electricity consumption allows proper estimation of energy demand in order avoid electrical shortage. The main challenge of this method was finding accurate data in an acceptable period of time [14].

Electrical consumption profiles of consumers varies by consumer groups and weather conditions and this variation which illustrates the stochastic nature of electricity consumption is proved by the result of this analysis. An important feature of the proposed model is that it requires only fundamental data as input and based on the statistical analysis, we show that the weather is a dominant factor that affect the power consumption pattern.

REFERENCES

[1] S. Depuru, W. Lingfeng, V. Devabhaktuni and N. Gudi "Smart meters for power grid challenges, issues, advantages and status" in Proc. Power System Conference and Exposition (PSCE), pp. 1-7, 20-23 March 2011.

[2] A. Barua and P. K. Goswami "Smart Metering Deployment Scenarios in India and Implementation Using RF Mesh Network" in Proc. IEEE International Conference on Smart Grid and Smart Cities (ICSGSC) pp. 243-247 July 2017.

[3] "Next Generation Smart Metering – IP Metering" Smart Grid Bulletin, volume: 2, Issue 9, 2015.

[4] F. Malandra and B. Sanso "Analytical performance analysis of a large scale RF mesh smart meter communication system" in Proc. IEEE Conference, 2015.

[5] G. Iyer, P. Agrawal, E. Monnerie and R. S. Cardozo "Performance Analysis of Wireless Mesh Routing Protocols for Smart Utility Networks" Communication Networks and Smart Grid Forum, 2011

[6] G. Leon, "Smart Planning for Smart Grid AMI Mesh Network", Technology White Paper, 2011, Available at https://www.smartgrid.gov/files/documents/Smart_Planning_for_Smart_Grid_AMI_Mesh_Networks_201109.pdf

[7] W. Sanford, Applied linear regression. 4th ed. Hoboken, New Jersey: Wiley; 2013.

[8] Weather database Available at: https://www.accuweather.com/en/in/guwahati/186893/june-weather/186893

[9] A. Viloria, H. Hernandez-P, O. B. P. Lezama and J. Vargas "Prediction of Electric Consumption Using Multiple Linear Regression Methods" Springer, 2020

[10] Available at: http://home.iitk.ac.in/~shalab/regression/Chapter2-Regression-SimpleLinearRegressionAnalysis.pdf

[11] S. Chatterjee, H. Ali S. Regression analysis by example. 5th ed. Hoboken, New Jersey: Wiley; 2012.

[12] N. Fumo, M.A.R. Biswas "Regression analysis for prediction of residential energy consumption" Elsevier Journal Renewable and Sustainable Energy Reviews 47(2015)332–343

[13] V. Bianco, O. Manca and S. Nardini "Linear Regression Models to Forecast Electricity Consumption in Italy" Energy Sources, Part B: Economics, Planning, and Policy. April, 2013 Available at: http://www.tandfonline.com/loi/uesb20

[14] M. Safaa, J. Allenb, and M. Safac, "Predicting Energy Usage Using Historical Data and Linear Models" The 31st International Symposium on Automation and Robotics in Construction and Mining, 2014