The Nucleus 53, No. 1 (2016) 14-25

www.thenucleuspak.org.pk

The Nucleus
ISSN 0029-5698 (Print)
ISSN 2306-6539 (Online)

Development of a Typical Hourly Electricity Consumption Profile for Student Residence Halls Based on Central Tendency Method

K.P. Amber[1, 2]*, W. Aslam[1] and M.A. Bashir[1]

[1]Mirpur University of Science and Technology, Mirpur, AK, Pakistan

[2]Faculty of Engineering, Science and the Built Environment, London South Bank University, London, SE1 0AA, UK

khuram.parvez@must.edu.pk; m.w.aslam@liverpool.ac.uk; anser.me@must.edu.pk

ARTICLE INFO

Article history :
Received : 06 January, 2016
Revised : 19 January, 2016
Accepted : 05 February, 2016

Keywords :
Hourly electricity profile,
Student residence hall,
Central tendency method,
Median

ABSTRACT

Actual measured hourly electricity consumption profile (HEP) helps the building engineer in numerous ways, e.g. for the optimum and accurate sizing of a solar PV for their building, for identifying abnormal peaks and drops, for negotiating with utility supplier etc. Conventional electricity meters do not have features to provide hourly consumption; instead these provide daily or monthly consumption. In such situation, it becomes difficult to make a precise estimate of building's hourly consumption and a reliable and quick method is desired to estimate hourly consumption. Using four years of measured hourly electricity consumption data for three residence halls, this paper aims to use the central tendency method to develop a dimensionless typical HEP for this building category. Based on the skewed nature of hourly data distributions, median was selected as suitable measure of central tendency. Therefore, using the hourly median values, a typical HEP was developed. The proposed HEP was tested and compared with the actual HEPs of three other similar buildings and it was found that the estimated HEPs matched very well with actual HEPs with a maximum hourly RMSE of 0.9%. Finally, limitations of the proposed HEP are discussed and some recommendations are made in this regard.

1. Introduction

This research highlights the practical advantages of having actual hourly electricity profile (HEP) and attempts to develop a typical dimensionless HEP of student residence halls of the UK's Higher Education sector by using the Central Tendency Method.

Section 1 presents the background. Data has been discussed in Section 2. Development of HEP is discussed in Section 3. Testing and validation of the proposed profile is discussed in Section 4. Section 5 presents the error analysis whereas Section 6 presents the limitations of the proposed profile. Section 7 presents the conclusion.

1.1 Background

Universities in the UK are required to reduce their carbon emissions by 34% by 2020 compared to the baseline year, i.e. 2005/2006 [1]. This carbon reduction target has accelerated their efforts of making their campuses greener through different initiatives such as energy conservation measures and installation of sustainable technologies such as Combined Heat and Power (CHP), wind turbine or solar PVs [2]. Optimum sizing of such alternative technologies highly depends on the actual, i.e. measured hourly electricity consumption profile (HEP) of the building. In the absence of an actual HEP, the optimum sizing of a technology (e.g. solar PV) becomes difficult and its economics will greatly depend on the accuracy of the assumed HEP based on which it would have been sized otherwise. In such a case, a solar PV, wind turbine or a CHP project may result in an under sized or over sized technology which will ultimately fail to achieve its economics and environmental targets [3].

Student residence halls are common in the Higher Education sector of the United Kingdom and generally are located within a University campus. Space heating is mainly provided by the electrical panel heaters or by the hot water radiators in these types of residence halls. Usually, each room is equipped with a dedicated panel heater or a hot water radiator with a thermostat. These electrical heaters consume massive amount of electricity in order to provide heating effect and therefore, are proving to be a major source of scope 2 emissions in the UK higher education sector.

Before an hourly profile development study could be undertaken, it is important to understand the significance of having a building's actual measured energy consumption profile at different intervals such as hourly, daily, monthly and annually. For this purpose it is essential to have actual measured hourly consumption data for that building. Hourly profile enables the building

*Corresponding author

operators to understand the hourly peaks and drops during different periods of the day. Daily consumption profile helps in understanding the difference between consumption on a working day or non-working day whereas monthly profile helps in understanding the buildings consumption pattern during different months and seasons of the year.

Of these, Hourly Electricity Profile (HEP) of a building carries higher significance due to the following:

i. If actual measured hourly consumption profiles are available, then daily, monthly and annual profiles could be calculated automatically.
ii. HEP helps building operator to identify abnormal peaks or drops during different hours of the weekdays and weekends.
iii. HEP helps to calculate an optimum size of an alternative energy source for the building.
iv. It helps in understanding the electricity consumption ratio during peak and off-peak hours and thus enables the building operator to set energy reduction targets and to negotiate with its electricity supplier for the electricity price.
v. The visualization of HEP improves the awareness among the building occupants as they could see how their building is performing on the hourly basis.

To get an actual HEP, a building should have a SMART electricity meter installed as these meters are capable of providing electricity consumption data on half hourly or hourly interval. In the UK, buildings with 100kW or more than 100kW connected load have to install SMART electricity meter under the building regulations [3]. In the absence of a SMART meter, the second option to have a measured HEP of the building is to physically record hourly meter reading which is obviously not a feasible option.

The conventional electricity meters only provide daily or monthly electricity consumption. For such situation where measured HEP is not available, an estimated HEP could be used provided that it carries considerable accuracy, reliability and has been validated.

In cases where only daily consumption figures are available and HEP has to be estimated, it becomes difficult to disaggregate daily total consumption into hourly consumption. A comprehensive knowledge of the buildings operation and different types of activities taking place during different periods of the day is required in order to model and estimate HEP of the building. Researchers have used different methods, mainly statistical and machine learning methods such as Multiple Regression (MR) and Artificial Neural Networks (ANN) for the development of estimated HEPs for different types of buildings.

Noren and Pyrko [32] developed dimensionless typical hourly load profile for Swedish schools based on hourly temperature values. They found that their proposed typical load shapes match well with the real data of both buildings categories with an average hourly error less than 20%.

They developed estimated hourly load profiles of schools using data from 26 schools and by applying Multiple Regression Method [MRM]. They considered three main independent variables in this study, i.e.

a. No. of meals cooked daily in the kitchen;
b. Relationship between sports centre area and floor area; and
c. Daily mean outdoor temperature.

Very low values of R^2 were found during the night hours whereas R^2 values were found between 0.26 and 0.47 during the day hours.

Other methods for predicting HEP include Neural Networks, Genetic Programming and Support Vector Machine methods [33]. These methods require significant amount of hourly data for different factors and variables affecting the hourly electricity consumption of the building plus a comprehensive training of these methods is required in order to develop models which could be then used to predict an estimated HEP of a building. The reliability of such estimated profiles is highly dependent on the accuracy and amount of the input data considered in the development of models. However, for an Energy Manager of a large organization, these methods may not be attractive methods for the estimation of HEP of their buildings as a significant amount of data, time and training would be required to use these methods.

Busy Energy Manager of a large organization would prefer a simple, quick, reliable and validated method to estimate the HEP of their building [4].

1.2 Central Tendency Method

Majority of the research studies are based on a large amount of data which should be effectively summarized to get some meaningful results. One such way is summarizing a huge dataset into a single value which should be the best representative of the entire dataset. This is called central tendency method in which a whole dataset is represented by a single value, known as measure of central tendency [5]. There are mainly three measures of the central tendency, i.e. mean, median and mode. Each of these measures is described below with their relevant advantages and disadvantages.

1.2.1 Mean

Mean or average is the most common measure of central tendency and could be calculated by dividing the sum of all individual observations by the total number of observation [5]. Its mathematical expression is shown in Eq. (1).

$$\bar{x} = \frac{\sum x}{n} \tag{1}$$

Mean has certain advantages over other measures of central tendency, e.g. it takes into account all the observations. [5]. However, this advantage becomes a disadvantage when there are outliers, particularly when the number of observations is small. [6, 18]. Its second disadvantage is that it does not represent the central tendency for the skewed data distributions [7, 8].

Skewness is a measure of symmetry and informs how much a data distribution is skewed, i.e. how much it varies from a normal distribution. For a normal distribution, its value should be zero. A negative value for skewness indicates that the data distribution is negatively skewed having a longer tail towards the left side of the mean whereas a positive value indicates that the data distribution is skewed towards right side of the mean. Skewness of a data distribution can be calculated using Eq. (2) as follows [9]:

$$Skewness = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3}{(N-1)s^3} \tag{2}$$

where $\bar{x}$ denotes the mean, *s* represents the standard deviation, and N is the sample size.

1.2.2 Median

Median is the value which is in the middle (for odd number of observations) when all the observations are organized in an ascending/descending order. It splits the data distribution into purely two halves which means that 50% of observations in a distribution have values at or below the median [10]. Major advantage of using median as a measure of central tendency is that it is not affected by outliers [11]. Its major disadvantage is that it does not consider the value of each observation and therefore, does not use all information available in the data [12].

1.2.3 Mode

The mode is the most frequently repeated value in a data set and therefore, it is worth calculating the mode and comparing it with mean and median values in order to find the most suitable measure of central tendency of hourly percentage values [12].

The mode of a grouped data may be calculated by using the following formula [13] as presented in Eq. (3).

$$M_0 = L + \frac{(f_i - f_{i-1})}{(f_i - f_{i-1}) + (f_i - f_{f+1})} b_i \tag{3}$$

Where;

L is the lower limit of modal class;

f_i is the frequency of the modal class;

f_{i-1} is the frequency of the class immediately below the modal class;

f_{i+1} is the frequency of the class immediately above the modal class; and

b_i is the width of class having the modal class.

Major advantage of Mode is that it cannot be influenced by the extreme value but on the other hand it has poor stability and non-uniqueness [14].

1.2.4 Selection of an appropriate measure of central tendency

Generally, mean is preferred as the measure of central tendency when the data distribution is normal. However, it may not be the best measure in certain situations, especially when the data distribution is skewed. In such situations median is preferred to mean. Mode is preferred when the data are measured in a nominal scale [15, 16, 17].

Zhao et.al [14] used central tendency method in their study to determine the energy usage intensity (EUI, kWh/m^2/year) of hotel and Government offices based on the dataset from 56 hotels and 128 Government office buildings. Due to the remarkable central tendency in their data distribution, they concluded that mode is the most appropriate measure to represent the central tendency and therefore, selected mode values of EUIs for both hotel and office buildings as their typical EUIs.

Sharp [18] used central tendency method to determine the energy consumption benchmark of office buildings using the energy data of 6,751 office buildings obtained through Commercial Buildings Energy Consumption Survey (CBECS). Median was found as the most suitable measure of central tendency and therefore, its value was selected as the benchmark for office buildings.

1.3 Methods of Summarizing a Large Dataset

1.3.1 Use of box plot

One way of summarizing a large dataset is the use of box plot. It is a common method to observe the nature of data distribution. It further helps in understanding the position of median, range and spread of data with upper and lower limits, called whiskers and helps in identifying whether the distribution is a skewed or not. For skewed distributions, usually many points exceed the whiskers; such data points are called outliers [19]. Other use of box plot is that it helps in comparing two or more sets of data distributions. Fig. 1 shows a box plot for a typical dataset.

Two ends of the box represent the lower and upper quartiles, i.e. Q1 (25th) and Q2 (75th), respectively. Median is represented by the bar inside the box and represents the 50th percentile of the data. The two lines attached to the two ends of the box are called ‘whiskers’. These whiskers represent the lowest and highest values (the range) of the data [20]. For a normal distribution, median is exactly in the middle of the box whereas for a skewed distribution, the majority of data are located on the high or low side of the graph and the median may not be exactly in the middle of the box. A number of researchers from energy sector [e.g. 21, 22, 23] in their research studies have successfully used box plot to describe and summarize their data distributions.

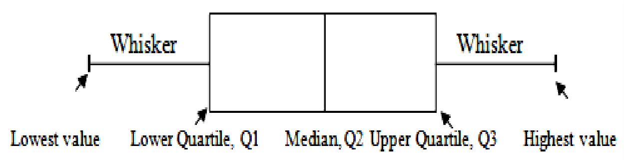

Fig. 1: A box plot of a typical data distribution

1.3.2 Use of frequency distribution

Another method of interpreting a large set of data easily is by use of frequency distribution. Frequency distributions present frequency counts of a parameter. Generally frequency distributions are presented in the form of frequency tables, histograms or bar charts [24]. A careful analysis of the histograms of frequency distribution of the data could reveal the nature of data distribution, i.e. whether it is a normal or skewed distribution and thus helps in the selection of an appropriate measure of central tendency. Ward, I., et.al [25] used the frequency distribution histograms to summarize and explain their collected energy consumption data for the HE sector of the United Kingdom.

From the above literature review, it is evident that central tendency method has been used by researchers in energy sector with major focus on determination of EUIs. No published literature was found which has used central tendency method to develop a typical hourly electricity consumption profile of any type of buildings. Therefore, this research takes an opportunity to fill this knowledge gap by employing this technique to develop a typical HEP for the student residence halls of the UK HE sector, which should be a reliable, quick and easy option for the Energy Managers to estimate the hourly load of their residence halls.

1.4 Multiple Regression Method

Multiple regression analysis is a statistical method for estimating the relationships among variables and has been widely used for the forecasting of energy consumption in buildings. It includes many techniques for modeling and analyzing several variables, when the focus is on the relationship between a dependent variable and one or more independent variables [26]. How good an MR model is, could be generally judged by its Coefficient of Correlation (R) which represents the proportion of the variation in the dependent variable due to variation in the explanatory variables. The value of R is between -1 < R < +1 If x and y have a strong positive linear correlation, R is close to +1. On the other hand, if x and y have a strong negative linear correlation, R is close to -1. If there is no linear correlation or a weak linear correlation, r is close to 0. A value near zero means that there is a random, nonlinear relationship between the two variables. Generally R> 0.5 is described as strong, whereas R< 0.5 is generally described as weak.

“R” can be computed using Eq. (4) :

$$R = \frac{n\sum xy - (\sum x)(\sum y)}{\left(\sqrt{n(\sum x^2)} - \sqrt{(\sum x)^2}\right) \times \left(\sqrt{n(\sum y^2)} - \sqrt{(\sum y)^2}\right)} \qquad (4)$$

where n presents the number of pairs of datasets [27]. In this study, R has been calculated using built in formula “CORREL” in the MS Excel ®.

1.5 Aims and Objectives

Aim of this study is the development of a dimensionless hourly electricity consumption profile (HEP) for student residence halls by using the central tendency method. To achieve this aim, the following objectives were set.

a. Data collection, screening and analysis;
b. Development of HEP based on central tendency method;
c. Testing and validation of proposed HEP through comparing it with the real/measured HEPs of other similar buildings; and finally,
d. To evaluate the reliability of proposed HEP using error analysis, i.e. Root Mean Square Error (RMSE)

2. Methods

2.1 Site Selection

For the purpose of this study, we selected the Mile End campus site of the Queen Mary University of London (QMUL) as this site has 18 student residence halls [28].

Hourly electricity consumption data of three residence halls i.e. Beaumont Court, Creed Court and Maurice Court were obtained from the office of Head of sustainability for the period 1 Jan 2007 to 31 Dec 2010. Other information about these buildings was gathered from the office of Buildings manager through a questionnaire. Fig.2. shows the location of these three electrically heated student residence halls within the student village of QMUL in London. Of these, the Lynden house receives electricity from the same meter as of the Maurice court and therefore, in this study it has been considered as a part of Maurice court. All these buildings have double glazed windows and all were built in 2004. Building occupancy rate is 100% throughout the year. Heating is mainly supplied by the electrical panel heaters. Each room has an electrical panel heater of 750W whereas staircases and kitchens have heaters with a rating of 1500W. Kitchens are fully electric.

It is important to mention here that all three residence halls have occupancy of 100% throughout the year and the collected electricity consumption data reveals that there is hardly a difference in the shape of HEPs on working and non working days. Further the change in the duration in the day is another factor that could affect the HEP. The data analysis shows that the difference in the hourly profiles of summer and winter months is in the range of 1 to 1.5% provided the hourly consumption data is in % instead of kWh.

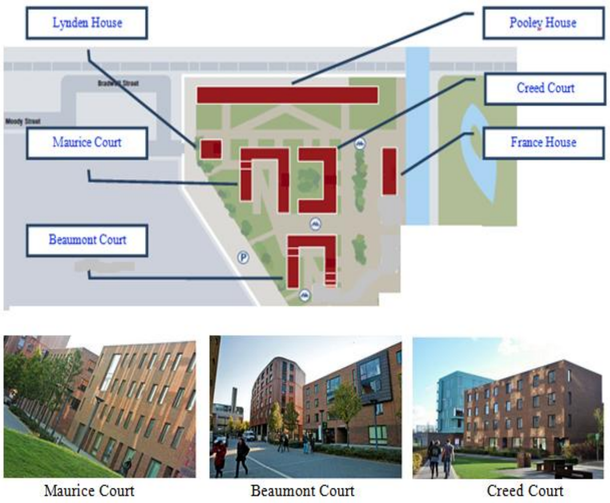

Fig. 2: Electrically heated student residence halls at the mile end campus of QMUL [26]

2.2 *Data Collection*

This section discussed different types of data collected for this research study and different methods used for obtaining the data.

2.2.1 *Building operational data*

Buildings information was obtained from the office of the Buildings Manager for student residences. A questionnaire was sent over to the office of Building Manager who returned the questionnaire with requested information filled in. Table 1 shows the information against each of these three buildings.

2.2.2 *Electricity consumption data*

Hourly electricity consumption data for the period 1 Jan 2007 to 31 Dec 2010 of three electrically heated residence halls, i.e. Creed Court, Pooley House and France House were received from the office of the Head of Sustainability in MS Excel ® files. Hourly ambient temperature data for the London region were downloaded from the website of the Environmental Research Group – Kings College [29].

Table 1: Building information for QMUL student residence halls

Parameter	Beaumont court	Maurice Court	Creed Court
Built Year	2004	2004	2004
GIA, m²	3,887	4,361	2,851
No. of rooms	167	197	124
Plant equipment age, years	3	8	3
No. of lifts	0	0	0
No. of A/C units	0	0	0
No. of kitchens	29	35	22
No. of fridges	196	204	151
Occupancy - Jan to Dec	100%	100%	100%

Fig. 3 shows the variation of hourly electricity consumption of Creed Court against hourly ambient temperature for four years, i.e. 2007, 2008, 2009 and 2010. It is apparent that with increase of ambient temperature, the hourly electricity consumption decreases and vice versa in this type of building category. This fact was further investigated using MR method which showed that there is a strong correlation between these two variables as R values were observed between -0.57 and - 0.65.

It can be further seen that this type of hourly plot helps in understanding the base and peak loads during different periods of a year. e.g. it is clear from Fig. 3 that the Creed Court building maintained a base electrical load of 12kW during the summer when there is no heating demand. During winter, the peak load raised to 75kW, 70kW, 75kW and 85kW in 2007, 2008, 2009 and 2010 respectively.

3. Development of a Dimensionless HEP

The received hourly electricity consumption data were in kWh. Hourly consumption in kWh could be different for different similar type of buildings based on a number of factors such as building area, building age, building's orientation, building location, type of equipment installed etc. Therefore, a dimensionless HEP is desired. This is done by transforming hourly kWh data into hourly % data as described below. First, the daily total electricity kWh were calculated by summing up the measured hourly consumption for each day for the period 1 Jan 2007 to 31 Dec 2010, as shown in Eq. (5).

$$E_d^{kWh} = \sum_{i=1}^{24} E_{hi}^{kWh} \quad (5)$$

where

E_d^{kWh} is the daily total electricity consumption in kWh

E_{hi}^{kWh} is the hourly electricity consumption in kWh

i is the hour and has values 1,2,3……24

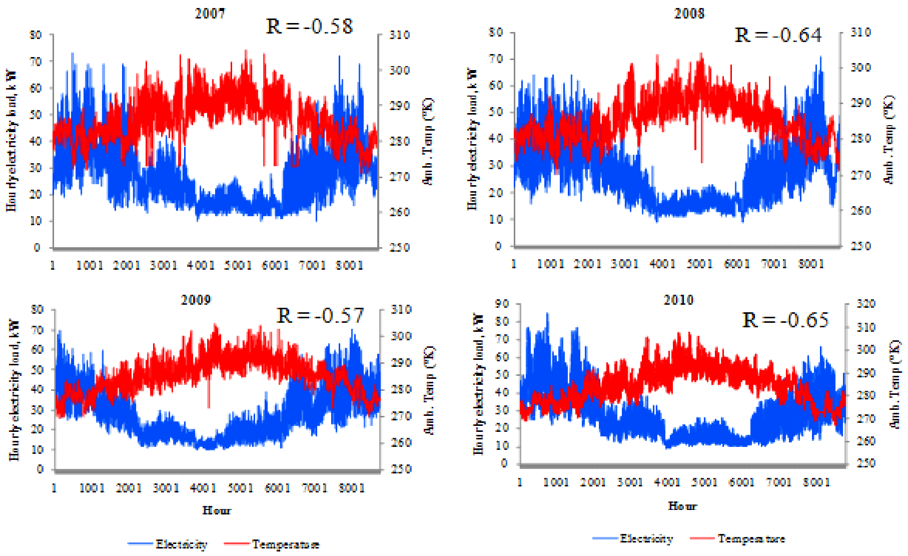

Fig. 3: Relationship between hourly electricity consumption and hourly mean ambient temperature

Now to calculate the percentage for every hour, we used the following formula as shown in Eq. (6).

$$E_{hi}^{\%} = \frac{E_{hi}^{kWh}}{E_{d}^{kWh}} \times 100 \quad (6)$$

$E_{hi}^{\%}$ is the hourly electricity consumption in %

E_{hi}^{kWh} is the hourly electricity consumption in kWh

E_{d}^{kWh} is the daily total electricity consumption

This hourly % was calculated for each hour for the period 1 Jan 2007 to 31 Dec 2010 for Creed Court, Pooley House and France House. Therefore, the total number of observations (N) for each hour of the day is 1,461 x 3 = 4,383.

After obtaining the values for $E_{hi}^{\%}$, the following were calculated for each hour's $E_{hi}^{\%}$;

a. Min
b. Max
c. Range
d. Std. deviation
e. Mean
f. Median
g. Mode
h. Skewness

Table 2, shows the hourly values for each of aforementioned parameters over the period of four years (1 Jan 2007 to 31 Dec 2010).

It could be seen from Table 2 that skewness for all 24 hours is between -1 to +1 range. This means that the hourly data distributions could not be considered as a normal distributions. For such skewed distributions, median is the appropriate measure of central tendency [15, 16, 17] and therefore, hourly values of median could be used as standard hourly % consumption.

The skewed nature of hourly distributions could be further demonstrated with the help of frequency distribution plots, also called histograms. A frequency histogram helps to observe whether the distribution is normal or skewed. It also shows the frequency of occurrence of different hourly % values. Fig. 4 shows the frequency histograms of each quarter of the day, i.e. 12 am to 6 am, 6 am to 12 pm, 12 pm to 6 pm and 6 pm to 12 am whereas frequency values for every hour of a typical day are presented in Table 3.

It can be seen from Fig. 4 that hourly consumption % is 3.5-4% during the first quarter of the day , i.e. night period. It increases to 4-4.5% during second and third quarters as the activities start taking place. During the fourth quarter, i.e. from 6pm to 12am, average hourly consumption increases to 4.5-5% and this is mainly due to high electricity usage in kitchens during the dinner period.

The frequency distribution of hourly consumption (%) is shown in Table 3 as below.

The hourly consumption % ranges with higher frequency are highlighted in red, intermediate in yellow whereas % ranges with lower frequencies are highlighted as green. If carefully observed, this table in fact also shows a reflection of the shape of hourly electricity consumption profile for the student residence halls through the pattern of red (high frequency) and yellow (intermediate frequency) colors. It can be seen that during the dinner hours, i.e. between 6 pm to 9 pm, hourly electricity consumption is highest, i.e. between 6.5% to 7.5% of the daily total. This is probably because of extensive usage of kitchen appliances during these hours. During day time, hourly consumption remains in the range of 3% to 5% with 4- 4.5% values having higher frequencies. During night hours, consumption drops to a range of 2.5 – 3.5%.

Table 2: Descriptive result for hourly electricity consumption (%)

Hour	Min (%)	Max (%)	Mean (%)	Median (%)	Mode (%)	St dev (%)	Skewness	N
1	2.35	6.13	4.20	4.23	3.85	0.47	0.06	4383
2	1.91	5.51	3.89	3.91	3.71	0.48	-0.16	4383
3	1.76	5.51	3.61	3.64	3.58	0.49	-0.18	4383
4	1.61	4.95	3.40	3.41	3.85	0.51	-0.08	4383
5	1.47	4.87	3.24	3.27	3.58	0.49	0.08	4383
6	1.74	5.91	3.23	3.21	3.58	0.55	0.29	4383
7	1.67	5.44	3.33	3.29	3.85	0.60	0.43	4383
8	1.98	5.37	3.55	3.52	3.85	0.58	0.26	4383
9	2.14	5.46	3.81	3.86	3.85	0.52	-0.17	4383
10	2.19	5.66	3.95	3.98	4.17	0.50	-0.18	4383
11	2.17	5.54	4.02	4.03	4.01	0.47	-0.09	4383
12	2.53	6.17	4.14	4.15	4.17	0.46	-0.07	4383
13	2.72	6.82	4.16	4.17	4.01	0.45	0.19	4383
14	2.70	5.64	4.04	4.03	3.85	0.42	0.18	4383
15	2.61	7.04	3.98	3.98	4.01	0.42	0.48	4383
16	2.49	7.33	4.12	4.09	3.85	0.45	0.67	4383
17	3.03	7.62	4.51	4.48	3.85	0.57	0.63	4383
18	3.23	7.92	4.99	4.94	4.73	0.76	0.50	4383
19	3.03	7.80	5.22	5.19	4.77	0.70	0.32	4383
20	3.34	7.87	5.21	5.20	4.55	0.61	0.41	4383
21	3.31	6.98	5.11	5.10	5.89	0.54	0.25	4383
22	3.34	6.70	4.95	4.96	4.89	0.48	0.27	4383
23	3.33	6.47	4.78	4.79	4.77	0.46	0.22	4383
24	3.23	5.95	4.55	4.57	4.77	0.37	0.05	4383

3.1 Proposed HEP

Based on the observed fact that the hourly data distributions are not exactly normal in shape as the hourly skewness values were observed above or below zero (Table 2), median is the suitable measure to represent the central tendency of the dataset [15, 16, 17] and therefore, in our case hourly median % values have been considered as the hourly electricity consumption (%) and have been used in the development of a typical HEP for the student residence halls. This profile is shown in Fig. 5. If the daily total electricity consumption (kWh) is known, hourly electricity consumption could be estimated using this proposed HEP. Fig. 5 also shows the minimum and maximum limits of the proposed HEP.

4. Testing and Validation of Proposed HEP

The proposed HEP for the student residence halls has been tested and validated against the measured HEPs for year 2011 of the following residence halls.

a. Beaumont Court
b. Maurice Court
c. McLaren House

Of these, Beaumont and Maurice courts are located in the Mile End Campus of the Queen Mary University of London and both are electrically heated whereas McLaren House is a part of London South Bank University and is located in its Southwark campus in London. Heating in McLaren house is provided by the hot water radiators. It is important to mention here that this 2011 data of above three halls has not been used in the development of proposed HEP and has been purely used for the testing and validation of the proposed HEP.

Fig. 6 shows the comparison between the estimated HEP and the measured HEPs of Beaumont Court, Maurice court and McLaren House. It is apparent that the actual profiles of Beaumont Court, Maurice court and McLaren House have fitted well.

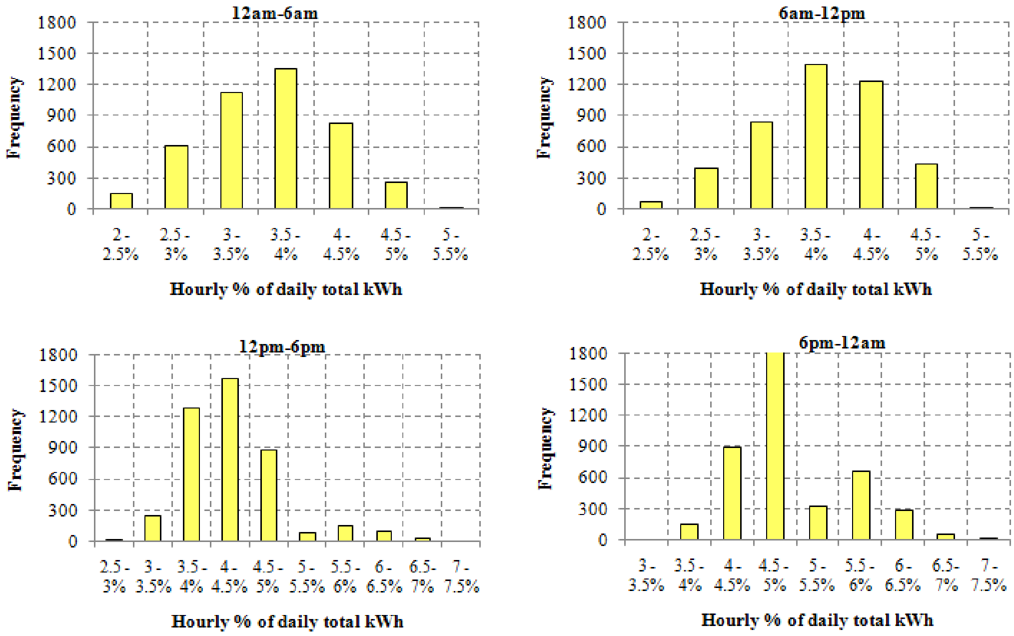

Fig. 4: Histograms of frequency distributions of electricity consumption (%) of four quarters of a day

Table 3. Frequency distribution of different ranges of hourly electricity consumption (%)

Hourly % range	Hour of the day																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
	Frequency of occurrence																							
7.5 - 8%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	7	7	4	0	0	0	0
7 - 7.5%	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	3	0	46	58	32	0	0	0	0
6.5- 7%	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	13	144	143	136	44	20	0	0
6 - 6.5%	6	0	0	0	0	0	0	0	0	0	0	3	3	0	0	9	63	483	658	553	339	141	60	0
5.5 - 6%	12	3	3	0	0	3	0	0	0	3	3	3	12	6	0	12	196	714	1071	1094	954	553	257	36
5 - 5.5%	50	21	3	0	0	3	3	12	21	6	25	22	49	15	21	28	153	266	314	389	479	424	257	108
4.5 - 5%	928	390	130	64	42	55	149	251	333	492	538	852	856	543	443	738	1453	1288	1338	1571	1976	2337	2390	2188
4 - 4.5%	1961	1338	739	411	209	306	450	738	1215	1530	1640	1804	1848	1683	1514	1715	1649	1005	669	525	487	783	1212	1728
3.5 - 4%	1090	1814	1784	1391	1070	1028	949	1192	1666	1554	1643	1347	1321	1723	1915	1595	750	403	113	72	100	121	196	303
3 - 3.5%	314	625	1221	1561	1684	1394	1427	1397	839	640	445	302	274	388	450	267	103	28	11	8	4	4	11	20
2.5 - 3%	19	183	427	774	1070	1237	1147	717	282	142	71	50	15	24	37	12	0	0	0	0	0	0	0	0
2 - 2.5%	3	9	76	181	308	358	259	76	27	15	18	0	0	0	0	3	0	0	0	0	0	0	0	0

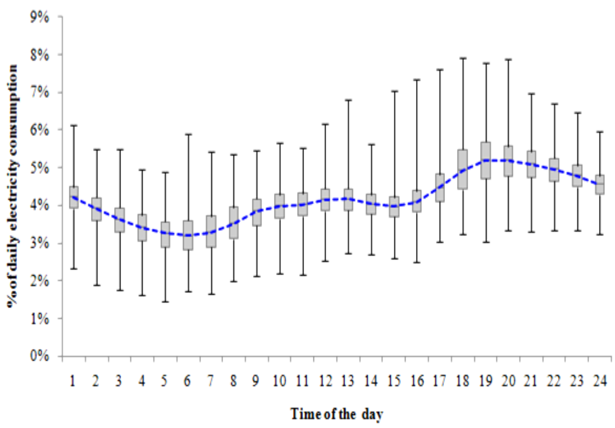

Fig. 5: Proposed typical HEP for student residence halls

By applying the hourly percentage values of proposed HEP on the measured daily total kWh of Beaumont Court, Maurice Court and McLaren House for 365 days of year 2011, 365 estimated HEPs were developed. These estimated HEPs were than compared with the measured HEPs using the hourly box plots. Box plots have been used to visualize and compare the distribution of measured and estimated HEPs for year 2011. Figures 7, 8 and 9 shows the comparison of measured, i.e. actual and estimated HEPs for year 2011 for the Beaumont Court, Maurice Court and McLaren House respectively in kW.

The following are the general observations made from the comparison between actual and estimated hourly profiles as shown in Figures 7, 8 and 9.

i. The hourly data distributions for both actual and estimated HEPs are not normal but are skewed as the hourly medians are not exactly in the centre of the box plots.

ii. The lower quartile and the median of actual and estimated HEPs have matched well for the Beaumont Court and Maurice Court.

iii. The spread of data towards the lower quartiles is compact showing a fairly lower variation in the minimum usage of electricity.

iv. For the McLaren House, both lower quartiles of actual HEPs have slightly lower values than same of estimated HEPs. This is probably because this residence hall has boilers for providing space heating.

v. The wider spread of data towards the upper quartiles shows that maximum hourly electricity load varies and this load variation is mainly due to the variation of day length. During the summer months, average day length is between 11 to 16 hrs whereas during the winter months, the average day length varies between 8 to 10 hrs.

vi. Overall, the spread of estimated HEPs have matched well with the same of actual HEPs.

In the light of above observations which have been drawn from the comparisons of the data distributions for both measured and estimated HEPs as shown in Figures 7, 8 and 9, it could be safely concluded that difference between the hourly measured data and hourly estimated data is as low as 0.9% which indicates the reliability of the proposed HEP. A detailed error analysis of the proposed HEP has been discussed in Section 5.

5. Evaluation of Proposed HEP

Researchers have used different error matrices to evaluate the performance of their models/methods. Such error analysis includes, Mean Absolute Error (MAE), the Root Mean Square Error (RMSE), and Total Absolute Error (TAE). Of these, RMSE has been extensively employed by researchers to evaluate the performance of their models/methods [23, 30, 31].

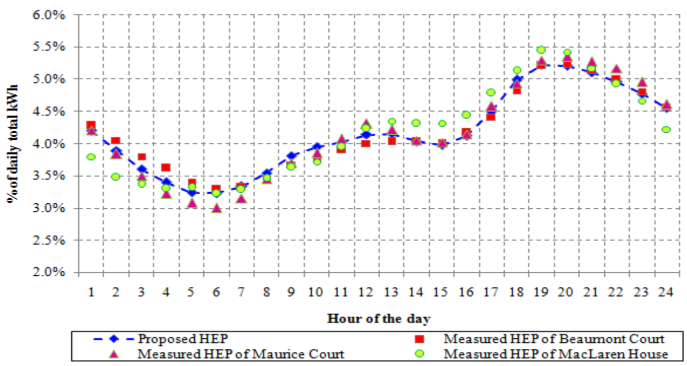

Fig 6: Actual vs. estimated HEPs for 2011

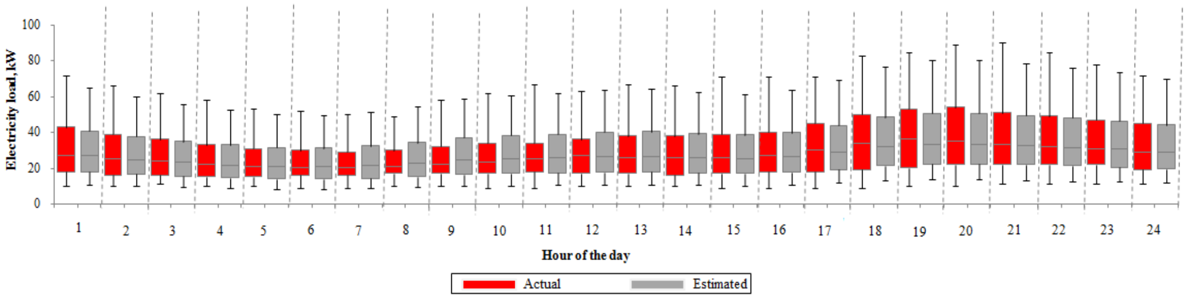

Fig. 7: Actual vs. estimated HEP for Beaumont court

Therefore, we have selected RMSE to analyse the performance and reliability of our proposed HEP for the residence halls. RMSE could be calculated using Eq. (7) as follows;

$$RMSE(N) = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - Y_i^*)^2}{N}} \tag{7}$$

Where *Yi* is the actual value and *Yi** is the estimated value. *N* represents the sample size, which is 365 in this case. We calculated RMSE for Beaumont Court, Maurice Court and McLaren House.

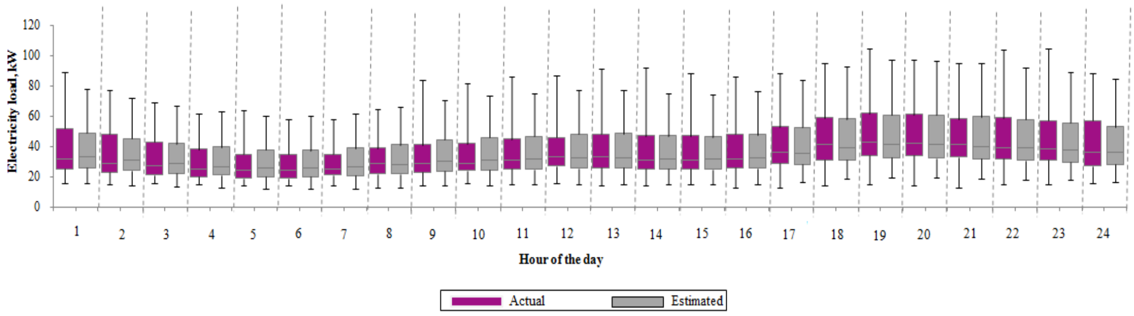

Fig. 8: Actual vs. estimated HEP for Maurice court

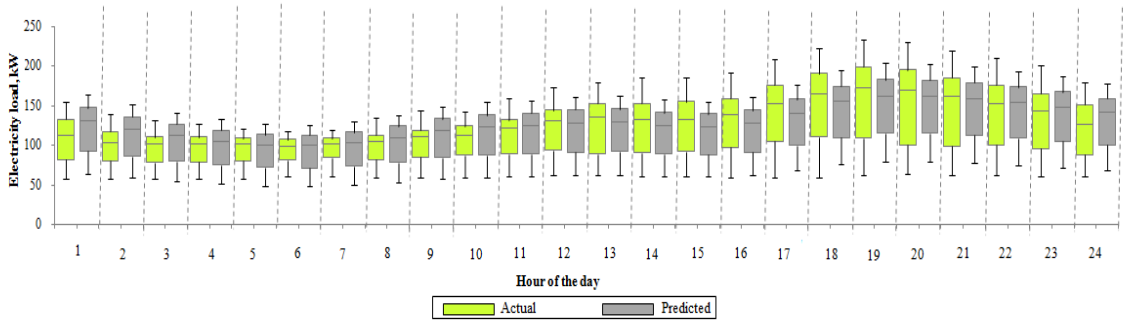

Fig. 9: Actual vs. estimated HEP for McLaren house

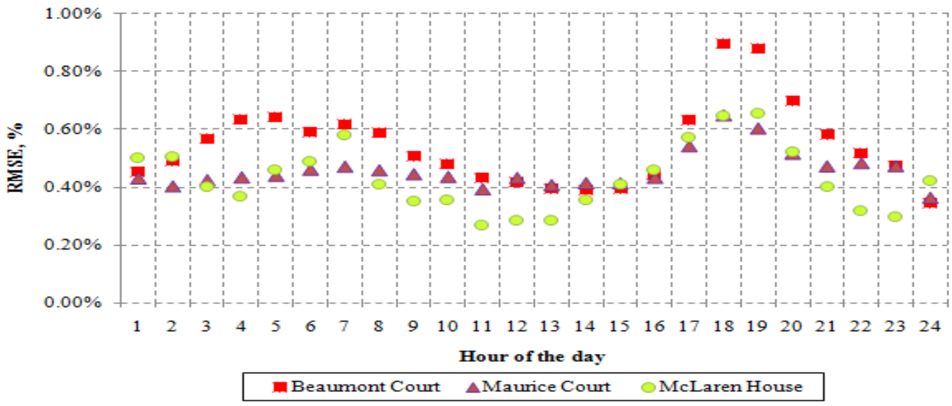

Fig. 10: Hourly RMSE for the estimated HEPs

Fig. 10 shows the plot of RMSE values in kW for each hour of the day for Beaumont Court, Maurice Court and McLaren House.

The hourly RMSE plot indicates that maximum errors for Beaumont Court, Maurice Court and McLaren house is 0.9%, 0.7% and 0.7% respectively. Maximum error occurred during evening period as kitchen equipment are extensively used for cooking during this period. Overall, hourly RMSE remains within a range of 0.2% to 0.9% which is low level of error and hence, it can be concluded that the proposed HEP (%) could be used to estimate hourly electricity load of student residence halls of the UK HE sector.

6. Limitations and Recommendations

Major limitation of this proposed hourly profile is that it is only applicable to the student residence hall buildings located in the United Kingdom as it has not been tested for similar type of buildings outside the UK. Secondly, due to unavailability of data, it has not been tested for the electrically heated residence halls which use battery storage system as such residence hall may have a complete different shape. A test of this profile for buildings outside the UK can give a better picture about its applicability at a larger scale.

7. Conclusions

Hourly consumption profile of a building is the mostly desired form of the electricity consumption data as it helps the Energy Managers in a number of ways such as indentifying abnormal peaks or drops and especially in the calculation of an optimum size of an alternative source of energy for the building. In this study, we collected actual hourly consumption data of three electrically heated student residence halls based in London for the period 1 Jan 2007 to 31 Dec 2010. Using the central tendency method, we analyzed 24 hourly distributions for this period of four years, i.e. one distribution for every hour's consumption and found that median is the most suitable measure of the central tendency for all 24 distributions due to their skewed nature. Therefore, based on the median value of each hourly distribution, we developed a standard HEP which provides hourly percentage values for the student residence halls. An Energy Manager could use this typical HEP to estimate the hourly load of their building.

We tested this proposed profile for three similar types of buildings using their measured data for 2011. Using the RMSE as a measure of error, we calculated values of RMSE for each hour for three buildings. It was observed that the proposed profile fits well with average RMSE for Beaumont Court, Maurice Court and McLaren House to be 0.54%, 0.46% and 0.43% respectively. These lower values of hourly RMSE show that the proposed profile carries a sufficient accuracy and therefore, could be used to estimate the hourly electricity load (kW) of a student residence hall building by multiplying the hourly % values with the daily electricity consumption (kWh).

Acknowledgement

Special thanks to Ms Rebecca Maiden, Head of Sustainability and Mr. John Alam, Building Manager for Student Residence Halls at Queen Mary University of London and Mr. Anuj Saush, Energy Manager at London South Bank University for providing data and information of their campus buildings for making this research successful. We would also like to thank the University of Azad Jammu and Kashmir, Pakistan for providing us the sponsorship for our PhD studies.

References

[1] HEFCE (Higher Education Funding Council for England), "Carbon Reduction Target and Strategy for Higher Education in England", 2010 [Online] Available at: http://www.hefce.ac.uk/ pubs/hefce/ 2010/10_01/10_01a.pdf.

[2] K.P Amber. and J. Parkin, "Barriers to the uptake of combined heat and power technology in the UK higher education sector", Int. J. Sustain. Energ., vol. 34, No. 6, pp. 406-416, 2015.

[3] K.P. Amber, "Development of a Combined Heat and Power Sizing Model for the Higher Education Sector of the United Kingdom", (Unpublished doctoral thesis), London South Bank University, London, United Kingdom, 2013.

[4] A. Saush, "Personal communication with Energy Manager of London South Bank University", June 17, 2011.

[5] S. Manikandan, "Measures of central tendency", J. Pharmacol. Pharmacother., vol. 2, no. 2, pp. 140-142, 2011.

[6] B, Dawson and R.G. Trapp, "Basic and Clinical Biostatistics". 4th ed. New York: Mc-Graw Hill, 2004.

[7] T.D. Swinscow and M.J. Campbell, "Statistics at square one", 10th Ed. New Delhi, India, Viva Books Private Ltd, 2003.

[8] Laerd Statistics, "FAQs - Measures of Central Tendency", 2014 [Online] Available at: https:// stat istics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median-faqs.php.

[9] NIST/SEMATECH, "E-Handbook of Statistical Methods", 2012 [Online] Available at: http:// www.itl.nist.gov/div898/handbook/.

[10] F.J. Gravetter and L.B. Wallnau, "Statistics for the behavioral sciences" 5th Ed. Belmont: Wadsworth – Thomson Learning, 2000.

[11] A. Petrie and C. Sabin, "Medical statistics at a glance" 3rd Ed. Oxford: Wiley-Blackwell, 2009.

[12] S. Manikandan, "Measures of central tendency: Median and mode", Pharmacol. Pharmacother., vol. 2, no. 3, 214-215, 2011.

[13] Vitutor, "Notes 2, Mode", 2010 [Online] Available at: http://www.vitutor.com/statistics/descriptive/mode.html.

[14] J. Zhao, Y. Xin and D. Tong, "Energy consumption quota of public buildings based on statistical analysis", Energy Policy, vol. 43, pp. 362-370, 2012.

[15] K.R. Sundaram, S.N. Dwivedi and V. Sreenivas, "Medical statistics principles and methods", 1st Ed. New Delhi: B.I Publications Pvt. Ltd, 2010.

[16] B. Dawson and R.G.Trapp RG. "Basic and Clinical Biostatistics" 4th ed. New York: Mc-Graw Hill, 2004.

[17] Wikipedia, "Median", 2014 [Online] Available at: http://en. wikipedia.org /wiki/Median.

[18] T. Sharp, "Energy benchmarking in commercial-office buildings". Proc. of the 1996 ACEEE Summer Study on Energy-Efficiency in Buildings, American Council for an Energy-Efficient Economy, Washington, DC, vol. 4, , pp. 321-329, 1996.

[19] M. Hubert and E. Vandervieren, "An adjusted box plot for skewed distributions", Comput. Stat. Data An., vol. 52, No. 12, pp. 5186-5201, 2008.

[20] G. Marshall and L. Jonker, "An introduction to descriptive statistics: A review and practical guide", Review Article, Radiography, vol. 16, No. 4, pp. 1-7, 2010.

[21] J. Keirstead, "Benchmarking urban energy efficiency in the UK", Energy Policy, vol. 63, pp. 575-587, 2013.

[22] V.M. Nik, A.S. Kalagasidisa and E. Kjellströmb, "Statistical methods for assessing and analysing the building performance in respect to the future climate", Build. Environ., vol. 53, pp-107-118, 2012.

[23] K.P. Amber, M.W. Aslam and K. Hussain, "Electricity consumption forecasting models for administration buildings of the UK Higher Education sector" Energ. Buildings, vol. 90, pp. 127-136, 2015.

[24] ABS, Australian Bureau of Statistics, "Frequency Distribution" [Online] Available from: http://www.abs.gov.au/websitedbs/ a3121120.nsf/home.

[25] I. Ward, A. Ogbanna and A. Altan, "Sector review of UK higher education energy consumption", Energy Policy, vol. 36, pp. 2939-2949, 2008.

[26] S. Katipamula, T.A. Reddy and D.E. Claridge, "Bias in predicting annual energy use in commercial buildings with regression models developed from short datasets" Pacific Northwest Laboratory, K5-20, P.O Box 999, Richland, Washington 99352, 1994.

[27] Mathbits, "Correlation Coefficient" Available at: http://mathbits. com/MathBits/TISection/Statistics2/correlation.htm.

[28] QMUL, Queen Mary University of London, "Mile End Accommodation" [Online] Available from: http://www.residences.qmul.ac.uk/college/qmaccommodation/mileend/.

[29] ERG, Environmental Research Group, Kings College London, Data download. [Online]. Available at: http://www.londonair.org.uk/ [Accessed on] November 25, 2012.

[30] Korolija, Y. Zhang, L.M. Halburd and V.I. Hanby, "Regression models for predicting UK office building energy consumption from heating and cooling demands", Energ. Buildings, vol. 59, pp. 214–227, 2013.

[31] V. Bianco, O. Manca and S. Nardini, "Electricity consumption forecasting in Italy using linear regression models", Energy, vol. 34, pp. 1413-1421, 2009.

[32] C. Noren and J. Pyrko, "Using Multiple Regression Analysis to Develop Electricity Consumption Indicators for Public Schools", Lund Institute of Technology, Sweden, 1997, [Online]. Available at: http://aceee.org/files/proceedings/1998/data/papers/0321.pdf.

[33] R. Edwards, J. New and L.E. Parker, "Predicting future hourly residential electrical consumption: A machine learning", Energ. Buildings, vol. 49, pp. 591-603, 2012.