Evaluation of time series techniques to characterise domestic electricity demand

Fintan McLoughlin [a,*], Aidan Duffy [a], Michael Conlon [b]

[a] School of Civil and Building Services and Dublin Energy Lab, Dublin Institute of Technology, Bolton St, Dublin 1, Ireland
[b] School of Electrical Engineering Systems and Dublin Energy Lab, Dublin Institute of Technology, Kevin St, Dublin 4, Ireland

ABSTRACT

This paper discusses time series approaches, often used by Transmission System Operators (TSOs) to forecast system demand, and applies them at an individual dwelling level. In particular, two techniques, Fourier transforms and Gaussian processes were evaluated and used to characterise individual household electricity demand. The performance of the characterisation approaches were evaluated based on Pearson correlation coefficient, descriptive statistics and paired sample *t*-tests for electrical parameters: Total Electricity Consumption, Maximum Demand, Load Factor and Time of Use of maximum electricity demand. Finally, a number of time series tests were carried out to ensure certain properties remained between the original and characterised series.

Both Fourier transforms and Gaussian processes were shown to be suitable techniques for characterising domestic electricity demand. Depending on customer demand load profiles, each approach has its own strengths and weaknesses. Fourier transforms are better at characterising the profiles of customers who consume electricity more evenly across the day (>1 h). In contrast, Gaussian processes are better at characterising customers whose demand is high for only short periods of time (≤ 1 h) throughout the day.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The European Union has set ambitious energy related targets by the year 2020: to cut greenhouse gas emissions by 20%, to improve energy efficiency by 20% and for 20% of EU energy demand to come from renewable energy resources. Advanced Metering Infrastructure (AMI) such as smart meters provide a mechanism to achieve this by:

- improving network operations and peak load management through time of use tariffs thus leading to more efficient electricity generation and therefore helping to reduce greenhouse gas emissions
- encouraging energy efficiency by providing real-time information on customer demand and;
- provide a smart platform for selling power back and forth to the grid thus facilitating the installation of renewable energy generation and electric vehicle integration;

The introduction of large scale AMI is a relatively new concept to the residential sector (within the last 5–10 years). Prior to this, AMI was usually limited to large electricity consumers such as commercial or industrial end-users, mainly due to cost. However, AMI technology costs have decreased in the last decade and this, combined with meeting EU 20/20/20 targets has encouraged interest in the area with national policy makers willing to support smart metering programmes for reasons outlined above. As a result a wealth of new information now exists giving detailed electricity consumption data for large sample sizes in the residential sector.

Also, advances in generation and storage technologies such as micro-generation and electric vehicles has meant new opportunities now exist for changing how energy is used in the home. However, to assess the impact of such generation and storage technologies a detailed understating of how energy is used in the home is necessary. The smart metering trial carried out by Electric Ireland (formally Electricity Supply Board) has made this possible by recording half-hourly electricity demand for over 4200 residential customers in Ireland.

This paper identifies and summaries time series approaches that can be used to characterise domestic electricity demand in the home. Time series methods have rarely been used at an individual

* Corresponding author. Tel.: +353 14023918; fax: +353 14024035.
E-mail address: fintan.mcloughlin@dit.ie (F. McLoughlin).

dwelling level before and have more commonly been used to characterise total system demand at an aggregate level. The paper discusses the advantages and disadvantages of each method as applied to the domestic sector and evaluates the most promising techniques.

1.1. Electricity demand profiling

A large body of literature already exists for characterising aggregate electricity system demand load profiles. However, characterising electricity demand at an aggregate as compared to an individual dwelling level are two very different tasks. Fig. 1 (dashed line) shows a typical aggregate profile shape for total system demand for the Irish Transmission System Operator (TSO), Eirgrid, on the 1st July 2009 over a 24-h period. The profile shows consistent peaks in the morning, lunch and evening times. The profile shape varies slightly for different days of the week and over the course of the year due to seasonality.

In contrast, Fig. 1 (continuous line) shows a distinctly different electricity load profile shape for a single random dwelling on the same day of the year. The profile shows a peak in the late morning around 10 am which lasts until 4 pm in the evening and a later peak at 10.30 pm that night. This pattern of electricity use across the day is very different to that of an aggregate system demand, where a more sporadic use is apparent rather than a gradual smooth profile.

Fig. 2 shows a standard load profile issued by the Meter Registration System Operator (MRSO) for the Irish domestic electricity market for the same day. Standard domestic load profiles are used for the purposes of settlement between suppliers in the electricity market and are normalised over the period of one year. There are certain similarities between Figs. 1 and 2 where the standard domestic electricity load profile is comparable to both total system demand and individual dwelling load profiles in terms of its shape. This is a result of the methodology used to derive standard load profiles which is based on a regression of mean electricity demand of a sample of dwellings against various environmental and weekday parameters [1].

However, typical domestic electricity load profiles are far more variable than that shown in Fig. 2 and can change in the time (on a day to day basis) and space domain (between customers). Fig. 3 shows a single customer over a weekly period from 1st to 7th July 2009. On a daily basis, the profile shape can change significantly from one day to the next in terms of the magnitude of electricity demand and the time at which it is used.

Similarly the profile shape can change significantly between customers. Fig. 4 shows nine different customer profiles at random for the 1st July 2009. The figure shows how the profile can change considerably between customers in both magnitude and time of use of electricity demand.

A seasonality component also exists within a domestic electricity load profile, mainly a result of changes in external temperature and daylight hours for heating (albeit small in Ireland due to limited penetration of electric heating) and lighting homes respectively. This trend is shown in Fig. 5 where half-hourly periods for an individual customer are plotted across the year (1st July 2009 to 30th June 2010). A quadratic polynomial function is fitted to the data in order to show the seasonality trend where average electricity demand increases (by as much as 200 W) for the winter period. Daily peaks are also evident from the figure, increasing in magnitude during the winter. This trend is characteristic of the typical change in profile shape over a yearly period for customers living in Ireland. Total annual electricity consumption for this particular customer was 4537 kWh over the period which corresponds with the average in Ireland. Fig. 5 also shows a period of approximately two weeks in April where electricity demand decreases to near zero. This period is very different to any other time of the year and most probably identifies a time when the dwelling was unoccupied.

Therefore, in order to characterise each individual domestic electricity load profile successfully, an approach needs to consider the following key factors:

○ diurnal variations in electricity demand (Figs. 3 and 4);
○ intra-daily variations in electricity demand (Fig. 3);
○ electricity demand variations between customers (Fig. 4); and
○ seasonal electricity demand effects (Fig. 5)

The scope of this paper is primarily focussed on characterising the diurnal variations for individual customer demand load profiles as their highly variable nature makes them the most difficult to characterise effectively.

2. Methodology

Time series approaches have been used extensively for modelling aggregate electricity system demand as shown in Fig. 1 (dashed line).
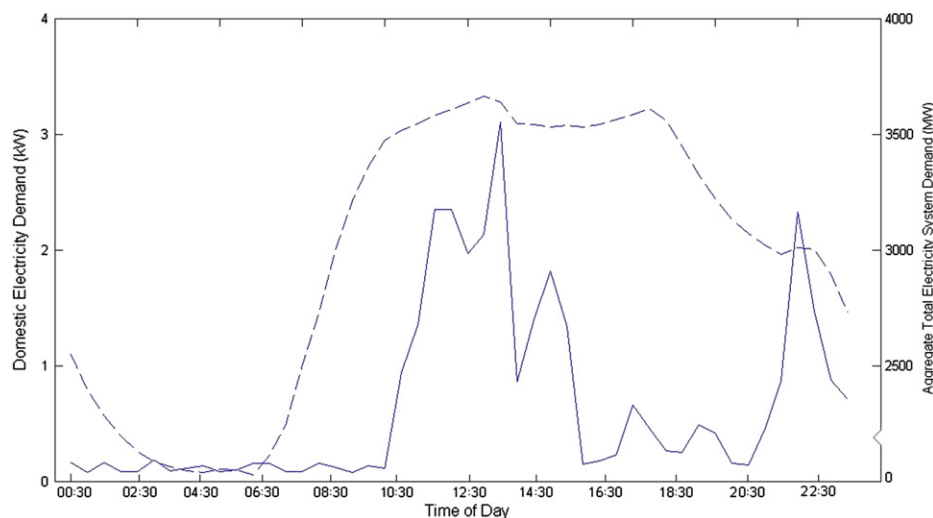

Fig. 1. Daily electricity load profile for total system demand (dashed line) and an individual dwelling (continuous line) across a 24-h period on 1st July 2009 (source: Eirgrid and Smart Metering Data).
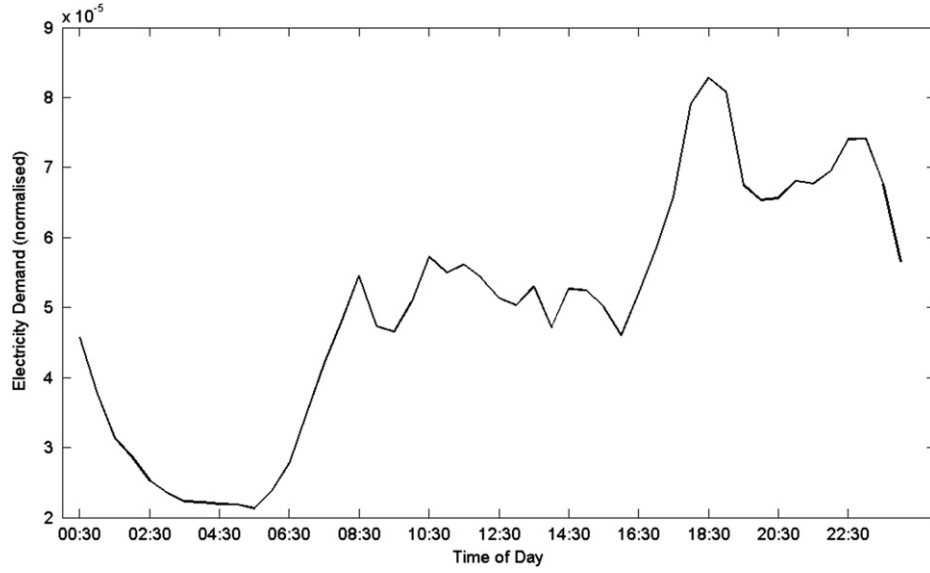
Fig. 2. Daily electricity standard load profile for urban domestic across a 24-h period on 1st July 2009 (source: MRSO).

However, their use at an individual domestic scale has been somewhat limited, mainly due to the historic lack of data at this level.

In order to characterise electricity demand in a presentable format, four electrical parameters were used from a previous paper by the same authors [2]. These were found to reflect the main characteristics of household electricity use such as Total Electricity Consumption, Maximum Demand, Load Factor and Time of Use of maximum electricity demand and are presented in Eqs. (1)–(4).

E_{TOTAL} is the total amount of electricity consumed over a yearly period in kWh where E_i is average electrical demand in kW for each half hour period and l is the total number of half-hourly periods over the year.

$$E_{Total} = 1/2 \sum_{i=1}^{l} E_i \qquad (1)$$

Eq. (2) describes mean daily maximum demand, E_{MD} over a yearly period in kW. E_{MD} refers to the largest value of electrical demand in a day, averaged over the year where E_i is average electrical demand in kW for each half hour period, n is the total number of periods in a day and m is the total number of days in the year.

$$E_{MD} = \frac{1}{m} \sum_{j=1}^{m} \max\{E_i, 1 \leq i \leq n\} \qquad (2)$$

Daily load factor, E_{LF} is a ratio and is shown in Eq. (3) below, where E_i is average electrical demand in kW over each half hour period, n is the total number of periods in a day and m is the total number of days over the year. It is a measure of daily mean to daily maximum electrical demand and is a measure of the "peakyness" of a customer's load profile.

$$E_{LF} = \frac{1}{m} \sum_{j=1}^{m} \frac{\frac{1}{n} \sum_{i=1}^{n} E_i}{\max\{E_i, 1 \leq i \leq n\}} \qquad (3)$$
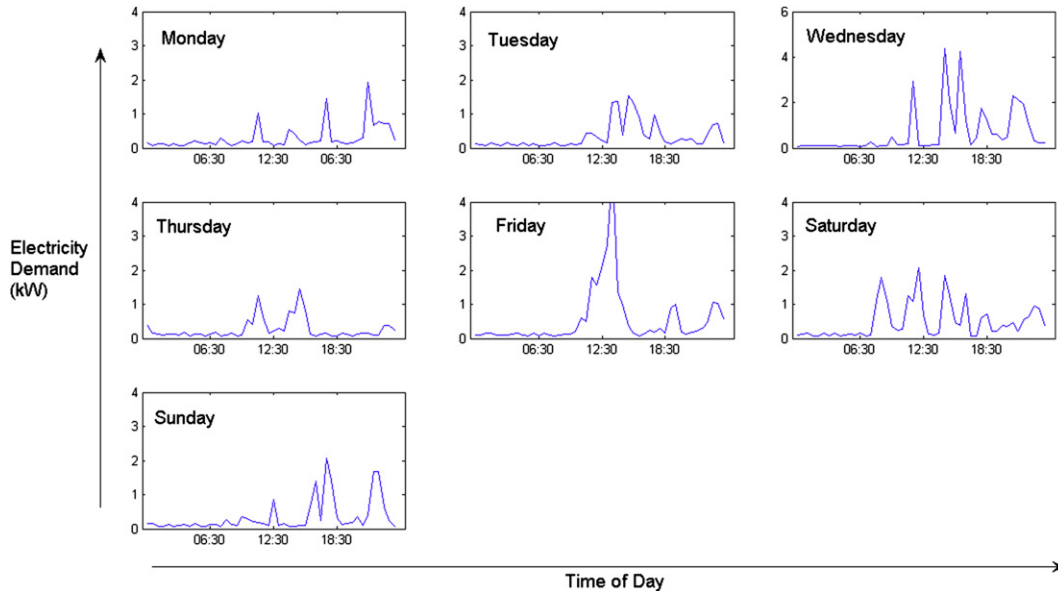

Fig. 3. Daily electricity load profiles for a single randomly chosen customer over a weekly period showing intra-daily variations.
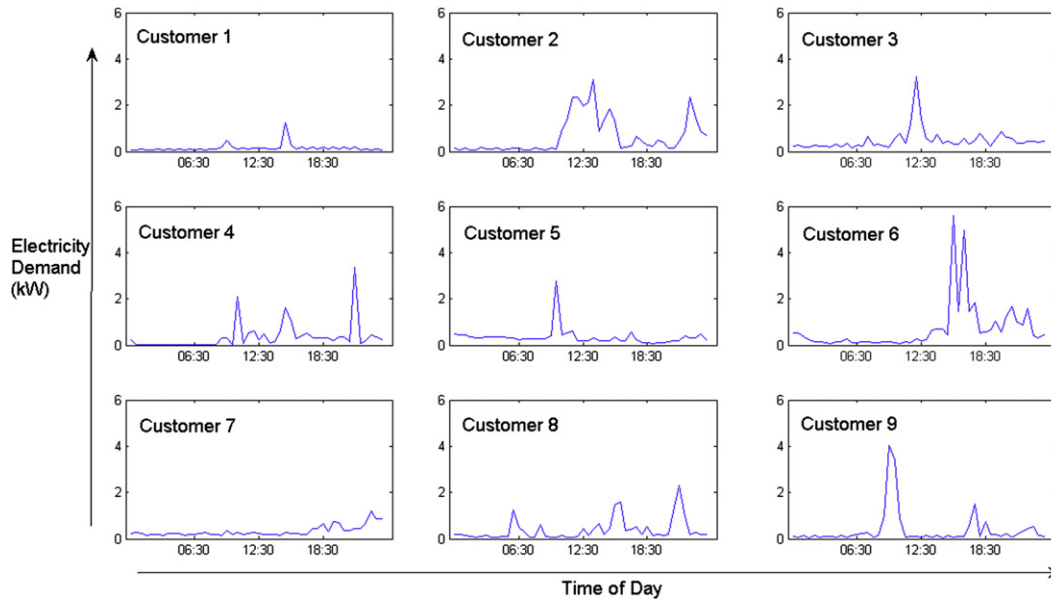
Fig. 4. Daily electricity load profiles for nine randomly chosen customers illustrating variation between dwellings.

A maximum Time of Use (ToU) parameter, E_{ToU} over a yearly period is defined by Eq. (4) below where E_i represents average electrical demand in kW over each half hour period and j_{max} corresponds to the time in which maximum electricity demand occurs across the daily period (where $1 = 00{:}00{-}00{:}30$ and $48 = 23{:}30{-}00{:}00$), n is the total number of intervals in a day and m is the total number of days in the year. ToU indicates the time of day at which maximum electricity demand occurs.

$$E_{ToU} = \text{mode}\left\{j_{max}\middle| E_{j_{max}} = \max\{E_i, 1 + n(j - 1) \le i \le n, 1 \le j \le m\}\right\}$$

(4)

A paired sample t-test was used to compare electrical parameters between the characterised and original time series. A number of assumptions were made such that: the observations are independent and that the differences between the paired samples are normally distributed. In addition, a number of tests were carried out to ascertain whether certain temporal properties remained between the original and characterised time series. These were first presented in a previous paper by the same authors [3] and include probability distribution function, autocorrelation function and a spectral periodgram. In conjunction with Eqs. (1)–(4), the authors found the tests to be central to establishing the suitability of each technique for electricity load profile characterisation.

When characterising aggregate electricity system demand load profiles a number of key performance indicators are used to measure the accuracy of approaches: mean absolute error, root mean squared error, mean absolute percentage error and root mean squared percentage error. The most commonly used performance indicator for measuring accuracy of electricity system demand
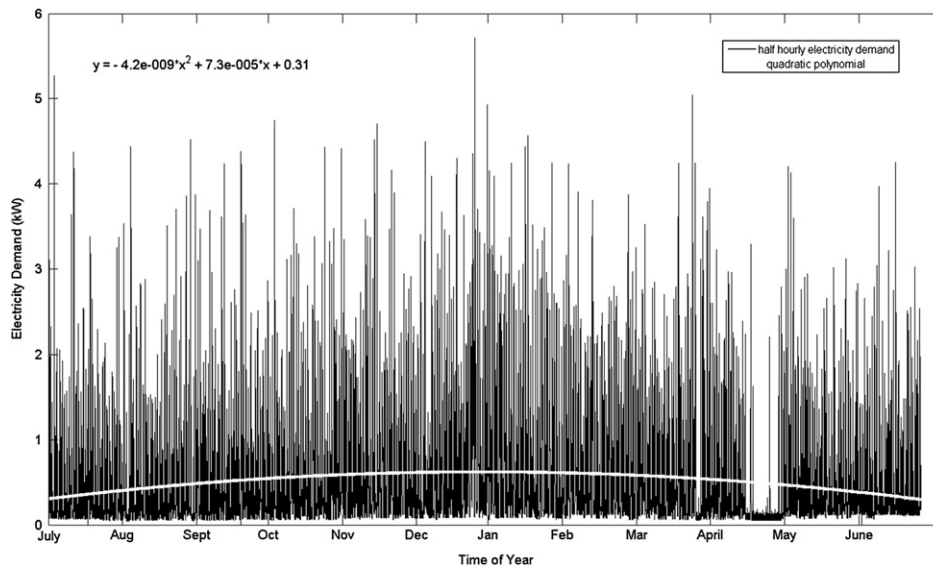

Fig. 5. Electricity load profile for a single randomly chosen customer over a yearly period.

forecasts is mean absolute percentage error (MAPE). However, this parameter is unsuitable for measuring the performance of the characterisation process with domestic loads due to the highly variable nature of individual dwelling level demand leading to very high errors. Instead, the performance of the time series technique is measured by correlating the original and characterised time series as described by Eq. (5) where $\rho_{x,Y}$ is the correlation coefficient between two variables X and Y with expected values μ_x and μ_y and standard deviations σ_x and σ_y.

$$\rho_{X,Y} = \text{corr}(X,Y) = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \quad (5)$$

The data used in the analysis was taken from a set of 345,645 residential households in Ireland. The set was divided into six groups based on total annual household electricity consumption to ensure an even spread of electricity consuming customers. An initial sample of 5574 was drawn on a randomised basis across all profiles. This was subsequently reduced to 5375 households by targeting certain groups to improve the representativeness of dwelling and socio-economic variables within the sample size. A control group sample size of 1170 customers was used to benchmark domestic electricity use across a yearly period. A final sample

size of 509 households from the control group was used, once non-continuous data (a result of technology communication errors) were removed. Anonymised half-hourly electricity demand data were recorded over a twelve month period from 1st July 2009 to 30th June 2010.

3. Method selection and discussion

Based on the literature, the majority of time series approaches for electricity load profiling can be grouped under the following headings: Fourier transforms, Neural networks, Gaussian processes, Autoregressive, Fuzzy logic, Wavelets and multiple Regression/Probabilistic. The literature associated with these methods is summarised in Table 1 where the main advantages and disadvantages for each approach is presented. Table 1 also indicates whether each approach was applied at an aggregate or individual dwelling level before and the time resolution at which it has been applied.

As discussed in Section 1.1, in order to characterise domestic electricity load profiles effectively an approach needs to consider the diurnal, intra-daily, seasonal and demand variations between customers. Therefore, ideally it would be advantageous to characterise domestic load profiles based on the temporal and magnitude

Table 1
Advantages and disadvantages by approach for electricity load profiling.

Characterisation type	Applied to aggregate demand	Applied to individual dwelling demand	Time resolution – High (≤1 h)	Time resolution – Low (>1 h)	Advantages	Disadvantages
Fourier Series	Yes [4–6]	No	Yes [4,5]	Yes [6]	Temporal and magnitude components represented in the variable coefficients with the latter scalable.	Fourier transforms are poor at characterising small "sharp" intervals of electricity demand.
Neural Networks	Yes [7–9]	Yes [10–12]	Yes [7–9]	Yes [10–12]	Good at characterising highly non-linear relationships such as domestic electricity load profiles.	Black box approach. Variable coefficients do not represent the temporal and magnitude components of an electricity load profile.
Gaussian Processes	Yes [13–15]	No	Yes [13–15]	No	Good at approximating small intervals of "sharp" electricity demand.	Less good at approximating "smother" average electricity demand profiles.
Autoregressive (incl. Markov chain)	Yes [16–19]	Yes [3,20]	Yes [3,16,18–20]	Yes [17]	Widely used in aggregate electricity system demand load profiling. Markov chains are able to characterise the variable component of domestic electricity load profiles.	Variable coefficients vary unpredictably with small changes in profile shape and don't represent temporal and magnitude components. Markov chains unable to characterise the temporal component unless each half hour period is characterised separately (i.e. leading to a large number of variables).
Fuzzy Logic	Yes [21–23]	No	Yes [21–23]	No	Cause and effect clearly defined between input and output.	Temporal component not characterised unless each half hour period characterised separately.
Wavelets	Yes [24–27]	No	Yes [24–27]	No	High and low frequency components represented by two different series analogous to base load and peak load for electricity load profiling.	The time series is effectively split in half, with each section characterised separately thus doubling the number of variables required.
Multiple Regression/ Probabilistic	Yes [28–31]	Yes [1,32–38]	Yes [1,28,29,32–37]	Yes [30,31,38]	Widely used for generating standard load profiles (as shown in Fig. 2)	Load profiles tend to be average rather than variable unless each half-hourly period is characterised separately.

Table 2
Descriptive statistics for mean Pearson's correlation coefficient.

Characterisation method	Mean	Median	Standard deviation	Maximum	Minimum
Fourier Transforms	0.8743	0.8761	0.0418	0.9878	0.6774
Gaussian Processes	0.9447	0.9473	0.021	0.9931	0.7843

Table 4
Descriptive statistics for paired sample *t*-test for total electricity consumption.

Paired samples	Mean	Standard deviation	Std. Error mean	t	Sig. (2-tailed)
Original-Fourier Time Series	0.0009	0.0164	0.0007	1.2140	0.2250
Original-Gaussian Time Series	98.5679	57.4282	2.5455	38.7230	0.0000

properties of the time series since this will result in cyclical patterns (i.e. diurnal, intra-daily, etc.) being captured in the variable coefficients. This would also allow demand variations between customers to be investigated by determining the influence of dwelling and occupant characteristics on the load profile shape throughout the day, however, this is beyond the scope of the paper. In addition, it would also be of benefit if the characterisation process were as simple as possible. The smaller the number of variables used to represent the electricity load profile, without sacrificing accuracy will ultimately lead to a simpler characterisation of the time series overall. Also, keeping the number of variables small, will generally lead to a less computationally demanding characterisation process.

The principal advantage of Fourier transforms is their ability to characterise the temporal and magnitude components within the data, with the latter scalable, an additional desirable quality [4]. This means that comparable profile customers that show similar patterns of electricity use can be grouped together. The disadvantage with Fourier transforms, as will be seen later, is that they have difficulty characterising small intervals of large electricity demand [39]. Neural networks are especially good at characterising non-linear relationships and are therefore well suited to the highly variable nature of domestic electricity load profiling. However, they are seen as a black box approach and it is often difficult to establish a relationship between input and output [23]. This is not ideal, particularly in this case where a clear understanding of the change in profile shape across the day is required. The structure is often quite complex too, involving multiple neurons and layers that require a significant number of variables to describe the daily load profile accurately. Nor do the variable coefficients reflect the temporal and magnitude components of the electricity load profiles; rather they represent the weights and biases of input to output for the time series. In contrast to Neural networks, Fuzzy logic has the advantage that the relationship between input and output is clearly defined [23]. However, the number of variables required to characterise the output is usually large, particularly when the load profile shape changes considerably across a daily period.

In contrast to Neural networks, Gaussian processes provide a much simpler representation of the load profile shape. Each profile is characterised by three moments: mean; variance; and a weight coefficient that describe each probability distribution [13]. Compared to Fourier transforms, Gaussian processes can sufficiently characterise small intervals of large electricity demand. However, it must be noted that the characterisation order needs careful consideration as if it is too high redundant distributions will lead to over fitting and if it is too low the profile peaks will not be

fully represented. Autoregressive approaches are well established in electricity system demand load profiling. They are usually combined with a moving average process to form an autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA). Due to the highly variable nature of individual domestic loads, autoregressive processes find it difficult to characterise without using high order methods, leading to a large number of variables. The variable coefficients also vary significantly with small changes in load profile shape and this makes it difficult to group or compare customers [4]. Another type of autoregression, Markov chains can characterise the highly variable nature of a domestic load profiles; however, they have difficulty capturing the temporal properties unless a large number of variables are used to represent each half hour period in a day [3].

Wavelets are similar to Fourier transforms as they apply the same spectral decomposition technique. However, their advantage over Fourier transforms is the separation of the electricity load profile into high and low frequency components before applying the transform. This results in two or more characteristic curves representing distinctly different patterns of electricity use for individual customers. The advantage in doing this is that certain dwelling and occupant characteristics have different periods of influence over electricity consumption in the home [26]. However, the disadvantage is that it effectively doubles the number of variables required to characterise the time series.

Finally, multiple regression is a technique that has been used extensively in electricity load profiling. Similar to autoregression it is most often applied at an aggregate level for system demand load forecasting and with time intervals greater than one day. It is the method of choice for the UK grid operator, National Grid, to develop standard load profiles for the purposes of electricity settlement as discussed earlier [1]. However, a large amount of variables are required to characterise these which are in effect average load profiles, such as that shown in Fig. 2. Monte Carlo analysis is the most common probabilistic approach to load profile characterisation. The advantage of this technique is that it is ideal for generating variable load profiles and so is well suited to domestic electricity load characterisation. However, using Monte Carlo to characterise domestic loads would require each half hour interval to be represented independently with a probability distribution function leading to a large number of variables.

Overall, Fourier transforms, Wavelets and Gaussian processes all appear to represent the temporal and magnitude components sufficiently within the variable coefficients. Fuzzy Logic,

Table 3
Descriptive statistics for mean total electricity consumption.

Characterisation method	Mean	Median	Standard deviation	Maximum	Minimum	Scale parameter (η)[a]	Shape parameter (β)[a]
Original Time Series	4146 kWh	4008 kWh	1870 kWh	9651 kWh	414 kWh	4687	2.38
Fourier Transforms	4146 kWh (0%)	4008 kWh (0%)	1870 kWh (0%)	9651 kWh (0%)	414 kWh (0%)	4687 (0%)	2.38 (0%)
Gaussian Processes	4047 kWh (−2.39%)	3903 kWh (−2.62%)	1835 kWh (−1.87%)	9462 kWh (−1.96%)	413 kWh (−0.24%)	4576 (−2.37%)	2.37 (−0.42%)

[a] Weibull Probability Distribution Function.

Table 5
Descriptive statistics for mean customer daily maximum demand.

Characterisation method	Mean	Median	Standard deviation	Maximum	Minimum	Scale parameter (η)[a]	Shape parameter (β)[a]
Original Time Series	2.34 kW	2.29 kW	0.92 kW	6.18 kW	0.14 kW	2.6293	2.7425
Fourier Transforms	1.68 kW (−28.21%)	1.66 kW (−27.51%)	0.68 kW (−35.29%)	3.89 kW (−58.87%)	0.09 kW (−55.56%)	1.8904 (−28.10%)	2.6885 (−1.97%)
Gaussian Processes	2.23 kW (−4.70%)	2.20 kW (−3.93%)	0.88 kW (−4.35%)	5.99 kW (−3.07%)	0.13 kW (−7.14%)	2.5082 (−4.61%)	2.7394 (−0.11%)

[a] Weibull Probability Distribution Function.

Table 6
Descriptive statistics for paired sample *t*-test for maximum demand.

Paired samples	Mean	Standard deviation	Std. Error mean	*t*	Sig. (2-tailed)
Original-Fourier Time Series	0.6554	0.3094	0.0137	47.7900	0.0000
Original-Gaussian Time Series	0.1053	0.0680	0.0030	34.9490	0.0000

Autoregressive, Neural Network and Regression also have this capability but would require each half hour period to be characterised independently by at least a single variable. This is a disadvantage as a minimum of forty eight variables would be required to characterise the temporal and magnitude components for each of these methods. Autoregressive approaches such as ARIMA have been used extensively in the past to forecast aggregate electricity system demand for markets all around the world. The Moving Average (MA) component lends itself well to characterising the smooth transitions between half-hourly periods which is typical of aggregate electricity system demand load profiles as shown in Fig. 1 (dashed line). However, this component is not as well suited to individual residential applications where electricity consumption changes very quickly over short periods of time. Regression, Probabilistic and Fuzzy logic techniques all take a descriptive approach and are deemed unsuitable in this instance as too many variables would be required to characterise the load profile shape. Neural networks are notoriously complex requiring a number of variables to represent the weights and biases at different layers to characterise the output successfully. As a result it is difficult to compare variable coefficients between customers because of this complicated architecture. Wavelets use Fourier transforms to decompose the time series into high and low frequency components so therefore there is some overlap between these two methodologies with the former requiring double the amount of variables.

Due to the fact that both Fourier transforms and Gaussian processes appear to be complimentary to each other, as well as being able to characterise the temporal and magnitude components with a relative small number of variables, both approaches were evaluated in the next section. In both cases an eighth order characterisation process was applied to the data.

4. Results and discussion

The following section presents characterisation results for both Fourier transforms and Gaussian process time series techniques. Both approaches were used to characterise half-hourly electricity demand on a daily basis over the entire year for each individual customer. Table 2 shows descriptive statistics for mean Pearson's correlation coefficient between the original and the two characterised time series for all customers over the entire year. Both characterisation approaches performed well with Gaussian processes on average 7% more accurate than Fourier transforms.

A number of electrical parameters were calculated using Eqs. (1)–(4) from the original data and both characterisation techniques, with the results shown in Tables 3–9. Percentage error is calculated with the following equation (%error=((Characterised−Original)/Original)*100%) between original and characterised parameter values and is shown in brackets in the following tables. A Weibull probability distribution function was found to be the best fit to parameters Total Electricity Consumption and Maximum Demand whereas a Log-Logistic probability distribution function was found to be the best fit to Load Factor.

Table 3 shows descriptive statistics for mean Total Electricity Consumption for all customers over the yearly period. Both Fourier transforms and Gaussian processes characterised parameter Total Electricity Consumption with less than 5% percentage error. Fourier transforms produced accurate results, with errors less than 1/10th of a percent. This is not surprising since Fourier transforms employ a data integration process. Gaussian processes on the other hand was less accurate but still within acceptable limits overall (<5% percentage error).

Table 4 shows results for a paired sample *t*-test between the original and characterised time series for parameter Total Electricity Consumption. A 2-tailed significance value of 0.225 for Fourier transforms indicates that there is little difference between the original and characterised parameters. This finding is supported by the small differences observed between the means and standard deviations of the Fourier transforms. In contrast, results for Gaussian processes indicate that there is a significant difference between the characterised and original time series for the same parameter.

Table 5 shows results for mean daily Maximum Demand for all customers over the year. Fourier transforms were poor at capturing the daily peak demands characteristic of almost all individual dwellings. Descriptive statistics in Table 5 show percentage errors in excess of 20% for Fourier transforms, with the largest errors at the extremities for Maximum and Minimum. In contrast, Gaussian processes were better at characterising this parameter with percentage errors of less than 5% in most instances. This highlights one of the more significant advantages of this particular characterisation method over Fourier transforms, which will be discussed further later.

Table 7
Descriptive statistics for mean customer load factor.

Characterisation method	Mean	Median	Standard deviation	Maximum	Minimum	Scale parameter (η)[a]	Shape parameter (β)[a]
Original Time Series	23.23%	22.35%	5.76%	48.69%	11.29%	−1.4935	0.1299
Fourier Transforms	31.79% (36.85%)	30.76% (37.63%)	6.59% (4.41%)	66.72% (37.03%)	18.05% (59.88%)	−1.1703 (−21.64%)	0.109 (−19.17%)
Gaussian Processes	24.74% (6.5%)	23.74% (6.22%)	6.54% (13.54%)	51.76% (6.31%)	11.89% (5.31%)	−1.434 (−3.98%)	0.138 (6.24%)

[a] Log − Logistic Probability Distribution Function.

Table 8
Descriptive statistics for paired sample *t*-test for load factor.

Paired samples	Mean	Standard deviation	Std. Error mean	*t*	Sig. (2-tailed)
Original-Fourier Time Series	−0.0829	0.0228	0.0010	−81.9350	0.0000
Original-Gaussian Time Series	−0.0123	0.0109	0.0005	−25.5020	0.0000

Table 9
Descriptive statistics for mean customer time of use (ToU) for daily maximum demand (where 1 = 00:00–00:30 and 48 = 23:30–00:00).

Characterisation method	Mean	Median	Standard deviation
Original Time Series	30.7	31.16	3.52
Fourier Transforms	31.44 (2.41%)	31.84 (2.18%)	3.62 (2.84%)
Gaussian Processes	29.63 (−3.49%)	29.91 (−4.01%)	3.3 (−6.25%)

Table 6 shows a paired sample *t*-test for Maximum Demand parameter. As discussed above, Gaussian Processes were better at representing the characteristics of this parameter. However, the results also show that the Maximum Demand parameter for both characterised time series were significantly different from that of the original time series at the 95% *p*-value level.

Table 7 presents results for mean Load Factor for all customers over the yearly period. It can be seen that, Fourier transforms were unable to accurately characterise mean customer Load Factor with percentage errors exceeding 30% in most instances. However, this is not surprising as Load Factor is a function of Maximum Demand.

Table 8 shows results for a paired sample *t*-test for the Load Factor parameter. Fourier transforms over estimated Load Factor compared to the original time series more than Gaussian processes. The results show that Gaussian processes are better at representing Load Factor but the *t*-test shows that both time series techniques were significantly different compared to the original data at the 95% *p*-value level.

Table 9 shows results for mean Time of Use (ToU) for daily Maximum Demand for all customers over the yearly period. Fourier transforms appear to be the more accurate of the two techniques, with less than 3% percentage error, but tended to overestimate its value indicating later use of maximum electricity demand.

Gaussian processes percentage errors on the other hand were slightly greater and on average tended to predict peak time electricity use slightly earlier in the evening. However, because the parameter has a bi-modal distribution such descriptive statistics do not show the full extent of the variation. As a result, a probability distribution function was not fitted to the Time of Use parameter and a paired sample *t*-test was not carried out as it would break the assumption of normality between the characterised parameter values.

Therefore, in order to test the temporal properties of the time series more rigorously a number of additional tests were also carried out. For example domestic electricity demand could be characterised by a simple Monte Carlo simulation where a customer's load over a daily period is represented by a probability distribution function and could theoretically produce accurate parametric results for: Total Electricity Consumption, Maximum Demand and Load Factor. However, equally important to characterising magnitude of domestic electricity demand is the timing of its use. The Time of Use (ToU) parameter tests for the occurrence of maximum electricity demand but does not test for other periods of use. The Autocorrelation function and Spectral Density Periodgram test these properties for the time series.

A period of one day, 1st July 2009, is chosen for two random customers to illustrate graphically typical characterising performance for both Fourier transforms and Gaussian processes and is shown in Fig. 6. Customer 1, shown on top in Fig. 6 shows both techniques replicating the time series within a reasonable degree of accuracy across the daily period. It is interesting to note that Gaussian processes were unable to sufficiently replicate the late peak at night at around 11 pm. In contrast, Customer 2, shown on the bottom of Fig. 6, shows a slightly different profile shape with three distinct short periods of electricity demand across the day. Gaussian processes almost identically replicate these peaks with Fourier transforms showing a smoother demand load profile at the same times. This shows the main difference between the two characterisation approaches already mentioned above. Due to their nature, Fourier transforms tend to be better at characterising profiles where electricity is consumed more evenly over a number of hours in the day. In contrast Gaussian processes are better at characterising short intervals of high electricity consumption (≤1 h) across the day. Therefore depending on the customer
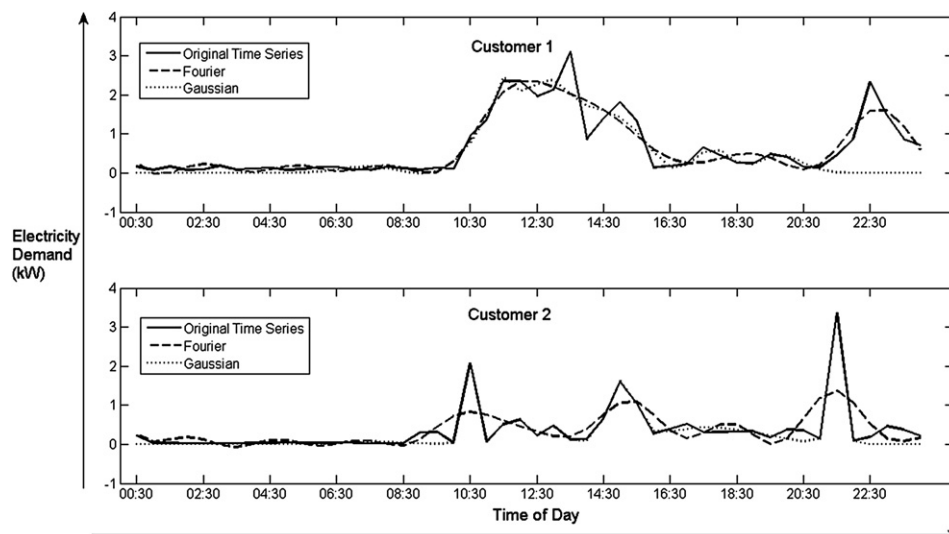

Fig. 6. Time series plot of daily electricity demand on the 1st July 2009 for two random customers.
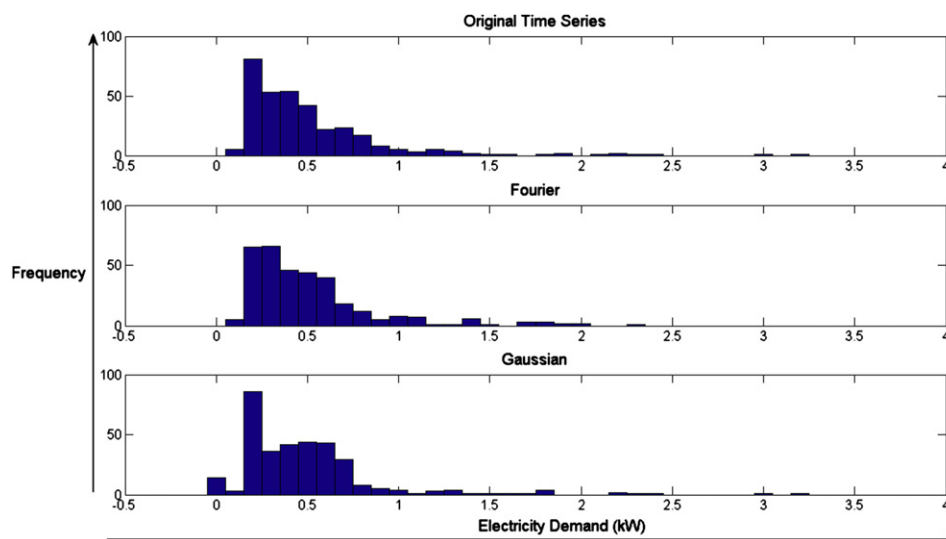
Fig. 7. Frequency histogram of electricity demand for a random customer between 1st and 7th July 2009.

electricity demand profile shape and end-user requirements each characterisation approach has its advantages.

Fig. 7 shows a frequency histogram over a weekly period for a single random customer. It is evident from Fig. 7 that Fourier transforms have difficulty replicating sharp high electricity peaks, as already discussed. However, aside from this both approaches replicate the magnitude component of the electricity load profile well. A disadvantage of both techniques is that they show negative values of electricity demand which is clearly an unrealistic situation where no on site generation exists. However, the frequency occurrence of negative values is small and where it does occur, is very low in magnitude.

In terms of temporal properties, both characterisation techniques performed well. This is illustrated in Fig. 8 with the auto-correlation function where both Fourier transforms and Gaussian processes follow the original data over the weekly period from 1st to 7th July 2009 for an individual random customer. The first autocorrelation coefficient is excluded in Fig. 8 as this represents perfect correlation when the time series is regressed onto itself

with a zero time lag. Subsequent coefficient values for the weekly period fall between ±0.4. A value of 1 represents perfect correlation of the time series, 0 indicates no correlation and −1 represents anti-correlation. As Fig. 8 shows, a highly cyclical pattern of electricity demand over a 24 h period is apparent. Both Fourier transforms and Gaussian processes were able to replicate this pattern, however, both approaches tended to either over estimate or underestimate the peak values.

Fig. 9 shows the Power Spectral Density Periodgram for 1st July 2009 for a random customer, as calculated by the Fast Fourier Transform. The figure illustrates that both Fourier transforms and Gaussian processes can represent the series in the frequency domain, thus confirming that the temporal properties are retained between the original and characterised times series. The high spectral density component near the origin is the daily period (1/(60 s × 30 min × 48 periods)) for the time series. The cyclical daily electricity demand pattern is typical of all dwellings with some customers having smaller frequency patterns throughout the day.
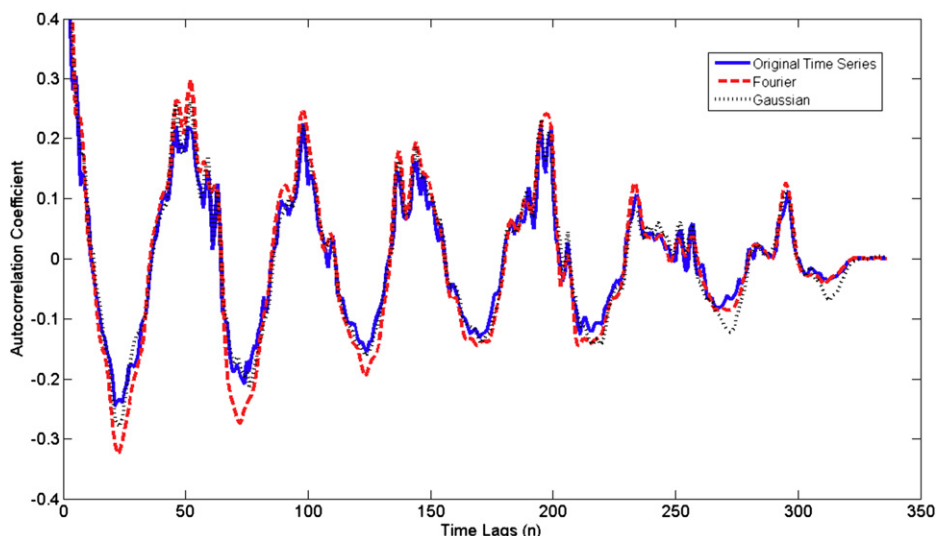

Fig. 8. Autocorrelation coefficients for electricity demand for a random customer between 1st and 7th July 2009.
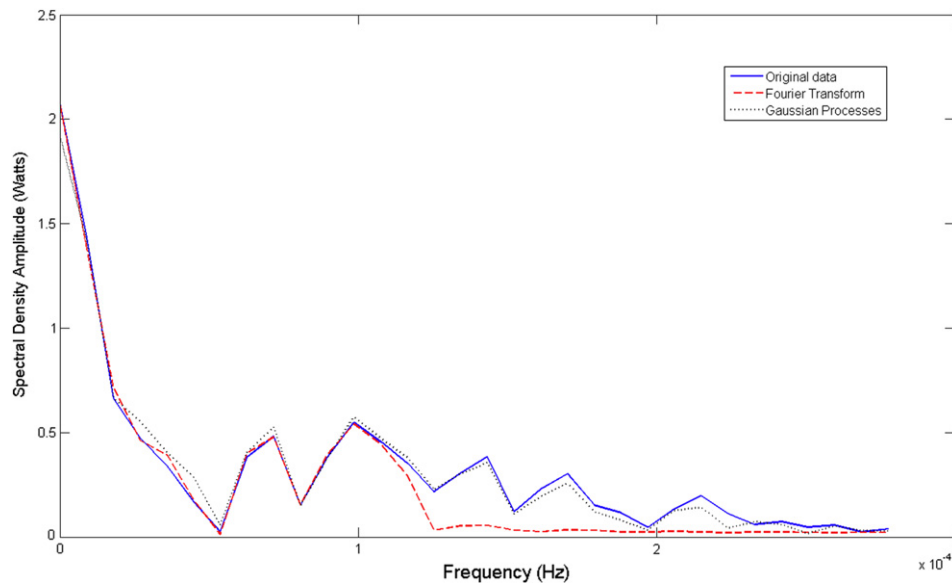
Fig. 9. Single-sided spectral density for electricity demand for a random customer on 1st July 2009.

Overall both Fourier transforms and Gaussian processes characterised the highly variable nature of domestic electricity demand profiles well. The most significant advantage to both these techniques is their ability to replicate the temporal characteristics of domestic electricity demand patterns with a minimum of half the number of variables compared to if each half hour period was characterised independently. In addition, both approaches have variable coefficients that are scalable which means they bear some resemblance to the magnitude and temporal characteristics of an electricity load profile. This effectively means that dwelling and occupant characteristics can be compared against variable coefficients for individual dwellings and their influence on the electricity load profile shape quantified, if required.

5. Conclusion

This paper discusses a number of time series approaches to electricity load profile characterisation at an individual dwelling level. The most promising time series techniques are evaluated and the results presented. This approach was taken so as to take advantage of the large body of literature that already exists in the area of aggregate electricity system demand load profiling but has yet to be applied at an individual dwelling level before.

Fourier transforms and Gaussian processes showed the greatest potential for characterising domestic electricity demand load profiles. Each technique was evaluated based on correlating the original and characterised time series as well as comparing certain electrical parameters. Descriptive statistics were presented for the parameters and a number of time series tests carried out. Furthermore, paired sample t-tests were also carried out in order to compare the differences between original and characterised electrical parameters.

Overall, Gaussian processes were found to be the more accurate of the two techniques when compared using mean Pearson's correlation coefficient. The two characterised time series were then compared using descriptive statistics and paired sample t-tests for four electrical parameters: Total Electricity Consumption, Maximum Demand, Load Factor and Time of Use. Total Electricity Consumption was successfully characterised using Fourier transforms, with very small percentage error (less than 1/10th of

a percent) and was found not to be significantly different at the 95% p-value level to that calculated from the original time series. Fourier transforms on the other hand were less successful at characterising Maximum Demand, where a mean percentage error of 28% was recorded. In contrast, Gaussian processes were more successful at replicating this parameter, with a much smaller percentage error of less than 5%, however, both were found to be significantly different than the original time series. Similar results were calculated for Load Factor for both Fourier transforms and Gaussian processes. Fourier transforms out-performed Gaussian processes marginally when calculating Time of Use of maximum electricity demand, with the former over-estimating the time (i.e. later in the day) and the latter underestimating (i.e. earlier in the day) however this is most likely skewed by the presence of a bi-modal distribution for the parameter.

In addition, a number of tests were also carried out and presented graphically to determine whether the characterised time series were temporally representative of the original data. Depending on the electricity demand pattern of the customer, each approach had individual strengths. Fourier transforms were better at characterising customers who consumed electricity more evenly across the day whereas Gaussian processes were superior at describing customers who consumed higher amounts of electricity over shorter time intervals. As a result Fourier transforms often underestimated Maximum Demand for individual customers. The autocorrelation and spectral properties remained between the original and characterised time series thus showing that the temporal properties remained for both techniques.

Acknowledgement

References

[1] Elexon. Load profiles and their use in electricity settlement. [Online]. Available from: http://www.elexon.co.uk/ELEXON Documents/load_profiles.pdf. [accessed 30.11.2011].

[2] McLoughlin F, Duffy A, Conlon M. Characterising domestic electricity consumption patterns by dwelling and occupant socio-economic variables: an Irish case study. Energy and Buildings Feb. 2012;2010.

[3] McLoughlin F, Duffy A, Conlon M. The generation of domestic electricity load profiles through Markov chain modelling. Euro-Asian Journal of Sustainable Energy and Development Policy 2010;3.

[4] Riddell AG, Manson K. Parametrisation of domestic load profiles. Applied Energy Jul. 1996;54(3):199–210.

[5] Moutter SP, Sc B, Be PSB, Ph D. Spectral decomposition and extrapolation of variations in electricity loading. IEEE Proceedings 1986;133(5):247–55.

[6] Gonzalez-Romera E, Jaramillo-Moran M, Carmona-Fernandez D. Monthly electric energy demand forecasting with neural networks and Fourier series. Energy Conversion and Management Nov. 2008;49(11):3135–42.

[7] Chen H, Canizares CA, Singh A. ANN-based short-term load forecasting in electricity markets. In: 2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194), 2001, vol. 2, pp. 411–415.

[8] Zadeh S, Masoumi A, Modelling residential electricity demand using neural network and econometrics approaches. In: 40th International Conference on Computers and Industrial Engineering (CIE), 2010, pp. 1–6, 2010.

[9] Ringwood JV, Bofelli D, Murray FT. Forecasting electricity demand on short, medium and long time scales using neural networks. Journal of Intelligent and Robotic Systems 2001;31:129–47.

[10] Aydinalp M, Ugursal VI, Fung AS. Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks. Applied Energy 2002;71:87–110.

[11] Aydinalp M. Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks. Applied Energy Oct. 2004;79(2):159–78.

[12] Aydinalp M, Ugursal V. Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector. Applied Energy Apr. 2008;85(4):271–96.

[13] Singh R, Pal BC, Jabr RA. Statistical representation of distribution system loads using Gaussian mixture model. IEEE Transactions on Power Systems 2010;25(1):29–37.

[14] D. J. Leith, M. Heidl, J. V. Ringwood, and S. Member, "Gaussian process Prior models for electrical load forecasting," 8th International Conference on Probabilistic Methods Applied to Power Systems, 2004.

[15] Lourenco J, Santos P. Short term load forecasting using Gaussian process models. Coimbra: Instituto de Engenharia de Sistemas e Comput adores de; 2010. pp. 1–12.

[16] Amaral L, Souza R, Stevenson M. A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting. International Journal of Forecasting Oct. 2008;24(4):603–15.

[17] Gabreyohannes E. A nonlinear approach to modelling the residential electricity consumption in Ethiopia. Energy Economics May 2010;32(3):515–23.

[18] Magnano L, Boland J. Generation of synthetic sequences of electricity demand: application in South Australia. Energy Nov. 2007;32(11):2230–43.

[19] Pappas S, Ekonomou L, Karamousantas D, Chatzarakis G, Katsikas S, Liatsis P. "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models. Energy Sep. 2008;33(9):1353–60.

[20] Ardakanian O, Keshav S, Rosenberg C. Markovian models for home electricity consumption. GreenNet Conference 2011, pp. 31–36, 2011.

[21] Hsu YY, Ho KL. Fuzzy expert systems: an application to short-term load forecasting. IEEE Proceedings-c 1992;139(6).

[22] Mastorocostas Pa, Theocharis JB, Bakirtzis a G. Fuzzy modeling for short term load forecasting using the orthogonal least squares method. IEEE Transactions on Power Systems 1999;14(1):29–36.

[23] Mori H, Kobayashi N. Optimal fuzzy inference for short-term load forecasting. Power 1996;11(1):390–6.

[24] Xu H, Member S, Niimura T. Short-term electricity price modeling and forecasting using wavelets and multivariate time series. In: Power Systems Conference and Exposition, pp. 1–5, 2004.

[25] Chen Y, Luh P, Guan C, Zhao Y, Michel L, Coolbeth M, et al. Short-Term Load forecasting: similar day-based wavelet neural networks. IEEE Transactions on Power Systems 2010;25(1):3353–8.

[26] Pahasa J, Theera-Umpon N. Short-term load forecasting using wavelet transform and support vector machines. In: International Power Engineering Conference, IPEC 2007, pp. 47–52, 2007.

[27] Nguyen HT, Nabney IT. Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models. Energy 2010;35(9):3674–85.

[28] McSharry PE, Bouwman S, Bloemhof G. Probabilistic forecasts of the magnitude and timing of peak electricity demand. IEEE Transactions on Power Systems May 2005;20(2):1166–72.

[29] Chicco ECG. Probabilistic characterisation of the aggregated residential load patterns. IET Generation, Transmission and Distribution 2008;2(3):373–82.

[30] Amarawickrama H, Hunt L. Electricity demand for Sri Lanka: a time series analysis. Energy May 2008;33(5):724–39.

[31] Bianco V, Manca O, Nardini S. Electricity consumption forecasting in Italy using linear regression models. Energy Sep. 2009;34(9):1413–21.

[32] Heunis SW, Herman R. Consumer loads. IEEE Transactions on Power Systems 2002;17(3):621–5.

[33] Cagni A, Carpaneto E, Chicco G, Napoli R, Elettrica I, Torino P. Characterisation of the aggregated load patterns for extra-urban residential customer groups. IEEE Melecon 2004;2004:1–4.

[34] Capasso A, Invernizzi A, Lamedica R, Prudenzi A. Probabilistic processing of survey collected data in a residential load area for hourly demand profile estimation. In: IEEE/NTUA Athens Power Tech Conference, pp. 866–870, 1993.

[35] McQueen DHO, Hyland PR, Watson SJ. Monte Carlo simulation of residential electricity demand for forecasting maximum demand on distribution networks. IEEE Transactions on Power Systems Aug. 2004;19(3):1685–9.

[36] Jardini JA, Tahan CMV, Gouvea MR, Ahn SU, Figueiredo FM, Member S. Daily load profiles for residential, commercial and industrial low voltage consumers. IEEE Transactions on Power Delivery 2000;15(1):375–80.

[37] Chen CS, Member S, Kang MS, Hwang JC, Huang CW. Temperature effect to distribution system load profiles and feeder losses. IEEE Transactions on Power Systems 2001;16(4):916–21.

[38] Parti M, Parti C. The total and appliance-specific conditional demand for electricity in the household sector. The Bell Journal of Economics 1980;11(1):309.

[39] Graps A. An introduction to wavelets. IEEE Computational Science and Engineering Summer 1995;2(2) [Online]. Available from: http://www.amara.com/IEEEwave/IEEEwavelet.html [accessed: 30.11.2011].