2013-2

Characterising Domestic Electricity Demand for Customer Load Profile Segmentation

Fintan McLoughlin
Technological University Dublin, fintan.mcloughlin@tudublin.ie

Characterising Domestic Electricity Demand for Customer Load Profile Segmentation

A thesis submitted to Dublin Institute of Technology in fulfilment of the requirements for the degree of Doctor of Philosophy

By

Fintan McLoughlin (BSc. Eng (Hons), MSc.)
School of Civil and Building Services Engineering
Dublin Institute of Technology
Bolton Street, Dublin 1, Ireland

February 2013

Supervisors:

Dr. Aidan Duffy (School of Civil and Building Services Engineering, Dublin Institute of Technology)

Dr. Michael Conlon (School of Electrical Engineering Systems, Dublin Institute of Technology)

Declaration

I certify that this thesis which I now submit for examination for the award of PhD, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.
This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for another award in any Institute.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for ethics in research.

The Institute has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature ___________________________________ Date _______________

Candidate

Abstract

The aim of this research was to characterise domestic electricity patterns of use on a diurnal, intra-daily and seasonal basis as a function of customer characteristics. This was done in order to produce a library of representative electricity demand load profiles that are characteristic of how households consume electricity. In so doing, a household's electricity demand can be completely characterised based solely on their individual customer characteristics.

A number of different approaches were investigated as to their ability to characterise domestic electricity use. A statistical regression approach was evaluated which had the advantage of identifying key dwelling, occupant and appliance characteristics that influence electricity use within the home. An autoregressive Markov chain method was applied which proved to be effective at characterising the magnitude component to electricity use within the home but failed to adequately characterise the temporal properties sufficiently. Further time series techniques were investigated: Fourier transforms, Gaussian processes, Neural networks, Fuzzy logic, and Wavelets, with the former two being evaluated fully. Each method provided disparate results but proved to be complimentary to each other in terms of their ability to characterise different patterns of electricity use. Both approaches were able to sufficiently characterise the temporal characteristics satisfactorily, however, were unable to adequately associate customer characteristics to the load profile shape.

Finally clustering based approaches such as: k-means, k-medoid and Self Organising Maps (SOM) were investigated. SOM showed the greatest potential and when

combined with statistical and regression techniques proved to be an effective way to completely characterise electricity use within the home and their associated customer characteristics. A library of domestic electricity demand load profiles representing common patterns of electricity use on a diurnal, intra-daily and seasonal basis within the home in Ireland and their associated household characteristics are then finally presented.

Acknowledgements

This research was carried out at the Dublin Energy Lab, Dublin Institute of Technology, Ireland. It was funded under the Department of Education and Science *Strand III* and Dublin Institute of Technology *Fiosraigh* scholarship programmes.

I would like to express my sincere gratitude to my supervisors Dr. Aidan Duffy and Dr. Michael Conlon for their consistent guidance and invaluable contributions which they have made throughout the course of my research. I would also like to thank the head of engineering research Dr. Marek Rebow for his assistance throughout and the head of school John Turner. In addition I would also like to express my thanks to Electric Ireland, particularly Brian Dowd and Frank Burke, for access to data and support over the term of my research.

I am sincerely thankful to my mother, Sheila, family and friends for their continued love and support without which my research would not have been possible. I would also like to thank my colleagues at the Dublin Energy Lab: Dr. L. Ayompe, Dr. M. Claudy, D. Reilly, R. Oliver, A. Famuyibo, B. Cleary and D. Wolff for their support over the years.

This work is dedicated to my late father, Seamus, whose honesty and integrity has been a constant role model throughout my life.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AMI	Advanced Metering Infrastructure
AR	Autoregression
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
APPT	Average Percentage Profile Time
BMU	Best Matching Unit
BER	Building Energy Rating
CDA	Conditional Demand Analysis
CER	Commission for Energy Regulation
CFL	Compact Fluorescent Lights
CIE	Chief Income Earner
CPI	Customer Profile Index
CO_2	Carbon Dioxide
DB	Davies Bouldin
DFT	Discrete Fourier Transform
DOC	Dwelling and Occupant Characteristics
DST	Daylight Saving Time
EA	Electrical Appliance
ESB	Electricity Supply Board
ETS	Emissions Trading Scheme
EU	European Union
FFT	Fast Fourier Transform
GDP	Gross Domestic Product
HBS	Household Budget Survey
HH	Household
HMode	Household Mode
HoH	Head of Household
ISSDA	Irish Social Science Data Archive
MA	Moving Average
NRS	National Readership Survey
NSHQ	Irish National Survey of Housing Quality
PAM	Partition Around Medoids

PSD	Power Spectral Density
RMDS	Retail Market Design Service
SETAR	Self Exciting Threshold Autoregressive
SOM	Self Organising Maps
STR	Smooth Transition Regression
STPAR	Smooth Transition Periodic Autoregression
ToU	Time of Use
TPER	Total Primary Energy Requirement
TSO	Transmission System Operator
UEC	Unit Energy Consumption
UNFCCC	United Nations Framework Convention on Climate Change

Nomenclature

a_i	Gaussian process amplitude coefficient
$APPT_{HMode}$	Average percentage time each household spends within each profile group (Chapter 8)
a_r	Fourier transform magnitude '*cos*' coefficient
a_0	Fourier transform constant coefficient
b_i	Gaussian process location (centroid) coefficient
b_r	Fourier transform magnitude '*sin*' coefficient
c_i	Gaussian process shape coefficient (peak width)
$CPI_{cust,\ day}$	Clustering matrix identifying a particular Profile group for individual dwelling *'cust'* on a particular *day* (Chapter 8)
DB index	Davies-Bouldin clustering index
E_i^j	Average electrical demand for each half hour period *i* and day number *j* in kilowatts
E_{LF}	Mean daily load factor (%)
E_{MD}	Mean daily maximum demand in kilowatts
E_{Total}	Total amount of electricity used during a period in kilowatt hours
E_{ToU}	Mean daily Time of Use of maximum electricity demand
Exp (β)	Odds or likelihood ratio for logistic regression
$HMode_{cust}$	The Profile group household's *'cust'* frequented most of the time during a period (Chapter 8)
$P_{1,1}$	Probability of going from one discrete state to the next used in Markov chain's (Chapter 6)
t	t-statistic from student's t-test
X_n	Explanatory variable pertaining to individual dwelling, occupant and appliance characteristics in multivariate and logistic regression

Greek symbols

α	Probability distribution shape parameter
β	Coefficient value calculated by multivariate regression
η	Probability distribution scale parameter

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

1.1 Overview

Throughout the EU, there has been a move towards smarter electricity networks, where increased control over electricity generation and consumption has been achieved with improvements in new technologies such as Advanced Metering Infrastructure (AMI). Smart metering is part of this and is seen as a necessary component to achieve EU energy policy goals 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 [1].

Advances in metering, data management and information services as well as the regulatory environment over the past decade has meant that smart metering programmes are more prolific of late, especially for the residential sector [2]. These advances, in combination with the mandate for European countries to collectively meet EU 20/20/20 targets has encouraged interest in the area, with policy makers and energy suppliers willing to support smart metering programmes [3]. As a result, a wealth of new information now exists giving detailed electricity consumption data for large sample sizes in the residential sector [4].

Also, advances in generation and storage technologies such as microgeneration and electric vehicles has meant new opportunities now exist for changing how energy is produced and consumed within 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.

In July 2009, the Commission for Energy Regulation (CER), in collaboration with the largest Irish electricity supplier – Electric Ireland (formally Electricity Supply Board) - commenced a smart metering trial for the residential sector and small-to-medium enterprises [5]. The trial was conducted between 2009 and 2010 and consisted of installing smart meters in over 5,000 residential dwellings in Ireland. Electricity demand at half hourly intervals as well as information on dwelling and occupant characteristics for a representative sample of dwellings in Ireland was recorded [5].

The dataset was kindly provided to the author for this research by Electric Ireland. Subsequently, it has been made publically available from the Irish Social Science Data Archive (ISSDA) [6]. The collection of such a detailed amount of data in the residential sector for such a large sample size, offers a unique opportunity to investigate the manner with which electricity is consumed in the home and the significant factors behind its use.

1.2 Problem Definition

Electricity is a unique form of energy such that its supply and use need to occur at the same time. As a result, a large amount of research has been focussed on characterising and forecasting electrical demand at a system demand level, in order to balance supply and demand [7]. Various mathematical techniques have been used to do this, each with their own strengths and weaknesses [8][9]. Recently, the availability of detailed electricity consumption for the domestic sector also means that it is now possible to apply these techniques to characterise individual dwelling electricity demand [6].

However, patterns of electricity use at a system demand level and at an individual dwelling level are very different. Figure 1.1 shows a typical system demand load profile for the Irish Transmission System Operator (TSO), Eirgrid, on the 1st July 2009 over a twenty four hour period [10]. The figure shows a smooth profile shape with relatively small amount of electricity consumption over the night time, a clearly defined peak in the morning time and a smaller defined peak in the evening time. Although not shown below, the profile shape changes slightly for different days of the week and over the course of the year due to fluctuating working patterns and seasonality respectively.

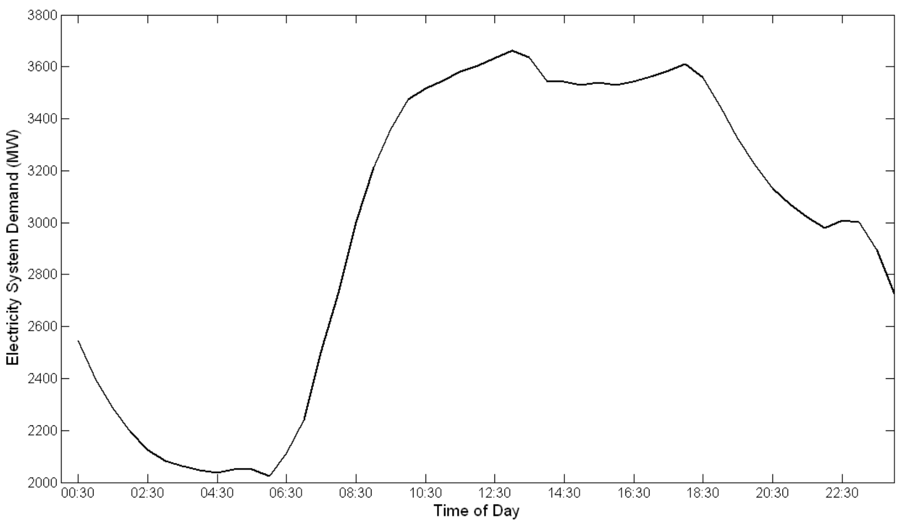


Figure 1.1: Daily electricity system demand load profile across a 24hr period on 1st July 2009 [10]

The system demand load profile shown in Figure 1.1 is classed as a diversified load. What this means is that as each individual user connected to the grid consumes electricity at different times of the day, the combined effect of a large number of households is in fact an averaging process. In contrast, Figure 1.2 shows a distinctly different pattern of electricity use from the smart metering dataset for a single random dwelling on the same day of the year [6]. The profile shows a peak in the late morning around 10am which lasts until 4pm in the evening and a later peak at 10.30pm that night. This pattern of electricity use across the day is very different to that of the system demand load profile, where a more sporadic use is apparent rather than a gradual smooth profile shape as one would anticipate.

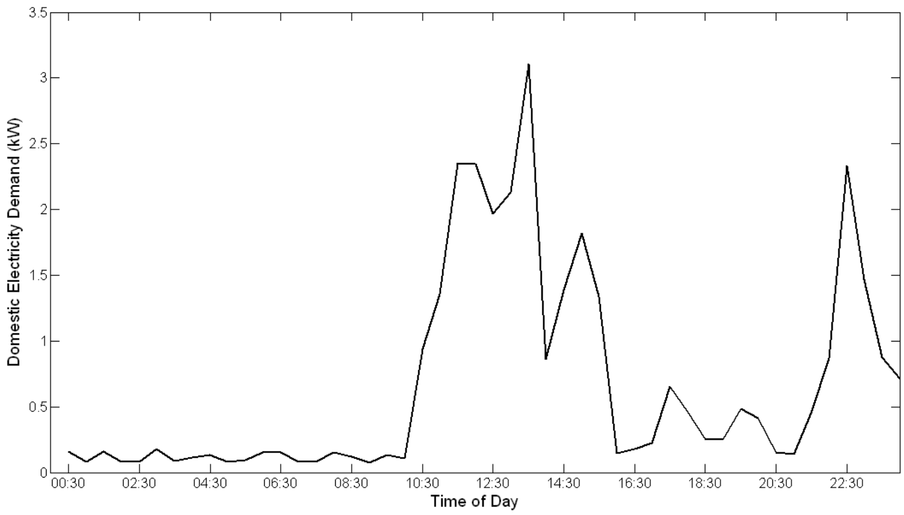


Figure 1.2: Daily electricity demand load profile for an individual dwelling across a 24hr period on 1st July 2009 [6]

Figure 1.3 shows a standard load profile issued by the Retail Market Design Service (RMDS) for the Irish domestic electricity market on the same day as above [11]. Standard domestic load profiles are used for the purposes of settlement between suppliers in the electricity market. They are normalised, with the summation of each interval across a day (96 intervals of fifteen minute periods) and for each day of the year summing to one. The methodology used to characterise the profiles is based on a regression on various parameters across a representative sample of domestic customers [12].

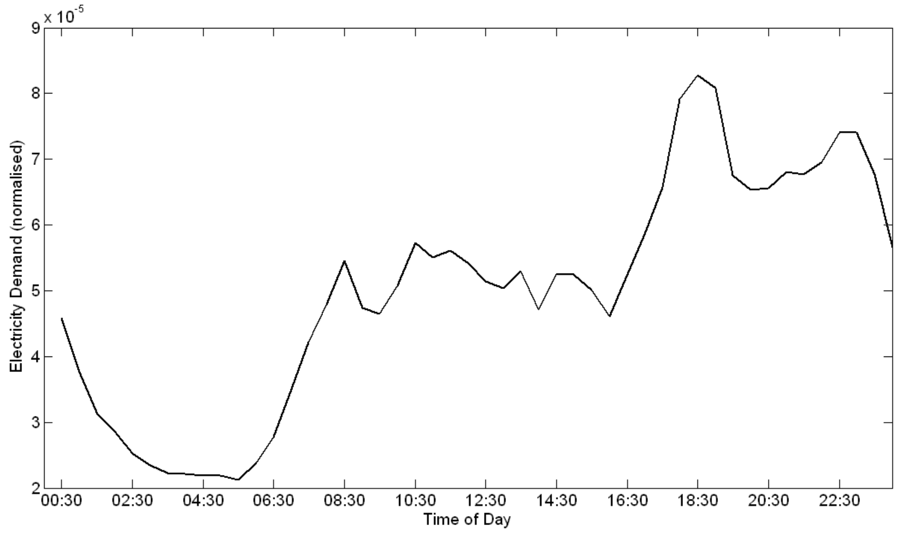


Figure 1.3: Electricity standard load profile for urban domestic across a 24hr period on 1st July 2009 [11]

Domestic standard load profiles, like that shown in Figure 1.3, reflect average electricity consumption for all households across a 24 hour period. They can be considered to be deterministic in nature and can be in part explained by three key events. Firstly, very little electricity is used over the night time period, a result of little or no activity within the household while occupants are sleeping. Secondly, there is an increase in electricity demand in the morning time as household occupants awaken and start to use electrical appliances. Finally, as people come home from work and start to cook the dinner there is a further increase in electricity demand before dropping off as occupants return to bed. This is characteristic of how individual households on average consume electricity across the day.

However, in practice electricity is consumed far more stochastically across a 24 hour period as was shown in Figure 1.2. There are similar characteristic deterministic patterns to that of the standard load profile, but these often change on a daily basis and between households. In this manner, a domestic electricity demand load profile can be thought of as a combination of both deterministic and stochastic processes. Therefore, it is apparent that the standard load profiles used by electricity suppliers to characterise domestic households is not an accurate reflection of how individual dwellings consume electricity and merely reflects a highly averaged usage pattern for all customers more in common with that of a system demand load profile shown in Figure 1.1.

As discussed, typical domestic electricity load profiles are far more variable than that shown in Figure 1.3 and can vary greatly in the time (on a day to day basis) and space domains (between customers). Figure 1.4 shows a single random household from the dataset [6] over a weekly period from 01st – 07th 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.

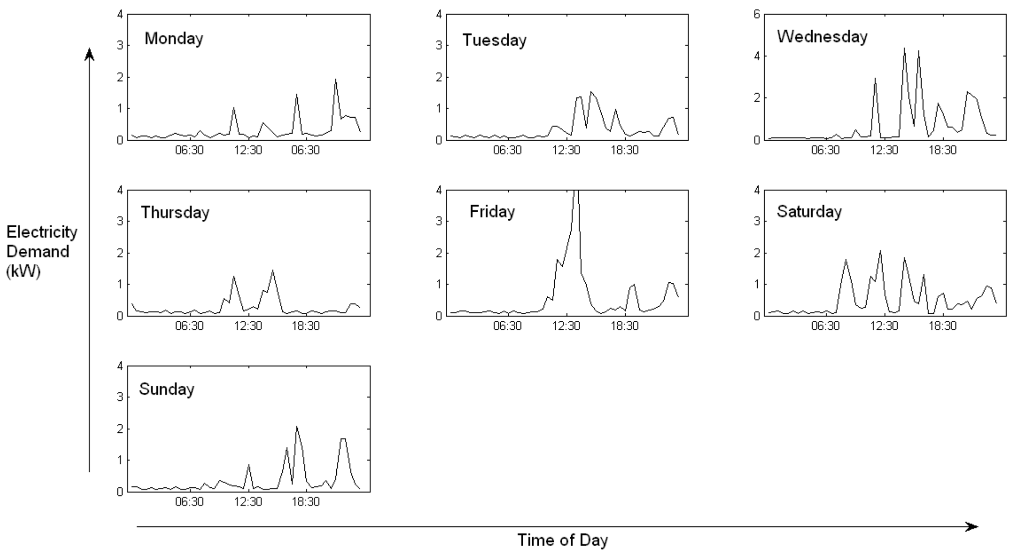


Figure 1.4: Daily electricity load profiles for a single randomly chosen household over a weekly period showing intra-daily variations [6]

Similarly, the profile shape can change significantly between households. Figure 1.5 shows nine different customer profiles from the dataset at random for the 1st July 2009 [6]. The figure shows how the profile can change considerably between households in both magnitude and Time of Use (ToU) of electricity demand.

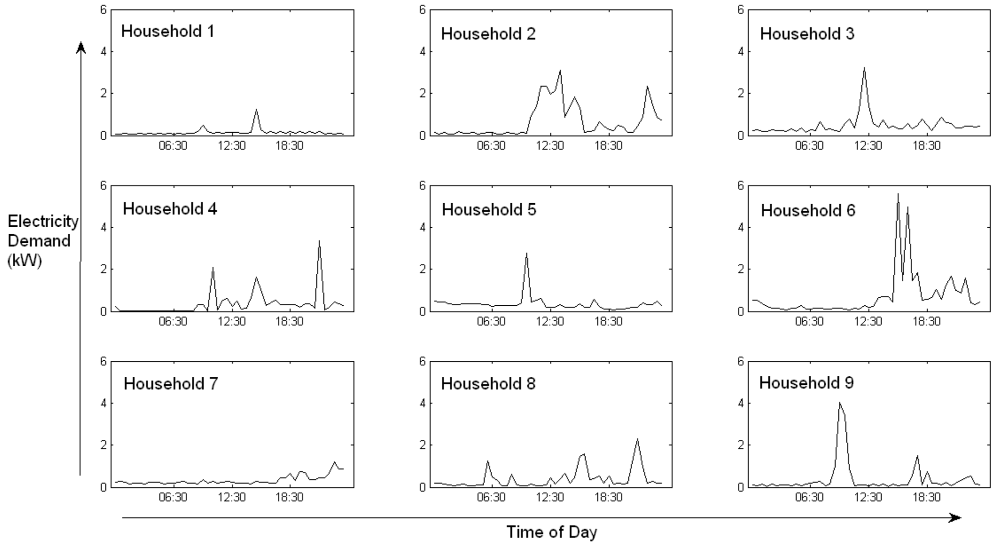


Figure 1.5: Daily electricity load profiles for nine randomly chosen households, illustrating variation between households [6]

A seasonality component also exists within a domestic electricity demand load profile, mainly as a result of changes in external temperature and daylight hours for heating (albeit small in Ireland due to limited penetration of electric heating – see Figure 3.5) and lighting homes respectively. This trend is shown in Figure 1.6 where half hourly periods for a random individual household from the dataset [6] are plotted across the year (01st July 2009 to 30th June 2010). When a trend line (quadratic polynomial function) is fitted to the data it shows an increase in electricity demand during the winter period of approximately 200 Watts compared to the summer time. Figure 1.6 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.

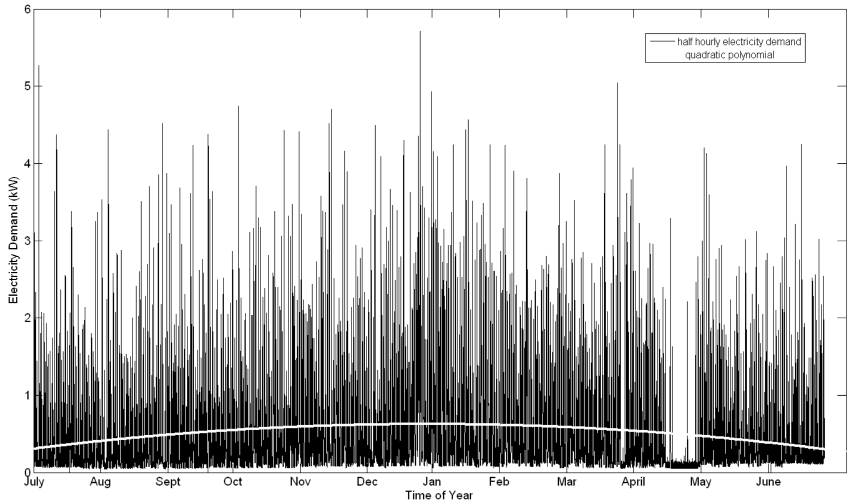


Figure 1.6: Electricity demand load profile for a single randomly chosen household over a yearly period [6]

Therefore, in order to characterise individual domestic electricity demand load profiles effectively an approach needs to consider the following key factors:

- diurnal variations in electricity demand;
- intra-daily variations in electricity demand (i.e. day of the week);
- seasonal electricity demand effects ; and
- electricity demand variations between households

1.3 Data Structure

In order to identify methods to characterise domestic electricity load profiles it is important to first define the data structure. Figure 1.7 shows how electricity demand can be broadly categorised into four groups based on the period of data collection and the level at which it is collected. The x-axis indicates the time interval at which electricity demand is collected (i.e. small time interval refers to ≤ 1 hour where as large time interval can refer to anywhere from one day to a year) and the y-axis indicates the level at which the data is collected (i.e. at an individual dwelling or at an aggregate system demand level).

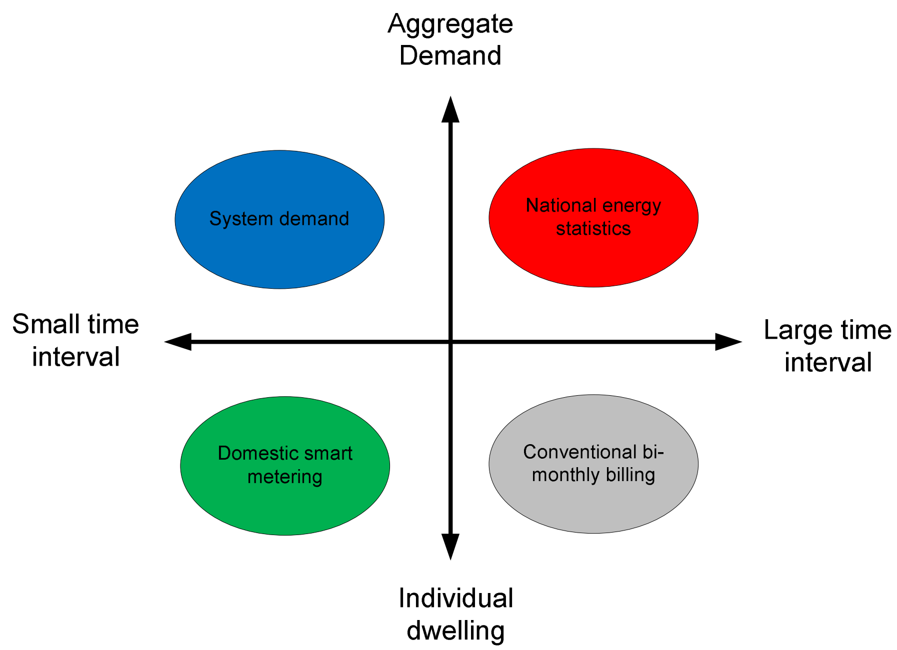


Figure 1.7: Taxonomy of data structure relating to electricity demand load profiling

For large time intervals, data collection for the domestic sector has historically relied upon manual meter readings taken every two months or more. This is then often aggregated together and/or combined with other data sources to provide electricity consumption statistics at a national level for the sector. Data collected at these time intervals obviously does not allow for the level of characterisation discussed above. However, although detail is lost by collecting data in this manner, it provides a relatively straightforward method of determining the factors that influence electricity consumption in the home (such as dwelling and occupant characteristics) due to the aggregated nature of the demand.

For smaller time intervals, studies involving metering individual households have usually been limited to small sample sizes due to the prohibitive cost of installing AMI [13]. This often results in electricity load profiles that are not representative and do not reflect common patterns of electricity use within the home. In recent years, smart meters have become more prevalent in the residential sector, providing large amounts of data at intervals of less than one hour. Historically, this period of data collection has been limited to other sectors such as commercial, industrial and at an electricity system demand level. In particular, there is a significant amount of literature and characterisation approaches that have been applied at an electricity system demand level as will be shown in the next Chapter. A large proportion of these methods have yet to be applied at an individual dwelling level. However, it must be stressed that patterns of electricity use at a system demand level and at an individual dwelling level are very different as was shown in Figure 1.1 and Figure 1.2.

1.4 Motivation

In the past, characterisation of domestic electricity demand at small time intervals has been limited to small sample sizes and hence often cannot be considered to be representative [13][14]. However, this has changed in recent years with large smart metering programmes being rolled out in most European countries [3]. Conversely where large samples sizes do exist, the approaches used to characterise domestic electricity demand often result in highly averaged load profiles, like that shown in Figure 1.3.

Therefore there is a need for approaches to be able to characterise domestic electricity demand based on the following criteria:

- describe household electricity demand on a diurnal, intra-daily and seasonal basis at small time intervals (i.e. at half-hourly time intervals or less);
- be representative of the national housing stock and population for a country or region;
- be characteristic of the way with which dwelling occupants consume electricity; and
- can be linked to household characteristics by one or more variables (i.e. dwelling type, head of household age, etc) in order to describe the factors influencing electricity use across the day within the home

The smart metering trial carried out by CER has provided the necessary data to carry out such a task. However, gathering data for individual households at such small time intervals also produces its own difficulties. Electricity suppliers are now being faced with a data tsunami where detailed information needs to be collected and stored efficiently. In addition to this, methods of extracting useful information from the raw data need to be found in order to help condense and present it in a meaningful way, thus making full use of the richness of the data source.

This research presents new methods for characterising individual domestic electricity demand and provides a number of representative load profile groups based on patterns of electricity use and common dwelling and occupant characteristics. A profile group can then be assigned to a particular household without any prior knowledge of their electricity consumption and solely based on their household characteristics. The electricity load profile groups can then be used to investigate various scenarios, some of which are described below:

1.4.1 Planning and Forecasting

Over the last decade, electricity markets are becoming increasingly competitive, mainly due to the liberalisation of the sector across the EU. As a result, new entrants are joining the market, eager to gain market share. In order to compete in such an environment, utilities need to have a better understanding of their customers to gain a competitive advantage over other market participants. In addition, detailed knowledge of their customer base will allow utilities to shape market strategies for their business and plan for future growth on their network.

Understanding customer electricity consumption and how this relates to dwelling and occupant characteristics can result in more direct marketing strategies for a electricity suppliers. For example, there are certain electricity customers that are more profitable than others and these can be targeted in order to maximise revenue.

1.4.2 New Technologies

The introduction of new technologies such as micro-generation and electric vehicles onto European networks is beginning to gather pace, as most member states try to reduce greenhouse gas emissions in order to meet EU 20/20/20 targets. In order to fully assess these from a technological, economic and environmental perspective a detailed understanding of individual customers demand is required. This research provides the means to carry out this by presenting a number of electricity load profiles that are representative in the manner with which homeowners consume electricity and therefore can be used to fully assess the performance of these new technologies.

1.4.3 Tariff Structure

Currently, electricity prices for the residential sector are constant across the day and therefore do not accurately reflect the actual cost of electricity generation at different times of the day [15]. This means that there is no incentive for dwelling occupants to consume electricity at off peak times when the cost of generation is more efficient for the supplier. Currently there are only four tariffs on offer in Ireland for the domestic sector, 24 hour and Nightsaver for both urban and rural customers [15]. A range of new pricing plans is expected in the future when a national roll out of smart meters occurs over the next 4 – 7 years. This will enable electricity suppliers to offer ToU tariffs at a domestic level in order that the true cost of generation across the day can be reflected within the price.

However, in order to build appropriate tariff structures knowledge of customers' electricity demand is required. Applying the profiles presented in this research will enable new pricing plans to be developed that are appropriate to certain types of customer. Designing new tariffs for domestic customers could potentially achieve savings for both the electricity supplier and the customer alike. The customer will benefit by choosing an electricity tariff that is suited to their lifestyle while at the same time the supplier will be able to diversify their customer base. This should allow electricity suppliers to purchase electricity more efficiently on the wholesale market by smoothing out demand and supply across the day.

1.4.4 Environment and Sustainability

Reducing Greenhouse gas emissions, of which Carbon Dioxide (CO_2) is the largest contributor, is a driving force within EU energy policy [16]. As a result a number of different policies are being implemented throughout member states within Europe (e.g. micro generation, electric vehicles, demand side management) that are changing how electricity has traditionally been consumed in the home [17]. In order to fully assess the impact of such policies to offset national CO_2 emissions a detailed understanding of domestic electricity consumption is required.

1.5 Aims and Objectives

Current approaches used to characterise domestic electricity consumption generally lead to a highly averaged load profile shape for all households like that shown in Figure 1.3 which is not an accurate reflection of how individual households consume electricity across the day. This research aims to characterise the different patterns of electricity use within the home and relate this to dwelling and household characteristics by:

- developing a methodology for characterising individual domestic electricity demand load profiles on a diurnal, intra-daily and seasonal basis;
- identifying a number of different electricity demand load profiles that are representative of how electricity is consumed within the home in Ireland;
- associate the electricity demand load profiles to dwelling and occupant characteristics so they can be used to test various scenarios such as those outlined in Section 1.4.

A number of different approaches are presented in this research such as statistical, autoregressive (Markov chain), time series (Fourier Transforms and Gaussian Processes) and clustering (Neural network - Self Organising Maps). Each method is applied to the dataset in hand and discussed in terms of its relative strengths and weaknesses. A series of electricity demand load profiles are then presented which reflect common patterns of electricity use within the home that are representative of the domestic building stock in Ireland.

1.6 Thesis Layout

The thesis is composed of nine main components which are as follows:

- Review of Literature (Chapter 2)
- Description of the Smart Metering Dataset (Chapter 3)
- Presentation of overall methodologies for domestic electricity load profile characterisation (Chapter 4)

- Application of statistical analysis and regression to electrical parameters for electricity load profile characterisation (Chapter 5)
- Application of autoregressive - Markov chains for electricity load profile characterisation (Chapter 6)
- Application of time series techniques (Fourier transforms and Gaussian processes) for electricity load profile characterisation (Chapter 7)
- Application of clustering techniques (Self Organising Maps) for electricity load profile characterisation (Chapter 8)
- Review of findings and recommendations (Chapter 9)

Chapter 2 contains a description of the literature available in the area. This chapter is split into four sections, detailing the different approaches to electricity load profile characterisation. First of all, characterisation approaches that have either been applied at an aggregate level or over large time intervals to domestic electricity consumption are discussed. The second section then introduces approaches that have used data collected at an individual dwelling level and for small time intervals. These mainly consist of engineering and statistical approaches that have been used to characterise domestic electricity consumption but traditionally have been limited to small sample sizes. These two sections together form a large percentage of the literature in the area to date. The next section discusses methods that have mainly been applied at an aggregate level to characterise electricity system demand at small time intervals, which mainly consist of time series approaches. Lastly, clustering methods to electricity load profile characterisation are presented which have been applied at an individual level and for small time intervals but mostly in the commercial and industrial sectors.

Chapter 3 provides a complete overview of the smart metering dataset used throughout the research. The dataset is split into three main categories: dwelling characteristics, occupant characteristics and appliance characteristics.

Chapter 4 gives a methodological overview of the how the approaches in the following chapters are applied. It introduces four electrical parameters and a number of time series tests that are used throughout the research to characterise and validate each characterisation approach. The chapter also describes the methods of regression which will be used to associate dwelling and occupant characteristics to electricity consumption in the home. Finally the individual methodologies that will be applied in each chapter are presented.

Chapter 5 presents a statistical and regression approach to characterising domestic electricity consumption. Four electrical parameters described in Chapter 4: total electricity consumption, maximum demand, load factor and ToU of maximum electricity demand are used to parameterise the dataset. These parameters are then linked to dwelling and occupant characteristics through multivariate regression.

Chapter 6 describes an autoregressive approach to characterising domestic electricity demand load profiles. A Markov chain process is presented and used to characterise electricity demand load profiles for four individual dwelling types chosen at random. A number of statistical and time series tests are performed on the characterised profiles in order to assess the accuracy of the characterisation process, particularly investigating the temporal properties.

Chapter 7 discusses a number of time series techniques, more often applied to characterise electricity system demand. In particular two techniques: Fourier transforms and Gaussian processes are used to characterise domestic electricity demand. The same electrical parameters and statistical tests used in the preceding chapters are evaluated and used to compare the accuracy of the characterisation processes.

Chapter 8 applies clustering methods to characterise individual households' electricity demand. A number of techniques are discussed such as: k-means, k-medoid and Self Organising Maps (SOM). SOM showed the greatest potential for domestic electricity load profile characterisation and are therefore evaluated further. A characterisation methodology is then applied to produce a series of representative electricity load profile groups for the domestic sector. Each profile group is presented and shows common patterns of electricity use within the home across the day. Finally a multi-nominal logistic regression is applied to each profile group in order to determine the dwelling and occupant characteristics that most likely describe each electricity load profile. Descriptive statistics are also presented in order to graphically present the results.

Chapter 9 provides final conclusions for the research presented and further recommendations for future work in the area.

1.7 Contribution to Knowledge

The contribution to knowledge for research in the area can be summarised as follows:

- The characterisation of domestic electricity demand in terms of four key electrical parameters, which when combined with statistical and time series tests are also used for validation purposes throughout the research.

- The first time application of various time series techniques such as Auto-regression (Markov chain), Fourier transforms and Gaussian processes to characterise domestic electricity demand.

- The application of statistical and probabilistic techniques to infer relationships between different patterns of electricity use within the home and dwelling, occupant and appliance characteristics.

- Through the application of clustering algorithms and regression, a library of representative electricity demand load profiles were produced that reflect common patterns of electricity use within the home and their associated household characteristics. In this way, a household and the manner with which they use electricity within the home can be identified based solely on their individual characteristics, without any prior knowledge as to how they may have consumed electricity in the past.

CHAPTER 2

CHARACTERISING DOMESTIC ELECTRICITY DEMAND

2 CHARACTERISING DOMESTIC ELECTRICITY DEMAND

2.1 Introduction

Over the past 5-10 years, electricity grids throughout the world have been going through a period of significant change. Smart grids are changing the way electricity has traditionally been generated, supplied and consumed. Part of the 'smartening' of electricity grid infrastructure, is the collection of large amounts of data that up until now had previously not existed. Smart meters provide such information, which deliver near real time electricity demand for individual dwellings, as well as other valuable pieces of electrical data such as voltage levels, power quality, etc.

However, up until recently, energy utilities relied on manual electricity readings for the most part, which varied in frequency anywhere from monthly to every six months. This is a dramatic shift in the period of collection for domestic electricity consumption and this is reflected in the literature to date [18]. Combining this information with new and existing data sources which describe individual dwelling and occupant characteristics has meant that novel approaches to characterising domestic electricity demand patterns are now possible.

In the past, much attention in the area has focussed on modelling domestic electricity demand [19][20][21]. However, modelling individual dwelling electricity demand is a complicated task, not least as it can in part be influenced by the physiological and behavioural decisions of the dwelling occupants [22]. The ability to predict electricity demand in advance at an individual dwelling level is also questionable, especially when it is considered to be highly variable in nature. However, despite this a number of approaches have been used to predict or simulate electricity demand in the home, mainly based on occupancy patterns and appliance holdings [23][24][25]. In contrast, electricity characterisation is a process which describes its use rather than attempting to predict its behaviour over time. Characterisation explains the manner with which electricity is used in the home and relates this to dwelling, occupant and environmental characteristics. In this way, a picture is built of a particular household and the manner with which they consume electricity.

A number of characterisation approaches have been used in the past; however, much of the literature has been focussed on small sample sizes [20][14][26]. An averaging effect of the electricity load profile shape also occurs throughout much of the literature, a result of combining similar dwelling and occupant characteristics but who differ completely in the manner with which they consume electricity [13]. This research differs from previous work in the area as currently representative load profiles that reflect different patterns of electricity use within the home for a large sample of dwellings and which do not reflect an averaged profile shape do not exist. In addition, by correlating dwelling, occupant and appliance characteristics to electricity use within the home means that a household's electricity consumption profile can be completely identified based solely on their household characteristics.

2.2 Characterisation Techniques

2.2.1 Statistical, regression and probabilistic techniques

Statistical, regression and probabilistic approaches are particularly useful when comparing household characteristics against electricity consumption patterns. Most approaches that choose this type of technique are based on aggregated or large time interval electricity demand [27][28]. ‘Top-down’ approaches take data collected at a high level such as demographics, housing statistics and Gross Domestic Product (GDP) to derive causal relationships between these and electricity consumption [29]. In contrast to this, ‘bottom up’ models use data collected at an individual dwelling level to determine casual relationships between household characteristics and electricity consumption [13].

A general expression for multivariate linear regression model is described in Equation 2.1 below [30]. A dependent variable (which in most cases throughout this thesis can be regarded as electricity consumption) is regressed against a set of explanatory variables to produce a series of coefficients where $y(x)$ is the electricity consumption, $X_1, X_2, \dots, X_n$ are the explanatory variables referring to dwelling, occupant, and appliance characteristics and $\beta_1, \beta_2, \dots, \beta_n$ are the regression coefficient values that explain the influence of each explanatory variable on $y(x)$ and β_0 is a constant.

$$y(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \tag{2.1}$$

These models give a good understanding of electricity consumption patterns but can be costly to implement due the amount of data collection required [27][28] which invariably leads to small sample sizes on occasions [13]. Often linear regression is used to determine the degree of correlation between just one explanatory variable [13] and electricity consumption as sometimes multicollinearity issues arise where similar variables are investigated together.

2.2.2 Neural Networks

Neural networks have historically been used to forecast electricity system demand [31], however, they have also been applied at a domestic level for large time intervals [32][33]. A mathematical expression for a single input neuron within a network is shown in Equation 2.2 below where three distinct functional operations are taking place [34]. First, the scalar input p is multiplied by the scalar weight w to form the product wp. Second, the weighted input wp is added to the scalar bias b to form the net input. Finally, the net input is passed through the transfer function f, which produces the scalar output a. The names given to these three processes are: the weight function, the net input function and the transfer function.

$$a = f(wp + b) \tag{2.2}$$

Although there are various different network architectures for neural networks, common networks consist of three layers: an input, a hidden and an output layer [32]. Their self learning capabilities can result in an accurate means of characterising electricity consumption within the home [19]. However, neural networks are often regarded as a black box approach to characterisation, which also means that it may disregard the influence of important structural information (such as individual dwelling, occupant and appliance characteristics) on the output [7].

2.2.3 Engineering

Engineering approaches can be considered to be a ‘bottom up’ approach to electricity characterisation. Engineering methods use information such as appliance power ratings or end-use characteristics to build up a description of electricity consumption patterns within the home. A series of statistical and mathematical functions are usually used to describe its use. One of the major strengths associated with such an approach is that it is the only methodology that can model electricity consumption without any historical information on electricity use [18]. However, engineering methods can be complex to implement and need to be validated [23].

Mathematical expressions differ considerably between one approach and the next as there is no generic structure. However an example of an engineering method developed by Aydinalp and Ugursal [19] is presented in Equation 2.3 for illustration purposes. The authors take a Conditional Demand Analysis (CDA) approach which use inputs on appliance ownership and end-use to describe household electricity consumption.

$$HEC_{i,t} = \sum_{j=1}^{n} UEC_{i,j,t} \, . \, s_{i,j} \tag{2.3}$$

where $HEC_{i,t}$ is the energy consumption by household i in period t, $UEC_{i,j,t}$ is end-use j unit energy consumption of household i in period t, n is the total number of appliance end-uses and s_{ij} is a binary indicator of household i's ownership of appliance j.

2.2.4 Fourier Transforms

Fourier approaches have been used previous for load forecasting at a Transmission System Operator (TSO) level [35][36][37]. However, their application at a domestic level has been somewhat limited, mainly due to the poor availability of data [38]. A Fourier series is a representation of a time series signal in the frequency domain. It is often used to model and characterise system demand as particular patterns exist on a daily, weekly and annual basis [39]. The time series is broken into its individual frequency components, where each signal is composed of a collection of sinusoids as shown in Equation 2.4 with Fourier coefficients defined in Equation 2.5 [40]. The individual coefficients; a_0 representing a constant, a_r and b_r, where $(r=1,2,3\ldots)$ correspond with the magnitude of each sinusoid at a particular frequency which when all summed together represent the original time series signal.

$$f(t) = \frac{a_0}{2} + \sum_{r=1}^{k} a_r cos(rt) + \sum_{r=1}^{k} b_r sin(rt) \qquad (2.4)$$

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)dt\,; \quad a_r = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)cos(rt)\,dt;$$

$$b_r = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t)sin(rt)dt \quad where\; r = 1,2,3 \ldots \qquad (2.5)$$

As discussed, Fourier transforms are useful for characterising time series that show certain patterns of reoccurrence due to their representation of the signal in the frequency domain. However, whether these properties exist and can be characterised effectively at an individual dwelling level has yet to be determined.

2.2.5 Gaussian Processes

Gaussian processes have been used to forecast electricity system demand but to date have not been applied at a domestic level [41][8]. Using this technique, an electricity load profile can be fully characterised by fitting a series of peaks like that shown in Equation 2.6 [42].

$$y = \sum_{i=1}^{n} a_i.\, e^{\left[-\left(\frac{x-b_i}{c_i}\right)^2\right]} \tag{2.6}$$

where a is the amplitude, b is the centroid (location), c is related to the peak width, n is the number of peaks to fit. In contrast to Neural networks, Gaussian processes provide a much simpler representation with only three moments required to represent each distribution for the load profile shape.

2.2.6 Autoregression

Autoregression (AR) is a time series approach that has often been applied to electricity system demand load forecasting but has not been directly applied at a domestic level before [39][43][9]. The AR process describes a time series y_t as a linear function of previous elements and an error term ε_t as shown in Equation 2.7 [44]. Variable coefficients A_i are calculated by regressing the time series onto itself where t is time, a is a constant and i is the time lag.

$$y_t = a + \sum_{i=1}^{p} A_i \,.\, y_{(t-i)} + \varepsilon_t \tag{2.7}$$

Box and Jenkins [45] developed a methodology that selects the most appropriate forecasting technique based on a combination of an Autoregression (AR) and a Moving Average (MA) process. The developed method is called an Autoregressive Moving Average (ARMA) process. Equation 2.8 shows the MA component where a is a constant, B_j are coefficient values and ϵ is the white noise error terms [44].

$$y_t = a + \sum_{j=1}^{q} B_j \varepsilon_{(t-j)} + \varepsilon_t \tag{2.8}$$

The ARMA process is described in Equation 2.9 which is a linear combination of Equations 2.7 and 2.8 [44].

$$y_t = a + \sum_{i=1}^{p} A_i \,.\, y_{(t-i)} + \sum_{j=1}^{q} B_j \,.\, \varepsilon_{(t-j)} + \varepsilon_t \tag{2.9}$$

For the ARMA process to be stable it is necessary for the time series to be stationary [45]. For this reason Box and Jenkins introduced a differencing component into their methodology so as to remove any seasonality components. The number of times the series is differentiated depends upon the extent of the seasonality component. This type of process is then referred to as an Autoregressive Integrated Moving Average (ARIMA) process.

A variant of the autoregressive approach are Markov chains. Markov chains have been used to forecast electricity system demand before [46]. They describe a stochastic process where the future value of a time series is calculated based on past probabilities of going from one discrete state to another. A time series can be described by Equation 2.10 as X(t), possessing discrete states space S={1,2,.,K} [47]. In general, for a given sequence of time points $t_1 < t_2 \ldots < t_{n-1} < t_n$, the conditional probabilities are:

$$\begin{aligned} Pr\{X(t_n) = i_n | X(t_1 = i_1, \ldots, X(t_{n-1}) = i_{n-1})\} \\ = Pr\{X(t_n) = i_n | X(t_{n-1}) = i_{n-1}\} \end{aligned} \tag{2.10}$$

The conditional probabilities $Pr\{X(t) = j | X_s = i\} = P_{ij}(s,t)$ are called transition probabilities from state i to state j for all indices $0 \leq s < t$, with $1 \leq i$ and $j \leq k$ [47].

2.2.7 Fuzzy Logic

Fuzzy logic approaches have been used to characterise electricity system demand [48][49] and are well suited to describing non-linear relationships [50]. Their application at a domestic level has not been done to date to the best of the authors knowledge. A fuzzy set is described by Equation 2.11 below [48].

$$A = \{[x, \mu_A(x)] | x \in X\} \tag{2.11}$$

where $\mu_A(x)$ is called the membership function of x in A. The membership function denotes the degree that x belongs to A and is normally limited to values between 0 and 1. For instance a high value of $\mu_A(x)$ implies that it is very likely that x is a member of A.

2.2.8 Wavelets

Wavelets take a time series signal and decompose it into high and low frequency components. Both components are characterised separately which has the added advantage over Fourier transforms in that the time series signal can be analysed at different resolutions.

A continuous wavelet transform is represented in Equation 2.12 below [51].

$$C(a,b;f(t),\psi(t)) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt \qquad (2.12)$$

where $f(t)$ is the original time series signal and ψ^* is the complex conjugate of the wavelet function defined by scale a and position b in Equation 2.12.

2.2.9 Clustering

Clustering has been used, by large electricity suppliers to group customers together which share similar electrical characteristics [52]. Its use at an individual dwelling level has been somewhat limited; the main focus to date being the commercial and industrial sectors to define particular customer groups [53][54]. There are numerous different methods of clustering; however, usually they can be grouped under the following two headings: Hierarchical and Partitional [55].

Hierarchical Clustering

Hierarchical clustering produces a set of nested clusters that can be visualised as a tree like structure such as a Dendogram shown in Figure 2.1 [56]. Either a top down (divisive) or a bottom up (agglomerative) approach to clustering the data can be taken. The advantage with this type of approach is that the number of clusters does not need to be defined, and the Dendogram can be cut at a particular height so as to define a specific number of clusters. However, the disadvantages are that it is more susceptible to outliers within the data and has difficulty dealing with clusters of different sizes [57].

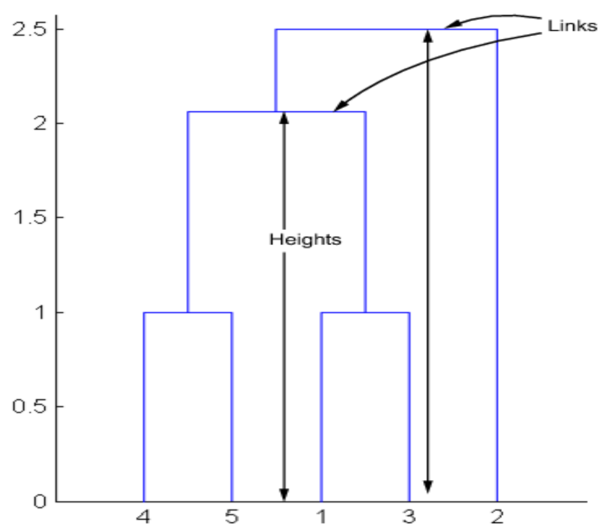


Figure 2.1: Dendogram for hierarchical clustering [56]

Hierarchical clustering uses a similarity matrix to define individual customer groups. Each customer is assigned to a particular cluster based on a linkage criteria and distance metric (shown in Figure 2.1 by the heights of the Dendogram) between itself and the similarity matrix. It is an iterative process by which each individual cluster is repeatedly merged with a larger one until all customers are represented by a single cluster. Figure 2.1 shows the individual cluster groups on the x-axis and on the y-axis the distance at which the clusters merge.

Equation 2.13 shows the equation for the Euclidean metric *d* [58] which is one of the most common forms of distance measurements between two points *x* and *y* used in clustering [59].

$$d = |x - y| = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \tag{2.13}$$

Partitional Clustering

In contrast partitional clustering divides the data into a predefined number of non-overlapping clusters. There are various different types of partitional clustering methods, however, some of the more common approaches include: k-means, k-medoid and Self Organising Maps (SOM) [60]. The advantages with these types of techniques is that the clusters are predefined and do not overlap [61]. However, this is also a disadvantage as one has to decide upon the number of clusters before the process is started [59].

K-means uses an iterative process that assigns customers into groups based on the distance between itself and a cluster centre. Initially, cluster centres are chosen at random within the sample data set. The distance (again usually Euclidian) is then calculated between the sample customer and the cluster's centres. The customer is then assigned to the cluster with the minimum distance to its centre. The cluster centre is then re-calculated based on the addition of a new customer.

Equation 2.14 describes the k-means algorithm where given a set of observations $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$, where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k subsets $(k \leq n)$ $\boldsymbol{S} = \{S_1, S_2, \ldots, S_j\}$ so as to minimize the within-cluster sum of squares and where $\boldsymbol{\mu}_j$ is the geometric centroid of the data points in S_j in order to achieve a global minimum for J [62].

$$J = \sum_{j=1}^{k} \sum_{n \in S_j} \left\| x_n - \mu_j \right\|^2 \tag{2.14}$$

K-medoid is similar to k-means except that each cluster is represented by one of its own points. It is less susceptible to outliers within the data compared to K-means because peripheral cluster points do not affect the cluster centres but as a result has a high computation cost [59]. The algorithm most often applied with this type of technique is Partitioning Around Medoids (PAM).

Self Organising Maps (SOM) or sometimes referred to as Kohonen maps are based on the principles of neural networks but can be considered to be a clustering technique in its own right [61]. Figure 2.2 shows the basic structure to a SOM [63].

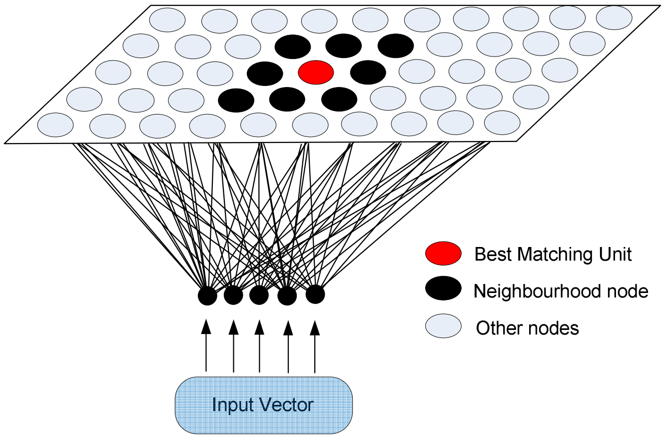


Figure 2.2: Self Organising Map (SOM) structure [63]

SOM apply a neural network process that uses unsupervised learning to divide the data. A rectangular or hexagonal lattice structure of nodes is usually used to segregate the data. Each hexagonal node is defined by a weight vector which consists of a series of different dimensions depending on the input vector. The mapping process is started by initialising weight vectors with random values at each node. As the network progresses each input vector is compared with the weights of each node and the node with the greatest similarity, called the Best Matching Unit (BMU), is assigned that particular vector. The weights are then adjusted at the BMU and neighbourhood nodes based on the input vector. The process is repeated until all input vectors have been categorised into groups [64].

2.3 Application of Techniques to Demand Load Profiling

This section describes the different approaches to characterising domestic electricity consumption patterns. In the past, literature for the area has been divided in different ways. Swan and Ugursal [18] separated the literature based on top-down and bottom-up approaches to modelling energy consumption within the home. Top-down techniques were divided into econometric and technological categories. Bottom-up approaches were divided into two categories: statistical and engineering. These two categories were further sub-dived into: regression, Conditional Demand Analysis (CDA) and neural network for statistical techniques, and population distribution, archtype and sample for engineering methods.

Time series approaches have mainly been applied to electricity system demand load forecasting and Alfares and Nazeeruddin [31] divide the literature into nine different categories: regression, exponential smoothing, iterative reweighted least-squares, adaptive load forecasting, autoregressive, fuzzy logic, neural networks and knowledge-based expert systems.

Based on the data structure presented earlier in Figure 1.7, the literature can be categorised into the following subsections 2.3.1 – 2.3.4:

- The first section deals with electricity consumption collected at an aggregate level or over large time intervals. These approaches have been used extensively in the past and are explained in detail in Section 2.3.1. These studies tend to pre-date smart metering programmes where more detailed information has become available.
- The second section describes approaches that use data collected at an individual dwelling level and for small time intervals. They mainly consist of engineering and statistical approaches which have been used to characterise electricity demand across the day but in the most part have been limited to small sample sizes as will be shown in Section 2.3.2.

- The third section describes methods that have mainly been applied at an aggregate level to characterise electricity system demand for small time intervals. These mainly consist of time series approaches which have mostly been applied at an electricity system demand level which will be shown in Section 2.3.3. However, with the introduction of smart meters and the availability of data at the same time resolution for the domestic sector as that at a system demand level, these approaches can now be applied to characterise individual dwellings.

- Finally, the last section describes clustering based approaches to load profile characterisation. These are applied at an individual level and for small time intervals. However, to date these methods have mainly been applied to non-residential electricity data such as commercial and industrial as will be shown in Section 2.3.4.

These methods and where they relate to the data structure presented earlier are shown in Figure 2.3 below. The decision to categorise the literature in this manner was done due to the overlap between different approaches but also due to wealth of literature available at a system demand level that can now be applied to an individual dwelling.

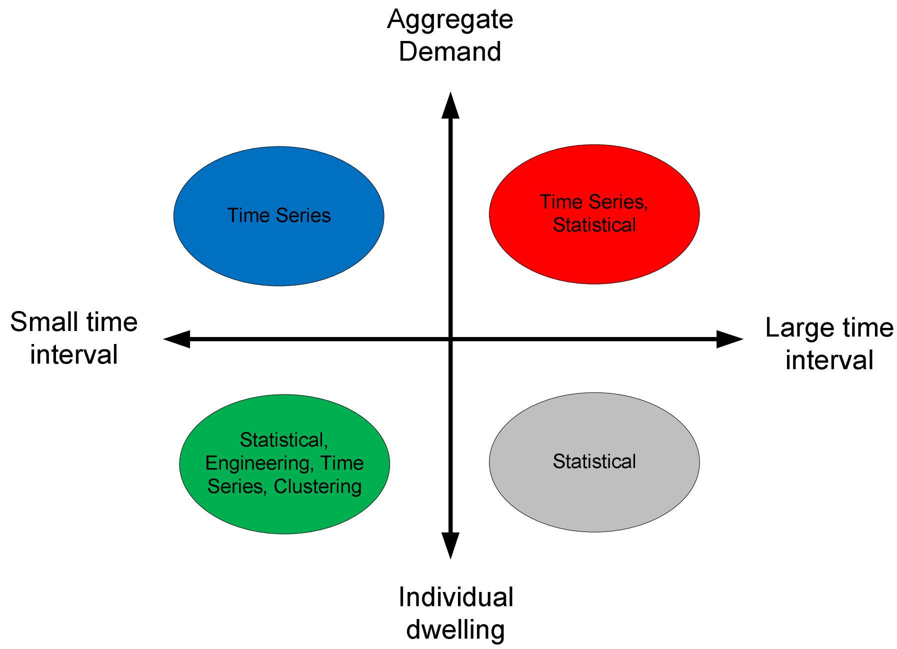


Figure 2.3: Taxonomy of characterisation approaches for electricity load profiling

2.3.1 Large time interval characterisation approaches

In the past, data has not been available at adequate sample sizes or at sufficiently high time resolution to enable detailed characterisation of domestic electricity demand. For that reason, a large number of methods used to characterise domestic electricity demand were based on what could be considered aggregate or large time interval demand approaches. Methods based on this approach tend to be easier to characterise in terms of dwelling and occupant characteristics, a result of a larger deterministic component caused by the averaging of the electrical demand over a period of time for each household.

A commonly used method to describe domestic electricity consumption in this manner is regression. Using regression in this manner also provides a method of determining the household characteristics that have a significant influence over electricity consumption within the home. O'Doherty et al. [27] used data from a National Survey of Housing Quality and applied a Papke-Wooldridge generalised linear regression model to infer a relationship between appliance ownership and household electricity consumption. Their analysis showed explanatory variables that had a high significance with respect to electricity consumption such as: dwelling characteristics; location, value and dwelling type as well as occupant characteristics; income, age, period of residency, social class and tenure type. Leahy and Lyons [28] also applied an ordinary linear least squares regression using data from the Irish Household Budget Survey. Disposable income, household size, dwelling age and socio-economic group were among the variables that were shown to influence electricity consumption in the home. Leahy et al. [65] extended their work and applied a similar regression methodology to the Irish smart metering dataset in order to examine the household characteristics that influence appliance ownership and use within Ireland. Among the authors findings were that the number of people living within a household had a positive association with appliance ownership and use and that the highest earning households tended to own more appliances but did not necessarily use them more often.

Baker and Rylatt [20] used regression to determine a relationship between household characteristics and annual electricity consumption for 148 dwellings in two major UK cities. In particular the characteristics that showed the greatest significance with respect to electricity consumption within the home were: floor area, home working, number of televisions, personal computers, digital boxes, portable electric heaters, storage heaters and showers per week.

A variant of the regression approach is a technique called Conditional Demand Analysis (CDA). CDA is a method by which total electricity consumption is disaggregated into separate components (e.g. lighting, cooking, refrigeration, etc) by knowing the appliance holdings and total consumed load within a dwelling. A regression analysis is carried out on the data to generate a series of equations that when combined together represent total energy end-use. A large sample size is generally required in order to conduct this form of analysis and produce accurate results. Parti and Parti [66] pioneered the method and used monthly electricity bills over a yearly period and appliance ownership figures and a number of demographic variables to disaggregate electricity demand into 16 different end-uses. This methodology showed the high significance of appliance ownership with respect to electricity consumption patterns.

Aydinalp et al. [32] applied a neural network to model electricity consumption for domestic appliances, lighting and space cooling in the home. Aydinalp [33] also extended this work to develop neural network models for space and domestic hot-water heating. In addition, Aydinalp [19] carried out a comparison of models: CDA, neural network and engineering approaches to modelling end-use energy consumption for the residential sector. The authors found that CDA was equally as good a method as neural network and engineering approaches, however, it was less able to model the effect of socio-economic factors due to the limited number of variables the CDA method could accommodate. Gabreyohannes [21] used a time series approach to model monthly domestic electricity consumption in Ethiopia. The author applied a self-exciting threshold autoregressive (SETAR) approach and a smooth transition regression (STR) to model demand over a two year period of which the SETAR model compared better to a simple AR model.

2.3.2 Small time interval at an individual dwelling level (small sample size)

For the most part, engineering approaches use data at less than or equal to an hourly time interval. However, there is a trade off between the level of detail at which the time interval of electricity consumption is recorded and the ability to manage and analyse vast amounts of data for large sample sizes. Research carried out by Wright and Firth [67] found that in certain circumstances, 30 minute time intervals is sufficient. The authors investigated the effect of time averaging of electricity consumption for the purposes of modelling on-site generation for seven domestic dwellings. Data was collected at time intervals of 1, 5, 15 and 30 minutes. The authors concluded that 30 minute data is adequate when investigating the percentage export of on-site generation to the grid. For import, results were less favourable with the authors suggesting 5 minute intervals would be preferable. However, the Irish Commission for Energy Regulation (CER) concluded that 30 minute data interval was adequate for electricity metering. This was reinforced by the French decision to roll out smart metering nationwide at 30 minute intervals stating confidentiality of consumer behaviour as the reason [68]

Yohanis el al. [13] applied a statistical approach to investigate patterns of electricity use in 27 representative dwellings in Northern Ireland. Electricity load profiles were characterised based on dwelling type, floor area, number of occupants, number of bedrooms, tenure, occupant age and household income. In particular, the authors found a significant relationship between domestic electricity consumption and floor area by using regression. Firth el al. [14] also used a statistical approach to characterise electricity consumption in 72 domestic dwellings in the UK over a two year period.

The author categorised domestic appliances based on four groups according to their pattern of electricity use: continuous appliances, standby appliances, cold appliances and active appliances. The focus of their research was to identify trends to explain changes in electricity consumption patterns over time. Wood and Newborough [26] characterised domestic electricity consumption using descriptive statistics for 44 households in the UK based on three categories: "predictable", "moderately predictable" and "unpredictable". "Predictable loads" consisted of small cyclic loads occurring when a dwelling is unoccupied or all the occupants are asleep. "Moderately predictable" related to the habitual behaviour of the occupants and "unpredictable" described the vast majority of electricity consumption within a dwelling.

Hart and de Dear [69] investigated the influence of external temperature on household electricity demand in Australia. Regression was used to establish the relationship between electricity use and degree days. Their research concluded that weather variables did influence electricity consumption and that this tended to be stronger during periods of cooler weather. Parker [70] also looked at the affect of external temperature on electricity consumption. Fifteen minute electricity consumption data was collected from 204 residences in central Florida, USA. A significant relationship was found by applying linear regression, between all electricity end-uses and external temperature. However, it is important to note that both preceding studies presented by Hart and de Dear and Parker were carried out in hot climates where electricity is commonly used to heat and cool homes, something which is not replicated in more temperate climates such as the UK and Ireland. Schick et al. [71] regressed household characteristics and weather variables onto hourly metered data in order to determine their influence on the load profile shape. The analysis is broken up into two sections, with the first regressing all variables onto the load profile. Variables that had little or no influence over the load profile shape are disregarded and the analysis is run again. The authors found this two stage process to be an effective method of modelling household electricity demand.

Aigner et al. [72] furthered the CDA approach developed by Parti and Parti mentioned in the last section. In this case, twenty-four regression equations were used to estimate electricity consumption for each hour of the day. Bartels et al. [73] developed a load forecasting model called DELMOD at half hourly intervals using CDA in New South Wales, Australia. Half-hourly electricity readings for up to sixteen different end-uses were taken from a sample of 250 households over fifteen months. DELMOD not only produced a number of average load profiles for each end-use but also estimated the influence of socio-demographic and weather related variables on its use.

Cross and Gaunt [74] used CDA and appliance holdings to characterise electricity demand for small rural villages in South Africa. CDA was used to disaggregate data collected form fifteen nearby villages where electricity was monitored and survey data collected. The CDA curves, along with appliance holdings for new villages were then used to estimate hourly load curves for newly electrified villages. Larsen and Nesbakken [75] used total electricity consumption, appliance holdings and household characteristics to develop a CDA approach based on data from a 1990 energy survey in Norway. The authors compare results to an engineering model ERAD [76] and concluded that there are drawbacks to both methods but comment that if survey data was collected with CDA analysis in mind the results would be more accurate. Tiedemann [77] presented a CDA methodology to estimate residential energy end-use and energy saving measures from data collected from an end-use study carried out in 2004 in Canada. A survey detailing all major end-uses for 791 customers were collected including billing information, housing characteristics, demographics and attitudes towards energy use. Regression coefficients were calculated representing unit energy consumption (UEC) for all major residential end-uses. A method to calculate the saturation rate is also presented which describes the average number of individual end-uses per household. The authors then use the product of UEC and the saturation rate to determine the average electricity consumption for individual end-uses across all households.

Yao and Steemers [23] developed a dynamic software engineering approach that produced load profiles based on occupancy patterns, appliance ownership and ratings. The authors categorised electricity consumption determinants based on two categories: behavioural and physical, both of which are strongly related to dwelling occupancy patterns. Behavioural determinants relate to decisions made on an hourly, daily and weekly basis regarding the use of particular appliances. Physical determinants relate to "fixed" variables that do not change often or at all with time such as dwelling size. Widen and Wackelgard [24] used an engineering method that used time-use data (i.e. occupant's schedule of living activities) as well as appliance holdings, ratings and daylight distributions to produce electricity load profiles. Three sets of Swedish time-use data and energy measurements were used to model and validate the results. The approach built upon previous research by the same authors but was performed at a higher time resolution [78]. Richardson et al. [25] developed occupancy and daily activity profiles based on time-use data and combined this with appliance profiles in order to characterise household electricity demand in the UK. The measured and generated electricity load profiles were compared based on a number of parameters including total electricity consumption on an annual, monthly and daily basis with small differences between the two. Shimoda et al. [79] also used an engineering approach to characterise household electricity demand for different dwelling and occupant characteristics in Japan. The authors showed that occupant's time-use, external temperature, appliance efficiencies and dwelling thermal characteristics all significantly influenced the electricity consumption pattern across the day. Capasso et al. [22] used homeowner's occupancy patterns as well as appliance ownership, usage and ratings to characterise electricity consumption patterns for the home in Italy. The author also incorporated various socioeconomic, demographic, psychological and behavioural characteristics for the dwelling occupants in order to determine the effect on the load profile shape. Walker and Pokoski [80] applied an engineering approach based on homeowner's psychological decisions (mental and behavioural) to describe daily load electricity patterns in the USA. A function was developed indicating if a person was available at home to use a particular appliance. The model was compared against Connecticut Light and Power Company data for residences having similar family size and stock of appliances with comparable results.

Jardine [81] used metered data and average appliance usage to produce electricity load profiles for one thousand homes. The principal part of this analysis is the development of an occupancy model that is constructed from a sample of one hundred domestic electricity load profiles. The occupancy model relies on the presence of non-baseload appliances to predict when an occupant is within the home. The method then assumes electricity use as a function of three separate factors; number of appliances owned, average rated power of appliances and duration of time appliances are used. Average load profiles are produced based on the occupancy model and appliance ownership levels along with load duration curves for domestic appliances.

Stokes [82] developed a high resolution domestic load profile method to investigate the impact of embedded generation on the low voltage distribution network. The author used a number of layers to build load profiles at one minute time intervals. Firstly, mean demands at half hourly time intervals were used which reflect an average load profile. The second layer is then added which introduces diversity into the profile by introducing dwelling and occupant characteristics and well as other economic factors. Finally, appliance duty cycles, triggered at random intervals introduce a variable element to produce one minute electricity load profiles. The author concludes that the profiles compare well with measured data, particularly when compared against derived parameters such as mean and peak electricity demand and load factor.

Prudenzi [83] used a neuron nets based procedure to identify appliance pattern of use from fifteen minute interval electricity data. The methodology uses data from nine domestic houses in Italy where total electricity consumption was recorded for large appliances, along with survey data indicating time-use for each appliance. The methodology contains three separate stages: pre-processing stage, time-use identification stage and a post-processing stage. The three stages are used to firstly analyse the load shape based on time of day, secondly to identify an appliance type to be used, and lastly to produce a load profile for a dwelling. The approach becomes less accurate when a large number of appliances are used simultaneously.

2.3.3 Small time interval at an aggregated level

Time series approaches have been extensively applied to characterise electricity system demand, similar to that shown in Figure 1.1. However, their use at an individual dwelling level has been somewhat limited, mainly due to the historic lack of data at this level. Based on the literature, the major time series approaches to electricity load profile characterisation can be grouped under the following headings: Fourier transforms, Neural networks, Gaussian processes, Autoregressive, Fuzzy logic, Wavelets, Multivariate regression and Probabilistic

There is some overlap between time series approaches and other methods presented earlier. However, the majority of techniques discussed in this section have been applied to datasets with time intervals of less than or equal to one hour.

Fourier Transforms

Fourier transforms have been applied extensively to characterise and forecast electricity system demand but have rarely been applied at an individual dwelling level before. Riddell and Manson [38] fitted polynomials and other mathematical functions to electricity demand load profiles before settling on a Fourier series to approximate the load profile shape. The authors applied a fourth order Fourier series to characterise electricity load profiles over a twenty four hour period. Although the data used in the analysis was taken from domestic households, the majority of it was collected at transformer level and hence in most cases represented anywhere between one to fifty households. Moutter el al. [35] used a Fourier transform approach to forecast medium (weekly) to short term (hourly) electricity system demand load profiles at half hourly intervals over a yearly period in New Zealand. Longer term forecasts over a year posed a problem to the authors due to the time series being non-stationary. Gonzalez-Romera et al. [37] used a hybrid approach to model electricity system demand at monthly intervals in Spain. The authors applied a Fourier series to model the periodic behaviour of the time series whilst the seasonal trend was modelled with a neural network. Satisfactory results were achieved with the approach out performing a neural network and autoregressive integrated moving average (ARIMA) techniques using the same dataset.

Neural Networks

Neural networks have not been used to model individual dwelling electricity demand for small time intervals across the day but have been applied to forecast system demand. Chen et al. [84] applied a neural network for short term load prediction using inputs such as load, day type, temperature and electricity price in Ontario, Canada. A three layer feed forward neural network with back propagation was shown to be successful at modelling the highly non-linear relationship. Zadeh and Masoumi [85] also used a neural network with back propagation to model aggregated residential electricity demand in Iran. The authors used the previous year's electricity prices and consumption data to forecast demand. Ringwood et al. [7] also examined the use of neural networks to forecast electricity system demand. In particular, the authors found that neural networks outperformed linear models for short to medium term time periods, with longer time periods better characterised by the latter.

Gaussian Processes

Singh et al. [86] applied a Gaussian process model to produce electricity load profiles for a generic distribution network in the UK. The authors used a parametric estimation technique known as Expectation Maximisation to obtain values for the mixture components (mean, variance and weight). A Chi-square goodness of fit test was used to determine the accuracy of the fitted distribution function and the original time series. Leith et al. [41] applied a Gaussian process to forecast electricity system demand over a yearly period in Ireland. The authors found that the Gaussian process performed better when compared against an integrated seasonal autoregressive approach and a basic structural model. Lourenco and Santos [8] also used a Gaussian process to forecast short term electricity system demand in Portugal. The authors used data collected over a three year period from three separate sub-stations representing non-residential, residential and services sector. Satisfactory results were achieved for different time periods and load profiles.

Autoregressive

Autoregression has commonly been used to forecast electricity system demand. The methods are often adjusted to improve the performance by adding additional mathematical functions but all rely on the same autoregressive principles. Pappas et al [43] choose a simple ARMA process to model electricity system demand in Greece. The authors found it to be successful for fitting the data and that a multi-model partitioning filter was the best selection criteria to determine the model order. Magnano and Boland [39] used a hybrid approach to forecast electricity system demand over a three year period in Australia. The authors found that the model performed better if the stochastic and deterministic components were modelled separately. The stochastic components were modelled using an ARMA process and the deterministic components with a Fourier series and polynomial functions. Amaral et al [9] applied a smooth transition periodic autoregressive (STPAR) model to forecast short term electricity system demand in Australia. The authors compared STPAR with a simple autoregressive and a neural network approaches over the forecast horizon of a week, with the STPAR model performing best. Ardakanian et al. [46] used a continuous time Markovian process to model home electricity consumption in Canada. The authors divide the day into three different time periods representing on-peak (7am-11am and 5pm-9pm), mid-peak (11am-5pm), off-peak (12am-7am and 9pm-12am). Twelve different profiles were constructed based on the three different time periods and four different customer classes which relates to the size of the dwelling and the type of heating and cooling system used. The authors found it to be a useful tool for transformer sizing for the electrical grid.

Fuzzy Logic

Hsu and Ho [48] applied a fuzzy logic process to forecast weather variables for a short term electricity system demand model in Taiwan. The authors used regression to determine the correlation between temperature and peak and trough electricity demand. As temperature forecasts are often inaccurate an error term was introduced to the electricity forecast model equations. The error term was modelled using a fuzzy logic process to take account of this uncertainty. Mastorocostas et al. [49] also applied a fuzzy logic process to predict short term electricity system demand in Greece. A number of fuzzy models were used to generate hourly loads for each day of the year. In total 28 different models were used, one for each day type (i.e. Monday, Tuesday, etc.) and season of the year. The model was compared against a neural network with similar results. Mori and Kobayashi [50] proposed a fuzzy logic model for short term electricity system demand load forecasting in Japan. Membership functions were optimally evaluated through a learning algorithm developed by the authors. The authors initially used two months data to train the model after which a single month's electricity demand was forecasted.

Wavelets

Xu and Niimura [87] used wavelets and autoregression to model short term electricity prices in the USA. Historical prices were decomposed by applying a wavelet transform to the time series. An autoregressive ARMA process was then applied to forecast wavelet coefficients for the next day. The forecasted electricity price was obtained by applying the inverse wavelet transform. Chen et al. [88] combine both wavelet and neural network methods for short term electricity system demand load forecasting in the USA. The authors use a wavelet transform to decompose a similar day's load into high and low frequency components. The two components are then adjusted for weekday and weather variables and modelled by two separate neural networks before the two components are added back together. Pahasa and Theera-Umpon [89] also used a wavelet transform to forecast short term electricity system demand in Thailand. The authors first decomposed the time series into high and low frequency components and then used support vector machines, a classification technique, to forecast each component separately. Each component is then summed together to determine the forecasted load. Nguyen and Nabney [90] applied a wavelet transform for predicting electricity system demand and gas forward price one day in advance in the UK. The authors showed that forecasting accuracy significantly improved when using wavelet transforms with a number of adaptive models (multi-layer perceptron, radial basis functions and linear regression).

Multivariate regression and probabilistic

McSharry et al. [91] applied regression to forecast the magnitude of peak time electricity system demand on a daily basis in the Netherlands. Variables used in the forecasting model were: weather variables (temperature and wind speed), luminosity, day of week and special event days such as Christmas. Yearly seasonality was modelled with a fourth order polynomial. The UK [12] forecast residential electricity demand, mainly for the purposes of electricity settlement, by applying regression to a number of variables. For each half hour time period regression coefficients are calculated based on variables such as temperature, sunset and day of the week. A similar process is used in the Republic of Ireland to produce standard load profiles, as shown earlier in Figure 1.3, also for the purposes of electricity settlement.

Heunis and Herman [92] used a probabilistic approach to simulate domestic electricity demand in South Africa. The authors used a beta probability distribution function and applied a Monte Carlo simulation to predict electricity consumption for individual households. In contrast Cagni et al. [93] applied a Gamma probability distribution function to characterise domestic electricity demand in Italy. The authors applied a Chi-square goodness of fit test to determine the most suitable distribution. Domestic electricity demand was then simulated for individual dwellings based on sampling the gamma probability distribution function. Capasso et al. [94] defined probability distributions for individual events such as cooking within the home. The data used to build probability distributions was taken from a customer survey where information on electrical appliances and use was recorded. The events were then simulated over a 24 hour period and the results compared against total electricity consumption for individual dwellings with reasonable accuracy. Carpaneto and Chicco [95] developed a probabilistic approach for characterising aggregated domestic electricity demand for a number of dwellings. Probability distributions were assigned to time intervals for individual customers. The evolution of mean and standard deviation for increasing numbers of customers was calculated and it was found that the gamma probability distribution fitted the aggregated domestic electricity demand best. McQueen et al. [96] used Monte Carlo analysis to simulate individual domestic electricity demand at half hourly intervals using a gamma distribution. The profiles were then diversified and used to predict maximum demand for a distribution network. Similarly Jardini et al. [97] used mean and standard deviation from a Gaussian distribution to represent domestic electricity demand at each half hour interval. The authors then generated average load profiles based on sampling the probability distribution functions. Chen et al. [98] applied the same method as Jardini et al. to describe different customer groups based on quarter hourly electricity demand. Individual domestic electricity demands were then integrated to derive a system demand load profile for the distribution system.

A table summarising each of the time series approaches above is shown in Appendix C, outlining the advantages and disadvantages of each method when applied to electricity load profiling. The table also indicates whether each method was applied at an aggregate or individual dwelling level.

2.3.4 Small time interval at an individual level (large sample size)

Clustering is a common technique used by electricity utilities to segment their customer base. Chicco et al. [53] characterised customer's electricity demand based on a set of indices representing their electrical behaviour throughout the day. A number of daily and weekly indices were defined such as load factor, night time load (between 23:00

and 06:00) for weekday and weekend, lunch time load between the hours of 12:00 to 14:00. An automatic clustering algorithm using unsupervised learning was then used to segment customers into groups based on these indices. Chicco et al. [99] then compared a number of common clustering techniques (modified follow the leader, k-means, fuzzy k-means, two types of hierarchical and self organising maps) to divide non-residential customers into groups. The authors found that hierarchical and modified follow the leader clustering approaches performed the best. Verdú et al. [100] applied Self Organising Maps (SOM) and a set of indices similar to Chicco et al.[53] as well as some frequency based indices to characterise customer electrical demand throughout the day. The author also included a further three indices into their analysis, particularly looking at electrical demand during daylight hours. Figueiredo et al. [52] applied clustering and just three indices, load factor, night impact between the hours 23:00 and 07:00, and lunch impact between 12:00 and 15:00 and used these to characterise a sample of 165 customers in Portugal. Pitt [101] applies an adaptive load profiling methodology that uses various clustering algorithms to relate weather, time and customer characteristics to the load profile shape. The author acknowledges the complications involved in clustering high dimensional data as well as the difficulties in dealing with heterogeneous data, where different customer's electricity demand may vary dramatically on any given day.

Gavrilas et al. [102] used a modified fuzzy SOM algorithm to produce nine typical electricity load profiles for commercial and residential data metered on the low voltage side of eleven substation transformers. The authors highlighted the small but important differences between weekdays and weekends. Tsekouras et al. [103] used a two stage methodology for classifying electricity customers. In the first instance, typical load diagrams for each customer are determined based on clustering. Secondly, customers are clustered again based on their typical pattern of electricity use to group similar customers together. This ensured representative classification of customers. The authors compared methods k-means, SOM, fuzzy k-means and hierarchical clustering by using six different adequacy indicators. Carpaneto et al. [104] used frequency based indices such as amplitude and phase of customer electricity demand rather than time domain data to characterise customers into groups. Espinoza et al. [105] used a periodic autoregression and k-means clustering to develop a short term load forecasting model for 245 transformer sub-stations in Belgium. The stationary properties were extracted from the autoregressive model, delivering individual daily load profiles for each transformer. These were then clustered using k-means to reduce the number of profile classes down to eight in total. Zhang et al. [106] investigates three methods k-mean, fuzzy k-means and SOM to segment large electricity consuming customers. A stability index is used to evaluate the most appropriate clustering method and a priority index to rank the number of clusters. Bidoki et al. [107] compared a number of clustering techniques for non-residential electricity customer classification. The authors found that modified follow the leader performed best for identifying the most distinct clusters. However, if more compact clusters were required, weighted fuzzy average k-means was found to be a better performing method.

2.4 Summary of Previous Work

The literature was divided up into four sections based on the level of aggregation and time interval applied. This approach was taken on account of the relatively recent widespread introduction of smart meters into domestic homes [3]. This has resulted in the availability of a large amount of data at small time intervals and for large sample sizes such as that recorded by the smart metering trial in Ireland. As a result, methods that have traditionally been applied to other sectors such as system demand load forecasting and clustering approaches for customer segmentation can now be applied to characterise domestic electricity consumption at small time intervals.

Section 2.3.1 described methods that characterised electricity consumption patterns at large time intervals and used data that is often aggregated between customers. These approaches are largely based on statistical and regression and have been used in the past to characterise domestic electricity consumption by engineers, economists and government officials for infrastructure planning and electricity load forecasting [21]. One of the strengths with these methods, lies in their ability to assess the influence of characteristics on electricity consumption patterns [28]. This approach was taken in Chapter 5, where parameters were used to aggregate the data in the time domain. Characteristics relating to dwelling, occupants and appliances were investigated based on their influence with respect to these parameters describing electricity use within the home.

Section 2.3.2 presented methods which have been applied most often at small time intervals for individual dwellings and mainly consist of engineering and statistical approaches. These techniques represent a significant proportion of the literature to date for characterising electricity consumption in the domestic sector. Engineering methods describe or model electricity demand as a function of variables such as occupancy or appliance holdings. They are the only approach that does not require any information about electricity demand. However, they need to be validated against data collected at individual dwelling level. As a result it is often the case that these types of approaches are carried out with small sample sizes [23]. Similarly statistical approaches used to characterise domestic electricity demand at this level are often limited to small sample sizes too [14][26][13]. This is a result of the difficulty with characterising many different patterns of electricity use and individual customer characteristics together at small time intervals.

It was found from literature that occupancy patterns are one of the main variables used with engineering models [24][25]. Although this variable was recorded by the Irish smart metering trial dataset, it was not recorded at sufficient time resolution to enable daily load patterns to be produced and therefore an engineering approach was not pursued. Similarly, the amount of data available from the smart metering data set means that a statistical approach similar to that described in literature would have been difficult to implement without a method of reducing the data first. In particular, Chapters 5 and 8 provide methods for reducing the data before any such analysis is carried out.

Section 2.3.3 discussed the next group of approaches, applied at small time intervals but mostly for aggregated loads such as system demand. The availability of smart metering data at a similar time resolution to that recorded for system demand has meant that an opportunity exists to apply similar methods used to characterise system demand and apply them at an individual dwelling level. These approaches mainly consist of time series methods. The main challenge with these techniques is that they have most often been used to characterise system demand which varies much less frequently between half hour periods compared to individual domestic customers (as was identified in Figure 1.1 and Figure 1.2). Chapter 7 provides an in depth discussion as to how these approaches may be applied to the domestic sector with two techniques being identified as the most suitable.

Finally, clustering approaches to characterising domestic electricity demand were presented in Section 2.3.4. These methods have been used to characterise electricity demand, often in combination with the calculation of indices [53]. The methods are mostly applied to characterise small time intervals of electricity demand and can vary in sample sizes from less than a hundred [100] to more than five hundred [97]. Their ability to handle large data sets makes them a suitable choice for characterising domestic electricity demand in this instance. However, similar to time series methods much of the literature to date in this area has focussed on other sectors such as industrial and commercial customers. Chapter 8 specifically applies clustering techniques to characterise electricity consumption for the domestic sector.

2.5 Conclusion

The literature presented in this chapter provides a review of existing approaches to electricity load profile characterisation. A gap in the literature exists mainly in terms of applying time series and clustering techniques to characterise electricity consumption at an individual dwelling level for small time intervals (≤1 hour). Previous research in this area has focussed on sample sizes of less than 1,000 and mainly for other sectors such as commercial and industrial customers. However, the research presented here for the most part is focussed on applying methods capable of characterising large samples of domestic electricity customers (≥1,000) at small time intervals.

In addition, identifying common representative patterns of electricity use within the home has previously never existed before. Most research to date has either focussed on providing highly averaged electricity load profiles for the sector using methods such as linear regression, or alternatively engineering and probabilistic techniques which often tend not to be representative of the general domestic building stock. In addition, the ability to correlate electricity load profiles with dwelling, occupant and appliance characteristics, thus allowing individual households to be identified as to how they consume electricity, based solely on these characteristics has also never been done previously before.

CHAPTER 3

OVERVIEW OF SMART METERING DATASET

3 OVERVIEW OF SMART METERING DATASET

3.1 Introduction

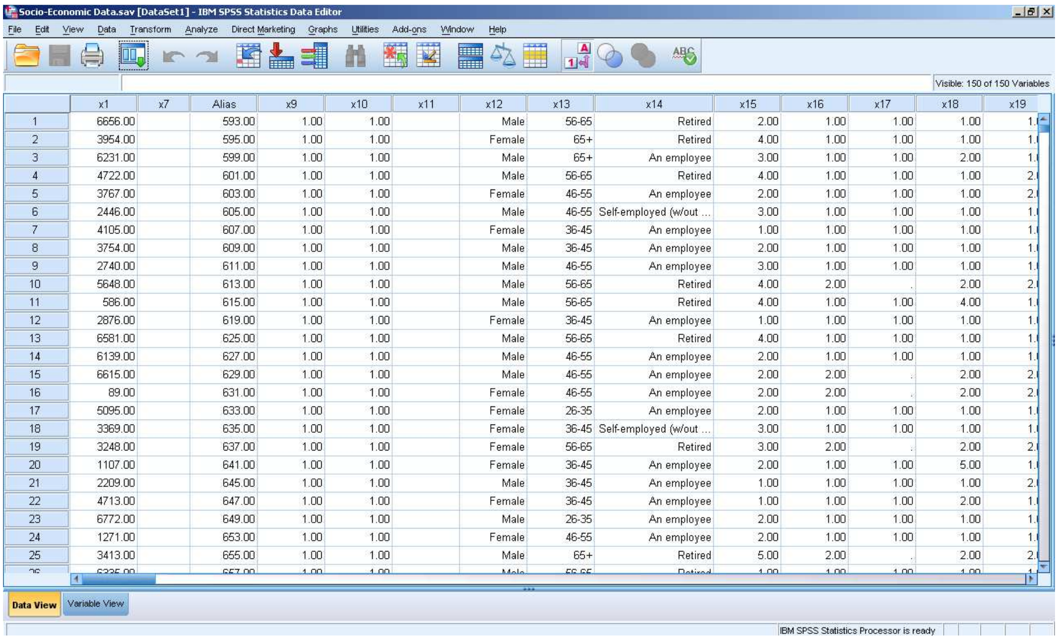
This chapter gives an overview of the smart metering dataset used throughout the research. A full description of the dataset in terms of sample sizes and the associated dwelling, occupant and appliance characteristics recorded by the survey are presented. The software packages (Matlab and SPSS) used to carry out the analysis, along with a sample of the data, is also presented.

3.2 Smart Metering Trial Overview

Smart metering is entering a new phase with completion of pilot projects and plans for national rollouts in a number of EU countries over the next few years. In particular countries such as Finland, France, Italy, Ireland, Netherlands, Norway, Spain, Sweden, UK and Malta have clear legal and regulatory frameworks for installing smart meters nationwide [3].

The smart metering pilot trial was carried out by the Commissioners for Energy Regulation (CER) and Electric Ireland and ended in December 2010. The overall objective was to conduct a nationally representative smart metering trial in order to assess the costs and benefits of smart meters and to inform decisions relating to the full rollout of a national smart metering programme. A series of reports were published presenting the results for customer behaviour trails, technology trials and a cost-benefit analysis for a nationwide rollout [108]. In July 2012, the CER decided to proceed with a national roll out to be completed by 2019 [68].

In order for the smart metering trial to be representative at a national level approximately 5,000 residential dwellings were metered at half hourly intervals as well as recording a detailed list of socio-economic, demographic and dwelling characteristics for each individual household. A full listing of the questions collected by the survey is shown in Appendix D. A sample of the socio-economic data recorded in SPSS (the statistical package used in the research analysis) is shown in Figure 3.1.



	x1	x7	Alias	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18	x19
1	6656.00		593.00	1.00	1.00		Male	56-65	Retired	2.00	1.00	1.00	1.00	1.
2	3954.00		595.00	1.00	1.00		Female	65+	Retired	4.00	1.00	1.00	1.00	1.
3	6231.00		599.00	1.00	1.00		Male	65+	An employee	3.00	1.00	1.00	2.00	1.
4	4722.00		601.00	1.00	1.00		Male	56-65	Retired	4.00	1.00	1.00	1.00	2.
5	3767.00		603.00	1.00	1.00		Female	46-55	An employee	2.00	1.00	1.00	1.00	2.
6	2446.00		605.00	1.00	1.00		Male	46-55	Self-employed (w/out ...	3.00	1.00	1.00	1.00	1.
7	4105.00		607.00	1.00	1.00		Female	36-45	An employee	1.00	1.00	1.00	1.00	1.
8	3754.00		609.00	1.00	1.00		Male	36-45	An employee	2.00	1.00	1.00	1.00	1.
9	2740.00		611.00	1.00	1.00		Male	46-55	An employee	3.00	1.00	1.00	1.00	1.
10	5648.00		613.00	1.00	1.00		Male	56-65	Retired	4.00	2.00	.	2.00	2.
11	586.00		615.00	1.00	1.00		Male	56-65	Retired	4.00	1.00	1.00	4.00	1.
12	2876.00		619.00	1.00	1.00		Female	36-45	An employee	1.00	1.00	1.00	1.00	1.
13	6581.00		625.00	1.00	1.00		Male	56-65	Retired	4.00	1.00	1.00	1.00	1.
14	6139.00		627.00	1.00	1.00		Male	46-55	An employee	2.00	1.00	1.00	1.00	1.
15	6615.00		629.00	1.00	1.00		Male	46-55	An employee	2.00	2.00	.	2.00	2.
16	89.00		631.00	1.00	1.00		Female	46-55	An employee	2.00	2.00	.	2.00	2.
17	5095.00		633.00	1.00	1.00		Female	26-35	An employee	2.00	1.00	1.00	1.00	1.
18	3369.00		635.00	1.00	1.00		Female	36-45	Self-employed (w/out ...	3.00	1.00	1.00	1.00	1.
19	3248.00		637.00	1.00	1.00		Female	56-65	Retired	3.00	2.00	.	2.00	2.
20	1107.00		641.00	1.00	1.00		Female	36-45	An employee	2.00	1.00	1.00	5.00	1.
21	2209.00		645.00	1.00	1.00		Male	36-45	An employee	1.00	1.00	1.00	1.00	2.
22	4713.00		647.00	1.00	1.00		Female	36-45	An employee	1.00	1.00	1.00	2.00	1.
23	6772.00		649.00	1.00	1.00		Male	26-35	An employee	2.00	1.00	1.00	1.00	1.
24	1271.00		653.00	1.00	1.00		Female	46-55	An employee	2.00	1.00	1.00	1.00	1.
25	3413.00		655.00	1.00	1.00		Male	65+	Retired	5.00	2.00	.	2.00	2.

Figure 3.1: Sample socio-economic data file in SPSS

The collection of electricity consumption data from individual households commenced on the 1st July 2009, in order to ensure all infrastructure was operating effectively. The trial officially started on the 1st January 2010 and lasted until the 31st December 2010. As a large amount of data was collected from individual households, Matlab was used to carry out any manipulation and analysis of the data. A sample of the electricity consumption data recorded in Matlab is shown in Figure 3.2. The link between the socio-economic SPSS file shown in Figure 3.1 and the electricity consumption Matlab file was maintained through a unique service ID (alias).

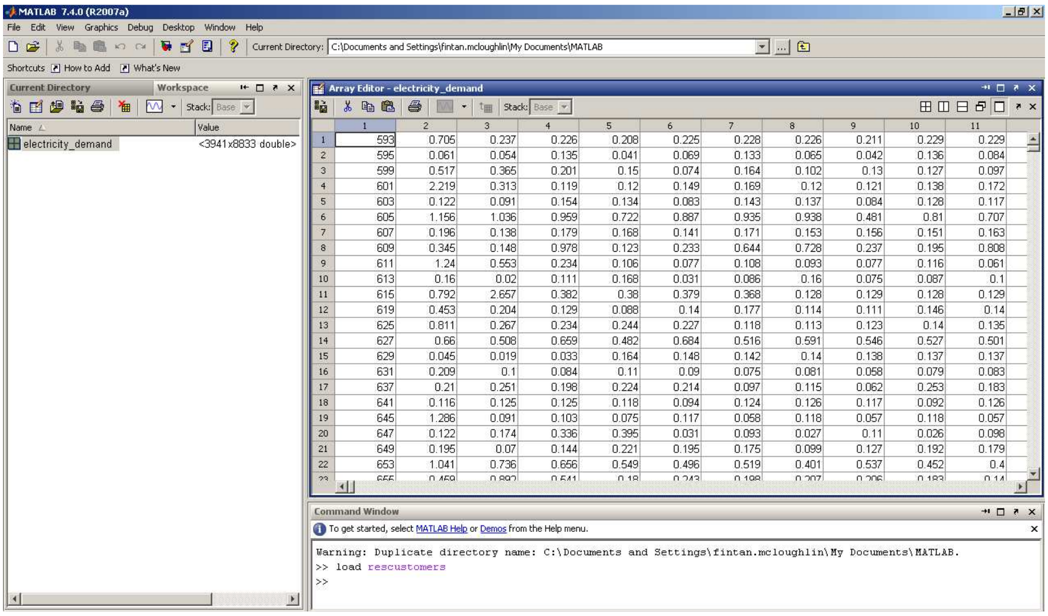


Figure 3.2: Sample electricity demand data file in Matlab

The survey questionnaire was designed to capture the main components within a home in Ireland such as household demographics, dwelling and appliance characteristics and household investment in a number of energy efficiency initiatives. Analysis of the participant responses by CER determined that the households were broadly representative of the national population [5].

During the smart metering trial period, households were subject to different tariffs and customer behaviour stimuli. This was done in order to test the effectiveness of the different tariff structures and behavioural stimuli at reducing overall and peak time usage. This is in contrast to the current offering of a single flat tariff rate to residential customers for electricity irrespective of the time of day at which it is consumed. The type of behavioural stimuli subjected to individual customers includes: bi-monthly detailed bill, monthly detailed bill, bi-monthly detailed bill + In-House Display (IHD), bi-monthly detailed bill + overall load reduction scheme. The breakdown and sample size for each category is shown in Table 3.1.

Table 3.1: Smart Metering Sample Sizes per Tariff Structure and Stimuli [5]

	Bi-monthly detailed bill	**Monthly detailed bill**	**Bi-monthly detailed bill + IHD**	**Bi-monthly detailed bill + overall load reduction scheme**	**Total**
Tariff A	342	342	342	342	1,368
Tariff B	127	129	127	128	511
Tariff C	342	342	343	343	1,370
Tariff D	127	129	126	127	509
Weekend	--	--	--	--	100
Control Group	--	--	--	--	1,170
Total	**938**	**942**	**938**	**940**	**5,028**

The effect of different tariffs and behaviour stimuli is not part of this research and is covered under the published reports from the CER [5][109]. However, it is the total sample sizes and the periods of collection which are important for the purposes of the research presented here. The final datasets used in the research are presented in Table 3.2.

Table 3.2: *Dataset I* and *Dataset II* used throughout the research

	No. of customers	**Period of collection**	**Array size**	**Total No. of entries**
Dataset I	3,941	1st July - 31st December 2009	3941 x 184	725,144
Dataset II	509	1st July 2009 - 30th June 2010	509 x 365	185,785

In order to use the full sample size, any analysis would need to be carried out prior to households being subjected to tariffs or behaviour stimuli, as these would influence households' behaviour enticing them to reduce overall usage and peak demand. *Dataset I* describes the full sample for the period (1st July 2009 to 31st December 2009) prior to any tariffs or stimuli being imposed on the household. *Dataset I* had to be trimmed in size from circa 4,928 households to just 3,941 in total. This was done in order to remove erroneous data for households where communication with the meter was lost for a period of time. In such instances, zero electricity consumption was recorded thus severely impacting individual households load profile.

Dataset II described above in Table 3.2 contains the control group for the sample. As these households were not subjected to any tariffs or behaviour stimuli over the period of the trial (1st July 2009 to 30th June 2010) the sub-sample could also be used in the analysis. Similar to *Dataset I*, erroneous periods of non-communication with the meter were removed from the dataset. This resulted in *Dataset II* being trimmed from 1,100 to 509 households in total. *Dataset II* had to be trimmed by a greater amount compared to *Dataset I* due to a doubling of the period of collection and hence there was a greater probability of a households' smart meter malfunctioning or a data communications breakdown.

Both datasets described Table 3.2 were used in the analysis (*Dataset I* was used in Chapters 5, 6, 8 and *Dataset II* was used in Chapter 7). The reason for this can be explained as follows:

- access to data - as the research was carried out in parallel with the collection of data, *Dataset I* was the earliest available. As a result of its early availability and large sample size this dataset was mostly used throughout the research

- sample size - *Dataset I* is the larger of the two sample sizes by a factor of seven compared to *Dataset II*. *Dataset II* is also fully representative at a national level on account of it being the control group. The smaller, *Dataset II*, was used in Chapter 7 on account of the time series characterisation processes being computationally demanding for *Dataset I*.

The following section presents tables and figures relating to the household characteristics. The results presented are based on own calculations from *Dataset I* using the software application SPSS.

3.3 Smart Metering Trial Dwelling Characteristics

The dwelling characteristics collected by the smart metering dataset include the following: dwelling type, tenure, period of construction, floor area, number of bedrooms, space heating fuel type, water heating fuel type, cooking fuel type, Building Energy Rating (BER), percentage Compact Fluorescent Lights (CFL), percentage double glazing, presence/absence of hot water lagging jacket, attic insulation and external wall insulation. The main dwelling attributes used throughout the research from the smart metering dataset [6] are presented in the following figures.

Figure 3.3 shows percentage penetration by dwelling type for the smart metering dataset. Detached dwellings, which includes bungalows, make up the majority of the domestic building stock within Ireland. Apartments only account for just 1.7% of the overall building stock within the smart metering survey. This category of dwelling is significantly under represented when compared against national census data from 2011, where apartments accounted for approximately 10% of the overall building stock [110]. This can be explained by the exclusion of short terms tenancies to reduce the probability of attrition from the smart metering trial, thus resulting in an under representation of apartment dwellings in the sample [5]. The remaining categories are broadly in line with that recorded by the 2011 census detached (42%), semi-detached (28%), terraced (17%) and apartment (10%) [110].

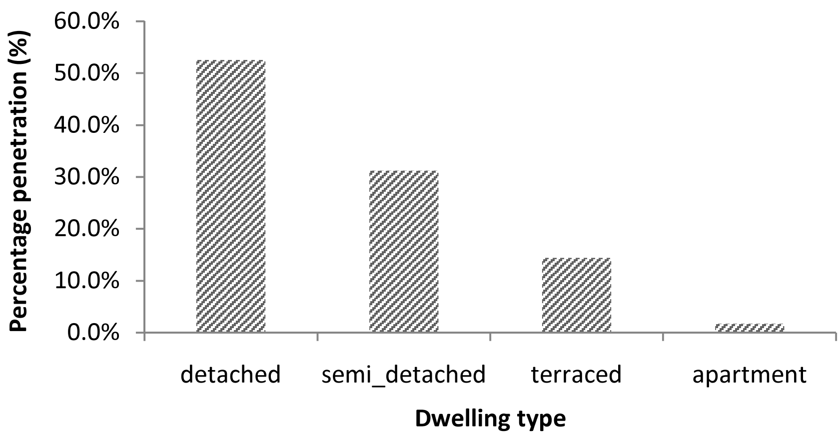


Figure 3.3: Percentage penetration by dwelling type

Figure 3.4 shows percentage penetration for dwelling number of bedrooms. Three bedroom dwellings are the most common, representing approximately 45% of the domestic building stock. The national census does not enquire as to the number of bedrooms (only the total number of rooms) within a dwelling. As a result, the Irish National Survey of Housing Quality (NSHQ) which was carried out in 2001/02 was used for comparison purposes. The number of one (1%) and two (8%) bedroom dwellings within the smart metering dataset compared to the NSHQ (3% and 11% respectively) shows these type of dwellings were slightly under represented [111]. Similar to that discussed above this was most likely a result of an under representation of apartment dwellings due to exclusion of short term tenancies. The remaining categories for three (46% and 45%), four (30% and 34%) and 5+ (9% and 11%) bedroom dwellings differed slightly between the smart metering and NSHQ datasets respectively.

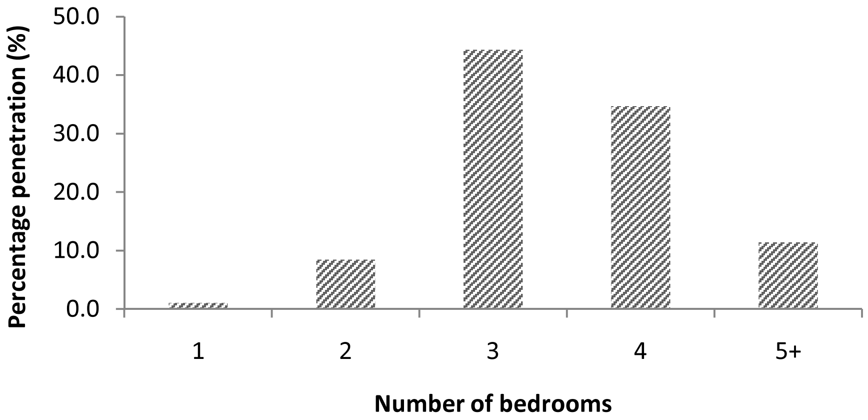


Figure 3.4: Percentage penetration by dwelling number of bedrooms

Figure 3.5 shows space heating fuel type penetration for the smart metering dataset. A comparison could not be made with another independent dataset as the census does not record this data and the NSHQ categories were somewhat different making a direct comparison unattainable. Oil, gas and solid fuel such as coal, peat or wood are the most common type of fuel source for space heating within Irish dwellings. Electric heating penetration was quite small with less than 5% used for central heating systems such as storage heaters and plug in convector type heaters in each instance. This is consistent with figures published through the Household Budget Survey (HBS) with only 3% of households using electricity for their household central heating system [112].

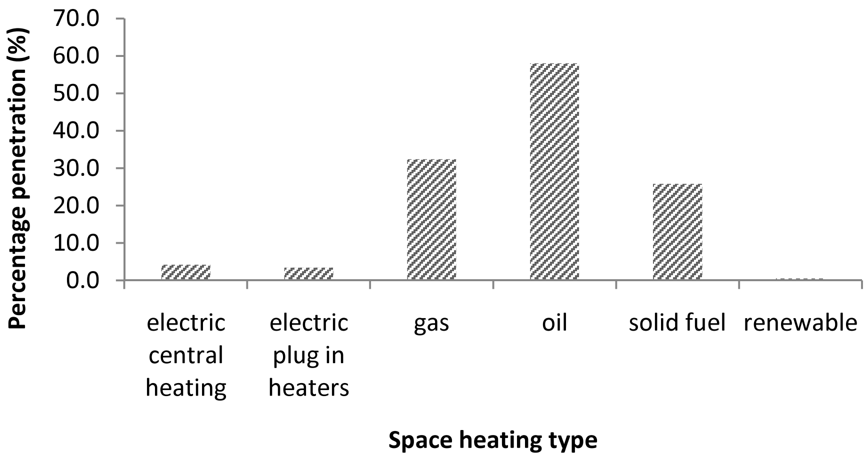


Figure 3.5: Percentage penetration by space heating fuel type

Figure 3.6 shows water heating fuel type penetration for the smart metering dataset. The smart metering survey allowed households to choose multiple heating types, therefore resulting in an overall percentage penetration of greater than 100%. This also explains the high percentage penetration for immersions making it the primary method for heating water in the home. However, the NSHQ survey showed that although a high percentage of households have an immersion system installed in their home (76%), only 10% of households use it as the primary method for heating water [111]. In contrast to the smart metering survey, the NSHQ reported that in 82% of households the main method of heating water was with a central heating system [111]. Although there appears to be a large discrepancy between both surveys, when oil and gas categories, which are primarily used in central heating systems in the home [112] are included as part of the central heating category the figures are comparable.

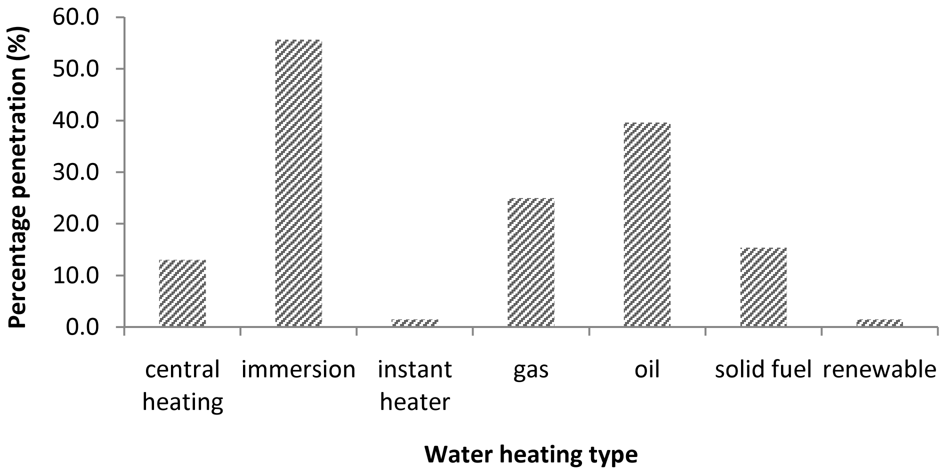


Figure 3.6: Percentage penetration by water heating fuel type

Figure 3.7 shows cooking fuel type penetration for the smart metering dataset. Electricity (70%) is by far the most common method used for cooking, with gas used in only about 20% of households. However, the manner with which the question was phrased within the smart metering survey did not allow for multiple fuels to be selected. Therefore respondents were only able to indicate which method they used most of the time, thus leading to a possible over estimation of electricity used for cooking. In contrast to space heating and water heating figures presented above, data was not available from another Irish source to check the representatively for fuel type end use for cooking within the home. Yao and Steemers [23] reported that electric ovens and hobs have a penetration of 56% and 37% in UK homes respectively. SEAI report on figures from the Department of Business, Enterprise and Regulatory Reform (BERR) in the UK that just 6% of electricity end-use is used for cooking [112]. Further data is required in Ireland on cooking fuel type end use so that the results in Figure 3.7 can be verified.

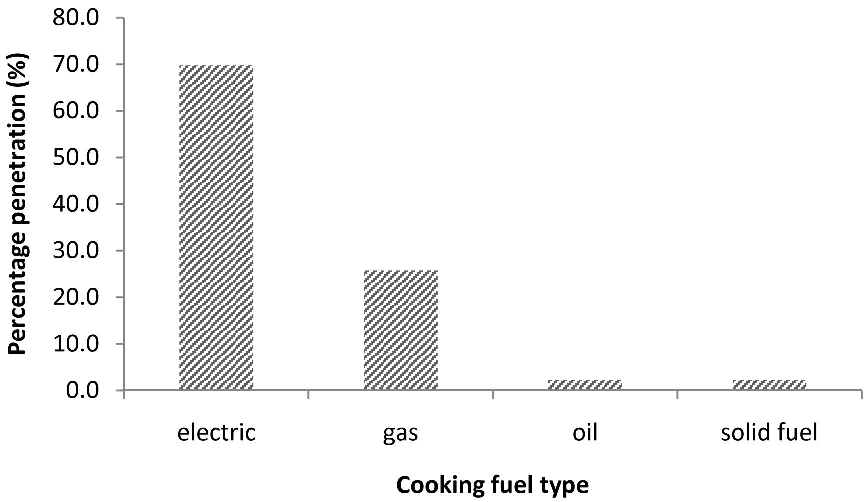


Figure 3.7: Percentage penetration by cooking fuel type

3.4 Smart Metering Trial Occupant Characteristics

The main occupant characteristics collected by the smart metering dataset include the following: gender, Head of Household (HoH) age, HoH employment status, social class, household composition, number of occupants, occupancy, HoH education level, household income.

Figure 3.8 shows HoH age penetration for the smart metering dataset. The number of age categories was reduced from six to three overall in order to have three larger groups representing young, middle aged and older HoH's. When compared against the NSHQ, the proportion of age group <36 years (12%) was slightly under represented in the smart metering trial (10%) [111]. This again is most likely related to the exclusion of short term tenancies from the trial [5]. The remaining categories for the smart metering trial, between 36 and 55 years (45%) and 56 years plus (45%) are in line with that recorded by NSHQ (44% and 45% respectively).

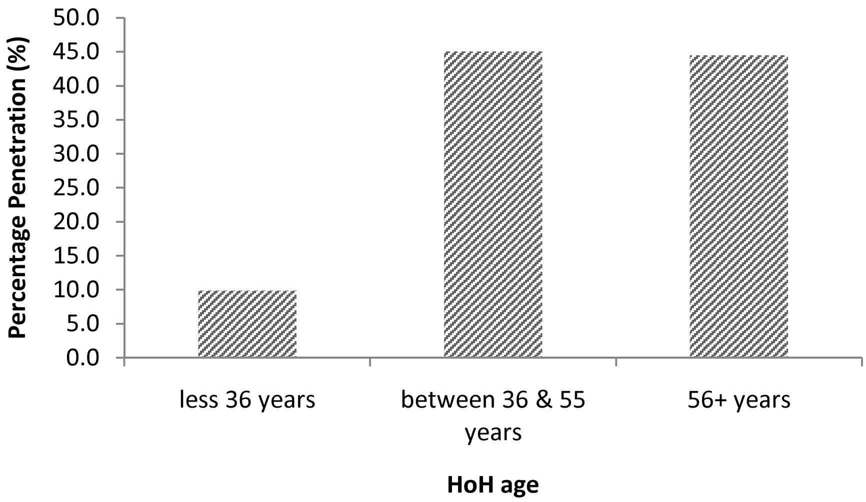


Figure 3.8: Percentage penetration by HoH age

Figure 3.9 shows social class for households within the smart metering dataset. Social class was based on the UK National Readership Survey (NRS) social grade system [113]. A full listing of the categories used and their descriptions are contained in Appendix E. Five categories were reduced to four from the smart metering survey by combining categories C1 and C2. Category AB corresponding with the middle class and upper middle class represent around 15% of the households. The largest category C (43%) corresponds to households of lower middle and skilled working classes. Category DE (38%) describes the working class and those on the lowest level of subsistence. Finally social class F (3%) corresponds to farmers which represent only a small proportion of the households within the survey. A direct comparison with the 2011 census results could not be made on account of the use of different categories for social class.

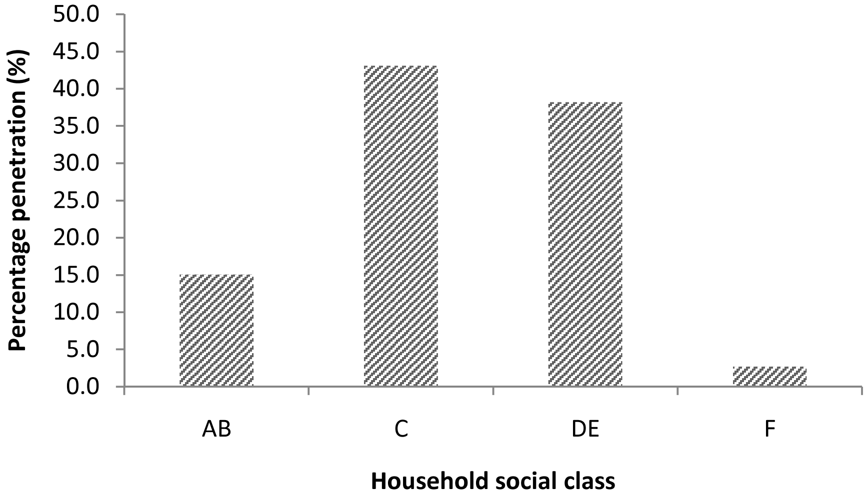


Figure 3.9: Percentage penetration by household social class

Figure 3.10 shows household composition for the smart metering survey. Adults only made up the largest category, however, there was also a high percentage of households with only one occupant living alone. When compared against the NSHQ the results were similar for an adult living alone (22% for NSHQ compared to 19% for the smart metering survey). For the other two categories for the smart metering survey, adults only (52%) living together and adults and children (28%) living together were very different to that recorded by the NSHQ (24% and 53% respectively). It is likely that there is some discrepancy between the results and that adults only category may be significantly over represented and adults and children under represented within the smart metering survey.

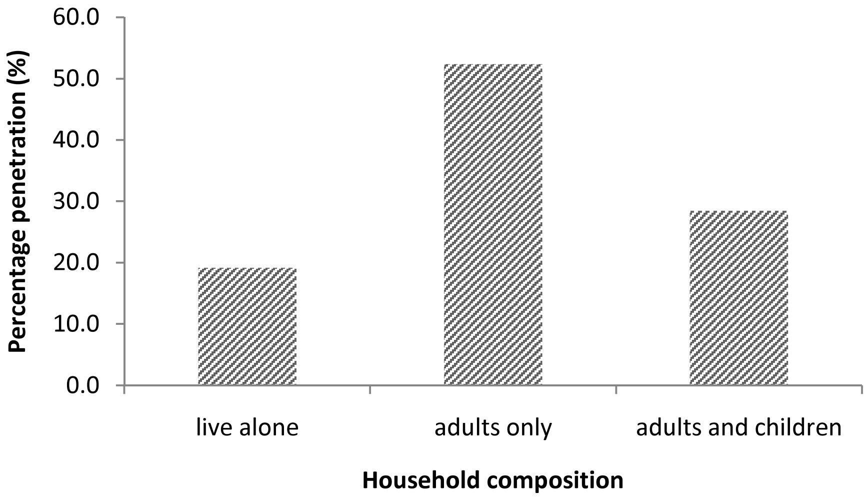


Figure 3.10: Percentage penetration by household composition

Figure 3.11 shows household number of occupants for the smart metering dataset. The most frequent number of occupants living within a household was two occupants. There was a slight difference of around 1% between Figure 3.10 and Figure 3.11 for one occupant households which was likely down to survey error. When the results are compared against that of the census 2011 data one occupant households are under represented by the smart metering survey (20% compared to 24% for the census) [114] where as two occupant households are over represented (32% compared to 29% respectively). The remaining categories: three (17%), four (18%), five (8%) and six plus (4%) occupants are similar between the smart metering survey and census results (18%, 16%, 9% and 4% respectively).

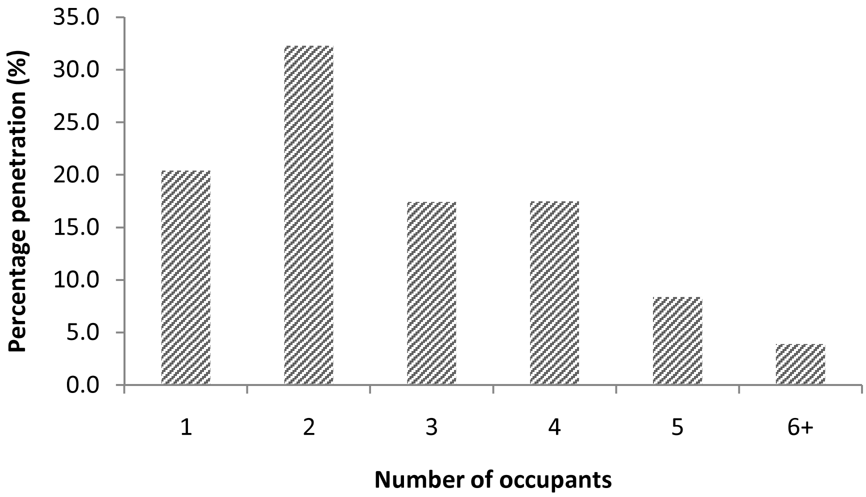


Figure 3.11: Percentage penetration by household number of occupants

Figure 3.12 shows an Efficiency Indicator for individual households. A measure to determine the perceived attitude within a household towards energy efficiency was sought. A question was used from the smart metering questionnaire (Question 36 in Appendix D) in order to infer energy efficiency behaviour within the home. Respondents were asked how much they believed they could cut their bills by making changes in the manner with which they use electricity within the home. The largest group representing over 50% of the sample believed that they could only cut their electricity bills by less than 10%.

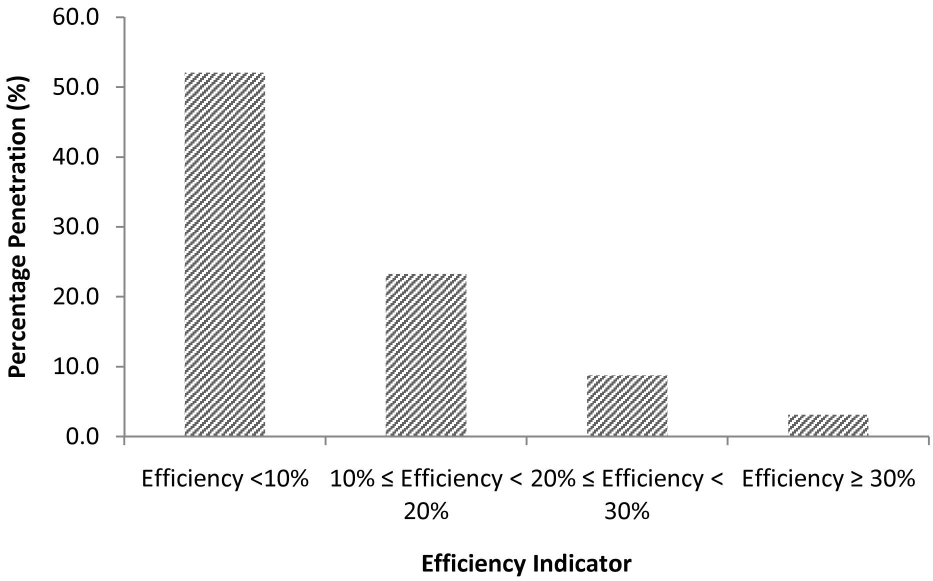


Figure 3.12: Percentage penetration by Efficiency Indicator showing what percentage households believed they could cut their electricity bills by

3.5 Smart Metering Electrical Appliance Characteristics

The number and frequency of use was recorded by the smart metering survey for the following appliances: washing machine, tumble dryer, dishwasher, electric shower (instant), electric shower (pumped), electric cooker, electric heater (plug in convector), stand alone freezer, water pump, immersion, televisions (< 21 inches), televisions (> 21 inches), desktop computer, laptop computer, game console.

Table 3.3 shows the penetration of common household electrical appliances for the smart metering dataset. Nearly every home in Ireland has a washing machine making it the most common appliance recorded by the smart metering survey. Tumble dryers and dishwashers had approximately the same level of penetration at 68% and 67% respectively. Instant electric showers are far more prevalent than pumped showers in Irish homes. Electrical cookers are also very common in Irish households with a penetration of 77%. Plug in electric heaters have one of the lowest penetrations of appliances along with water pump with respective penetrations of 31% and 20%. Approximately half of all Irish households own a stand-alone freezer. The penetration of immersions used for domestic hot water heating has a high penetration of 77% but are not necessarily used all the time. Larger televisions (greater than 21 inches) have an almost 20% higher penetration than smaller ones (less than 21 inches) in Irish households. Multiple televisions are also a common feature in Irish homes. Both desktop and laptop computers had similar penetrations, with multiple laptops per household more common than desktop computers. Finally, game consoles had an overall penetration of 33% within Irish dwellings. The HBS survey 2009/10 also recorded the penetration of some of the same electrical appliances within the home as the smart metering survey [115]. The results are consistent with those presented for the smart metering survey, with only major differences between appliance type freezer (35.3%) and computer (77.3%) recorded by the HBS. The difference in percentage penetration between appliance type computer was probably contributed too by the smart metering survey collecting information on both desktop and laptop computers thus splitting the category.

Table 3.3: Penetration of common household electrical appliances

Appliance type	**Percentage penetration for number of appliances (%)**				
	1	**2**	**3**	**4**	**Total**
Washing machine	97.6	0.6	--	--	98.2
Tumble dryer	68.2	0.1	--	--	68.3
Dishwasher	66.7	0.2	--	--	66.9
Electric shower (instant)	63.6	5.1	0.5	--	69.2
Electric shower (pumped)	26.6	2.1	0.4	--	29.1
Electric Cooker	76.8	0.3	--	--	77.1
Heater (plug in convective)	23.5	5.2	1.8	--	30.5
Freezer (stand alone)	47.9	1.8	0.1	--	49.8
Water pump	19.1	0.4	0.1	--	19.6
Immersion	76.4	0.3	--	--	76.7
TV (<21 inches)	39.8	17.9	5.8	2.1	65.6
TV (>21 inches)	50.7	25.0	6.0	2.3	84.0
Computer (desktop)	44.6	2.3	0.3	0.2	47.4
Computer (laptop)	42.2	8.4	2.0	0.9	53.5
Game console	22.2	8.3	2.1	0.7	33.3

3.6 Conclusion

This chapter presented an overview of the smart metering dataset. A description of the dataset was given along with a sample of the data and the software packages used in the analysis. The chapter showed that in general the dataset can be regarded as representative at a national level and compares well with other similar studies carried out in the past. However, an abnormality with the dataset was identified in terms of the exclusion of short term tenancies which had a knock on effect as to the representatively of certain categories such as number of apartments and younger HoH's being under represented within the sample.

CHAPTER 4

METHODOLOGY

4 METHODOLOGY

4.1 Introduction

This chapter outlines the techniques, methodologies and validation approaches used throughout the research to characterise domestic electricity demand. The methods used were selected based on the literature review presented in Chapter 2 and also on account of the ability to meet the objectives outlined in Section 1.5. An initial overview of the methodological process involved is first given in Sections 4.2 to 4.5 before a more in depth description of each characterisation approach applied in the following chapters is presented in Section 4.6.

4.2 Data and Averaging

The datasets used throughout the research were described in Section 3.2. *Dataset I* was used for all but one characterisation approach on account of the larger sample size. This resulted in larger sample sizes for each dwelling and occupant category thus helping to improve representation, particularly for smaller uncommon categories. *Dataset II* was used in Chapter 7 due to its smaller size and hence was computationally less demanding.

The research was primarily focussed on characterising domestic electricity demand over a 24 hour period. Due to the size of the datasets (*Dataset I* – 35 million entries, *Dataset II* – 8.5 million entries) some level of averaging was required in order to reduce the data to a suitable format for the presentation of results. The data averaging was carried out in two ways:

Time Domain Averaging (Longitudinal)

Time domain averaging consists of averaging across a particular time period such as a day, week, month, six-months or a yearly period. In doing so, information is lost on the intra-daily and seasonality components to the electricity demand load profile. This type of averaging was mostly applied throughout the research in order to reduce the data.

Space Domain Averaging (Cross-sectional)

Another approach is to average across individual households within the dataset. This enables the intra-daily and seasonality components to remain but information is lost on each individual household. This type of averaging was carried out in Chapter 8 as part of the clustering approach to electricity load profile characterisation.

4.3 Characterisation

A number of different approaches were used to characterise household electricity consumption each with their own advantages and disadvantages. The methods presented in the following chapters are:

- parameterisation - a number of electrical parameters were used to characterise daily domestic electricity demand. A multivariate linear regression was used to link parameters to dwelling and occupant characteristics;
- autoregressive - a Markov chain approach was used to characterise daily domestic electricity demand in terms of a probability transitional matrix;

- time series - Fourier transforms and Gaussian processes were used to characterise daily domestic electricity demand. A multivariate linear regression was used to determine the influence of dwelling and occupant characteristics on patterns of electricity use; and
- clustering - a data mining process was used to segment daily domestic electricity demand into groups based on similar patterns of electricity use throughout the day. A multi-nominal logistic regression was used to link patterns of electricity use to dwelling and occupant characteristics.

A more detailed explanation as to how each one of these processes was applied in later chapters is presented in Section 4.6.

4.4 Validation

4.4.1 Electrical Parameters

A number of electrical parameters were used to characterise domestic electricity demand in Chapter 5. These same parameters were also used to validate the approaches applied in Chapters 6, 7 and 8. The parameters describe the main features of domestic electricity load profiles and include Total Electricity Consumption, Maximum Demand, Load Factor and ToU of maximum electricity demand. These are commonly used in the electricity industry for billing and describing profile characteristics.

Equation 4.1 shows total electricity consumption, E_{TOTAL} which is the total amount of electricity consumed over a period in kWh where E_i^j is average electrical demand in kW for each half hour period on day j, n is the total number of periods in a day and m is the total number of days.

$$E_{Total} = \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{n} E_i^j \qquad 4.1$$

This parameter characterises the total amount of electricity used across a time period. It was chosen as this parameter is currently used for billing domestic customers, which is done on a bi-monthly basis in Ireland.

Equation 4.2 describes mean daily maximum demand, E_{MD} over a time period in kW. E_{MD} refers to the largest value of electrical demand in a day, averaged over a time period where E_i^j is average electrical demand in kW for each half hour period on day j, n is the total number of periods in a day and m is the total number of days.

$$E_{MD} = \frac{1}{m}\sum_{j=1}^{m} max\{E_i^j, 1 \leq i \leq n\} \qquad 4.2$$

This parameter was chosen as it characterises the largest value of electricity demand across a 24 hour period. This parameter is often used as part of billing non-domestic electricity customers. However, it is also of use in the domestic case as it is a defining characteristic of the load profile shape.

Daily load factor, E_{LF} is a ratio and is shown in Equation 4.3 below where E_i^j is average electrical demand in kW over each half hour period on day *j*, n is the total number of periods in a day and m is the total number of days.

$$E_{LF} = \frac{1}{m}\sum_{j=1}^{m} \frac{\frac{1}{n}\sum_{i=1}^{n} E_i^j}{max\{E_i^j, 1 \le i \le n\}} \qquad 4.3$$

It is a measure of daily mean to daily maximum electrical demand and is a measure of the "peakiness" of a households load profile. Typically, larger values of E_{LF} correspond to households who consume electricity more evenly across the day whereas a low E_{LF} indicates small intervals of large electricity consumption.

A maximum ToU parameter, E_{ToU} over a period is defined by Equation 4.4 below where E_i^j represents average electrical demand in kW over each half hour period on day *j*, n is the total number of periods in a day (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00) and m is the total number of days.

$$E_{ToU} = mode\{max\{E_i^j, 1 \le i \le n\}, 1 \le j \le m\} \qquad 4.4$$

ToU indicates the time of day at which maximum electricity demand occurs and is important as this parameter characterises the most likely time at which they will consume most electricity demand.

These parameters are used for validation purposes, in order to assess the performance of each characterisation technique. Paired sample t-tests are used to determine whether there are significant differences between the sample and the characterised profiles for each individual approach. In addition these parameters are also used to determine the main drivers of domestic electricity demand in terms of dwelling and occupant characteristics in Chapter 5.

4.4.2 Time Series Tests

A number of time series tests are used to interrogate the temporal properties of each characterisation approach further. These ensure that the characterised time series not only accurately reflects the magnitude of electricity demand across a 24 hour period but also that it occurs at appropriate times of day. Four tests were used: time series plot, frequency histogram, auto-correlation and power spectral density functions.

A time series plot is used to visually compare sample data with that of characterised electricity load profiles. It is a less analytical approach to the other methods; however, it gives a good indication of the performance of each technique at charactering the magnitude and timing of electricity use within the home.

A frequency histogram is used for visually comparing the performance of different characterising techniques. In particular certain characterisation methods may be better at characterising small values of electricity demand across the day whereas other approaches are better at characterising larger values of demand. A frequency histogram visually indicates this clearly.

The autocorrelation function is used for investigating the temporal properties of an electricity load profile. The function regresses a time series $f(t)$ onto itself with a time lag τ in order to determine whether a pattern exists. The autocorrelation function is shown in Equation 4.5 for a function *f* and where $\bar{f}$ is the complex conjugate of the time series signal [116].

$$f * f = \int_{-\infty}^{+\infty} \bar{f}(\tau)f(\tau + t) * d\tau \qquad 4.5$$

As domestic electricity load profiles sometimes exhibit cyclical patterns on an intra-daily basis, this function will show whether the properties are transferred between sample and characterised demand load profiles.

The Power Spectral Density (PSD) function, $P_y(v)$ as calculated by the Fast Fourier Transform (FFT) and was used to describe the temporal properties of the original sample and characterised time series *y(t)*. The function is shown in Equation 4.6 where *v* is the frequency in rad/sec, *t* is the time interval in half hour periods and *T* is the length of the time series [117].

$$P_y(v) = \lim_{T\to\infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} [y(t) - \bar{y}] e^{-2\pi i v t} dt \right|^2 \qquad (4.6)$$

Evaluating the PSD of a time series identifies periodicities within an electricity load profile. The function quantifies the exact frequency of occurrence of a pattern and the overall contribution made by each individual frequency component.

.

4.5 Associating dwelling and occupant characteristics through regression

Regression is used to link dwelling and occupant characteristics to electricity load profiles. Depending upon the characterisation approach this takes on two different forms. In Chapters 5 and 7, multivariate linear regression (described earlier in Equation 2.1) is used as the dependent variable is of type interval. In Chapter 5 the dependent variables is a series of electrical parameters where as in Chapter 7 it is represented by a series of characterisation coefficients. The influence of explanatory variables (X_1, X_2, etc) such as dwelling and occupant characteristics are then investigated as to their association over each dependent variable. The significant influence of each explanatory variable is measured using a p-value. P-values of 0.1, 0.05 and 0.01 represent significance levels at 90% 95% and 99% respectively. The coefficient or β value shows the level of influence over the dependent variable and whether it is positive or negative.

Where the dependent variable is of type nominal (i.e. qualitative as opposed to quantitative) as is the case in Chapter 8, either binary logistic (output has two categories) or multi-nominal logistic regression (output has more than two categories) is used. The expression for binary logistic regression is shown in Equation 4.7 where *p(x)* ranges from 0-1 and *logit [p(x)]* is called the odds or likelihood ratio that the dependent variable is 1 [118].

$$logit[p(x)] = ln\left[\frac{p(x)}{1-p(x)}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \quad (4.7)$$

Similar to multivariate linear regression, explanatory variables represent the dwelling and occupant characteristics that are being investigated for their influence over the dependent variable. The same measure of significance for each variable is used and instead of using the beta coefficients (i.e. β_1, β_2, etc), its exponential value ($Exp(\beta)$) is used to gauge influence of a particular characteristic on the dependent variable.

4.6 Methodologies

The following sub-sections 4.6.1 to 4.6.4 describe the specific techniques and methodologies applied in each chapter throughout the research.

4.6.1 Statistical Analysis using Daily Parameters

Statistical approaches were described in Chapter 2 and have been widely used in past literature to describe household electricity use [27][28][65]. The statistical approach applied in Chapter 5 uses the same four electrical parameters defined in Equations 4.1 – 4.4 above to characterise domestic electricity demand. Each electrical parameter is evaluated for *Dataset I* and include: E_{Total}, E_{MD}, E_{LF} and E_{ToU}. As discussed earlier, the parameters describe important characteristics of domestic electricity consumption across a daily period. A longitudinal averaging process was applied to each daily parameter except for E_{Total} where no averaging was required.

A multivariate linear regression, shown in Equation 2.1 was applied to associate electrical parameters to dwelling and occupant characteristics. Two different approaches were used: first looking at dwelling and occupant characteristics and secondly looking at individual appliances that influenced electricity consumption patterns in the home. The first approach, Dwelling and Occupant Characteristics (DOC), describes the variables that influence electricity use in the home such as HoH age and number of bedrooms etc. These variables do not "consume" electricity but serve to influence occupants demand within the home and may help explain the underlining causes of different patterns of electricity use. The second approach, Electrical Appliances (EA), looks directly at the individual appliances and describes the direct relationship between their ownership and use on electricity consumption within a household. The EA approach helps to further understanding to electricity use within the home but does not explain underlining causes.

The decision to carry out the analysis using two separate models was taken on the grounds of the evaluation of a reduced coefficient of determination when all variables were lumped together into the same model. This meant that the model was less able to explain the variation in electricity consumption than when two separate models were used. This was likely down to the effect of multi-collinearity between variables causing a reduction in the coefficient of determination. Therefore the decision was made to examine the effect of appliances and dwelling and occupant characteristics separately as to their influence on home electricity use.

4.6.2 Autoregressive (Markov Chain)

Although highly variable, electricity consumption in the home does follow certain patterns. For example when an electrical appliance is switched on, often it is left on for a period of time before it is switched off again [25]. Therefore past values of electricity demand should be a good indication of future values of electricity demand (most of the time). It is based on this observation that Markov chains were investigated in order to characterise domestic electricity consumption patterns. However, traditionally Markov chains have been applied in situations where the times series is stationary (i.e. is independent of time).

Autoregressive Markov chains were presented in Chapter 2 Section 2.2.6 and are evaluated in Chapter 6. They are based on the construction of a transitional probability matrix where the transition from one discrete state to another is characterised in terms of its probability. A first order Markov chain compares the current state and the one immediately preceding it to calculate the probability of going to the next state. A second order Markov chain compares the two previous states with the current state to determine the next state. For a first order Markov chain, the transitional probability matrix, P, can be defined with $p_{k,k}$ probabilities for k states as shown in Equation 4.8 [47].

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,k} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,k} \end{bmatrix} \tag{4.8}$$

Each element in the matrix *P* represents the probability of going from one discrete state to the next. For each group of states shown in the transitional matrix (i.e. each row) the cumulative probability equals one. This represents the relative probability of changing from the current to every other state including the present state at the next time interval.

In order to characterise each individual electricity load profile successfully the following three processes need to be carried out:

- an individual dwelling's electricity demand load profile is sampled;
- a transitional probability matrix is derived; and
- finally a load profile is generated.

The execution of all three processes requires a significant amount of computer processing time to characterise each individual electricity demand load profile. Therefore the decision was made not to use the full sample in the following analysis and instead test the methodological approach on a much smaller sample size. Four individual demand load profiles for each dwelling type (detached, semi-detached, terraced and apartment) were selected at random from *Dataset I.* The disadvantage with this approach means that the four profiles selected may not be representative of either the entire sample or the dwelling type with which they are associated. However, by carrying out the afore mentioned processes, gives a good indication as to the ability of Markov chains to sufficiently characterise domestic electricity demand load profiles and whether any further analysis is warranted on the full sample.

Initially a 12x12 matrix was used to sample the data, however, this resulted in too few values of electricity demand to adequately characterise each household. Therefore a 24x24 transitional probability matrix was applied with bin sizes based on mean (0.5525 kW) and standard deviation (0.0837 kW) for the entire sample. Individual transitional probability matrices were produced for each of the four random dwelling types.

The first state of the Markov chain sequence is generated by a Gaussian distributed random number generator with values between 0 and 1. After the initial state is chosen, the transitional probability matrix was then used to select every other consecutive state after this. The state with the highest probability, which is usually the same state, will be selected most often but will depend upon the probability matrix. This is reflected in the transitional matrix where the largest probabilities are located along the diagonal. A uniformly distributed random number generator is used to choose values between each bin width so that the same value of electricity demand is not repeatedly selected.

Descriptive statistics are presented for each of the four dwelling types. Electrical parameters described above are evaluated and paired sample t-tests used to compare original and characterised profiles. Similar to Chapter 5, a longitudinal averaging process was applied across daily parameters: E_{MD}, E_{LF} and E_{ToU}. Time series tests are used to test the temporal properties of characterisation process. An attempt to correlate dwelling and occupant characteristics through regression was not pursued on the grounds that the Markov chain process failed to adequately characterise the temporal properties for domestic electricity demand load profiles sufficiently as will be shown in Chapter 6.

4.6.3 Time Series Approaches

Time series approaches to electricity load profile characterisation were first presented in Chapter 2 Sections 2.2.4 to 2.2.8. Chapter 7 compares and contrasts these methods, specifically looking at: Fourier transforms, Neural networks, Gaussian processes, Auto-regression, Fuzzy logic, Wavelets and multivariate regression. The most applicable methods are then applied to characterise domestic electricity demand load profiles in order to meet the objectives outlined in Section 1.5.

In order to evaluate the performance of the characterisation process, *Dataset II* was used in the analysis. The decision to use *Dataset II* was made on account of its smaller sample size. The time series techniques evaluated in Chapter 7 were computationally intensive and therefore required a considerable amount of time to characterise individual customers. As *Dataset II* was smaller by a factor of seven but still representative for all customers it was used in the analysis.

Electrical parameters and time series tests were evaluated to determine the accuracy of the characterisation techniques. A multivariate linear regression was then used to associate dwelling and occupant characteristics to domestic electricity demand load profiles, where a longitudinal averaging process was applied across the coefficient values. Median values instead of mean were used to minimise the influence of large outliers produced by the characterisation process.

4.6.4 Clustering

Clustering was discussed earlier in Section 2.2.9 and is applied in Chapter 8. The overall methodology used in Chapter 8 can be broken down into the following three sub-sections:

- Evaluation of clustering methods and number of clusters
- Characterising domestic electricity demand load profiles

- Profile classification by dwelling and occupant characteristics

Evaluation of clustering methods and number of clusters

In order to evaluate the most appropriate clustering approach and the number of clusters a 'validity' index was used. The Davies-Bouldin (DB) index is a ratio of the intra cluster distance (i.e. average distance of all patterns in a cluster to the cluster centre) divided by the inter cluster distance (i.e. the distance between different cluster centres). It is a measure of how compact individual clusters are while maximising the distance between each cluster centre. An expression describing the DB index is shown in Equation 4.9 [119].

$$DB\ index\ (\bar{R}) = \frac{1}{N}\sum_{i=1}^{N} R_i \tag{4.9}$$

where N is the number of clusters, R_i is the similarity measure of cluster i with its most similar cluster. The best choice for number of clusters, is the one that minimises the average system wide similarity $\bar{R}$.

Clustering algorithms: k-mean, k-medoid and SOM are evaluated for *Dataset I*; these use unsupervised learning to segment a dataset into clusters. Unsupervised learning requires the number of clusters to be pre-defined before the process is carried out. Therefore, in order to establish an appropriate number of clusters, each method was evaluated for a range of different values (clusters 2-16).

In order to evaluate the best performing clustering approach a single random day was clustered. Each clustering method was applied for varying number of clusters and the DB index was evaluated. The best performing clustering technique along with the number of clusters was identified by a minimum value for the DB index.

Characterising domestic electricity demand load profiles

The same process described above (for the chosen clustering method and number of clusters) was applied to each day over the six month period for *Dataset I*. Figure 4.1 outlines the overall methodology employed to characterise domestic electricity load profiles. This enables the diurnal, intra-daily and seasonal patterns of electricity demand, as well as the variations between households to be characterised to fulfil the objectives set out in Section 1.5. The process is divided into three stages:

- clustering
- electricity load profile characterisation
- electricity load profiles by day type (weekday, Saturday and Sunday)

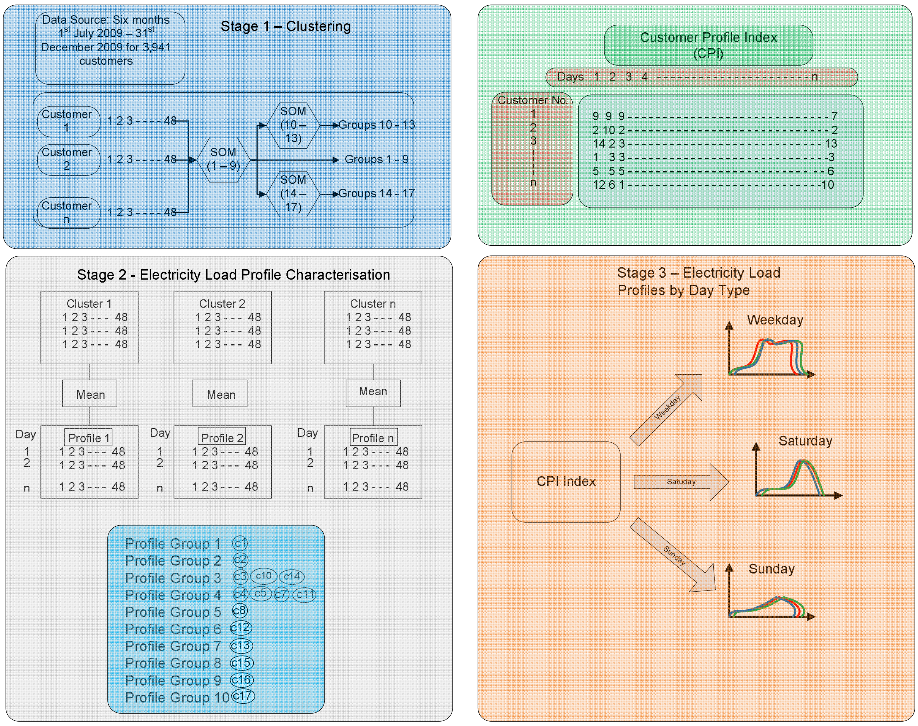


Figure 4.1: Clustering methodology for electricity load profile characterisation using SOM

Clustering

As will be shown later Self Organising Maps (SOM) provided the lowest DB index overall and hence this was evaluated further. The clustering method applied resulted in two very large clusters representing over two thirds of the entire sample. As a result, when dwelling and occupant characteristics were investigated, the level of significance for each category was lost within the two larger clusters. This is not very useful from a practical point of view and therefore a sub-clustering approach was adopted in order to divide the two largest clusters. This technique has been used before most recently by Lo et al. and Zainal et al. [120][121].

Electricity Load Profile Characterisation

Mean half-hourly electricity demand from each cluster was used to produce individual load profiles based on all households using that particular profile on that day. Profiles that showed similar patterns of electricity use were grouped together. This was done in order to reduce the number of similar profiles that differed only slightly in terms of timing and magnitude of electricity use. This resulted in ten electricity load profile groups overall, indicating different representative types of electricity use within the home over the six month period.

It is important to note that as households tend to use electricity differently on an intra-daily basis (as identified earlier in Figure 1.4) each household may use different profile groups over the six month period. In order to show this, a Customer Profile Index (CPI) was produced indicating which profile (i.e. P1 – P10) each household used on a particular day over the six month period. The Household Mode (HMode) of the CPI was calculated using Equation 4.10 and shows which profile group households most frequented for the majority of the time over the six month period.

$$\{HMode_{cust} | 0 < cust \leq 3941\} = mode\{CPI_{cust,day} | 0 < cust \leq 3941, 0 \leq day \leq 184\} \tag{4.10}$$

The Average Percentage Profile Time (APPT) each household spends within each profile group is calculated in Equation 4.11. Individual households are assigned to the HMode profile group and the time spent within this and all other groups over the six month period is calculated. This indicates how often a household will use a particular profile group over the six month period.

$$\{APPT_{HMode} | 0 < HMode \leq 10\} = freq\{CPI_{HMode,Profile\ Group}, 0 < Profile\ Group \leq 10\} \tag{4.11}$$

Electricity load profiles by day type (weekday, Saturday and Sunday)

Finally the CPI index was split by day types: weekday, Saturday and Sunday. This was done in order to identify whether different patterns of electricity use were apparent during the weekdays, where occupancy and behavioural patterns may differ compared to the weekend. Appendix A presents each of the ten characterised electricity load profile groups based on a diurnal, intra-daily and seasonal basis over the six month period.

Profile classification by dwelling and occupant characteristics

In order to establish a relationship between each electricity load profile group and individual dwelling and occupant characteristics a multi-nominal logistic regression was applied as shown in Equation 4.7. The dependent variable is represented by the HMode profile group number (P1 – P10) and the explanatory variable by each dwelling and occupant characteristic. The characteristics used in the analysis are the same as those used in Chapter 5 where Dwelling and Occupant Characteristics (DOC) and Electrical Appliances (EA) were investigated.

The odds or likelihood ratio (Exp(β)) was shown earlier in Equation 4.7. This represents the likelihood of a household using a particular profile group based on their dwelling and occupant characteristics. It is important to note that for small sample sizes within a variable category, the multi-nominal logistic regression sometimes produces an unusually large odds ratio. Consequently if one plots odds ratio for these small sample size categories it generally results in an exceptionally large influence when compared against other categories. Therefore, in addition to carrying out the multi-nominal logistic regression, percentage penetrations for each explanatory variable and profile group are presented in the following figures in order to graphically show the household characteristic composition for each profile group.

4.7 Conclusion

This chapter presented the techniques, methodologies and validation approaches applied throughout the research. Firstly the types of averaging processes applied across the datasets were discussed in order to reduce the data. Secondly, the characterisation processes were introduced which will be used throughout the research. The validation methodology was then presented which consisted of two parts: calculating the electrical parameters and evaluating a number of time series tests in order to determine how successful each method was at characterising domestic electricity load profiles. The methodology of associating dwelling and occupant characteristics to the load profile shape were then presented. Finally the individual methodologies used in each of the forthcoming chapters were presented in detail.

CHAPTER 5

STATISTICAL ANALYSIS USING DAILY PARAMETERS

5 STATISTICAL ANALYSIS USING DAILY PARAMETERS

5.1 Introduction

A statistical regression approach is presented in this chapter to characterise *Dataset I*. Total Electricity Consumption across a six month period as well daily parameters, which describe an electricity demand load profile across a 24 hour period are used. These parameters are then linked to dwelling and occupant characteristics by multivariate linear regression. This process was also used to identify the most significant dwelling and occupant characteristics that influence domestic electricity consumption across the day.

5.2 Discussion and Results

There are two main advantages to regressing the electrical parameters against dwelling and occupant characteristics rather than individual half hourly demand:

- Firstly it removes the need to have multiple equations representing each individual half hourly period, therefore simplifying the analysis and interpretation of results. Instead of having 48 individual equations representing each individual half hourly period, each household is represented by a single value for each parameter. Note that as a result of this the time series tests outlined in the methodology Chapter 4 were not carried out for this particular approach.

- Secondly as electricity use within the home is aggregated in the case of E_{TOTAL} and averaged for the remaining parameters E_{MD} E_{LF} and E_{ToU}, the probability of a significant relationship increases. This is due the highly variable use of electricity across the day as was described in Figure 1.2. These parameters help smooth out this variability thus making the relationship between household characteristics and electricity use within the home less susceptible to chance.

Electrical Parameters

Descriptive statistics are presented for each electrical parameter in Table 5.1. A Weibull and Log-logistic probability distribution function were found to be the best fit for the parameters, with scale and shape values presented in the table below.

Table 5.1: Descriptive statistics for electrical parameters

Parameter	**Mean**	**Median**	**Standard Deviation**	**Max**	**Min**	**Probability Distribution Scale Parameter (η)**	**Probability Distribution Shape Parameter (α)**
E_{TOTAL} (kWh)	2,261	2,142	1,108	10,065	99	2,555	2.15
E_{MD} (kW)	2.50	2.49	1.01	7.36	0.07	2.81	2.65
E_{LF} (%)	23.43	22.53	6.33	82.00	8.13	-1.49 *	0.14 *
E_{ToU} **	31.40	35.00	9.85	n/a	n/a	n/a	n/a

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \text{ where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

* Log-logistic Probability Distribution Function

$$f(T) = \frac{e^{z}}{\alpha T(1+e^{z})^{2}} \text{ where } z = \frac{T' - \eta}{\alpha}, \quad T' = \ln(T), \quad 0 < T < \infty, \quad -\infty < \eta < \infty, \quad 0 < \alpha < \infty$$

**where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00

Regression

A multivariate linear regression was carried out using the following variables: dwelling type, number of bedrooms, head of household (HoH) age, household (HH) composition, HoH social class, water heating type and cooking type. An efficiency indicator variable (defined in Section 3.4) was also included to measure potential household electricity savings by asking those surveyed to quantify how much they believed they could cut their electricity consumption by changing their behaviour.

All the variables listed above were found to have the greatest significance on dwelling and occupant characteristics without causing multi-collinearity. Other independent variables tested for significance included dwelling age, number of occupants, HoH employment status, tenure type, HoH education level and space heating type. These variables were omitted from the analysis since they either showed little or no significance over the tested parameters or showed a high degree of multi-collinearity with other independent variables. In particular, HoH age showed strong collinearity with dwelling age and tenure type with Pearson correlation coefficients exceeding 35% in both cases. This can be explained by younger HoH's having a higher percentage of mortgages and occupying newer dwellings. In comparison, a higher percentage of older HoH's have their mortgage paid off and live in older dwellings. Similarly number of occupants showed a high degree of collinearity with dwelling number of bedrooms and household composition, with a Pearson correlation coefficient also exceeding 35% in both instances. It was therefore decided to use household composition only in this instance. HoH employment status and education level had little effect on the parameters and showed high collinearity to HoH social class with Pearson correlation coefficients exceeding 25%. Space heating type (electric, non-electric) had no significance at all over the four parameters, due to the very low penetration of electric central heating (less than 5%) in Ireland and therefore was excluded from the analysis.

A full listing of the independent variables and sample sizes used in the analysis is shown in Table 5.2, with the base variable highlighted in bold italics where dummy categorical variables are used.

Table 5.2: List of independent variables used in regression model

Variable name	Variable explanation	Sample Size (N)
Detached	***Dwelling is detached (includes bungalows)***	***2068***
Semi-detached	Dwelling is semi-detached	1230
Terraced	Dwelling is terraced	569
Apartment	Dwelling is apartment	67
No. of bedrooms	Number of bedrooms within dwelling	3941
18 ≤ HoH Age < 36	***Head of household age between 18 & 35***	***390***
36 ≤ HoH Age < 56	Head of household age between 36 & 55	1776
56 ≤ HoH Age	Head of household age above 55	1753
HH Comp. - Live Alone	***Household composition - live alone***	***756***
HH Comp. - with Adults	Household composition - live with adults	2064
HH Comp. - with adults and children	Household composition - live with adults and children	1121
HoH Social Class - AB	***High and intermediate managerial, administrative or professional***	***593***
HH Social Class - C	Supervisory and clerical and junior managerial, skilled manual workers	1697
HH Social Class - DE	Semi-skilled and unskilled manual workers, state pensioners, unemployed	1505
HH Social Class - F	Farmers	107
Water Heat - Non-electric	***Water is heated by other (oil, gas, solid fuel)***	***3144***
Water Heat - Electric	Water is heated by electricity	771
Cooking Type - Non-electric	***Cooking is mostly done by non-electric means (oil, gas, solid fuel)***	***1192***
Cooking Type - Electric	Cooking is mostly done by electricity	2749
Efficiency <10%	***HoH who believe they can cut electricity consumption by 10%***	***1950***
10% ≤ Efficiency < 20%	HoH who believe they can cut electricity consumption by between 10% & 20%	916
20% ≤ Efficiency < 30%	HoH who believe they can cut electricity consumption by between 20% & 30%	345
Efficiency ≥ 30%	HoH who believe they can cut electricity consumption by more than 30%	123

Table 5.3 shows results for the linear regression for the DOC model and each of the four dependent electrical parameters. Pearson's coefficient of determination for each parameter as well as regression coefficients β (which indicates the magnitude of influence of each variable on the parameters and was defined in Equation 2.1) as well as standard errors are shown in the table below. The significance of variables on each parameter is shown by way of a p-value, indicating 90%, 95% and 99% significance levels.

Table 5.3: Regression results for dwelling and occupant characteristics model (DOC)

	E_{TOTAL} (kWh)		E_{MD} (kW)		E_{LF} (%)		E_{ToU}	
Pearson's Coefficient of Determination (%)	32%		33%		9%		2.60%	
	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error
(Constant)	18.6055	101.633	0.6388***	0.092	0.2169***	0.0068	29.4659***	1.0786
Semi-detached	-175.6725***	34.1701	-0.0766**	0.0309	-0.0082***	0.0023	-0.414	0.3626
Terraced	-147.045***	45.9229	-0.0583	0.0416	-0.0114***	0.0031	-1.2872**	0.4874
Apartment	-245.5571**	119.4231	-0.2963**	0.1081	0.0084	0.008	0.1958	1.2674
No. of bedrooms	349.036***	19.9182	0.2365***	0.018	0.0089***	0.0013	0.6785***	0.2114
36 ≤ HoH Age < 56	282.8721***	51.7462	0.0722	0.0468	0.0171***	0.0034	-0.9431*	0.5492
56 ≤ HoH Age	212.0358***	57.7676	-0.1515***	0.0523	0.0318***	0.0038	-2.0417***	0.6131
HH Comp. - with Adults	730.9512***	40.7046	0.7036***	0.0368	-0.0022	0.0027	1.2854***	0.432
HH Comp. - with adults and children	1083.688***	50.2313	0.9853***	0.0455	0.0043	0.0033	2.0295***	0.5331
HH Social Class - C	-73.6939*	44.1127	0.0407	0.0399	-0.0134***	0.0029	1.2344**	0.4682
HH Social Class - DE	-132.952**	48.522	-0.0146	0.0439	-0.0155***	0.0032	0.8489	0.515
HH Social Class - F	-370.2021***	98.0024	-0.2591***	0.0887	-0.0016	0.0065	-2.8708**	1.0401
Water Heat - Electric	148.9229***	29.5042	0.2379***	0.0267	-0.0077***	0.002	-1.3368***	0.3131
Cooking Type - Electric	185.6567***	32.2118	0.3896***	0.0292	-0.0241***	0.0021	0.1381	0.3419
10% ≤ Efficiency < 20%	142.7689***	37.6209	0.1139***	0.0341	0.0015	0.0025	-0.4104	0.3993
20% ≤ Efficiency < 30%	188.2471***	54.1685	0.1638***	0.049	0.0021	0.0036	-0.2999	0.5749
Efficiency ≥ 30%	274.1978***	85.5507	0.1476*	0.0774	0.0089	0.0057	-0.57	0.908

P values: *** p<0.01, ** p<0.05, * p<0.1

Base variables: *Detached, 18 ≤ HoH Age < 36, HH Comp. - Live Alone, HoH Social Class - AB, Water Heat - Non-electric, Cooking Type - Non-electric, Efficiency <10%*
E_{ToU} (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00)

Multivariate linear regression was carried out for the EA model with the same four dependent parameters as before and fifteen common household appliances as explanatory variables. The results are presented in Table 5.4 alongside household appliance penetration levels. The base variable chosen for the analysis was *washing machine* due to its high penetration level of 98.3% within the survey.

Table 5.4: Regression results for electrical appliances model (EA)

		E_{TOTAL} (kWh)		E_{MD} (kW)		E_{LF} (%)		E_{ToU}	
Pearson's Coefficient of Determination (%)		33.0%		31.0%		11.1%		2.4%	
	Appliance Penetration (%)	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error	Coeff. (β)	Std. Error
(Constant)		656.9107***	51.3526	0.8771***	0.0472	0.2444***	0.0035	29.8274***	0.5578
Tumble Dryer	68%	375.3768***	33.5586	0.3951***	0.0309	-0.0045*	0.0023	-0.1742	0.3645
Dishwasher	67%	406.0503***	33.7939	0.2894***	0.0311	0.0128***	0.0023	1.4145***	0.3671
Shower (instant)	69%	44.0911	32.8842	0.2557***	0.0302	-0.0189***	0.0022	-1.1625***	0.3572
Shower (pumped)	29%	34.5628	33.0484	-0.0159	0.0304	0.0025	0.0022	-0.2293	0.359
Electrical Cooker	76%	182.6508***	34.2263	0.3758***	0.0315	-0.0241***	0.0023	0.5208	0.3718
Heater (plug in convective)	30%	56.5369*	31.4838	-0.0339	0.029	0.008***	0.0021	-1.1678***	0.342
Freezer (stand alone)	50%	198.131***	29.6764	0.0775***	0.0273	0.0129***	0.002	0.0618	0.3224
Water pump	20%	208.1565***	36.7427	0.0902**	0.0338	0.0063**	0.0025	0.7612*	0.3991
Immersion	77%	73.4666**	34.6355	0.1701***	0.0319	-0.0068***	0.0023	-0.4635	0.3762
No. TV<21inch	66%	100.8994***	15.8887	0.1059***	0.0146	-0.0017	0.0011	0.434**	0.1726
No. TV>21inch	84%	197.2184***	18.4409	0.1393***	0.017	0.0026**	0.0012	0.5456**	0.2003
No. computer (desktop)	48%	287.3278***	26.4866	0.1626***	0.0244	0.0095***	0.0018	0.6874**	0.2877
No. computer (laptop)	54%	135.1009***	19.7789	0.0978***	0.0182	0.0042***	0.0013	0.2103	0.2149
No. game consoles	33%	193.1296***	20.7689	0.1953***	0.0191	0.0017	0.0014	0.2495	0.2256

P values: *** $p<0.01$, ** $p<0.05$, * $p<0.1$
Base variable: *Washing Machine*
E_{ToU} (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00)

5.2.1 E_{Total}

E_{Total} was regressed against dwelling and occupant characteristics described in Table 5.2 and a coefficient of determination of 32% was recorded for the DOC approach. This indicated that only 32% of the variation could be explained by the variables listed in Table 5.3 with the remainder due to other factors that may not have been able to be measured. This highlights the highly variable nature to domestic electricity demand even when daily electrical parameters are used.

All dwelling types had a negative effect on E_{Total} when compared to the base variable detached dwelling, which included bungalows. Apartments had significantly lower E_{Total} than all other dwelling types, a result of their smaller size and fewer number of occupants. For each additional bedroom, E_{Total} on average increased 349 kWh over the six month period. On a per capita basis, E_{Total} for the residential sector accounted for 948 kWh over the six month period.

Electricity consumption for younger HoH's was significantly lower when compared to the other two HoH age categories, 36-55 and 56 plus. This could be attributed to middle aged HoH's having more children living at home (thus having a higher number of occupants) and increased occupancy patterns (i.e. dwelling occupants at home for longer periods of the day). This is also apparent when looking at household composition: adults living with children consume considerably more electricity than those living alone or with other adults. HoH Social class had a negative effect on E_{Total} when compared against the base category AB, representing Higher Professionals. Social class was used as a proxy in the absence of reliable data on household income. This suggests that Higher Professionals are inclined to consume more electricity than Lower Professionals with the former tending to live in larger dwellings and have a greater number of electrical appliances, suggesting a possible income effect.

The efficiency indicator variable showed strong positive correlation with increasing electricity savings (i.e. respondents with higher electricity consumption believed they could make greater electricity savings than those who consumed less). This suggests that larger electricity consumers are wasteful (i.e. leave lights on in unoccupied rooms) and hence believe they can cut back on their electricity use. In contrast, those who consume less may believe that they are efficient in their use of electricity and cannot make further substantial cuts.

Table 5.4 showed regression results for the EA approach, where a coefficient of determination of 33% was recorded. Similar to the DOC method, this suggests that a large part of the variation could not be explained by the ownership of particular appliances and are a result of other factors. Tumble dryers, dishwashers, cookers, freezers, water pumps (used in low water pressure residential areas) and brown goods (televisions, computers, game consoles) were all significant at the 99% level. Showers showed no significance at all and immersions were only significant at the 90% level. It is also important to note that the analysis above is independent of lighting, which is a significant contributor to electricity consumption. The effect of lighting demand could not be distinguished in the survey as the number of fittings was not recorded. Similarly, electrical appliance refrigerator was not recorded as part of the survey.

5.2.2 E_{MD}

E_{MD} was regressed against the variables listed in Table 5.2 and a coefficient of determination of 33% was recorded for the DOC approach. E_{MD} was significantly influenced by semi-detached and apartment dwellings at the 95% level as was shown in Table 5.3. When compared against the base variable (detached dwelling) each had a negative influence on E_{MD}, particularly apartments. Number of bedrooms was significant at the 99% level and serves to increase E_{MD} by 0.23kW for every additional bedroom within a dwelling. Similarly, household composition significantly influenced E_{MD}, with adults and children consuming nearly an extra kilowatt compared to the base variable (adult living alone). Apartment dwellings tend to be smaller in size, have fewer occupants and have a smaller stock of appliances than other dwelling types, all of which are drivers of E_{MD}. As one would anticipate, homes with electric water heating and cooking also had higher E_{MD} compared to those that use other methods to heat water and to cook.

The EA approach recorded a coefficient of determination of 31%. Almost all household appliances showed significant influence on E_{MD} at the 99% level. Pumped showers and plug in convective heaters were the only appliances not to show any significance at all, possibly due to their low penetration within the sample. The three largest contributors to E_{MD} were tumble dryers, dishwashers and electric cookers which all have significant heating components to their operation. Instant electric showers and immersions, both used for heating water were the next largest contributors to E_{MD}.

Electricity generated at peak times such as early morning and evening times is far less efficient than electricity generated at other times of day. This is a direct result of running expensive peaking generation plant such as open cycle gas turbines to respond to quick changes in system demand, which are less efficient than other types of generation. Shifting demand away from peak times will result in a more efficient electricity system and as a consequence reduce greenhouse gas emissions for the sector. In particular, tumble dryers and dishwashers offer the best opportunity for shifting demand away from peak time use compared to electric cookers as they are less dependent on the timing of high priority household routines such as cooking. The introduction of ToU tariffs for the residential sector, so that electricity consumed at peak times reflects the true cost of generation, may encourage homeowners to shift non-essential appliance use to off peak times when electricity is cheaper.

5.2.3 E_{LF}

A significantly lower coefficient of determination, 9%, was recorded for E_{LF} for the DOC approach compared to the previous two parameters. E_{LF} changes only slightly between households as indicated by the low standard deviation for the parameter (6%)

as shown in Table 5.1. However, the parameter is useful for describing the electricity load profile shape for individual households. A low value for E_{LF} indicates households whose electricity consumption pattern is high for short periods of time whereas a higher value for E_{LF} indicates a more constant use of electricity across the day.

Semi-detached and terraced dwellings had a significant impact on E_{LF} compared to the base variable (detached dwelling). Larger dwellings such as detached and semi-detached homes had a positive effect on E_{LF}. For each additional bedroom, E_{LF} on average increased by 1%. HoH age also strongly influenced E_{LF} in a positive manner with younger HoH groups having slightly lower E_{LF} representing a more "peaky" load across the day. In contrast, older HoH age groups have a larger E_{LF}, indicating a smoother electricity consumption pattern across the day. This is most likely due to older HoH's possibly being more active in the home during the day. Water heating and cooking type influenced E_{LF} in a negative manner and therefore households that use electricty to heat water and cook will therefore tend to have lower values for E_{LF}.

The EA approach recorded a coefficient of determination of 11.1% for E_{LF}. Most household appliances were significant at the 99% level except for tumble dryers, electric showers (pumped), water pumps, televisions and game consoles. When compared against the base variable washing machine, appliances with negaitive coefficients decrease E_{LF} and correspond with high power devices that are not used continiously for long periods of time. In particular, electric showers (instant), cookers and immersions, which are all significant at the 99% level, tended to decrease E_{LF} due to their high power requirements and result in a more "peaky" domestic load profile.

Dishwashers and stand alone freezers on the other hand had a significant positive effect on E_{LF} as they are switched on for longer periods of time.

5.2.4 E_{ToU}

A poor coefficient of determination of 2.6% was recorded for E_{ToU} in the DOC method shown in Table 5.3. ToU showed high significance for household composition and HoH age. For HoH age, the older the head of the household the more negative the influence on the parameter indicating earlier use of E_{ToU} during the evening. Household composition had a positive effect on E_{ToU} with adults and children tending to use electricity demand later in the evening compared to occupants living alone. Younger and middle aged groups correspond to households with young families and therefore tend to have a greater number of occupants. These groups are inclined to use E_{ToU} later in the evening, most likely a result of increased number of active occupants later in the evening. Households with older HoH's tend to have fewer number of occupants, as children may have vacated the home and are also closer to retirement age and hence tend to be active earlier in the evening possibly due to lighter work commitments or retirement. Hence these groups are more likely to use E_{ToU} earlier in the evening.

5.3 Conclusion

Results were presented linking dwelling and occupant characteristics to electrical parameters: E_{Total}, E_{MD}, E_{LF} and E_{ToU}. Dwelling number of bedrooms, which was used as a proxy for dwelling size, was found to strongly influence E_{Total}. Apartment dwellings, which are proportionally smaller and have less occupants and appliances, consumed the least electricity when compared to other dwelling types. HoH age group

36 – 55 were found to be the largest consumers of electricity, probably due to the prevalence of children living at home among this age group. Social class was used as proxy for household income due to unreliable data recorded for this variable within the survey. Household social class was significant with Higher Professionals consuming more electricity than middle or lower classes, reflecting a possible income effect. Dwellings that used electricity for water heating and cooking also used a larger amount of electricity as would be expected. An efficiency variable also indicated the potential for reducing household electricity demand which showed significant positive correlation with the parameter, possibly indicating that larger electricity consumers are more wasteful of electricity than those who consumed less. Appliances that consumed the most electricity were tumble dryers and dishwashers.

Household composition, number of bedrooms, water heating and cooking type were the most significant variables to influence E_{MD}. It was also shown that the majority of common household electrical appliances included in the survey influenced E_{MD}. However, three appliances in particular: tumble dryer, dishwasher and electric cooker, contributed significantly more than the base variable washing machine. The introduction of ToU tariffs should discourage the use of non-high priority household tasks such as clothes and dish washing at peak times. E_{LF} was influenced by independent variables dwelling type and number of bedrooms. HoH age was also significant, with younger HoH's having a smaller E_{LF} representing a more "peaky" load profile shape. Water heating and cooking by electricity had the effect of lowering the overall E_{LF} as these appliances tend to consume large amounts of electricity for relatively short periods of time. This was also apparent from the EA model where the three most significant appliances to reduce E_{LF} were: electric shower (instant), cooker and immersion.

E_{ToU} was influenced more by occupant rather than dwelling characteristics as one would expect. Older head of households are more likely to use E_{ToU} earlier in the day. This was also reflected in the household composition variable where adults and children, which correspond with younger HoH's, tending to use E_{ToU} later in the day. Appliances that influenced ToU were dishwashers, electric showers, plug in convective heaters, televisions and computer desktops. The appliance that showed the greatest potential for shifting demand away from peak time use was the dishwasher due to its high power requirement and frequent use. This suggests the potential for the introduction of ToU and/or smart appliances for the domestic sector.

CHAPTER 6

AUTOREGRESSIVE (MARKOV CHAIN)

6 AUTOREGRESSIVE (MARKOV CHAIN)

6.1 Introduction

A Markov chain probabilistic approach to characterise domestic electricity demand load profiles is evaluated in this chapter. The technique has been used in the past to characterise various applications such as rainfall [122] and wind speed at specific locations [47]. However, to the best of the authors knowledge its application at a domestic level has never been done before. In this chapter Markov chains are applied to a small sample of dwellings in order to determine how effective they are for characterising domestic electricity demand.

6.2 Discussion and Results

6.2.1 Electrical Parameters

Electrical parameters were calculated from Equations 4.1 – 4.4 for each dwelling type and results presented in the following tables. Table 6.1 shows E_{Total} for original sample and characterised profiles as well as percentage error over the six month period. For three out of four cases, the percentage error was less than 5% indicating that the parameter was reproduced within a high level of accuracy. In each case, it is interesting to note that the Markov chain process over estimated the parameter. This is due to sampling error and is a result of having too few sample bins at the lower end of the electricity load profile.

Table 6.1: E_{Total} for each dwelling type over six month period (July – December 2009)

Dwelling Type	Original	Characterised	Percentage Error
Detached	2,163 kWh	2,236 kWh	3.37%
Semi-detached	2,574 kWh	2,593 kWh	0.74%
Terraced	2,872 kWh	3,065 kWh	6.72%
Apartment	616 kWh	634 kWh	2.92%

Table 6.2 presents results for electrical parameter E_{MD} as described by Equation 4.2. Standard deviation varies slightly between original and characterised load profiles and is a result of using a random number to generate a value between the two bin end points. The Minimum value for parameter E_{MD} varies more considerably between original and characterised profiles. This variation can again be attributed to the number of sample bins at the lower end of the electricity demand load profile.

Table 6.2: Descriptive statistics for original and characterised profiles by dwelling type for E_{MD} over six month period (July – December 2009)

Dwelling Type	Mean	Median	Std. Dev.	Max	Min	Scale Parameter (η)	Shape Parameter (α)
Detach original (kW)	2.71	2.69	0.95	5.98	0.77	2.68	0.54
Detach characterised (kW)	2.52	2.54	1.15	5.76	0.26	2.51	0.67
Semi-detach original (kW)	2.73	2.66	1.05	5.44	0.08	2.77	0.55
Semi-detach characterised (kW)	2.92	2.72	0.82	5.38	0.78	2.87	0.46
Terraced original (kW)	3.07	3.00	0.72	5.60	0.19	3.04	0.39
Terraced characterised (kW)	3.13	3.07	0.77	6.10	1.28	3.10	0.44
Apartment original (kW)	1.03	0.92	0.54	3.70	0.40	0.95	0.28
Apartment characterised (kW)	1.02	0.82	0.63	3.89	0.21	0.94	0.34

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \text{ where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

A paired sample t-test is shown in Table 6.3 and indicates that there is no significant difference between daily E_{MD} for original and characterised profiles for each dwelling type.

Table 6.3: Paired sample t-test for E_{MD} for each dwelling type over six month period (July – December 2009)

Paired Samples	**Mean**	**Standard Deviation**	**Std. Error Mean**	**t**	**Sig. (2-tailed)**
Detached	0.19	1.51	0.11	1.70	0.09
Semi-detached	-0.18	1.29	0.10	-1.89	0.06
Terraced	-0.06	1.02	0.08	-0.78	0.44
Apartment	0.00	0.85	0.06	0.06	0.95

Descriptive statistics for E_{LF} as described by Equation 4.3 is presented in Table 6.4. E_{LF} is a function of E_{MD} and therefore one would expect similar results to those presented in Table 6.2. However, E_{LF} varied slightly between original and characterised time series more so than the other two parameters presented in Table 6.1 and Table 6.2. E_{LF} is a ratio of maximum to average electricity demand during a 24 hour (00:30 – 00:00) period. The variation between original and characterised time series can be explained as follows. Firstly, as discussed the number of sample bins at the lower end of the load profile influences the average value of electricity demand evaluated across the day. Secondly, Markov chains are stochastic processes which are also independent of time. As a result, over the course of a 24hour period a typical peak value for E_{MD} like that shown in Figure 1.2 may or may not occur. Similarly, it is possible that two such typical peaks may occur during a single 24 hour period. This has a bearing on the calculated value for E_{LF}.

Table 6.4: Descriptive statistics for original and characterised profiles by dwelling type for E_{LF} over six month period (July – December 2009)

Dwelling Type	Mean	Median	Std. Dev.	Max	Min	Scale Parameter (η)	Shape Parameter (α)
Detach original	19.29	18.69	4.88	35.31	9.58	0.19	0.03
Detach characterised	21.89	20.87	6.99	46.37	8.04	0.21	0.04
Semi-detach original	25.77	22.00	14.97	81.46	9.36	0.23	0.06
Semi-detach characterised	20.21	19.85	5.73	39.25	8.09	0.20	0.03
Terraced original	21.99	21.88	6.08	79.00	11.66	0.22	0.03
Terraced characterised	22.73	22.16	5.65	42.29	12.14	0.23	0.03
Apartment original	15.39	15.08	5.17	30.72	6.22	0.15	0.03
Apartment characterised	17.47	16.36	7.47	44.68	6.26	0.17	0.04

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \quad \text{where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

The difference in mean E_{LF} values is presented in Table 6.5 where a paired sample t-test also shows that detached, semi-detached and apartment dwelling types were significantly different for parameter E_{LF} when compared against the original data.

Table 6.5: Paired sample t-test for E_{LF} for each dwelling type

Paired Samples	Mean	Standard Deviation	Std. Error Mean	t	Sig. (2-tailed)
Detached	-0.03	0.08	0.01	-4.08	0.00
Semi-detached	0.06	0.16	0.01	5.00	0.00
Terraced	0.00	0.09	0.01	-0.57	0.57
Apartment	-0.02	0.09	0.01	-3.26	0.00

Table 6.6 shows descriptive statistics for E_{ToU}. Maximum and Minimum values were not shown below as this would indicate times of 1 and 48 corresponding with 00:30 and 00:00 respectively. Significant differences exist between original and characterised profiles presented in Table 6.6, indicating the occurrence of E_{ToU} at different times of day.

Table 6.6: Descriptive statistics for original and characterised profiles by dwelling type for E_{ToU} (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00) over six month period (July – December 2009)

Dwelling Type	**Mean**	**Median**	**Std. Dev.**	**Scale Parameter (η)**	**Shape Parameter (α)**
Detach original	35.21	36.00	8.10	36.09	4.02
Detach characterised	24.18	23.50	14.13	24.10	8.56
Semi-detach original	29.20	34.00	10.70	30.26	6.20
Semi-detach characterised	24.69	26.00	14.77	24.58	9.00
Terraced original	31.90	35.00	7.15	32.97	3.76
Terraced characterised	25.37	24.00	14.45	25.21	8.83
Apartment original	22.38	19.00	7.87	21.35	4.41
Apartment characterised	25.31	26.50	14.28	25.43	8.68

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \quad \text{where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

A paired sample t-test shown in Table 6.7 confirms the results presented in Table 6.6. The table shows that E_{ToU} is significantly different for each dwelling type. Therefore we can conclude that the temporal characteristics of the original time series were not transferred to the characterised time series. In addition, the Shape parameter fitted to E_{ToU} shows significantly different values to the original data indicating that the distribution for electricity demand across the day was more random than anything else.

Table 6.7: Paired sample t-test for E_{ToU} for each dwelling type

Paired Samples	**Mean**	**Standard Deviation**	**Std. Error Mean**	**t**	**Sig. (2-tailed)**
Detached	10.93	15.45	1.14	9.60	0.00
Semi-detached	3.98	18.00	1.33	3.00	0.00
Terraced	30.83	6.76	0.50	61.83	0.00
Apartment	-3.13	14.59	1.08	-2.91	0.04

6.2.2 Time Series Tests

The following figures present illustrative results for a single dwelling and therefore cannot be considered to be representative for each dwelling. However, they give a good visual representation as to the characterisation performance and build on the results presented in Table 6.1 to Table 6.7. Figure 6.1 shows a time series plot for the detached dwelling described above for both original (solid blue line) and characterised profiles (dashed red line) across a 24 hour period. It is apparent from Figure 6.1 that daily peaks for original sample and characterised profiles do not coincide on a temporal basis. For example, the original profile produces daily peaks at 17:00 and 00:00 whereas the characterised profile shows daily peaks at 04.30 and 21:00. This confirms what was presented in Table 6.6 and Table 6.7 where E_{ToU} was shown not to occur at the same time between original and characterised profiles.

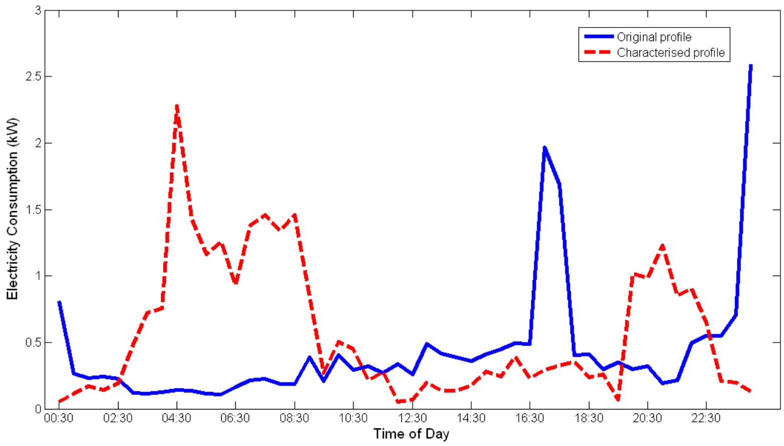


Figure 6.1: Original and characterised profiles for detached dwelling over a 24 hour period (1st July 2009)

Figure 6.2 show a frequency histogram for the same detached dwelling over a one week period from 1st – 7th July 2009. The characterised profile slightly under estimates the frequency of smaller values of electricity consumption (< 0.6 kW) within the home and over estimates larger values. This was also indicated in Table 6.1 where the Markov chain process resulted in an over estimation for parameter E_{Total} for the characterised time series.

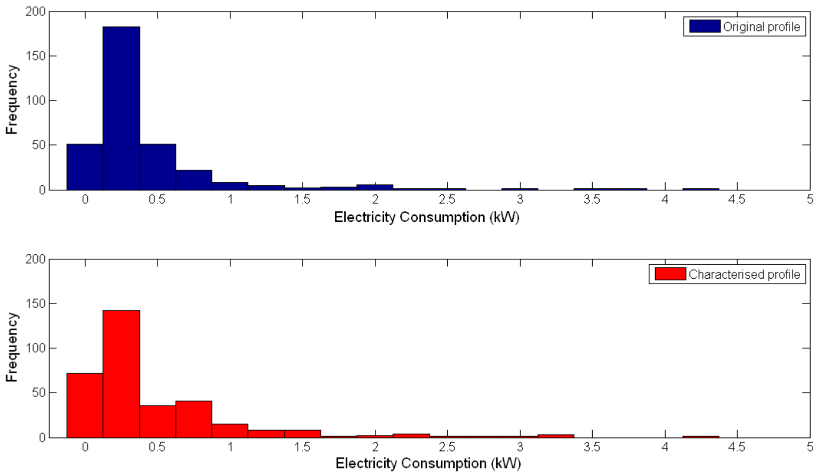


Figure 6.2: Frequency histogram for original and characterised profiles for detached dwelling over a weekly period (1st – 7th July 2009)

Figure 6.3 shows the autocorrelation function as described by Equation 4.5 for the same detached dwelling and period above. The original profile shows a clear cyclical pattern on a daily basis indicating high correlation between electricity consumed at the same time interval each day. In contrast, the characterised profile shows little or no correlation with the same time periods for each day as indicated by the autocorrelation function. In addition this figure highlights the attribute that even though domestic electricity load profiles can be considered to be stochastic on an intra-daily basis (as was shown in Figure 1.4) patterns of use are visible when the autocorrelation function is applied.

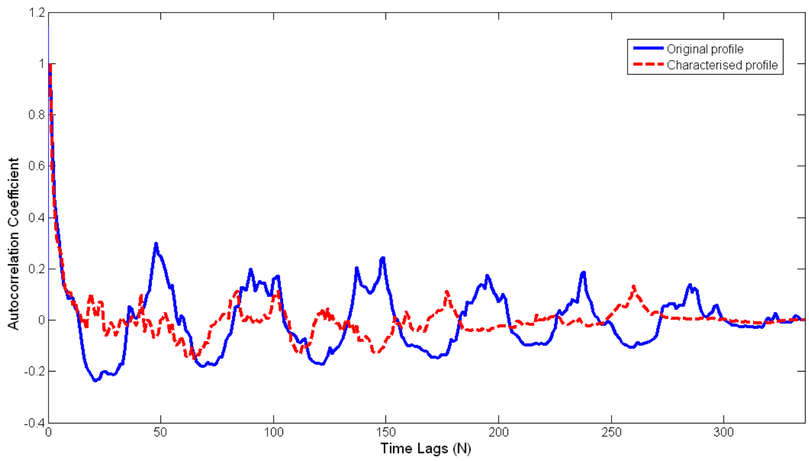


Figure 6.3: Autocorrelation function for original and characterised profiles for detached dwelling over a weekly period (1st – 7th July 2009)

Figure 6.4 shows PSD periodgram for the detached dwelling as calculated by Equations 4.6 over a weekly period. The original and characterised profiles have very different frequency components further indicating that the temporal properties are significantly different between the time series. Furthermore, Figure 6.4 indicates that the characterised profile reflects non-cyclical patterns of electricity use within the home.

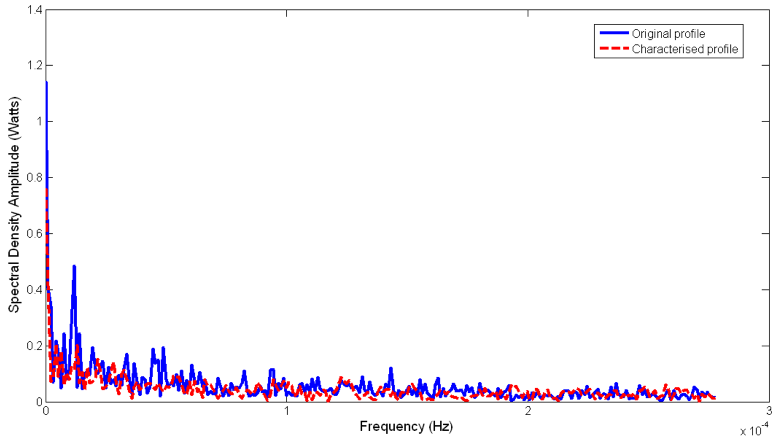


Figure 6.4: PSD function for original and characterised profiles for detached dwelling over a weekly period (1st – 7th July 2009)

Figure 6.1 to Figure 6.4 presented similar findings to the Tables presented above. Markov chains have shown to be effective at reproducing the magnitude component to domestic electricity demand load profiles as indicated by replicating parameters E_{Total}, E_{MD} and to a lesser extent E_{LF}. This is a result of their ability to model the stochastic component which is a common feature of domestic electricity demand load profiles as indicated earlier in Figure 1.4 and Figure 1.5.

The disadvantage with Markov chains when applied to characterising domestic electricity demand load profiles was their inability to sufficiently capture the temporal components. This was clearly shown by parameter E_{ToU} and in Figure 6.1 to Figure 6.4. In particular, the time series plot (Figure 6.1) and the autocorrelation function (Figure 6.3) show the timing of electricity demand not to coincide at similar intervals to the original profile. This is a result of the Markov chain process being independent of time. The intention of using Markov chains to characterise domestic loads was to use the principle that future values of electricity demand is highly correlated to past and present values. However, the analysis has shown that even though this is the case it is not enough alone to be able to sufficiently characterise individual domestic electricity demand load profiles. Therefore the decision was made to investigate other time series methods that could characterise these properties as will be shown in the next chapter.

6.3 Conclusion

A Markov chain process was used to characterise domestic electricity demand for four individual dwelling types chosen at random from *Dataset I*. Descriptive statistics, alongside electrical parameters were presented and used to compare original and characterised profiles. Time series tests were also used to interrogate the time series properties between profiles more rigorously.

Electrical parameters, E_{Total}, E_{MD} and E_{LF} were all reproduced within a reasonable degree of accuracy. Markov chains proved to be very effective at reproducing the magnitude components to domestic electricity load profiles as indicated by low percentage errors and small differences between mean values for parameter values in the statistical t-tests.

Electrical parameter E_{ToU} was not so well reproduced. The time series tests also showed significant differences in timing between ToU of electricity demand. This was shown to be most evident by the autocorrelation function where the cyclical daily pattern of electricity demand was not characterised adequately between profiles and similar results were obtained from evaluating the PSD function.

Markov chains were therefore deemed to be unsuitable to characterise domestic electricity demand in their current form. In addition, it would have proved difficult to link transitional probability matrices to dwelling and occupant characteristics as discussed in the objectives outlined in Section 1.5. As a result, it was decided not to pursue this approach further and concentrate on methods that could sufficiently characterise the temporal properties of domestic electricity demand load profiles adequately.

CHAPTER 7

TIME SERIES APPROACHES

7 TIME SERIES APPROACHES

7.1 Introduction

Time series approaches to characterising domestic electricity demand have been somewhat limited in the past. However, these methods have been used extensively in the electricity supply industry for characterising system demand like that shown earlier in Figure 1.1. This section discusses time series approaches applied to domestic electricity demand load profile characterisation by comparing and contrasting each method in order to meet the objectives outlined in Section 1.5. The following techniques are discussed Fourier transforms, Neural networks, Gaussian processes, Auto-regression, Fuzzy logic, Wavelets and multivariate regression.

7.2 Discussion and Results

7.2.1 Discussion

The principal advantage of Fourier transforms over other methods are their ability to represent the temporal and magnitude components within the characterisation coefficients, with the latter scalable [38]. This means that comparable households that show similar patterns of electricity use can be grouped together. However, the disadvantage with Fourier transforms, as will be discussed later, is their difficulty in characterising small intervals of large electricity demand [123].

Neural networks are especially good at characterising non-linear relationships and are therefore well suited to the variable nature of domestic electricity load profile characterisation. However, they are seen as a black box approach where it is often difficult to visualise a relationship between input and output [50]. In addition, the characterisation structure is often quite complex, 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 [50]. 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 variables: amplitude, centroid and peak width (shown in Equation 2.6) and that describe each probability distribution [86]. 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.

An autoregressive, Markov chain approach was applied in the last chapter to characterise domestic electricity demand. More common approaches include autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA), however, these have only been used to characterise electricity system demand load profiles like that shown in Figure 1.1. Due to the variable nature of individual domestic loads, ARMA and ARIMA find it difficult to characterise without using higher 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 households [38].

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 households. The advantage in doing this is that certain dwelling and occupant characteristics have different periods of influence over electricity consumption in the home [89]. However, the disadvantage is that it effectively doubles the number of variables required to characterise the time series.

Finally, multivariate linear regression is a technique that has been used extensively in electricity load profiling. Similar to autoregression it is most often applied to characterise and forecast system demand. It is the method of choice for UK grid operator, National Grid, to develop standard load profiles for the purposes of electricity settlement as discussed earlier [12]. However, a large number of variables are required to characterise the standard load profiles which reflect a single average electricity load profile across the day, such as that shown in Figure 1.3. Monte Carlo analysis is the most common probabilistic approach to load profile characterisation. The advantage with this technique is that it is ideal for generating variable load profiles and therefore is well suited to domestic electricity load characterisation. However, using Monte Carlo analysis to characterise domestic electricity demand requires each half hour period 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 within the variable coefficients. Fuzzy Logic, autoregression, neural network and multivariate regression also have this capability but require each half hour period to be characterised independently by a minimum of a single variable. This is a disadvantage as a minimum of forty eight variables would be required to characterise the temporal and magnitude components sufficiently for each household. Autoregressive approaches such as ARIMA have been used extensively in the past to forecast 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 electricity system demand load profiles as shown in Figure 1.1. However, this component is not as well suited to individual residential applications where electricity demand 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 electricity load profile. 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 households 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. Therefore, due to the fact that Fourier transforms and Gaussian processes both represent the temporal and magnitude components with a relatively small number of variables, both approaches are evaluated in the next section. A table summarising the principal advantages and disadvantages outlined above for each time series technique and the time interval at which it has been applied for electricity load profile characterisation is given in Appendix C.

7.2.2 Results

Electrical Parameters

The following section presents characterisation results for both Fourier transforms and Gaussian process time series techniques. In both cases an eighth order characterisation approach was applied which was the highest order accommodated by the Matlab toolbox software. Table 7.1 shows descriptive statistics for mean E_{Total} for all households over the yearly period. Both Fourier transforms and Gaussian processes characterised parameter E_{Total} with less than 5% percentage error across each descriptive statistic in Table 7.1. Fourier transforms produced accurate results with mean errors less than 0.1%. Gaussian processes on the other hand were less accurate but still within acceptable limits overall (<5% percentage error). A Weibull probability distribution function was found to be the best fit to the parameter, with scale and shape values also included in Table 7.1 below.

Fourier transforms essentially apply a data integration process to the time series. As parameter E_{Total} is a summation of the total amount of electricity consumed over a period of time this resulted in the Fourier based approach characterising the parameter with a high degree of accuracy. In contrast, Gaussian Processes characterise by applying a series of probability distribution functions to the time series which resulted in a slightly higher percentage error for the parameter.

Table 7.1: Descriptive statistics for mean E_{Total}

Characterisation Method	**Mean**	**Median**	**Standard Deviation**	**Maximum**	**Minimum**	**Scale Parameter (η)**	**Shape Parameter (α)**
Original Time Series (kWh)	4,146	4,008	1,870	9,651	414	4,687	2.38
Fourier Transforms (kWh)	4,146 (0%)	4,008 (0%)	1,870 (0%)	9,651 (0%)	414 (0%)	4,687 (0%)	2.38 (0%)
Gaussian Processes (kWh)	4,047 (-2.4%)	3,903 (-2.6%)	1,835 (-1.9%)	9,462 (-2.0%)	413 (-0.2%)	4,576 (-2.4%)	2.37 (-0.4%)

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \text{ where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

Table 7.2 shows results for a paired sample t-test between the original and characterised time series for parameter E_{Total}. 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 in Table 7.1 for 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 7.2: Descriptive statistics for paired sample t-test for E_{Total}

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

Table 7.3 shows results for mean daily E_{MD} for all households over the year. Similar to parameter E_{Total}, a Weibull probability distribution function was found to be the best fit for the parameter. Fourier transforms were poor at capturing the daily peak demands characteristic of almost all individual dwellings. Descriptive statistics presented in Table 7.3 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 able to characterise this parameter with percentage error of less than 5% in most instances.

Table 7.3: Descriptive statistics for mean daily E_{MD}

Characterisation Method	Mean	Median	Standard Deviation	Maximum	Minimum	Scale Parameter (η)	Shape Parameter (α)
Original Time Series (kW)	2.34	2.29	0.92	6.18	0.14	2.6293	2.7425
Fourier Transforms (kW)	1.68 (-28.2%)	1.66 (-27.5%)	0.68 (-35.3%)	3.89 (-58.9%)	0.09 (-55.6%)	1.8904 (-28.1%)	2.6885 (-2.0%)
Gaussian Processes (kW)	2.23 (-4.7%)	2.2 (-3.9%)	0.88 (-4.4%)	5.99 (-3.1%)	0.13 (-7.1%)	2.5082 (-4.6%)	2.7394 (-0.1%)

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \text{ where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

The Fourier Transform characterisation process smoothes out the load profile shape thus resulting in large electricity spikes not being characterised adequately. In contrast Gaussian Processes are better able to describe these electricity spikes as indicated by smaller percentage errors in Table 7.3 for parameter E_{MD}. This highlights one of the more significant advantages of using Gaussian Processes to characterise domestic electricity demand load profiles compared to Fourier transforms, which will be discussed again later.

Table 7.4 shows a paired sample t-test for E_{MD} parameter. As discussed above, Gaussian Processes were better at representing the characteristics of this parameter. However, the results also show that the E_{MD} for both characterised time series were significantly different from that of the original time series at the 95% p-value level.

Table 7.4: Descriptive statistics for paired sample t-test for E_{MD}

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

Table 7.5 presents results for mean E_{LF} for all households over the yearly period. A log-logistic probability distribution function was found to be the best fit for the parameter with scale and shape values also shown in Table 7.5. Similar to Table 7.3, Fourier transforms were also unable to accurately characterise E_{LF} with percentage

errors in excess of 30% in most instances. However, this is not surprising as E_{LF} is a function of E_{MD} as was shown in Equation 4.3. Gaussian Processes showed much smaller characterisation errors for parameter E_{LF} and can be attributed to being better able to characterise short electricity spikes in the load profile shape.

Table 7.5: Descriptive statistics for mean E_{LF}

Characterisation Method	**Mean**	**Median**	**Standard Deviation**	**Maximum**	**Minimum**	**Scale Parameter (η)**	**Shape Parameter (α)**
Original Time Series	23.23%	22.35%	5.76%	48.69%	11.29%	-1.4935	0.1299
Fourier Transforms	31.79% (36.9%)	30.76% (37.6%)	6.59% (4.4%)	66.72% (37.0%)	18.05% (59.9%)	-1.1703 (-21.6%)	0.109 (-19.2%)
Gaussian Processes	24.74% (6.5%)	23.74% (6.2%)	6.54% (13.5%)	51.76% (6.3%)	11.89% (5.3%)	-1.434 (-4.0%)	0.138 (6.2%)

Log-logistic Probability Distribution Function

$$f(T) = \frac{e^z}{\alpha T(1+e^z)^2} \text{ where } z = \frac{T' - \eta}{\alpha}, \quad T' = \ln(T), \quad 0 < T < \infty, \quad -\infty < \eta < \infty, \quad 0 < \alpha < \infty$$

Table 7.6 shows results for a paired sample t-test for E_{LF} parameter. Fourier transforms over estimated E_{LF} compared to the original time series more than Gaussian processes. The results illustrate that Gaussian processes were better at characterising the time series in terms of E_{LF} but the t-test showed that both time series techniques were significantly different compared to the original data at the 95% p-value level.

Table 7.6: Descriptive statistics for paired sample t-test for E_{LF}

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 7.7 shows results for mean E_{ToU} for all households over the yearly period. Fourier transforms performed slightly better when this parameter was evaluated with less than 3% percentage error, but tended to overestimate its value indicating later use of E_{ToU}. However, the difference is small and E_{ToU} on average still occurs within the same half hour period (15:30). Gaussian processes percentage errors were slightly greater and on average tended to predict peak time electricity use slightly earlier (15:00). The results indicate that both Fourier transforms and Gaussian process techniques were able to sufficiently characterise the time series temporal properties within a certain degree of accuracy.

Table 7.7: Descriptive statistics for mean household E_{ToU} (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 (2.4%)	31.84 (2.2%)	3.62 (2.8%)
Gaussian Processes	29.63 (-3.5%)	29.91 (-4.0%)	3.3 (-6.3%)

The performance of each characterisation technique differed slightly depending upon which parameter was evaluated. Fourier Transforms showed to be adequate when household electricity demand was aggregated over a period. Gaussian Processes were better at characterising individual electricity spikes across a day. However, in contrast to the Markov chain method presented in the previous chapter both characterisation approaches were able to reproduce the temporal properties of the original time series as shown by the ToU parameter. Similar to the previous chapter a number of time series tests will now be applied to graphically illustrate the results presented above in Table 7.1 to Table 7.7.

Time Series Tests

A period of one day, 1st July 2009, was chosen for two random households to illustrate graphically typical characterising performance for both Fourier transforms and Gaussian processes and is shown Figure 7.1. As stated in the previous chapter the profiles presented cannot be considered to be representative as they were chosen at random, however, they give an insightful visual representation as to the characterisation performance for each method. Household 1 (shown on top) shows both techniques replicating the time series within a reasonable degree of accuracy across a daily period. It is interesting to note that Gaussian processes were unable to sufficiently replicate the late peak at night at around 11pm. In contrast, Household 2 (shown on bottom) 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.

As shown in Section 2.2.4 Fourier Transforms describe a time series as a combination of sinusoids. Low and high frequency components combine together resulting in a smoothing out of the electricity load profile. This results in sharp electricity peaks not being sufficiently characterised by the Fourier Transform process. Gaussian Processes on the other hand apply individual probability density functions, which similar to Fourier Transform method, are combined together to characterise the time series. However, unlike Fourier Transforms, Gaussian Processes have a location variable which enables small periods of electricity demand to be individually characterised. These periods are characterised in terms of amplitude and width for each individual probability density function at the appropriate time during the day. This demonstrates the main difference between the two characterisation approaches. Fourier transforms tend to be better at characterising profiles where a large amount of electricity is consumed over a number of hours in the day. In contrast Gaussian processes are better at characterising short intervals of high electricity consumption (≤1hour) across the day. Therefore depending on the household electricity demand load profile each characterisation approach has its own advantages.

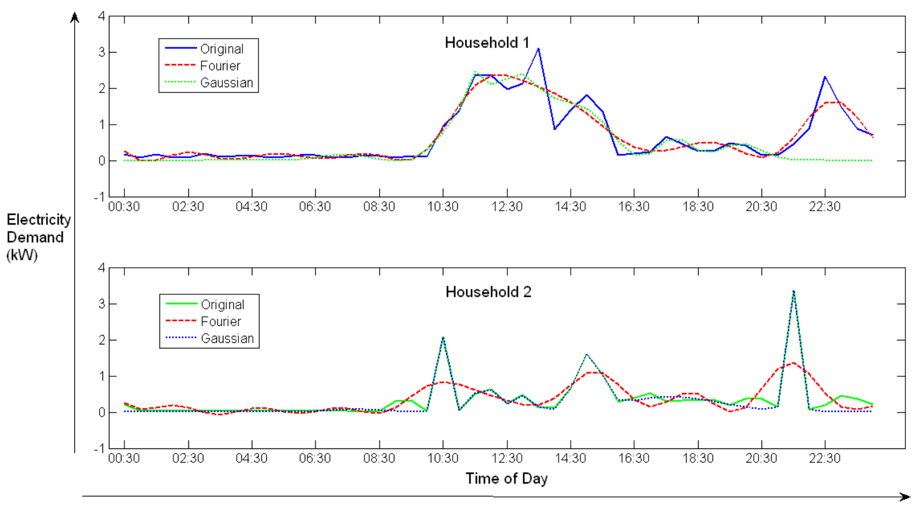


Figure 7.1: Time series plot for original and characterised load profiles for the 1st July 2009 for two random customers

Figure 7.2 shows a frequency histogram over a weekly period for a single random household. It is evident 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. The frequency occurrence of negative values is small and where it does occur, is very low in magnitude. This is a feature of both approaches and mainly occurs for profiles that use very little electricity demand across the daily period.

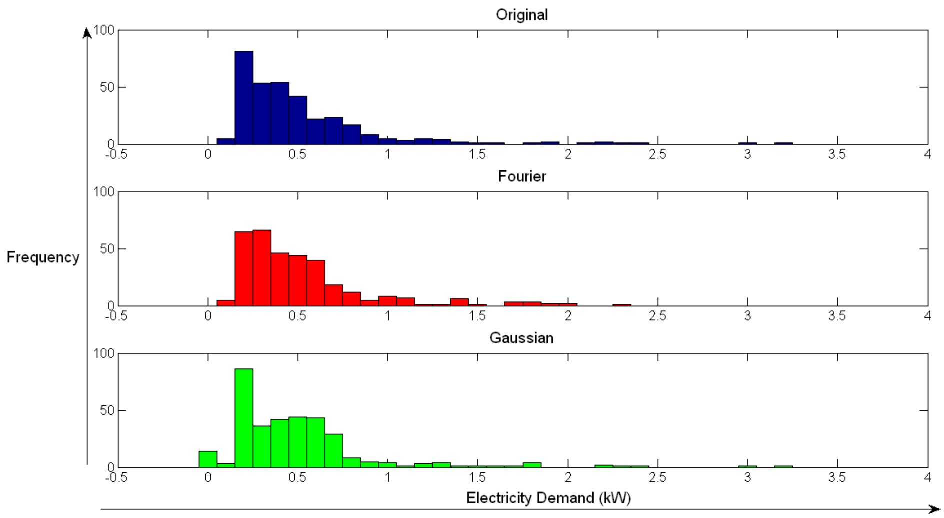


Figure 7.2: Frequency histogram for original and characterised load profiles for a random customer between 1st – 7th July 2009

Figure 7.3 shows the autocorrelation function where both Fourier transforms and Gaussian processes follow the original data over the weekly period 1st – 7th July 2009 for an individual random household. The first autocorrelation coefficient is excluded in Figure 7.3 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 Figure 7.3 shows, a highly cyclical pattern of electricity demand over a twenty four hour 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. In most cases in Figure 7.3 Gaussian Processes were shown to fit closer to the original time series for the autocorrelation function peaks and troughs. This confirms what has been presented earlier regarding their superior ability to characterise electricity load profile peaks when compared against Fourier Transforms.

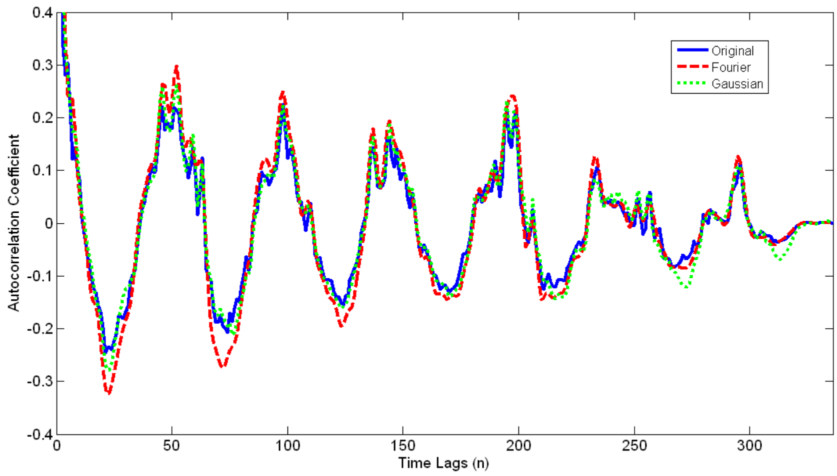


Figure 7.3: Autocorrelation coefficients for original and characterised load profiles for a random customer between 1st – 7th July 2009

Figure 7.4 shows the PSD periodgram for an individual random household over a weekly period 1st – 7th July 2009. 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 cyclical daily electricity demand pattern is typical of all dwellings with some households having smaller frequency patterns throughout the day.

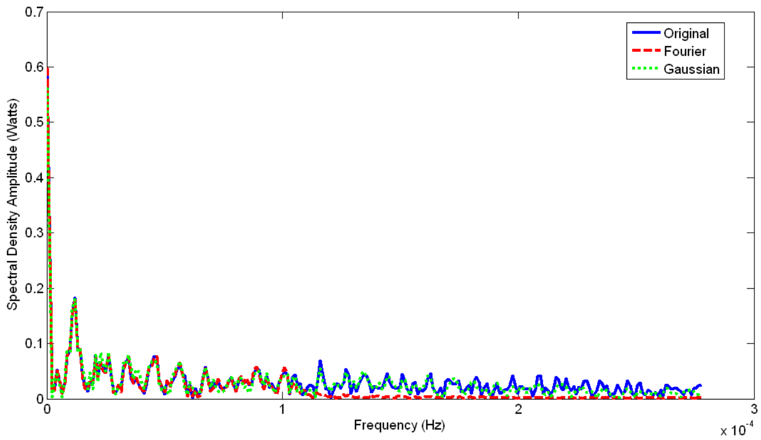


Figure 7.4: PSD function for original and characterised load profiles for a random customer between 1st – 7th July 2009

Fourier transforms - Regression

A multivariate linear regression was applied to the Fourier transform and Gaussian process characterised time series in order to determine the influence of dwelling and occupant characteristics on the electricity load profile shape. Table 7.8 shows regression results for Fourier transforms using the same dwelling and occupant variables applied in the statistical DOC approach in Chapter 5. The results show the significance (p-value) for each dwelling and occupant characteristic on the Fourier coefficients and also the magnitude of its influence indicated by β.

Overall, dwelling and occupant characteristics that were most influential over Fourier transform coefficients were number of bedrooms, household composition and whether electric water heating and cooking were present in the household. In particular, coefficient -*a0*- is a constant and hence does not contain any frequency components. Therefore the variable is largely influenced by characteristics that affect the magnitude component of the electricity load profile during the day such as number of bedrooms, household composition, presence of electric water heating and cooking and household efficiency indicator. The characteristics that influenced the smaller frequency components (such as *a1, b1, a2, b2, a3* and *b3*) tended to be greater in number, particularly at the 99% significance level compared to the larger frequency components. As the coefficients increase (i.e. a4 to b8) the dwelling and occupant characteristics become less influential as the load profile becomes less deterministic over shorter time intervals.

Table 7.8: Regression results for dwelling and occupant characteristics on Fourier transform coefficients

Fourier Coefficient Values	a0		a1		b1		a2		b2		a3		b3		a4		b4	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.0828	0.0569	0.1019***	0.0283	0.0830**	0.0388	0.0131	0.0182	-0.0682**	0.0277	0.0195	0.0123	-0.0318**	0.0134	-0.0176*	0.0104	-0.0081	0.0093
Dwell_type_semi_d	-0.0297	0.0194	-0.0026	0.0097	-0.0086	0.0133	-0.0071	0.0062	0.0037	0.0095	-0.0096**	0.0042	-0.0062	0.0046	0.0006	0.0035	0.0023	0.0032
Dwell_type_terr	-0.0262	0.0249	-0.0373***	0.0124	-0.0115	0.0170	-0.0079	0.0080	0.0079	0.0121	-0.0012	0.0054	0.0066	0.0058	-0.0067	0.0045	-0.0056	0.0041
Dwell_type_apt	-0.0711	0.0556	-0.0210	0.0277	0.0086	0.0380	-0.0215	0.0178	0.0195	0.0271	-0.0016	0.0120	0.0036	0.0131	-0.0008	0.0101	0.0060	0.0091
No_bedrooms	0.0504***	0.0106	0.0065	0.0053	-0.0356***	0.0073	0.0051	0.0034	-0.0141**	0.0052	-0.0039*	0.0023	-0.0084***	0.0025	0.0027	0.0019	-0.0047**	0.0017
HoH_age_36_55	0.0149	0.0301	-0.0222	0.0149	-0.0154	0.0205	-0.0113	0.0096	0.0196	0.0146	0.0136**	0.0065	0.0217***	0.0071	0.0038	0.0055	0.0131**	0.0049
HoH_age_56_plus	0.0182	0.0328	-0.0514***	0.0163	-0.0194	0.0224	-0.0072	0.0105	0.0334**	0.0160	0.0109	0.0071	0.0316***	0.0077	0.0026	0.0060	0.0134**	0.0053
HH_comp_with_adults	0.1389***	0.0207	-0.0005	0.0103	-0.0860***	0.0141	0.0107	0.0066	-0.0416***	0.0101	-0.0045	0.0045	0.0011	0.0049	0.0003	0.0038	0.0016	0.0034
HH_comp_with_adults_children	0.2449***	0.0272	-0.0116	0.0135	-0.1589***	0.0185	0.0143	0.0087	-0.0712***	0.0132	0.0027	0.0059	0.0053	0.0064	-0.0072	0.0050	-0.0012	0.0044
Social_class_C	-0.0132	0.0281	-0.0225	0.0140	-0.0204	0.0192	0.0054	0.0090	0.0183	0.0137	-0.0027	0.0060	0.0074	0.0066	0.0045	0.0051	-0.0041	0.0046
Social_class_DE	-0.0044	0.0301	-0.0501***	0.0150	-0.0335	0.0206	0.0223**	0.0097	0.0336**	0.0147	-0.0127*	0.0065	0.0088	0.0071	0.0099*	0.0055	-0.0113**	0.0049
Social_class_F	-0.0502	0.0540	-0.0793***	0.0269	0.0314	0.0369	0.0264	0.0173	0.0578**	0.0263	-0.0149	0.0116	0.0181	0.0127	0.0008	0.0099	-0.0145	0.0088
Water_heat_electric	0.0376**	0.0164	-0.0229**	0.0082	-0.0057	0.0112	0.0115**	0.0053	-0.0243***	0.0080	-0.0042	0.0035	0.0096**	0.0039	-0.0013	0.0030	-0.0002	0.0027
Cooking_type_electric	0.0730***	0.0176	-0.0063	0.0087	-0.0717***	0.0120	-0.0018	0.0056	-0.0121	0.0086	-0.0051	0.0038	-0.0048	0.0041	0.0057*	0.0032	0.0008	0.0029
Efficiency_betw_10_20	0.0315	0.0206	0.0077	0.0102	-0.0122	0.0140	0.0046	0.0066	-0.0075	0.0100	-0.0067	0.0044	-0.0007	0.0048	0.0036	0.0038	0.0062*	0.0034
Efficiency_betw_20_30	0.0921***	0.0324	0.0237	0.0161	-0.0553**	0.0221	-0.0030	0.0104	-0.0266*	0.0158	0.0022	0.0070	-0.0005	0.0076	0.0019	0.0059	-0.0053	0.0053
Efficiency_more_30	0.0997*	0.0531	0.0666**	0.0264	-0.0608*	0.0363	-0.0044	0.0170	-0.0951***	0.0259	0.0009	0.0114	-0.0190	0.0125	-0.0054	0.0097	0.0053	0.0087

	a5		b5		a6		b6		a7		b7		a8		b8		w	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	-0.0048	0.0068	0.0004	0.0098	0.0079	0.0062	-0.0002	0.0065	-0.0039	0.0047	-0.0076	0.0056	-0.0009	0.0054	0.0013	0.0079	0.1350***	0.0015
Dwell_type_semi_d	0.0013	0.0023	0.0024	0.0034	-0.0013	0.0021	0.0006	0.0022	0.0014	0.0016	-0.0020	0.0019	-0.0001	0.0019	0.0003	0.0027	0.0003	0.0005
Dwell_type_terr	0.0004	0.0030	0.0042	0.0043	-0.0020	0.0027	0.0051*	0.0028	0.0028	0.0021	-0.0007	0.0024	0.0054**	0.0024	-0.0032	0.0035	0.0010	0.0006
Dwell_type_apt	-0.0006	0.0067	0.0015	0.0096	-0.0069	0.0060	0.0000	0.0064	0.0003	0.0046	0.0036	0.0055	0.0032	0.0053	0.0000	0.0077	0.0018	0.0014
No_bedrooms	0.0023*	0.0013	-0.0038**	0.0018	-0.0016	0.0012	-0.0026**	0.0012	0.0003	0.0009	-0.0032***	0.0010	-0.0004	0.0010	-0.0042**	0.0015	-0.0003	0.0003
HoH_age_36_55	0.0021	0.0036	-0.0050	0.0052	0.0029	0.0033	0.0013	0.0034	0.0002	0.0025	0.0059**	0.0030	-0.0008	0.0029	0.0017	0.0042	0.0002	0.0008
HoH_age_56_plus	0.0061	0.0039	-0.0050	0.0057	0.0007	0.0035	0.0063*	0.0037	0.0029	0.0027	0.0070**	0.0032	0.0000	0.0031	0.0064	0.0046	0.0004	0.0008
HH_comp_with_adults	0.0074***	0.0025	-0.0115***	0.0036	0.0032	0.0022	-0.0065**	0.0024	0.0030*	0.0017	-0.0034	0.0020	-0.0002	0.0020	-0.0098***	0.0029	-0.0005	0.0005
HH_comp_with_adults_children	0.0089**	0.0033	-0.0161***	0.0047	0.0096***	0.0029	-0.0026	0.0031	0.0003	0.0023	-0.0051*	0.0027	-0.0013	0.0026	-0.0192***	0.0038	-0.0003	0.0007
Social_class_C	-0.0045	0.0034	0.0002	0.0048	-0.0025	0.0030	-0.0029	0.0032	0.0007	0.0023	0.0050*	0.0028	-0.0001	0.0027	-0.0011	0.0039	0.0007	0.0007
Social_class_DE	-0.0019	0.0036	0.0026	0.0052	-0.0073**	0.0033	-0.0017	0.0034	-0.0006	0.0025	0.0014	0.0030	0.0018	0.0029	0.0013	0.0042	0.0011	0.0008
Social_class_F	0.0039	0.0065	0.0099	0.0093	-0.0020	0.0059	-0.0014	0.0062	-0.0071	0.0045	0.0014	0.0053	0.0093*	0.0052	0.0099	0.0075	0.0026*	0.0014
Water_heat_electric	0.0009	0.0020	-0.0021	0.0028	0.0012	0.0018	-0.0045**	0.0019	-0.0034**	0.0014	-0.0007	0.0016	-0.0005	0.0016	-0.0003	0.0023	0.0009**	0.0004
Cooking_type_electric	0.0059**	0.0021	-0.0081**	0.0030	0.0011	0.0019	-0.0015	0.0020	0.0045***	0.0015	-0.0049***	0.0017	-0.0014	0.0017	-0.0106***	0.0024	0.0015***	0.0005
Efficiency_betw_10_20	0.0028	0.0025	0.0014	0.0036	0.0025	0.0022	0.0023	0.0024	-0.0009	0.0017	-0.0006	0.0020	0.0005	0.0020	-0.0046	0.0029	-0.0002	0.0005
Efficiency_betw_20_30	0.0010	0.0039	-0.0111*	0.0056	0.0020	0.0035	-0.0033	0.0037	-0.0021	0.0027	-0.0034	0.0032	-0.0012	0.0031	-0.0088*	0.0045	-0.0005	0.0008
Efficiency_more_30	0.0028	0.0064	-0.0120	0.0092	0.0134**	0.0058	-0.0069	0.0061	-0.0091**	0.0044	-0.0046	0.0052	-0.0030	0.0051	-0.0117	0.0074	-0.0017	0.0014

Base variables: Dwelling type detach, HoH age 18 35, HH comp live alone, HoH social class AB, Water heat non electric, Cooking type non electric, Efficiency less 10.

* p < 0.1; ** p < 0.05; *** p < 0.01

Figure 7.5 shows two individual load profiles for typical low and high electricity consumption households as described by their characteristics. Each coefficient value was calculated based on the presence or absence of particular household characteristics. The coefficient values were used to derive the load profile by applying the inverse FFT. Household 1 is a typical two bedroom apartment with adults only, HoH age less than 36 years and of social class C with electric water heating and cooking and an efficiency indicator variable of less than 10%. In contrast, Household 2 is a five bedroom detached dwelling with adults and children, HoH age between 36 to 55 years and of social class AB with both electric water heating and cooking and an efficiency indicator variable of greater than 30%. The figure shows the difference between the two profiles with Household 1 using much less electricity across the day compared to Household 2, however, the general shape of the two profiles are similar.

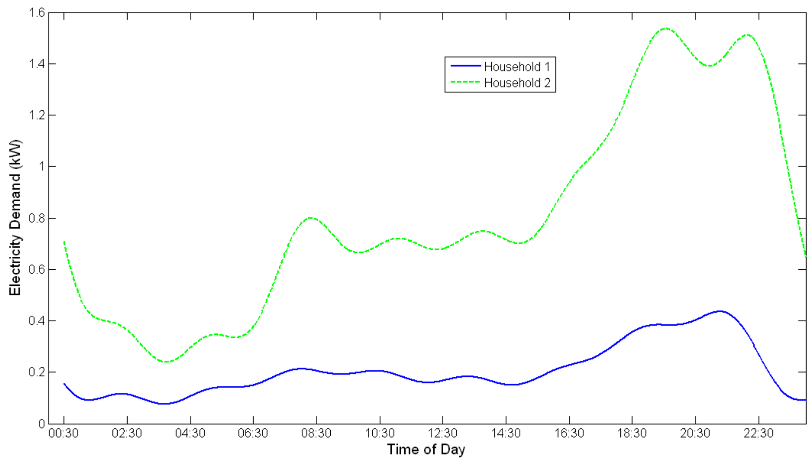


Figure 7.5: Household 1 and 2 electricity load profiles as calculated by regression of dwelling and occupant characteristics on Fourier transform coefficients respectively (low and high electricity user)

Figure 7.6 shows the influence of changing HoH age on electricity demand across the day with all other characteristics held constant at their base variable. Younger head of households use less electricity during the day and more in the evening times. This group also tend to use electricity earlier in the morning and later in the evening time. In contrast, older head of households use more electricity during the day and less in the evening time. These results are consistent with those presented earlier for the statistical characterisation approach shown in Chapter 5.

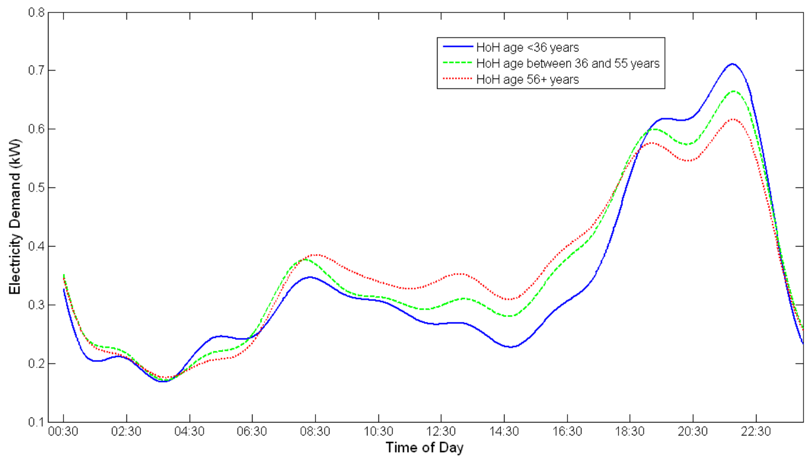


Figure 7.6: Load profiles by HoH age as calculated by regression of dwelling and occupant characteristics on Fourier transform coefficients

Figure 7.7 shows the influence of household composition on the daily load profile with all other variables held constant at their base category. The difference between the profiles is almost linear between the groups with adults and children almost consuming double that of a person living alone. These results are also consistent with those presented in Chapter 5.

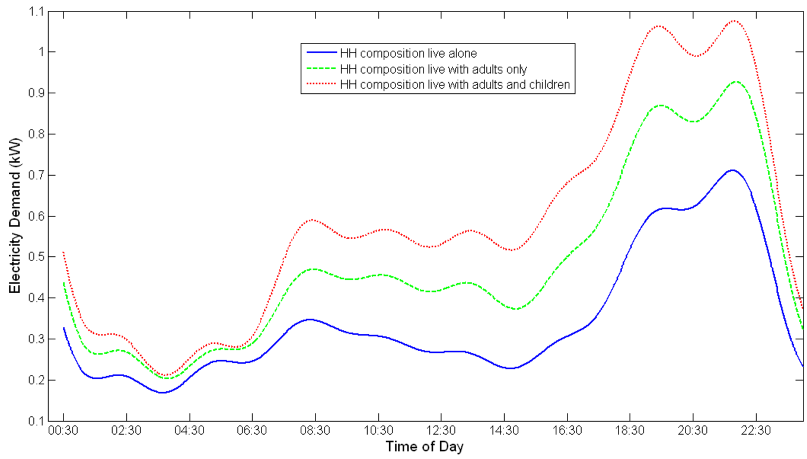


Figure 7.7: Load profiles by household composition as calculated by regression of dwelling and occupant characteristics on Fourier transform coefficients

Although the Fourier method of characterisation was able to quantify the dwelling and occupant variables as a function of the load profile shape across the day, the process resulted in a highly averaged profile similar to that presented earlier in Figure 1.3 for the standard load profile. This is the result of two factors. Firstly, a longitudinal averaging process was applied to each household before the regression was carried out. Secondly, the application of multivariate regression results in a smoothing out of the electricity load profile shape as individual characteristics are regressed against the Fourier coefficients as was shown in Figure 7.5 to Figure 7.7.

The same multivariate linear regression applied to electrical appliances (EA) approach in Section 5.2 is now carried out and results presented in Table 7.9. The appliances that influenced the -a0- Fourier coefficient term, at 95% or higher significance level were: tumble dryer, dishwasher, cooker, freezer (stand alone), televisions greater and smaller than 21 inch's, desktop computer and game consoles. These appliances all served to increase the constant value of the load profile. Similar to the results presented in Table

7.8, the number of appliances that significantly influenced frequency components of the electricity load profile is greater for the smaller coefficient values (a1 – b3). Therefore, this also suggests that it is the smaller, more deterministic components of the electricity load profile shape that are better characterised by appliance type.

Table 7.9: Regression results for appliance type on Fourier transform coefficients

Fourier Coefficeint Values	a0		a1		b1		a2		b2		a3		b3		a4		b4	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.1595***	0.0252	0.0073	0.0134	-0.0358**	0.0176	-0.0358**	0.0176	-0.0412***	0.0131	0.0005	0.0055	-0.0172**	0.0063	-0.0065	0.0047	-0.0116**	0.0043
Tumble dryer	0.0685***	0.0179	-0.0034	0.0096	-0.0468***	0.0125	-0.0468***	0.0125	-0.0124	0.0093	0.0028	0.0039	-0.0034	0.0045	0.0055	0.0034	-0.0022	0.0030
Dishwasher	0.0843***	0.0180	0.0064	0.0096	-0.0507***	0.0125	-0.0507***	0.0125	-0.0391***	0.0094	-0.0050	0.0039	0.0048	0.0045	0.0013	0.0034	-0.0012	0.0030
Shower (instant)	0.0174	0.0184	-0.0072	0.0098	-0.0059	0.0128	-0.0059	0.0128	-0.0029	0.0095	0.0035	0.0040	-0.0067	0.0046	0.0059*	0.0034	-0.0038	0.0031
Shower (pumped)	-0.0010	0.0183	0.0069	0.0098	-0.0043	0.0128	-0.0043	0.0128	-0.0056	0.0095	0.0039	0.0040	-0.0027	0.0046	-0.0015	0.0034	-0.0017	0.0031
Electric cooker	0.0525**	0.0191	-0.0043	0.0102	-0.0601***	0.0133	-0.0601***	0.0133	-0.0060	0.0099	-0.0075*	0.0042	-0.0121**	0.0048	0.0017	0.0036	0.0013	0.0032
Heater (plug in convective)	0.0152	0.0167	-0.0180**	0.0089	0.0041	0.0116	0.0041	0.0116	0.0159*	0.0086	0.0073*	0.0036	0.0100**	0.0042	0.0055*	0.0031	0.0030	0.0028
Freezer (stand alone)	0.0628***	0.0162	0.0061	0.0086	-0.0173	0.0112	-0.0173	0.0112	-0.0092	0.0084	-0.0032	0.0035	-0.0001	0.0041	0.0016	0.0030	-0.0021	0.0027
Water pump	0.0144	0.0201	-0.0045	0.0107	0.0081	0.0140	0.0081	0.0140	-0.0047	0.0104	0.0049	0.0044	0.0034	0.0050	-0.0023	0.0037	-0.0023	0.0034
Immersion	0.0222	0.0182	0.0114	0.0097	-0.0038	0.0127	-0.0038	0.0127	-0.0316***	0.0095	0.0010	0.0040	0.0030	0.0046	-0.0012	0.0034	-0.0001	0.0031
No. TV <21 in.	0.0179**	0.0087	0.0054	0.0046	-0.0227***	0.0061	-0.0227***	0.0061	-0.0013	0.0045	-0.0026	0.0019	0.0000	0.0022	0.0007	0.0016	-0.0019	0.0015
No. TV >21 in.	0.0264**	0.0098	0.0057	0.0052	-0.0233***	0.0068	-0.0233***	0.0068	-0.0024	0.0051	-0.0070***	0.0021	0.0008	0.0025	0.0024	0.0018	-0.0022	0.0016
No. computer (desktop)	0.0523***	0.0157	0.0093	0.0084	-0.0149	0.0109	-0.0149	0.0109	-0.0192**	0.0081	0.0075**	0.0034	-0.0066*	0.0039	-0.0030	0.0029	0.0010	0.0026
No. computer (laptop)	0.0164	0.0115	0.0149**	0.0061	0.0063	0.0080	0.0063	0.0080	-0.0160**	0.0060	0.0034	0.0025	0.0002	0.0029	-0.0050**	0.0021	0.0078***	0.0019
No. game consoles	0.0585***	0.0129	0.0053	0.0069	-0.0372***	0.0090	-0.0372***	0.0090	-0.0088	0.0067	0.0108***	0.0028	-0.0006	0.0032	-0.0007	0.0024	-0.0019	0.0022

Fourier Coefficeint Values	a5		b5		a6		b6		a7		b7		a8		b8		w	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.0035	0.0031	-0.0042	0.0044	0.0002	0.0028	-0.0035	0.0030	-0.0003	0.0022	-0.0061**	0.0026	0.0010	0.0025	-0.0029	0.0037	0.1362***	0.0007
Tumble dryer	0.0007	0.0022	-0.0076**	0.0031	0.0012	0.0020	-0.0049**	0.0021	0.0018	0.0015	-0.0025	0.0018	-0.0015	0.0018	-0.0054**	0.0026	0.0001	0.0005
Dishwasher	0.0055**	0.0022	-0.0064**	0.0031	0.0017	0.0020	-0.0032	0.0021	-0.0038**	0.0016	-0.0011	0.0018	-0.0024	0.0018	-0.0043	0.0026	-0.0018***	0.0005
Shower (instant)	0.0028	0.0023	-0.0063*	0.0032	0.0008	0.0021	-0.0028	0.0022	0.0009	0.0016	-0.0029	0.0019	-0.0018	0.0018	-0.0052*	0.0027	0.0010**	0.0005
Shower (pumped)	-0.0006	0.0023	-0.0018	0.0032	-0.0009	0.0021	-0.0002	0.0022	0.0008	0.0016	0.0009	0.0019	-0.0003	0.0018	-0.0047*	0.0027	-0.0002	0.0005
Electric cooker	0.0040*	0.0024	-0.0063*	0.0033	0.0010	0.0021	-0.0009	0.0023	0.0035**	0.0016	-0.0057***	0.0019	-0.0021	0.0019	-0.0109***	0.0028	0.0014***	0.0005
Heater (plug in convective)	-0.0029	0.0021	-0.0047	0.0029	0.0004	0.0019	0.0033*	0.0020	-0.0013	0.0014	0.0012	0.0017	0.0008	0.0016	0.0062**	0.0024	-0.0001	0.0004
Freezer (stand alone)	0.0023	0.0020	0.0013	0.0028	-0.0009	0.0018	0.0008	0.0019	0.0019	0.0014	-0.0007	0.0016	-0.0001	0.0016	-0.0005	0.0024	0.0004	0.0004
Water pump	-0.0010	0.0025	0.0064*	0.0035	0.0033	0.0023	-0.0047*	0.0024	-0.0007	0.0017	0.0009	0.0020	-0.0008	0.0020	0.0007	0.0029	0.0003	0.0005
Immersion	0.0013	0.0023	-0.0007	0.0032	0.0004	0.0020	-0.0031	0.0021	-0.0012	0.0016	0.0000	0.0018	-0.0001	0.0018	-0.0022	0.0027	0.0002	0.0005
No. TV <21 in.	0.0007	0.0011	-0.0033**	0.0015	-0.0018*	0.0010	-0.0002	0.0010	-0.0004	0.0007	-0.0013	0.0009	0.0005	0.0009	-0.0002	0.0013	-0.0001	0.0002
No. TV >21 in.	0.0008	0.0012	-0.0004	0.0017	-0.0008	0.0011	0.0000	0.0012	0.0018**	0.0008	-0.0004	0.0010	0.0009	0.0010	-0.0011	0.0014	-0.0001	0.0003
No. computer (desktop)	0.0024	0.0019	-0.0087***	0.0027	0.0025	0.0018	-0.0011	0.0019	0.0006	0.0013	-0.0020	0.0016	0.0005	0.0016	-0.0073***	0.0023	-0.0004	0.0004
No. computer (laptop)	0.0012	0.0014	0.0008	0.0020	0.0022*	0.0013	-0.0018	0.0014	-0.0009	0.0010	-0.0009	0.0012	0.0001	0.0011	0.0003	0.0017	-0.0004	0.0003
No. game consoles	-0.0019	0.0016	-0.0040*	0.0022	0.0057***	0.0015	0.0010	0.0015	-0.0023**	0.0011	-0.0008	0.0013	-0.0018	0.0013	-0.0051**	0.0019	0.0000	0.0003

Base variables: Washing machine

* p < 0.1; ** p < 0.05; *** p < 0.01

Figure 7.8 shows load profiles for four common household appliance types: electric cooker, dishwasher, freezer (stand alone) and electric shower. Similar to the dwelling and occupant characteristics investigated earlier, the load profiles reflect highly averaged electricity use by appliance type. The dishwasher profile shows the greatest amount of electricity use, particularly at peak times. As discussed earlier in Section 5.2, this is an appliance that shows potential for shifting domestic electricity demand away from peak time use.

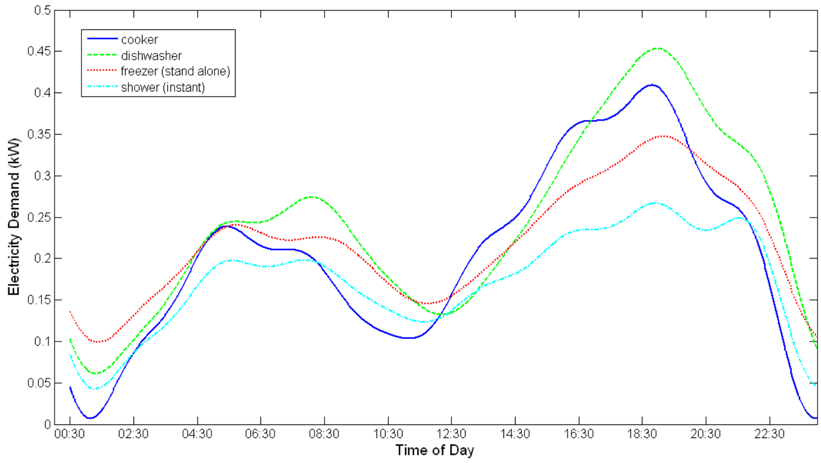


Figure 7.8: Load profiles by appliance type as calculated by regression of appliance characteristics on Fourier transform coefficients

Regression – Gaussian processes

The same multivariate linear regression was applied to the Gaussian process characterised time series. Similar to Fourier transforms, median values of Gaussian processes coefficients over the yearly period are used in the regression. Table 7.10 shows results for each Gaussian process coefficient and dwelling and occupant characteristics.

Table 7.10: Regression results for dwelling and occupant characteristics on Gaussian process coefficients

	a1		b1		c1		a2		b2		c2		a3		b3		c3		a4		b4		c4	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	-0.0211	0.0389	23.6033***	0.1440	1.6071***	0.1498	-0.1380	0.1062	24.9305***	0.0920	2.9115***	0.2650	-0.1755**	0.0665	23.35890***	0.0720	1.6466***	0.1121	0.4170***	0.1471	42.1470***	0.2856	3.3148***	0.3590
Dwell_type_semi_d	-0.0278**	0.0133	0.0458	0.0491	-0.0858*	0.0511	0.0171	0.0362	0.0255	0.0314	-0.0542	0.0904	-0.0012	0.0227	0.0299	0.0246	-0.0065	0.0382	-0.0397	0.0502	-0.2088**	0.0974	-0.1398	0.1225
Dwell_type_terr	-0.0342**	0.0170	0.0363	0.0629	-0.0023	0.0654	-0.0390	0.0464	0.0388	0.0402	-0.0996	0.1157	-0.0168	0.0291	0.0160	0.0314	0.0053	0.0490	-0.1089*	0.0643	-0.2713**	0.1248	-0.1094	0.1568
Dwell_type_apt	-0.0500	0.0380	0.0868	0.1407	-0.1086	0.1464	-0.0610	0.1038	0.0266	0.0900	-0.2451	0.2590	-0.0853	0.0650	-0.0134	0.0704	-0.0432	0.1096	-0.1618	0.1438	-0.2660	0.2792	-0.3107	0.3509
No_bedrooms	0.0141*	0.0073	-0.0068	0.0269	0.0305	0.0280	0.0429**	0.0199	0.0122	0.0172	0.0337	0.0496	0.0181	0.0124	0.0082	0.0135	0.0049	0.0210	0.1204***	0.0275	0.0647	0.0534	0.0780	0.0672
HoH_age_36_55	0.0159	0.0205	0.0146	0.0760	0.1238	0.0791	0.0459	0.0560	-0.0239	0.0486	0.0700	0.1399	0.0626*	0.0351	0.0113	0.0380	0.0704	0.0592	-0.1751**	0.0777	0.0071	0.1508	0.4386**	0.1895
HoH_age_56_plus	0.0322	0.0224	-0.0516	0.0828	0.1010	0.0862	0.1138*	0.0611	-0.0158	0.0530	-0.2185	0.1524	0.1185***	0.0383	0.0738*	0.0414	0.0495	0.0645	-0.2696***	0.0846	-0.0067	0.1643	0.7504***	0.2066
HH_comp_with_adults	0.0829***	0.0141	0.0393	0.0523	0.1118**	0.0544	0.1787***	0.0386	0.0808**	0.0335	-0.0901	0.0963	0.1384***	0.0242	0.0017	0.0262	-0.0582	0.0407	0.2403***	0.0535	-0.0155	0.1038	0.0374	0.1305
HH_comp_with_adults_children	0.1306***	0.0186	0.0885	0.0687	0.1495**	0.0715	0.2557***	0.0507	0.0879*	0.0439	-0.2095	0.1265	0.2366***	0.0318	-0.0093	0.0344	0.0114	0.0535	0.4594***	0.0702	-0.2946**	0.1363	-0.0849	0.1714
Social_class_C	0.0395**	0.0192	0.0508	0.0710	0.0010	0.0739	0.0311	0.0524	0.0187	0.0454	-0.1514	0.1307	0.0827**	0.0328	-0.0472	0.0355	-0.1201**	0.0553	-0.0712	0.0725	-0.1611	0.1409	-0.2291	0.1771
Social_class_DE	0.0633***	0.0206	0.0618	0.0762	-0.1216	0.0793	0.1303**	0.0562	-0.0175	0.0487	-0.5176***	0.1402	0.1312***	0.0352	-0.0512	0.0381	-0.2131***	0.0593	-0.1057	0.0778	-0.1104	0.1511	-0.1221	0.1900
Social_class_F	0.0364	0.0369	-0.0040	0.1366	-0.2480*	0.1421	0.0844	0.1007	-0.0946	0.0873	-0.7063**	0.2514	0.1855***	0.0631	0.0084	0.0683	-0.1690	0.1063	-0.3047**	0.1396	-0.1443	0.2710	0.0454	0.3406
Water_heat_electric	0.0150	0.0112	-0.0195	0.0416	-0.0680	0.0432	-0.0029	0.0307	-0.0413	0.0266	-0.1085	0.0765	0.0618***	0.0192	-0.0422**	0.0208	-0.0030	0.0324	0.0354	0.0425	-0.2249**	0.0825	-0.1974*	0.1037
Cooking_type_electric	0.0345***	0.0120	0.1128**	0.0444	0.0393	0.0462	0.1292***	0.0328	0.0506*	0.0284	-0.0997	0.0817	0.0761***	0.0205	0.0458**	0.0222	-0.0035	0.0346	0.1351***	0.0454	-0.4574***	0.0881	-0.4397***	0.1108
Efficiency_betw_10_20	0.0129	0.0141	-0.0500	0.0520	-0.0095	0.0542	0.0154	0.0384	-0.0075	0.0333	0.0239	0.0958	0.0259	0.0241	-0.0227	0.0260	-0.0441	0.0405	0.0707	0.0532	-0.0934	0.1033	0.0540	0.1298
Efficiency_betw_20_30	0.0480**	0.0221	-0.1349	0.0818	0.1549*	0.0852	-0.0035	0.0604	-0.0067	0.0523	0.2907*	0.1506	0.0476	0.0378	-0.0563	0.0409	0.0719	0.0637	0.1782**	0.0836	0.0011	0.1624	0.0546	0.2041
Efficiency_more_30	0.0108	0.0363	0.0624	0.1344	-0.1556	0.1398	-0.0226	0.0991	0.0770	0.0859	-0.2997	0.2473	0.0006	0.0621	-0.0540	0.0672	-0.1677	0.1046	0.4558***	0.1373	0.0154	0.2666	-0.3665	0.3351

	a5		b5		c5		a6		b6		c6		a7		b7		c7		a8		b8		c8	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	-0.3903**	0.1826	35.7805***	0.1404	1.5567***	0.1035	-0.3363***	0.1167	28.5307***	0.0857	2.0087***	0.1244	-0.1847	0.1302	20.6011***	0.1051	1.4367***	0.1172	0.2253	0.1493	14.0926***	0.2832	4.4074***	0.5993
Dwell_type_semi_d	-0.0042	0.0623	0.0654	0.0479	-0.0026	0.0353	-0.0107	0.0398	-0.0086	0.0292	-0.0382	0.0424	-0.0414	0.0444	-0.0060	0.0358	-0.0098	0.0400	-0.0331	0.0509	0.0863	0.0966	-0.4403**	0.2044
Dwell_type_terr	-0.0307	0.0798	-0.0734	0.0613	0.0008	0.0452	0.0104	0.0510	-0.0053	0.0374	-0.0644	0.0543	0.0239	0.0569	-0.0350	0.0459	0.0335	0.0512	0.0644	0.0652	0.2891**	0.1237	-0.8216***	0.2618
Dwell_type_apt	0.1639	0.1785	-0.0647	0.1372	-0.1399	0.1012	-0.0341	0.1140	0.0301	0.0838	-0.1909	0.1216	-0.1219	0.1272	0.0132	0.1027	-0.0355	0.1145	-0.0727	0.1459	0.0408	0.2768	-0.7194	0.5857
No_bedrooms	0.1192***	0.0342	0.0169	0.0263	-0.0017	0.0194	0.0797***	0.0218	-0.0053	0.0160	-0.0138	0.0233	0.0345	0.0243	0.0217	0.0197	0.0162	0.0219	-0.0019	0.0279	-0.0079	0.0530	0.0018	0.1121
HoH_age_36_55	0.1663*	0.0964	-0.1781**	0.0741	-0.0790	0.0547	0.1384**	0.0616	0.0241	0.0452	-0.0537	0.0657	0.0381	0.0687	0.0196	0.0555	0.0083	0.0619	0.1333*	0.0788	0.0585	0.1495	-0.1368	0.3163
HoH_age_56_plus	0.1546	0.1050	-0.2737***	0.0808	-0.0298	0.0596	0.1789**	0.0671	-0.0670	0.0493	-0.0619	0.0716	0.1439*	0.0749	-0.1164*	0.0604	0.0156	0.0674	0.1309	0.0859	0.1514	0.1629	-0.0583	0.3448
HH_comp_with_adults	0.4294***	0.0664	0.0129	0.0510	-0.0645*	0.0376	0.1932***	0.0424	-0.0083	0.0311	-0.0832*	0.0452	0.1843***	0.0473	0.0238	0.0382	0.0027	0.0426	0.1152**	0.0543	0.2281**	0.1029	-0.3580	0.2178
HH_comp_with_adults_children	0.6460***	0.0872	-0.1107	0.0670	0.0600	0.0494	0.3504***	0.0557	0.0079	0.0409	-0.0903	0.0594	0.2923***	0.0621	0.0125	0.0501	0.0834	0.0559	0.2654***	0.0713	0.5367***	0.1352	-0.5910**	0.2861
Social_class_C	-0.0153	0.0900	-0.0132	0.0692	0.0089	0.0511	0.0891	0.0575	-0.0274	0.0423	-0.0445	0.0614	0.0826	0.0642	-0.0117	0.0518	-0.0751	0.0578	-0.1259*	0.0736	0.1800	0.1396	-0.4915	0.2955
Social_class_DE	0.0082	0.0966	-0.1162	0.0743	-0.0445	0.0548	0.1275**	0.0617	-0.0652	0.0453	-0.1667**	0.0658	0.1750**	0.0689	-0.0171	0.0556	-0.1819***	0.0620	-0.2216**	0.0790	0.2889*	0.1498	-0.1096	0.3171
Social_class_F	-0.3681**	0.1732	0.1041	0.1332	-0.0276	0.0982	0.0325	0.1107	-0.0804	0.0813	-0.0987	0.1180	0.2192*	0.1235	-0.1989*	0.0997	-0.2431**	0.1112	-0.3034**	0.1416	0.0654	0.2686	0.6410	0.5685
Water_heat_electric	0.0975*	0.0527	0.0471	0.0405	-0.0011	0.0299	-0.0132	0.0337	-0.0538**	0.0247	0.0743**	0.0359	0.1385***	0.0376	-0.0412	0.0303	0.0022	0.0338	0.1278***	0.0431	0.1828**	0.0818	-0.3716**	0.1730
Cooking_type_electric	0.4578***	0.0563	0.0530	0.0433	-0.1501***	0.0319	0.1752***	0.0360	-0.0056	0.0264	-0.0630	0.0384	0.1243***	0.0402	0.0161	0.0324	0.0354	0.0362	0.0371	0.0461	0.2001**	0.0874	-0.3458*	0.1849
Efficiency_betw_10_20	0.0655	0.0660	0.0251	0.0508	-0.0007	0.0374	-0.0147	0.0422	-0.0405	0.0310	0.0266	0.0450	0.0309	0.0471	0.0347	0.0380	-0.0043	0.0424	-0.0088	0.0540	-0.0751	0.1024	-0.0846	0.2166
Efficiency_betw_20_30	0.2765**	0.1038	-0.0994	0.0798	0.0056	0.0589	0.0613	0.0663	0.0106	0.0487	0.1054	0.0707	0.0431	0.0740	0.0409	0.0597	0.0473	0.0666	-0.0122	0.0849	0.1095	0.1610	-0.0192	0.3407
Efficiency_more_30	0.2778	0.1704	0.1186	0.1311	-0.0584	0.0966	-0.0026	0.1089	-0.0524	0.0800	-0.0769	0.1161	-0.1066	0.1215	0.0629	0.0981	0.0567	0.1094	0.0802	0.1393	0.3842	0.2643	-0.7594	0.5593

Base variables: Dwelling type detach, HoH age 18 35, HH comp live alone, HoH social class AB, Water heat non electric, Cooking type non electric, Efficiency less 10.

* p < 0.1; ** p < 0.05; *** p < 0.01

Similar to the Fourier transform characterisation, dwelling and occupant characteristics that had a significant influence over the Gaussian process coefficients were household composition, electric cooking and to a lesser extent number of bedrooms and electric water heating. The Gaussian process coefficient terms a1 - a8 had the most significant influence over the characteristics which is not surprising as they correlate with the magnitude component for the electricity load profile.

Figure 7.9 shows an electricity load profile for the same typical low and high electricity households presented in Figure 7.5, as characterised with Gaussian processes. The profile shape is quite different to that presented in Figure 7.5 and looks less like the standard load profile presented in Figure 1.3 and more like that of an individual dwelling electricity load profile shown in Figure 1.2.

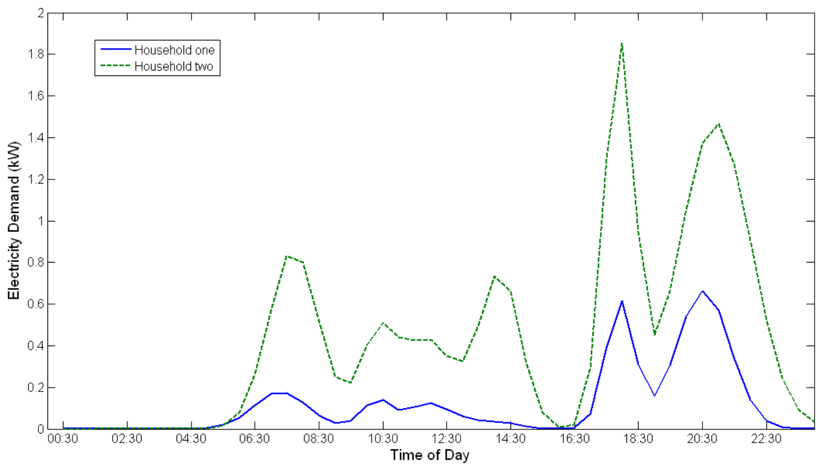


Figure 7.9: Household 1 and 2 electricity load profiles as calculated by regression of dwelling and occupant characteristics on Gaussian process coefficients respectively (low and high electricity user)

Figure 7.10 shows the influence of HoH age while keeping all other parameters constant at their base category. Consistent with results presented earlier in Figure 7.6 younger HoH's (less than 36 years old) use less electricity in the morning but significantly more in the evening time. Also older (HoH age 56 years old plus) use more electricity across the day and the least amount in the evening time.

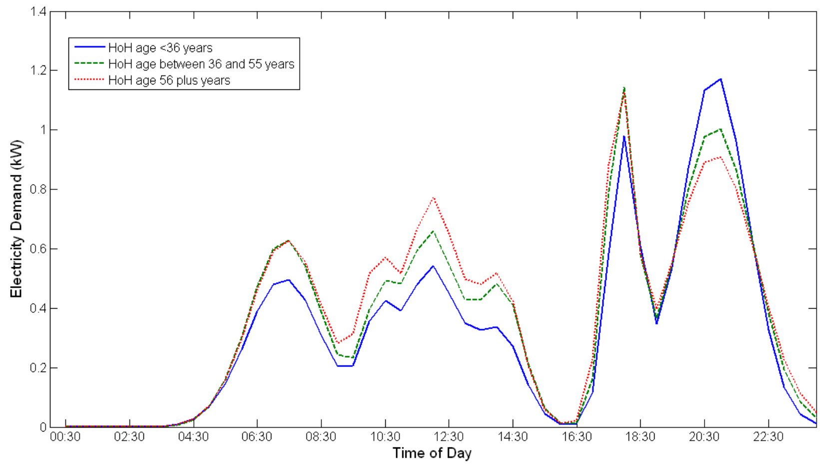


Figure 7.10: Load profiles by HoH age as calculated by regression of dwelling and occupant characteristics on Gaussian process coefficients

Figure 7.11 shows the influence of household composition on the daily load profile shape with all other independent variables held constant at their base category. Similar to Figure 7.7 the difference between profiles is almost linear between the different household composition groups. However, it is interesting to note the small lunch time peak for persons living alone compared to the other groups which was not identified by the Fourier transforms characterisation process. This is most probably due to a smaller probability of a person being home at this time for this category.

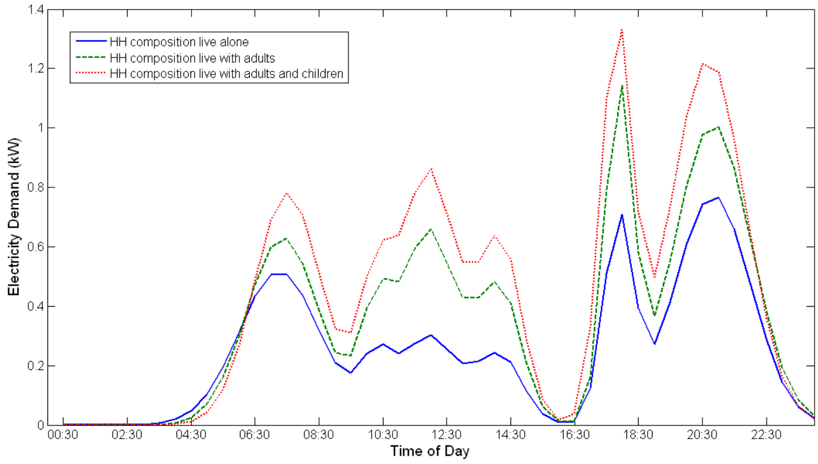


Figure 7.11: Load profiles by household composition as calculated by regression of dwelling and occupant characteristics on Gaussian processes coefficients

The same multivariate linear regression EA approach applied above with Fourier transforms is now used with Gaussian process coefficients and results presented in Table 7.11. In contrast to the Fourier transform approach there wasn't a single or group of electrical appliances that dominated significantly in the multivariate regression for Gaussian processes. However, the constant term in the regression was nearly always significant at the 99% level suggesting that appliances not included in the survey may also be significantly influencing the load profile shape throughout the day.

Table 7.11: Regression results for appliance type on Gaussian process coefficients

	a1		b1		c1		a2		b2		c2	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.0663***	0.0178	23.6726***	0.0658	1.6003***	0.0673	0.1551***	0.0491	24.9582***	0.0412	2.1393***	0.1232
Tumble dryer	0.0445***	0.0127	0.0469	0.0468	-0.0092	0.0478	0.0636*	0.0349	0.0728**	0.0293	-0.0129	0.0876
Dishwasher	0.0229*	0.0127	-0.0264	0.0470	0.0025	0.0481	0.0577	0.0351	-0.0633**	0.0294	0.0246	0.0880
Shower (instant)	0.0196	0.0130	0.1229**	0.0480	-0.0538	0.0491	0.0790**	0.0358	0.0272	0.0300	-0.1868**	0.0898
Shower (pumped)	-0.0086	0.0129	0.0335	0.0478	-0.0329	0.0489	-0.0241	0.0357	0.0576*	0.0300	-0.1008	0.0895
Electric cooker	0.0166	0.0135	0.0587	0.0498	0.0842	0.0509	0.0909**	0.0371	0.0319	0.0312	-0.0010	0.0932
Heater (plug in convective)	0.0224*	0.0118	-0.0002	0.0434	0.0170	0.0444	0.0777**	0.0324	-0.0810***	0.0272	-0.0185	0.0813
Freezer (stand alone)	0.0200*	0.0114	-0.0579	0.0421	0.0878**	0.0431	0.0375	0.0314	-0.0062	0.0264	0.1054	0.0789
Water pump	0.0032	0.0142	0.0356	0.0523	0.0109	0.0535	-0.0457	0.0390	0.0571*	0.0328	0.1538	0.0979
Immersion	-0.0109	0.0128	-0.0016	0.0475	-0.0175	0.0486	-0.0796**	0.0354	-0.0005	0.0297	0.1176	0.0889
No. TV <21 in.	0.0120*	0.0061	-0.0060	0.0227	0.0471**	0.0232	0.0182	0.0169	0.0077	0.0142	-0.0004	0.0425
No. TV >21 in.	0.0130*	0.0069	-0.0262	0.0255	0.0083	0.0261	0.0349*	0.0190	0.0129	0.0160	0.0224	0.0478
No. computer (desktop)	0.0238**	0.0111	-0.0548	0.0409	0.1234***	0.0418	0.0257	0.0305	0.0053	0.0256	0.2176***	0.0766
No. computer (laptop)	0.0054	0.0081	-0.0252	0.0300	0.0435	0.0306	-0.0150	0.0223	0.0208	0.0188	0.1752***	0.0561
No. game consoles	0.0278***	0.0091	0.0027	0.0337	0.0582*	0.0345	0.0422*	0.0251	-0.0036	0.0211	0.0237	0.0631

	a3		b3		c3		a4		b4		c4	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.0998***	0.0317	23.3798***	0.0344	1.4491***	0.0518	0.2987***	0.0701	42.1262***	0.1313	3.7481***	0.1657
Tumble dryer	0.0668***	0.0226	0.0024	0.0245	-0.0174	0.0368	0.0978*	0.0498	-0.2107**	0.0934	-0.1985*	0.1178
Dishwasher	0.0355	0.0227	0.0067	0.0246	0.0021	0.0370	0.2060***	0.0501	0.1674*	0.0938	0.0993	0.1184
Shower (instant)	0.0331	0.0231	-0.0206	0.0251	-0.0598	0.0377	0.0284	0.0511	-0.0846	0.0957	-0.2564**	0.1208
Shower (pumped)	-0.0027	0.0231	-0.0349	0.0250	-0.0547	0.0376	0.0162	0.0510	-0.0098	0.0955	0.0169	0.1205
Electric cooker	0.0502**	0.0240	0.0654**	0.0260	0.0527	0.0392	0.1298**	0.0530	-0.3875***	0.0994	-0.4827***	0.1254
Heater (plug in convective)	0.0371*	0.0210	0.0208	0.0227	0.0214	0.0342	-0.0673	0.0463	0.1354	0.0867	0.2820**	0.1094
Freezer (stand alone)	0.0141	0.0203	-0.0144	0.0220	0.0026	0.0332	0.0693	0.0449	-0.0591	0.0841	0.1030	0.1061
Water pump	0.0001	0.0252	-0.0267	0.0273	0.0709*	0.0412	0.0139	0.0557	-0.0674	0.1044	-0.0144	0.1318
Immersion	-0.0059	0.0229	-0.0473*	0.0248	0.0121	0.0374	0.0658	0.0506	-0.2656**	0.0948	-0.0978	0.1196
No. TV <21 in.	0.0221**	0.0109	-0.0074	0.0119	-0.0046	0.0179	0.0438*	0.0242	0.0334	0.0453	0.0792	0.0572
No. TV >21 in.	0.0232*	0.0123	0.0048	0.0133	0.0078	0.0201	0.0431	0.0272	-0.0409	0.0509	0.1728**	0.0643
No. computer (desktop)	0.0065	0.0197	0.0643***	0.0214	0.0865**	0.0322	0.1147**	0.0436	0.0912	0.0816	0.0706	0.1030
No. computer (laptop)	-0.0327**	0.0144	-0.0060	0.0157	0.0460*	0.0236	0.0607*	0.0319	0.0372	0.0598	0.0093	0.0755
No. game consoles	0.0536***	0.0163	-0.0362**	0.0176	-0.0119	0.0265	0.1094***	0.0359	-0.0172	0.0673	-0.1144	0.0849

	a5		b5		c5		a6		b6		c6	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	-0.0611	0.0824	35.6144***	0.0651	1.4689***	0.0471	0.0934*	0.0533	28.3687***	0.0393	1.6693***	0.0575
Tumble dryer	0.2942***	0.0586	0.0163	0.0463	-0.0871**	0.0335	0.1530***	0.0379	0.0302	0.0279	-0.0151	0.0408
Dishwasher	0.1731***	0.0589	0.0101	0.0465	0.0635**	0.0336	0.0520	0.0380	0.0222	0.0281	0.0156	0.0410
Shower (instant)	0.0636	0.0601	-0.0115	0.0475	-0.0568	0.0343	0.0543	0.0388	-0.0583**	0.0286	-0.0632	0.0419
Shower (pumped)	0.0337	0.0599	0.0007	0.0473	-0.0395	0.0342	-0.0002	0.0387	-0.0647**	0.0286	-0.0458	0.0418
Electric cooker	0.3492***	0.0624	0.0833*	0.0493	-0.0859**	0.0356	0.1286***	0.0403	0.0100	0.0297	-0.0185	0.0435
Heater (plug in convective)	-0.0296	0.0544	-0.0501	0.0430	0.0700**	0.0311	0.0683*	0.0352	-0.0204	0.0260	0.0141	0.0379
Freezer (stand alone)	0.1101**	0.0528	-0.0843**	0.0417	0.0297	0.0301	-0.0058	0.0341	0.0581**	0.0252	0.0741**	0.0368
Water pump	-0.0467	0.0655	-0.0143	0.0518	-0.0006	0.0374	0.0068	0.0423	0.0426	0.0313	-0.0016	0.0457
Immersion	0.1251**	0.0595	0.0468	0.0470	-0.0387	0.0340	-0.0172	0.0384	-0.0196	0.0284	0.0407	0.0415
No. TV <21 in.	0.0836***	0.0284	-0.0193	0.0225	-0.0083	0.0162	0.0350*	0.0184	-0.0033	0.0136	-0.0307	0.0198
No. TV >21 in.	0.0686**	0.0320	-0.0373	0.0253	-0.0023	0.0183	0.0167	0.0207	-0.0084	0.0152	0.0322	0.0223
No. computer (desktop)	0.1330**	0.0512	0.0266	0.0405	-0.0236	0.0293	0.0612*	0.0331	0.0143	0.0244	0.0286	0.0357
No. computer (laptop)	-0.0105	0.0375	0.0596*	0.0297	0.0213	0.0214	-0.0324	0.0243	0.0374**	0.0179	0.0660**	0.0262
No. game consoles	0.0995**	0.0422	-0.0534	0.0334	0.0529**	0.0241	0.1009***	0.0273	0.0399*	0.0201	0.0272	0.0294

	a7		b7		c7		a8		b8		c8	
	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error	Coef.	Std. Error
(Constant)	0.2007***	0.0615	20.5402***	0.0477	1.2539***	0.0533	0.1637**	0.0686	14.5093***	0.1308	3.7281***	0.2805
Tumble dryer	0.0612	0.0437	0.0496	0.0339	0.0688*	0.0379	0.0649	0.0487	0.0828	0.0930	-0.3323	0.1994
Dishwasher	0.0895**	0.0439	0.0093	0.0341	-0.0123	0.0380	0.0385	0.0490	-0.0161	0.0934	0.3423*	0.2003
Shower (instant)	0.0692	0.0448	-0.0333	0.0348	-0.0308	0.0388	0.0313	0.0500	0.2195**	0.0953	-0.6464***	0.2044
Shower (pumped)	-0.0448	0.0447	-0.0123	0.0347	-0.0099	0.0387	-0.0105	0.0498	0.0547	0.0951	-0.2566	0.2039
Electric cooker	0.0693	0.0465	0.0420	0.0361	0.0760*	0.0403	0.0045	0.0519	0.1095	0.0989	-0.1492	0.2122
Heater (plug in convective)	0.0581	0.0406	-0.0027	0.0315	0.0088	0.0352	0.0631	0.0453	0.0605	0.0863	0.0842	0.1852
Freezer (stand alone)	0.0347	0.0394	0.0092	0.0305	0.0140	0.0341	-0.0118	0.0439	-0.1012	0.0837	0.3707**	0.1796
Water pump	0.0221	0.0489	0.0089	0.0379	-0.0137	0.0423	-0.0275	0.0545	-0.0380	0.1040	0.3758*	0.2230
Immersion	0.0402	0.0444	-0.0260	0.0344	0.0214	0.0384	0.1186**	0.0495	0.1273	0.0944	-0.4049*	0.2025
No. TV <21 in.	0.0029	0.0212	0.0000	0.0165	0.0001	0.0184	-0.0372	0.0236	0.0789*	0.0451	-0.0722	0.0967
No. TV >21 in.	0.0256	0.0239	0.0141	0.0185	0.0095	0.0206	-0.0413	0.0266	0.0095	0.0507	0.1006	0.1088
No. computer (desktop)	-0.0140	0.0382	0.0175	0.0297	0.0803**	0.0331	0.1039**	0.0426	-0.0050	0.0813	-0.0648	0.1744
No. computer (laptop)	-0.0389	0.0280	0.0153	0.0217	0.0401	0.0242	0.0731**	0.0312	-0.0811	0.0595	-0.0698	0.1277
No. game consoles	0.0319	0.0315	0.0496**	0.0244	0.0286	0.0273	0.0912**	0.0351	0.0435	0.0670	-0.0191	0.1437

Base variables: Washing machine

* p < 0.1; ** p < 0.05; *** p < 0.01

Figure 7.12 shows individual electricity load profiles by common household appliance types: electric cooker, dishwasher, stand alone freezer and electric shower. Each individual profile is similar in shape, however, it is interesting to note a number of differences. Firstly, for appliance type (cooker) there is large surge in electricity demand in the early evening which typically corresponds with dinnertime in most households. Secondly, the appliance identified which uses the most electricity at peak times (i.e. morning 6.30am – 9.00am and late evening 19:30 – 20:30) is the dishwasher. This was already mentioned earlier, where in particular the dishwasher, offers good potential to shift electricity demand away from peak time use. Lastly, electric shower contributes significantly to morning and lunch time electricity use.

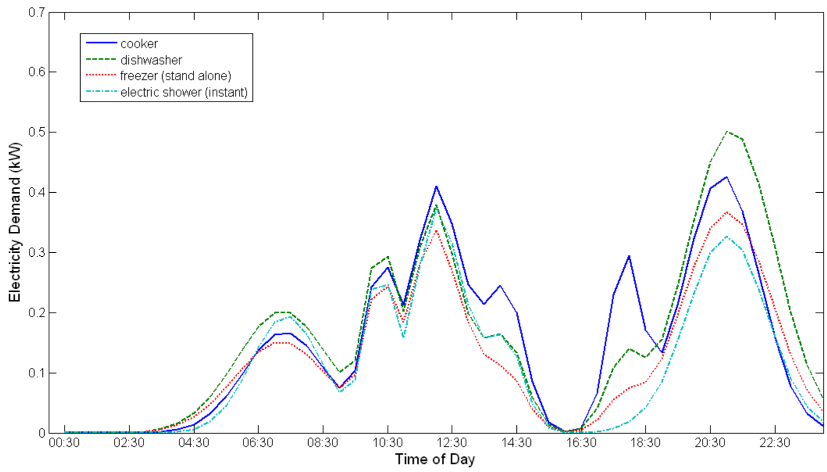


Figure 7.12: Load profiles by appliance type as calculated by regression of appliance characteristics on Gaussian process coefficients

7.3 Conclusion

A number of time series approaches to electricity load profile characterisation at an individual dwelling level were presented. Fourier transforms and Gaussian processes showed the greatest potential for domestic electricity load profile characterisation in order to meet the objectives outlined in Section 1.5 and hence these techniques were investigated further. Similar to the previous approaches, each technique was assessed based on evaluating and comparing electrical parameters between original and characterised profiles as well as carrying out a number of time series tests.

When assessing each characterisation technique in terms of electrical parameters, there were differing results. E_{Total} was successfully characterised using Fourier transforms, with a very small percentage error. Fourier transforms were less successful at characterising parameter E_{MD} where a significant percentage error was recorded. In contrast, Gaussian processes were able to sufficiently replicate parameter E_{MD} and less able to re-produce parameter E_{Total}. Fourier transforms performed slightly better than Gaussian processes when compared against parameter E_{ToU}, with the former over-estimating the time (i.e. later in the day) and the latter underestimating (i.e. earlier in the day). The time series tests showed similar results to those calculated by evaluating each electrical parameter. Depending upon the electricity demand load profile within the household, each approach had individual strengths. Fourier transforms were better able to characterise households who consumed larger amounts of electricity over longer intervals of time (>1hr), whereas Gaussian processes characterised households consuming higher amounts of electricity over shorter time intervals.

A multivariate linear regression was used to associate the influence of dwelling and occupant characteristics to the characterised electricity load profile shapes. The two statistical approaches, DOC and EA already used in Section 5.2 were applied to the characterised time series and results presented. The results showed that it was possible to associate dwelling and occupant characteristics to both Fourier and Gaussian process coefficients through multivariate regression and extract load profiles by applying the associated inverse transforms. A number of electricity load profiles were evaluated based on different dwelling and occupant characteristics. However, profiles tended to represent highly averaged load profile shapes (particular for Fourier transforms), which changed only marginally between varying dwelling and occupant characteristics and hence did not reflect the variation between household's and how they consumed electricity differently. This was a result of two factors: the longitudinal averaging applied and the regression process. Therefore a method of reducing the data first before any averaging is applied was sought and will be addressed in the next chapter.

The results presented in this chapter have shown that both methods; Fourier Transforms and Gaussian Processes are an effective method to characterise domestic electricity demand. It also showed that by applying both of these methods leads to a reduction in the number of variables required (18 values for Fourier Transforms and 24 for Gaussian Processes compared to the original time series consisting of 48) to sufficiently characterise a domestic electricity demand load profile. This reduces the amount of data required to describe electricity demand for individual households. Finally, investigating the influence of dwelling and occupant characteristics on specific electricity demand load profile patterns (described by characterisation coefficients) through multivariate regression although possible proved not to be an accurate reflection as to the manner with which electricity is consumed within the home.

CHAPTER 8

CLUSTERING

8 CLUSTERING

8.1 Introduction

Clustering (referred to as 'data mining' in the information services sector) is a useful tool for analysing large amounts of data. It is used to group data that show certain characteristic similarities together. Its use of late is becoming ever more prevalent, especially with the increasing number of new devices connecting to the internet on a daily basis, delivering large amounts of data. However, the availably of what has been described as a 'data tsunami' also poses its own problems. The collection of such a detailed amount of specific data means that traditional ways of analysing information such as statistical analysis has become increasingly more difficult, mainly due to their computationally intensive requirements and interpretation of the data. In addition, specific characteristics within the data often become lost, particularly when averages are applied. Therefore a method of filtering, to extract the most relevant pieces of information from the data, before such statistical analysis is applied is sometimes a necessary step. This not only simplifies the analysis but also diminishes the chances of losing vital characteristic information.

The application of clustering in the electricity industry is not a new concept. Historically, it has been used for customer segmentation, mainly for the purposes of tariff design [53]. However, its use at a domestic level has been somewhat limited to date. This chapter investigates three of the most widely used partitional clustering methods: k-means, k-medoid and Self Organising Maps (SOM). The best performing technique is evaluated in order to segment individual households into clusters based on their pattern of electricity use across the day. The process is repeated for each day over a six month period in order to characterise the diurnal, intra-daily and seasonal variations of domestic electricity demand. Finally a multi-nominal logistic regression is used to associate and determine the probability of a household with certain characteristics using a particular profile group. In addition a number of graphs are presented showing the percentage penetration of each characteristic for each profile group.

8.2 Discussion and Results

8.2.1 Evaluation of clustering techniques and number of clusters

The DB index was presented in Equation 4.9 and is used to evaluate an appropriate clustering technique and number of clusters. Smaller values of DB index indicate compact clusters with cluster centres further apart. In choosing a clustering method associated with a low DB index, this ensures that the characterisation process optimises the following two properties for the sample:

- firstly it ensures that the clustering method segmented households that used electricity most similarly into the same cluster; and
- secondly households that use electricity most differently were assigned to different clusters

Figure 8.1 shows the DB index for each clustering technique and varying number of clusters. K-medoid showed a comparatively larger DB index when compared to the other two methods. K-mean and SOM had similar results, however, SOM had a consistently lower DB index overall and hence was evaluated in the next section. Figure 8.1 also shows eight to ten clusters to be a good choice for segmenting the dataset. After this point the decrease in DB index is marginal for any further increase in the number of clusters. The levelling off in DB index indicates that the sample has been divided into an optimum number of disparate clusters.

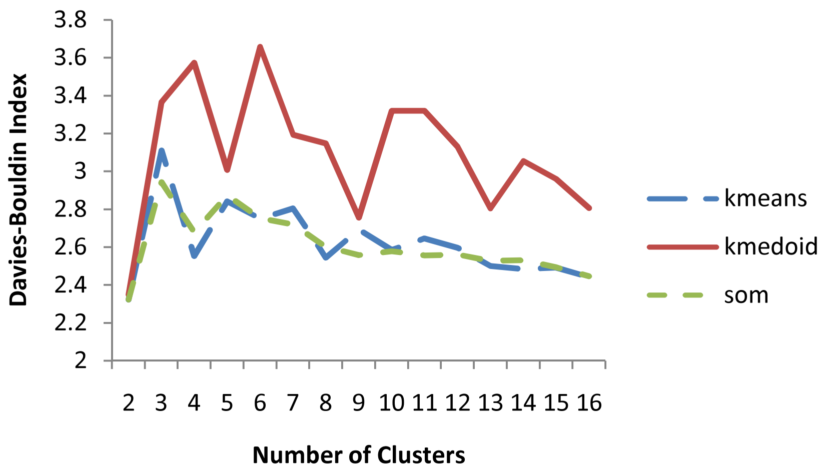


Figure 8.1: Davies-Bouldin index for partitional clustering methods k-means, k-medoid, and SOM

8.2.2 Characterising domestic electricity demand load profiles

Each cluster is defined by a weight vector which consists of 48 different dimensions, representing half hourly time intervals across a day. The mapping process is started by initialising weight vectors with random values at each cluster centre. As the network progresses each input vector is compared with the weights of each cluster centre and the one with the greatest similarity (called the Best Matching Unit) is assigned that particular vector. The weights are then adjusted at the cluster centre based on the input vector. The process is repeated until all input vectors have been assigned to clusters for an entire day. This procedure is repeated over the six month period until each daily electricity load profile for each household has been clustered.

Each daily profile for each household from *Dataset I* was segmented using SOM into nine different clusters (indicated by c1, c2, etc) based on a 3x3 hexagonal lattice structure. The number of clusters and structure was chosen based on the results from the DB index and in order for each cluster centre to be separated by a maximum of one cluster as is shown in Figure 8.2. This limited the potential for creating clusters with very small sample sizes which represented very uncommon patterns of electricity consumption throughout the day.

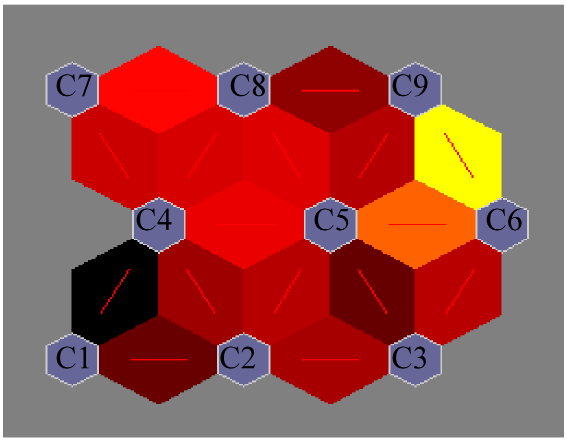


Figure 8.2: Hexagonal (3x3 lattice structure) for SOM clusters

Figure 8.2 shows the 3x3 hexagonal lattice structure for the nine different clusters. The cluster centres are shown to be visually separated by Euclidean distance which is also the metric used to compare individual electricity load profiles to the weight vectors of each cluster. The Euclidean distance was described earlier in Equation 2.13. The brighter colours in Figure 8.2 show clusters that are close together whereas the darker colours indicate cluster centres that are further apart. It can be seen that clusters c6 and c9 are similar to each other compared to any other cluster pair.

As discussed earlier, when applied to the entire dataset this produced two very large clusters, representing approximately two thirds of the entire sample. Therefore sub-clustering was used to divide clusters 6 and 9 into four clusters each. Figure 8.3 and Figure 8.4 shows the relative Euclidean distances (with brighter colours representing smaller and darker colours larger) for each of the two sub-clusters (c6 and c9).

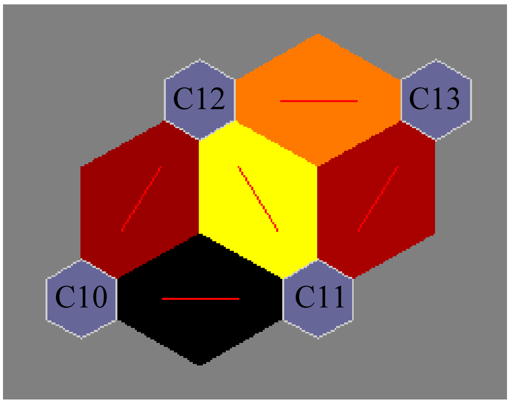


Figure 8.3: Sub-clustered hexagonal (2x2 lattice structure) for cluster 6

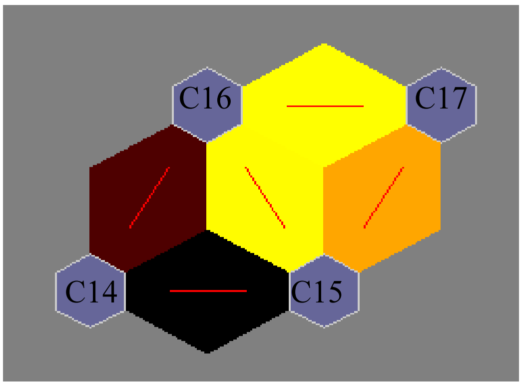


Figure 8.4: Sub-clustered hexagonal (2x2 lattice structure) for cluster 9

Ten profile groups in total were produced based on DB index results presented in Figure 8.1 and by combining profiles that differed only slightly in terms of magnitude and timing of electricity use within the home. The rational in doing this was to have a number of profiles that represented distinctly different patterns of electricity use within the home either in terms of timing or magnitude of electricity consumption throughout the day

Figure 8.5 shows the diurnal patterns of electricity use by day type for the ten profile groups (P1 – P10) where longitudinal averaging over the six month period was applied. A detailed description of each electricity load profile group is provided in Appendix A. A distinction between weekdays, Saturdays and Sundays is apparent. As one would anticipate, for the majority of profiles, there is an earlier morning peak for weekdays compared to weekends, as on average dwelling occupants get up earlier for work and school commitments. Similarly, less electricity is used throughout late morning to late afternoon during the weekdays as the majority of dwelling occupants are less likely to be at home at these times compared to the weekend. Finally, over the entire day dwelling occupants tend to use more electricity on Sundays compared to Saturdays or weekdays.

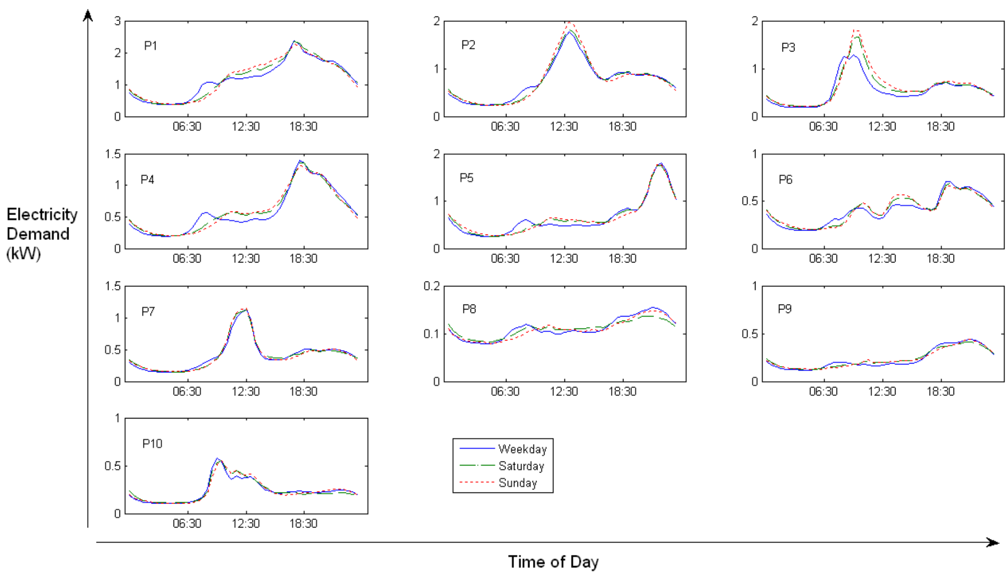


Figure 8.5: Mean electricity load profiles by day type over the six month period.

Table 8.1 shows the percentage Household Mode (HMode) as calculated by Equation 4.10. As stated earlier this reflects the percentage number of households that used a particular profile group most often over the six month period. Profile group P4 is the largest representing just less than one third of the entire sample.

Table 8.1: Percentage Household Mode (HMode) for the sample over the six month period

Profiles	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Total
Percentage HMode (%)	6.5%	8.2%	5.0%	31.1%	4.9%	13.0%	4.5%	12.1%	10.5%	4.2%	100%

As was shown earlier in Figure 1.4, patterns of electricity use within the home can change considerably on an intra-daily basis. Therefore, a particular household often uses more than one profile group across a period of time depending upon various factors within the home. This is shown in detail in Table 8.2 where Average Percentage Profile Time (APPT) over the six month period is calculated from Equation 4.11. However, each household on average will use a single profile group for the majority of time (as indicated by the diagonal in Table 8.2), with the remainder spread across a number of different profile groups. For example, a household that uses electricity within the home in a similar manner to Profile 1 (46.4% of the time) will also use Profile 4 (20.3% of the time). More detailed tables where a Gaussian probability distribution function is fitted to APPT for weekday, Saturday and Sunday are shown in Appendix B.

Table 8.2: Average Percentage Profile Time (APPT) over the six month period

Profile	Average Percentage Profile Time (APPT - %)									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
HMode (P1)	**46.4%**	12.7%	6.5%	20.3%	8.3%	2.7%	0.7%	1.2%	1.0%	0.2%
HMode (P2)	11.4%	**35.4%**	8.9%	18.7%	5.8%	8.9%	5.6%	2.2%	2.2%	1.1%
HMode (P3)	4.1%	8.7%	**36.6%**	16.8%	2.8%	9.5%	4.8%	6.4%	6.2%	3.9%
HMode (P4)	7.4%	10.1%	6.8%	**43.0%**	6.8%	12.3%	3.7%	3.4%	5.0%	1.5%
HMode (P5)	10.6%	7.1%	3.1%	24.0%	**37.5%**	9.1%	1.5%	3.0%	3.4%	0.6%
HMode (P6)	1.5%	5.3%	4.8%	18.9%	5.1%	**38.9%**	7.3%	3.4%	11.3%	3.5%
HMode (P7)	1.4%	9.7%	4.6%	11.3%	0.8%	13.8%	**36.8%**	4.7%	8.3%	8.6%
HMode (P8)	0.6%	1.4%	3.0%	6.2%	1.1%	4.6%	2.1%	**57.9%**	16.5%	6.6%
HMode (P9)	0.3%	1.2%	3.3%	10.6%	1.2%	13.4%	3.8%	16.2%	**42.0%**	7.9%
HMode (P10)	0.3%	2.2%	7.4%	6.8%	0.6%	10.7%	8.6%	15.5%	13.8%	**34.1%**

Electrical Parameters

The electrical parameters discussed in the methodology Chapter 4 were calculated by combining the CPI with each profile group for each day over the six month period. A paired sample t-test was used to compare the characterised profiles shown earlier in Table 5.1 to that of the original sample data with the results presented in

Table 8.3.

Table 8.3: Paired sample t-test for electrical parameters for original sample and characterised profiles

Paired Samples	**Mean**	**Standard Deviation**	**Std. Error Mean**	**t**	**Sig. (2-tailed)**
E_{TOTAL} (kWh)	8.13	387.97	6.18	1.32	0.1880
E_{MD} (kW)	1.33	0.64	0.01	130.47	0.0000
E_{LF} (%)	-0.244	0.0797	0.0013	-192.32	0.0000
E_{ToU} *	-4.61	3.73	0.06	-77.72	0.0000

*E_{ToU} (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00)

Table 8.4 presents electrical parameter characterisation results and can be compared against the original sample data presented earlier in Table 5.1. A Weibull and a Log-logistic probability distribution were fitted to the parameters.

Table 8.4: Descriptive statistics for electrical parameters for characterised profiles

Parameter	Mean	Median	Std Dev.	Max	Min	Scale Parameter (η)	Shape Parameter (α)
E_{TOTAL} (kWh)	2,269	2,292	864	5,497	498	2,546	2.85
E_{MD} (kW)	1.30	1.36	0.52	2.70	0.17	1.45	2.75
E_{LF} (%)	44.04%	42.03%	6.23%	66.86%	34.15%	-0.8464*	0.0701*
E_{ToU}**	35.66	35.66	3.13	n/a	n/a	n/a	n/a

Weibull Probability Distribution Function

$$f(T) = \frac{\alpha}{\eta}\left(\frac{T}{\eta}\right)^{\alpha-1} e^{-\left(\frac{T}{\eta}\right)^{\alpha}} \text{ where } f(T) \geq 0, \quad T \geq 0, \quad \alpha > 0, \quad \eta > 0$$

* Log-logistic Probability Distribution Function

$$f(T) = \frac{e^z}{\alpha T(1+e^z)^2} \text{ where } z = \frac{T' - \eta}{\alpha}, \quad T' = \ln(T), \quad 0 < T < \infty, \quad -\infty < \eta < \infty, \quad 0 < \alpha < \infty$$

**E_{ToU} (where 1 = 00:00 - 00:30 and 48 = 23:30 - 00:00)

E_{Total}

The results presented in Table 8.3 show that for parameter E_{Total} there was no significant difference between the original and characterised profiles. However, standard deviation differed substantially to that reported in Table 5.1 for the parameter which is a result of reducing the number of possible electricity load profiles from the sample size (3,941) to just ten profiles in total. The reduction in standard deviation to 864 kWh in Table 8.4 (from 1,108 kWh in Table 5.1) is indicative of this as there is less variation between individual households electricity demand for the characterised profiles.

E_{MD}

Table 8.3 showed that there was a significant difference for parameter E_{MD} (1.33 kW) between original and characterised profiles. When compared against Table 5.1 shown earlier the difference was about half that of the original sample data. This was a result of the cross-sectional averaging process applied which inevitably reduces the peak demand for large electricity users and increases the minimum for low electricity households.

E_{LF}

Similarly, as E_{LF} is a ratio of E_{MD} to average electricity consumption across a 24 hour period, results were comparable to those presented for E_{MD}. As E_{MD} decreases and assuming average electricity consumption remains the same across a 24 hour period, E_{LF} will increase. The results presented in Table 8.3 show that mean E_{LF} was significantly different to the original data shown in Table 5.1. The figures also showed that it was nearly double that presented in Table 5.1. This can be explained by the loss in variation by replacing the original sample data with a limited number of characterised profiles.

E_{ToU}

Finally, results presented for E_{ToU} in Table 8.3 showed that there was a significant difference between original and characterised profiles. In particular Table 8.4 shows that the characterised profiles estimated this parameter later in the evening compared to the original sample data in Table 5.1 (18:00 instead of 15:30). However, even though the characterised value is significantly different to that of the original sample data it is a more realistic value to that presented in Table 5.1. This may be perceived as odd but due to the variable nature of domestic electricity demand a certain amount of information is lost when the data is averaged across a six month period in order to calculate the electrical parameters. For example an individual household may predominantly use E_{ToU} in the evening time. However, every so often an infrequent event may occur such as a dwelling occupant being home sick from work during the day may cause electricity to be consumed much earlier in the day. This results in the mean value showing an earlier E_{ToU} than might be expected. This problem is overcome by comparing modal values between the original and characterised profiles for this parameter; when this is done, ToU is found to occur at the same time of 18:00.

Time Series Tests

Figure 8.6 shows an electricity load profile for a random household taken across a weekly period. Both original sample data and characterised profiles (obtained from the CPI) for the same household are shown. As one would expect there are differences between the original sample and characterised profiles. However, in general the two time series are similar in terms of their timing and magnitude of electricity use. The main notable differences between the two profiles are at the extremities such as daily E_{MD} and the minimum use of electricity over the night time period. The figure graphically shows the findings presented in Table 5.1 and Table 8.3 - Table 8.4 in relation to the characterised profiles often underestimating E_{MD} and over estimating minimum values such as at the night time period and the loss of variability (standard deviation) in the profile shape over a 24 hour period.

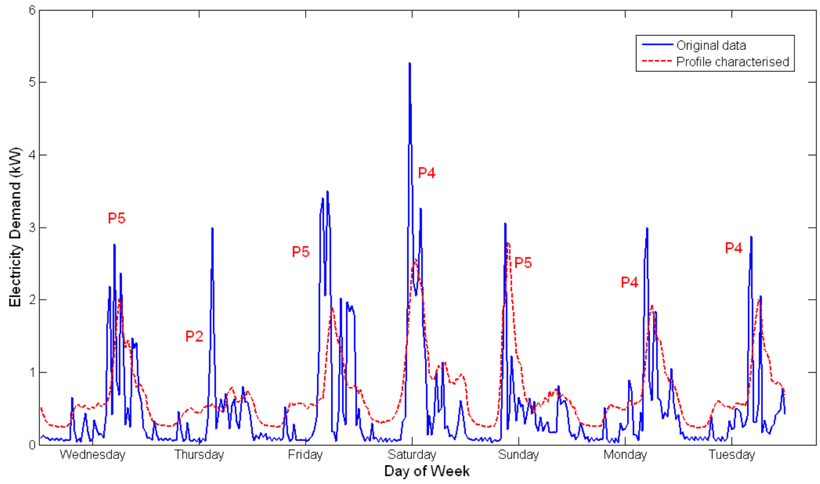


Figure 8.6: Original sample and characterised profile time series for a random household across a weekly period

Figure 8.7 shows a frequency histogram for the same household presented above. As discussed, the main differences between the characterised profile and the original sample relate to the under estimation of Maximum and Minimum values. This is clearly shown in the below figure and shows graphically the reduction in Maximum and standard deviation and increase in Minimum values of electricity demand.

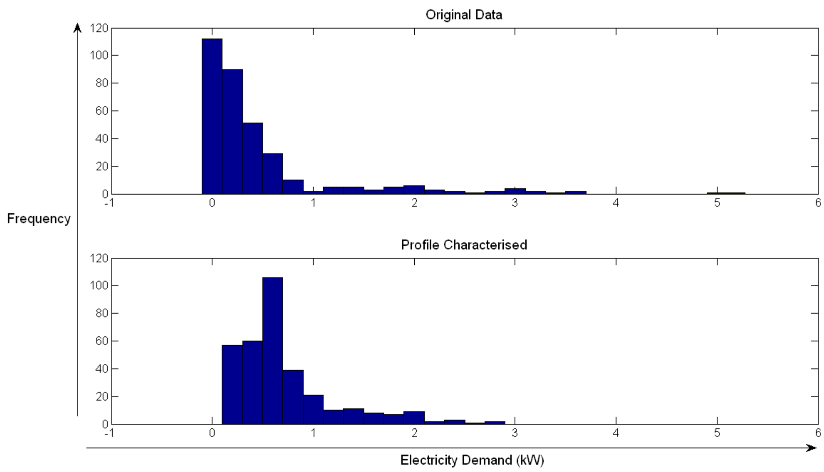


Figure 8.7: Frequency histogram for profile characterised and original data for a random household across a weekly period

Figure 8.8 shows the autocorrelation function for both profile characterised and original sample data over a weekly period for the same household above. The two autocorrelation functions are similar over the period shown indicating that the temporal properties remain between each time series. The main difference between the two autocorrelation functions is that the characterised profile is slightly smoother than the original data, indicating less variation within the former as discussed earlier. In contrast there is more variation on an intra-daily basis with the original sample data as electricity is used more unpredictably and hence the autocorrelation function appears less smooth as the time lag increases.

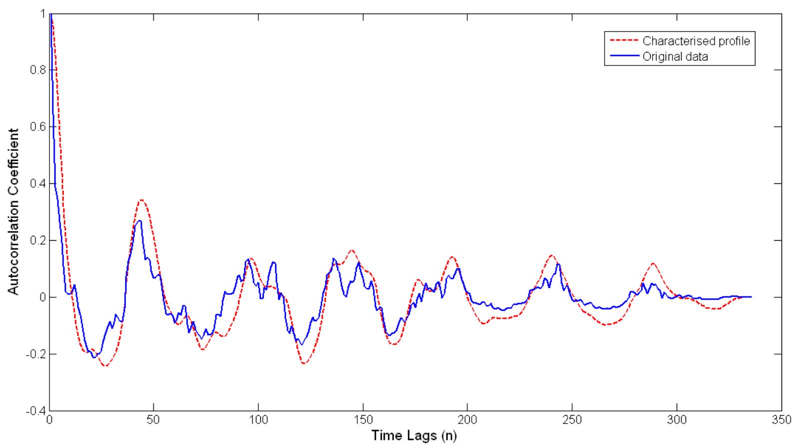


Figure 8.8: Autocorrelation coefficient for profile characterised and original data for a random household across a weekly period

Figure 8.9 shows PSD periodgram for the same household presented above over the weekly period as calculated by the FFT. The large spectral component near the origin is the daily period. The characterised profile has less frequency components particularly at high frequencies on account of the smoother profile shape compared to the original sample data.

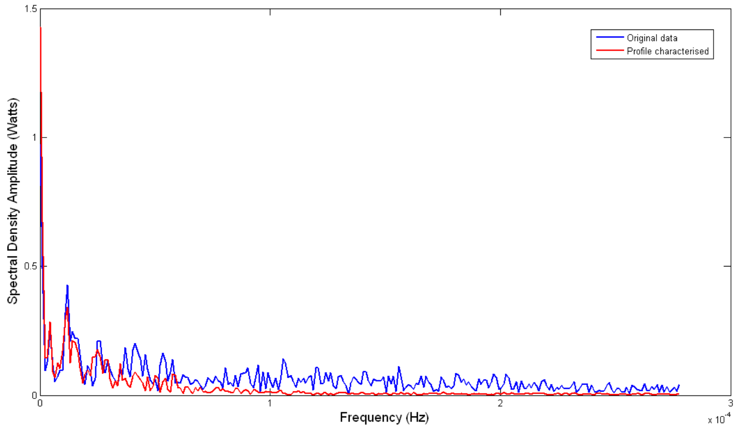


Figure 8.9: PSD function for profile characterised and original sample data for a random household across a weekly period

8.2.3 Profile classification by dwelling and occupant characteristics

A multi-nominal logistic regression was used to associate dwelling and occupant characteristics to each profile group based on the Household Mode (HMode). Table 8.5 provides a brief explanation highlighting the main characteristic traits for each profile group. As stated earlier, detailed descriptions for each electricity load profile group and the factors influencing its use are presented in Appendix A. As the difference between weekdays, Saturdays and Sundays profiles were shown to be marginally different in terms of the magnitude and timing of electricity use across the day, only regression results for weekdays are presented in the following sections so as to avoid repetition.

Table 8.5 Electricity load profile groups main characteristics description

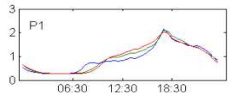
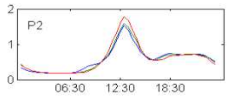
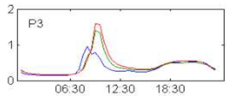
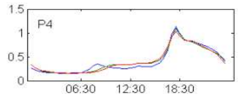
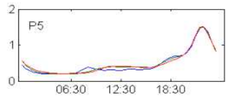
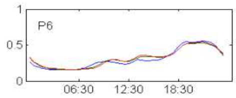
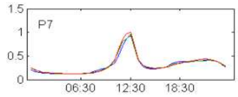
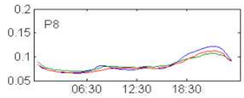
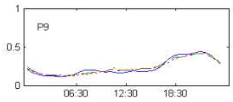
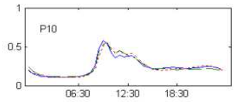
Profile Number	Profile Shape (kW)	Main Characteristic	Secondary Characteristic (if any)
Profile 1		Large evening peak	Small morning peak
Profile 2		Large lunchtime peak	Small evening peak
Profile 3		Large morning peak	Small evening peak
Profile 4		Evening peak	Small morning peak
Profile 5		Late evening peak	Small demand across the day
Profile 6		Small average load profile across the day	
Profile 7		Late morning peak	Small evening peak
Profile 8		Very small average load profile across the day	
Profile 9		Small late evening peak	
Profile 10		Small late morning peak	

Table 8.6 shows the results for the multi-nominal logistic regression applied to the Dwelling and Occupant Characteristics (DOC) and the profile groups. Profile group P4 was used as the reference category as it represented over 30% of households within the sample as was shown in Table 8.1.

Table 8.6: Multi-nominal logistic regression results for DOC model

Dwelling and occupant characteristics	Profiles																	
	p1		p2		p3		p5		p6		p7		p8		p9		p10	
	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error	Exp(B)	Std. Error
Intercept	**	.511	***	.670	*	.641		.847		.373	***	1.386	***	.396	***	.419		3863.812
Dwell_type_semi_d	0.700**	.176	0.490***	.240	1.320	.234	.601	.342	.946	.127	.583	.416	1.004	.145	1.130	.145	1.432	.401
Dwell_type_terr	1.278	.248	.608	.337	1.986**	.292	.424	.631	.879	.169	1.249	.456	.966	.184	1.310	.177	2.006	.428
Dwell_type_apt	6.342E-08	5245.821	1.989E-08	6343.375	1.874	1.113	3.411E-08	0.000	1.750	.473	7.050E-08	7945.919	1.399	.423	1.398	.443	6.571E-09	0.000
No_bedrooms_1	9.039E-09	7267.240	2.436	1.145	3.839E-09	0.000	2.093E-08	0.000	.493	.852	6.954E-08	0.000	2.293	.565	3.168*	.594	1.153E+08	3863.812
No_bedrooms_2	0.141***	.540	.790	.582	.622	.580	.467	1.096	1.985**	.282	1.267	1.254	6.480***	.309	5.718***	.340	1.657E+08	3863.812
No_bedrooms_3	0.186***	.211	.974	.301	.814	.336	.704	.494	1.413*	.199	3.048	.762	1.811**	.262	2.537***	.290	7.333E+07	3863.812
No_bedrooms_4	0.337***	.166	.826	.285	.701	.330	1.217	.430	.860	.200	1.738	.770	.993	.267	1.272	.297	1.085E+07	3863.812
HoH_age_36_55	1.786**	.266	.557	.351	1.079	.399	.876	.430	1.060	.224	1.807E+07	0.000	0.555**	.214	0.571**	.217	.513	.843
HoH_age_56_plus	1.579	.308	1.777	.373	1.557	.425	.568	.529	2.018***	.239	.693E+07***	.459	0.657*	.231	.781	.232	1.991	.842
HH_comp_with_adults	.810	.352	1.095	.380	0.240***	.260	0.347**	.510	0.357***	.168	1.125	.547	0.062***	.155	0.102***	.157	0.128***	.363
HH_comp_with_adults_and_children	1.160	.366	1.006	.445	0.113***	.355	0.294**	.548	0.228***	.211	.448	.835	0.021***	.234	0.032***	.235	0.123***	.660
HoH_social_class_C	.899	.179	.812	.296	.667	.284	.568	.342	.851	.182	1.087	.804	0.640**	.195	.747	.199	1.568	1.070
HoH_social_class_DE	.849	.219	1.201	.304	.659	.312	.667	.412	1.331	.188	3.100	.768	.778	.207	.843	.211	4.138	1.049
HoH_social_class_F	1.671	.397	3.176**	.481	1.450	.666	.574	1.065	1.880*	.368	5.364	1.051	3.682***	.366	1.733	.433	19.941**	1.301
Water_heat_electric	1.131	.144	1.394	.201	1.811**	.225	.695	.279	0.601***	.109	1.030	.335	0.776**	.124	0.542***	.124	.839	.331
Cooking_type_electric	0.699**	.162	1.123	.245	0.526***	.220	0.287***	.283	0.220***	.112	5.270**	.732	0.338***	.131	0.270***	.129	.582	.360
Efficiency_betw_10_20	1.220	.171	1.152	.230	.834	.271	1.797*	.335	1.047	.139	.666	.468	0.709**	.166	.804	.166	.858	.472
Efficiency_betw_20_30	1.662**	.218	.940	.343	1.734*	.304	2.007	.438	.810	.216	.496	.745	.711	.252	.789	.250	2.609E-08	3611.378
Efficiency_more_30	1.191	.342	.460	.734	.531	.736	1.568E-08	5914.582	0.504*	.364	2.139E-08	5666.841	.547	.407	.589	.397	1.829E-08	5798.637

P values: *** p<0.01, ** p<0.05, * p<0.1

Base variables: *Dwelling_type_detach, No. bedrooms_5, HoH_age_less_36, HH_comp_live_alone, HoH_social_class_AB, Water_heat_non_electric, Cooking_type_non_electric, Efficiency_less_10*

Detached dwellings typically represent larger dwellings while terraced and apartments are usually smaller in size. Semi-detached usually fall somewhere between the two in terms of size. Detached dwelling was used for the base variable and hence is not shown in Table 8.6. However, Figure 8.10 shows that profile groups P1, P2 and P5 have the largest penetration of detached dwellings with over 70% on average of households. The remaining profile groups are made up of a mixture of semi-detached, terraced and apartments. Apartments are predominantly characterised within profile groups P3, P6, P8 and P9 with a minimum odds ratio of 1.3 for these groups. This means that people living in apartment dwellings are 1.3 times more likely to use electricity in a way similar to that of these profile groups. Occupants living in semi-detached dwellings are 1.3 times more likely to use electricity similar to that of profile group P3 and P10. Finally terraced dwelling owners are most likely to use electricity in a manner similar to that of profile groups P3 and P10.

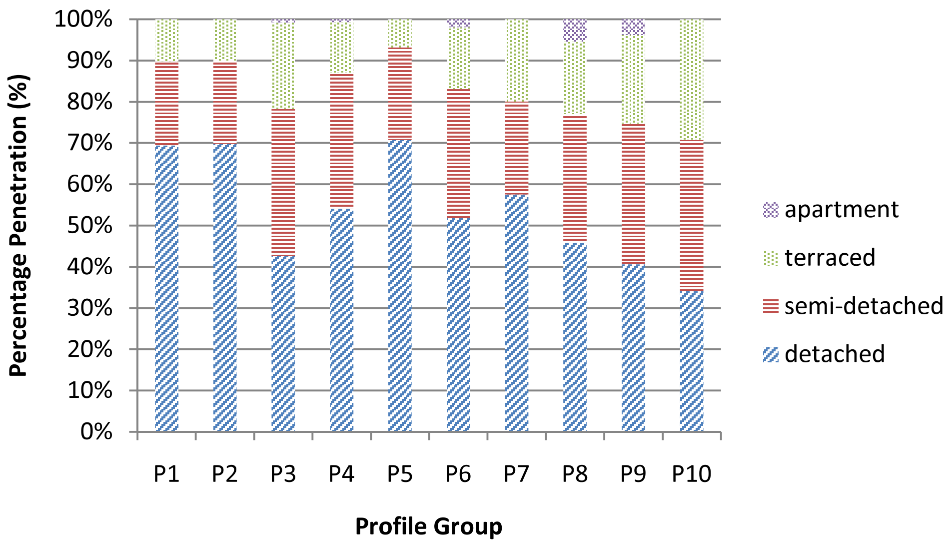


Figure 8.10: Percentage penetration of dwelling type by profile groups

The number of bedrooms is also a good indication of dwelling size and should correspond with the results presented in Figure 8.10. Five plus bedrooms was used as the base variable and hence is not shown in Table 8.6. Figure 8.11 shows profile groups P1, P2, P4 and P5 are characteristic of four and five plus bedroom dwellings, which make up over 50% of households within these groups. Profile groups P6 and P7 consist of more mid-size dwellings with three and four bedrooms. This is also reflected in the odds ratio which shows that occupants in 3 and 4 bedroom dwellings are 1.5 and 3 times respectively more likely to use electricity in this way compared to the other profile groups. Lastly, smaller dwellings of one, two and three bedrooms are most likely to use electricity similar to that of profile groups P8, P9 and P10 as indicated by high odds ratios in Table 8.6.

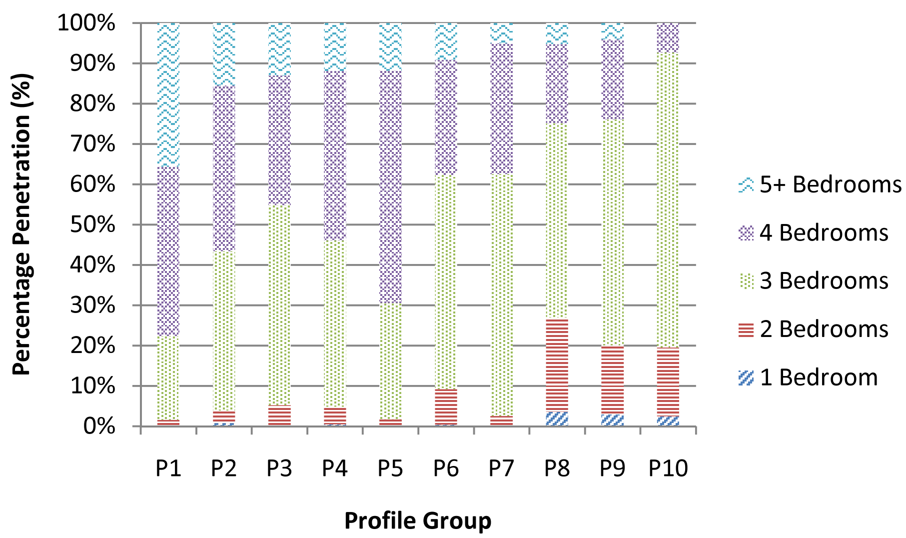


Figure 8.11: Percentage penetration of dwelling number of bedrooms by profile groups

Profile group P1 is mostly characteristic of dwellings with a HoH of between 36 and 55 years with a very high relative odds ratio of 1.786. Similarly Figure 8.12 shows that profile groups P3, P4 and P5 also has a high percentage penetration for this category.

These categories also correspond with larger dwellings as already shown in Figure 8.10 and Figure 8.11. Older HoH's are most likely to use electricity in manner similar to that of profile groups P6, P7 and P10 compared to the base category.

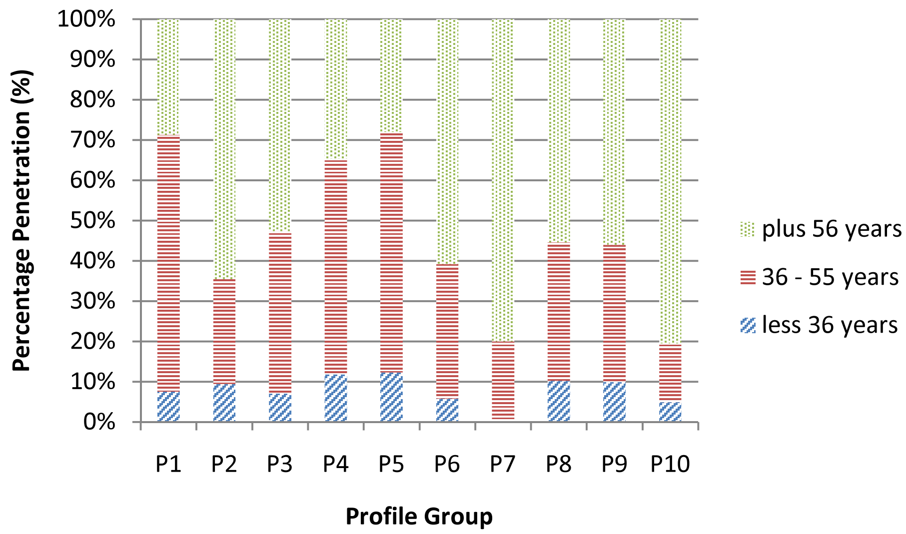


Figure 8.12: Percentage penetration of HoH age by profile groups

Figure 8.13 shows that profile groups P8, P9 and P10 have the highest percentage for occupants living alone which was used as the base variable and hence is not shown in Table 8.6. In contrast, profile groups P2 and P7 are more likely to characterise electricity consumption patterns for adults living together. Adults and children living together are most likely use electricity in a similar manner to that of profile groups P1, P4 and P5, which also correspond with the largest dwelling types.

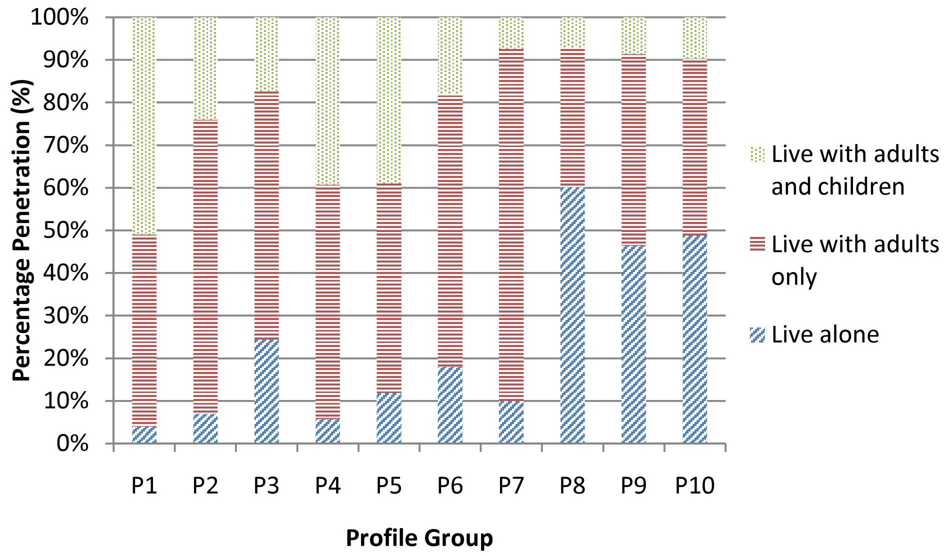


Figure 8.13: Percentage penetration of household composition by profile groups

Figure 8.14 shows that the higher social classes of AB (which was used as the base variable in Table 8.6) are most likely to use electricity in the home similar to that of profile groups P1 and P5. This also corresponds with the larger dwelling types shown earlier. Figure 8.14 also shows that profile groups P1 and P4 are largely characteristic in the manner with which the middle class C use electricity in the home. The lower classes of DE will tend to use electricity similar to that shown in profile groups P6 – P10 which is also reflected in higher odds ratios for these groups. Social class F is most closely associated with profile groups P2, P7, P8 and P10 with a minimum odds ratio of 3 shown for these groups in Table 8.6.

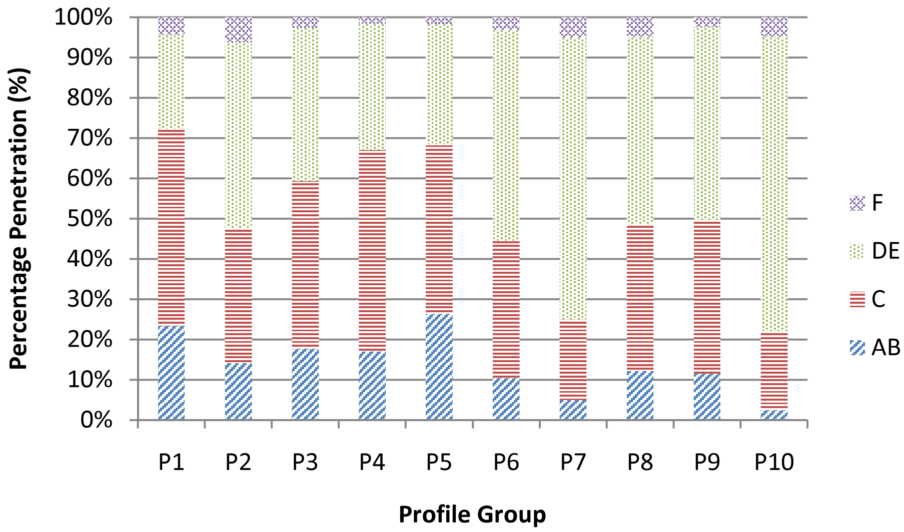


Figure 8.14: Percentage penetration of HoH social class by profile groups

Households that use electricity to heat water are most likely to use electricity as characterised by profile groups P1 – P3, indicated by a comparative high odds ratio within this category. Profile groups P2 and P7 are mostly characteristic of households who use electricity to cook in the home. This is also shown in Figure 8.15, with electric cooking showing an unusually high penetration for profile group P7. Electric space heating is also shown in Figure 8.15 for comparison purposes but was not included in the multi-nominal logistic regression due to its small sample size. As shown below, its penetration is relatively evenly spread across all profile groups, with profile group P10 having the largest penetration.

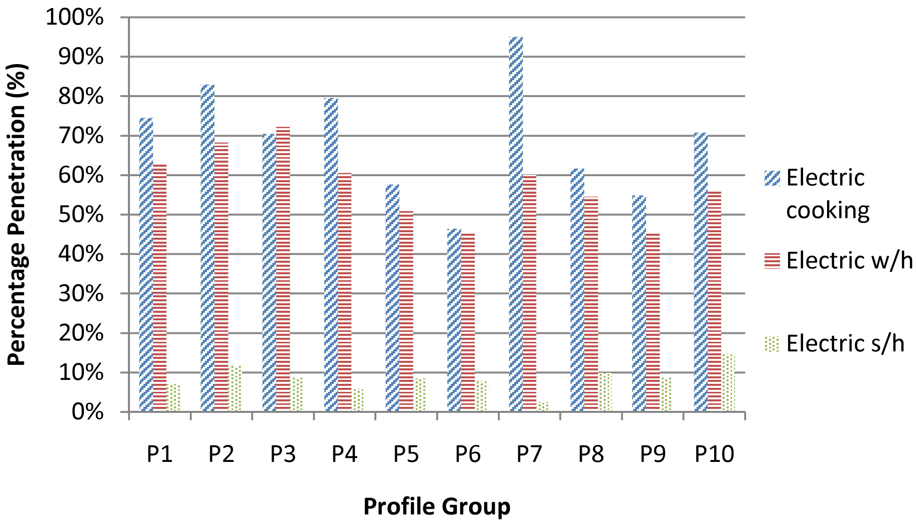


Figure 8.15: Percentage penetration of electric cooking, water and space heating by profile groups

Finally, Figure 8.16 shows the Efficiently Indicator which shows the percentage at which customers believe they can cut their household electricity consumption by making changes to the manner with which they use it in the home. An efficiency Indicator <10% was used as the base variable and hence is not shown in Table 8.6. Dwelling occupants who used electricity in a similar manner to profile groups P1 – P5 were less efficient, with the majority of these householders believing that they could cut 10% or more off their electricity bill. In contrast the majority of householders who used electricity in a similar manner to profile groups P6 – P10 believed they could only cut 10% or less off their electricity bill.

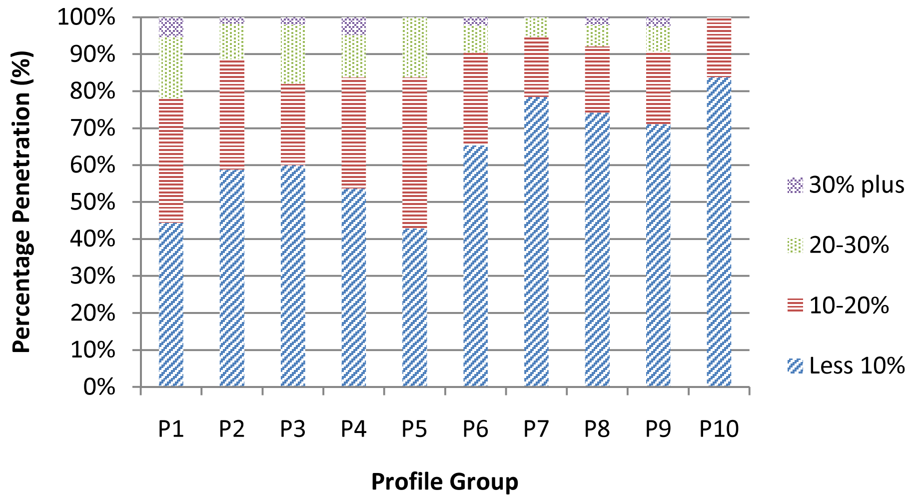


Figure 8.16: Percentage penetration of Efficiency indicator by profile groups

The results presented in Table 8.6 and Figure 8.10 to Figure 8.16 are summarised in one Table 8.7. The table shows the most common dwelling and occupant characteristics associated with each electricity load profile group. Water heating and cooking is indicated by high, medium and low relative percentage penetrations. Similarly, Efficiency Indicator is characterised by high, medium and low and indicates the level of savings a household believes they can achieve by cutting their household electricity bill.

Table 8.7: Profile group descriptions by dwelling and occupant characteristics

Dwelling & Occupant Characteristics	Profile Group									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Dwell_type_detached	*	*			*		*			
Dwell_type_semi_d			*	*		*				
Dwell_type_terr			*				*		*	*
Dwell_type_apt								*	*	*
No. bedrooms - 1								*	*	*
No. bedrooms - 2										*
No. bedrooms - 3			*	*		*	*			*
No. bedrooms - 4	*	*	*	*	*	*	*			
No. bedrooms - 5+	*	*			*					
HoH_age_less_36					*			*	*	
HoH_age_36_55	*		*	*						
HoH_age_56_plus		*	*			*	*	*	*	*
HH_comp_live_alone			*					*	*	*
HH_comp_with_adults		*	*	*	*	*	*			
HH_comp_with_adults_children	*			*	*					
Social_class_AB	*				*					
Social_class_C		*	*	*		*		*	*	
Social_class_DE		*	*			*	*	*	*	*
Social_class_F		*					*			*
Water_heat_electric Cooking_type_electric	High	High	High	Med	Low	Low	High	Med	Low	Med
Efficiency_less_10 Efficiency_betw_10_20 Efficiency_betw_20_30 Efficiency_more_30	High	Med	Med	Med	High	Med	Low	Low	Low	Low

Table 8.8 presents results for the Electrical Appliance (EA) with odds ratio Exp(B) and significance levels presented. The table shows the likelihood of each profile group containing a particular type of electrical appliance. Therefore in the following analysis the dependent variable is the profile group number and the explanatory variables are the appliance types.

Table 8.8: Multi-nominal logistic regression results for EA model

Appliance type	p1 Exp(B)	p1 Std. Error	p2 Exp(B)	p2 Std. Error	p3 Exp(B)	p3 Std. Error	p5 Exp(B)	p5 Std. Error	p6 Exp(B)	p6 Std. Error	p7 Exp(B)	p7 Std. Error	p8 Exp(B)	p8 Std. Error	p9 Exp(B)	p9 Std. Error	p10 Exp(B)	p10 Std. Error
Intercept	***	1.134	***	.510		1.126		.889	***	.586	***	1.176	***	.449	***	.480		1.168
Tumble Dryer	1.889***	.224	1.402	.250	0.494***	.210	.654	.312	0.700***	.120	.649	.351	0.415***	.124	0.376***	.127	0.509*	.342
Dishwasher	2.326***	.241	.819	.221	.794	.220	1.062	.344	0.570***	.118	0.373**	.351	0.364***	.126	0.421***	.128	0.270***	.372
Shower (instant)	.864	.159	1.307	.233	.942	.226	1.001	.309	0.749**	.117	.931	.372	0.806*	.129	0.722**	.130	.957	.366
Shower (pumped)	1.112	.146	.881	.217	.965	.230	1.376	.288	1.116	.121	1.121	.384	.956	.141	.998	.141	.622	.469
Electrical Cooker	.825	.190	1.046	.278	0.491***	.235	.394	.306	0.220***	.117	6.455*	1.019	0.358***	.135	0.296***	.134	0.498*	.372
Heater (plug in convective)	1.260	.147	1.038	.200	1.669**	.202	1.406	.284	1.171	.116	1.808*	.329	1.141	.127	1.073	.130	1.520	.337
Freezer (stand alone)	1.221	.142	1.113	.190	.854	.199	.974	.273	1.136	.108	1.340	.336	0.547***	.123	0.786*	.122	0.366**	.391
Water pump	1.471**	.154	1.459*	.213	.744	.281	1.175	.318	1.124	.131	.767	.459	0.707**	.170	.926	.159	.744	.494
Immersion	.975	.178	1.061	.238	1.731*	.288	.830	.326	0.540***	.118	.891	.394	.905	.138	0.691**	.135	0.557*	.345
No. TV <21inch	1.039	.154	1.127	.206	.979	.222	.897	.284	.879	.118	.947	.361	0.485***	.130	0.624***	.130	.665	.368
No. TV>21inch	1.078	.261	1.195	.306	0.585**	.256	3.354	.735	0.664**	.150	.881	.437	0.373***	.154	0.619***	.163	.518	.407
No. computer (desktop)	1.395**	.149	.802	.191	0.586**	.206	1.315	.283	.878	.109	0.451**	.363	0.395***	.131	0.528***	.128	0.348**	.409
No. computer (laptop)	1.299*	.153	0.628**	.191	.851	.204	1.249	.290	0.789**	.110	0.202***	.431	0.646***	.124	0.743**	.125	0.353**	.400
No. game consoles	1.924***	.147	0.465***	.223	0.518**	.243	1.036	.282	0.546***	.124	0.153**	.745	0.323***	.170	0.283***	.171	0.167**	.750

Base variables: Washing machine

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

The high penetration of washing machines throughout all homes in Ireland means that it makes it a good choice as the base appliance category. Households which own a tumble dryer are 1.8 and 1.4 times more likely to use electricity in a similar manner to profile groups P1 and P2 respectively than the base profile group. Similarly by owning appliance type dishwasher, households are twice as likely to use electricity in a manner similar to that of profile group P1 compared to the base profile group.

Figure 8.17 shows percentage penetrations for the same appliances discussed above with similar results presented. All three appliances have high penetrations in profile groups P1, P2, P4 and P5. Conversely, percentage penetrations of these appliances is lowest within profile groups P8, P9 and P10.

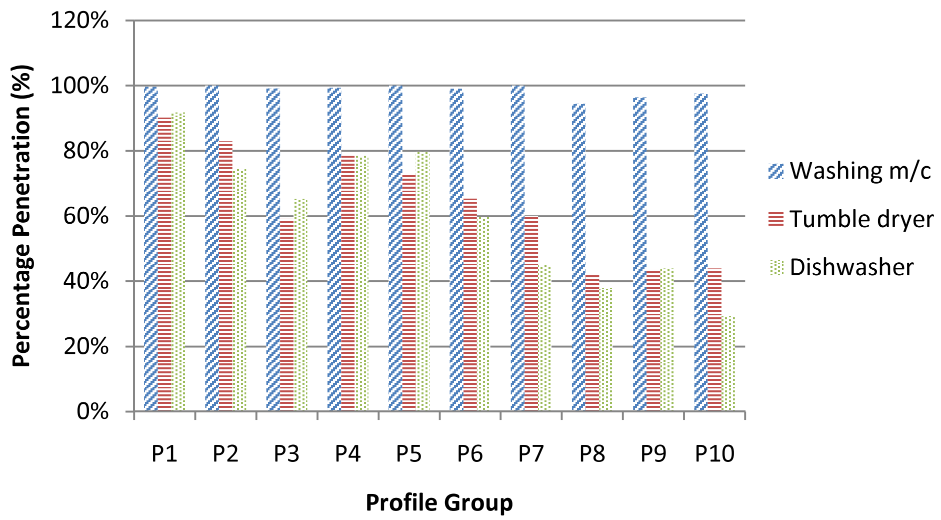


Figure 8.17: Percentage penetration of washing m/c, tumble dryer and dishwasher by profile group

Households who use electricity in a similar manner to profile groups P2, P4 and P5 were most likely to own an instant electrical shower. Profile group P5 households were most likely to own a pumped shower compared to any other profiles, however, penetration of this type of appliance in Irish households is much lower than that of instant showers as indicated in Figure 8.18. The likelihood of households who used electricity in a similar manner to profile group P3 was most likely to own an immersion compared to any other profile group. Profile groups P1 and P2 were more likely to own appliance type water pump (used in low water pressure residential environments).

Figure 8.18 shows percentage penetration of appliances; instant electric and pumped showers, water pumps and immersions used to heat water. Electric showers and immersion percentage penetration is highest for profile groups P1 – P4. Pumped showers have their highest penetration in for profile groups P1 and P5, with water pumps highest for profile groups P1 and P2.

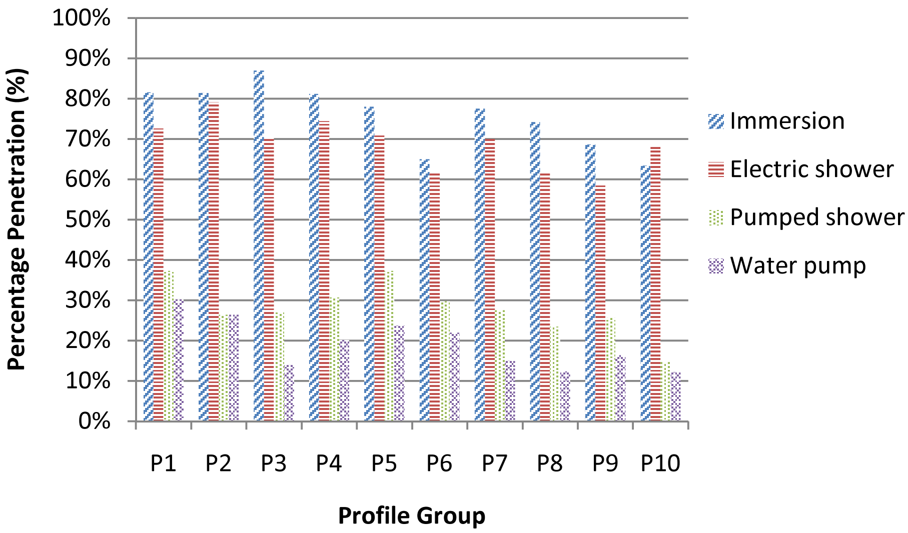


Figure 8.18: Percentage penetration of electric shower, pumped shower, water pump and immersion by profile group

The likelihood of owning a cooker was high for profile groups P2 and P7, with the latter particularly high. Profile groups P6, P8 and P9 were the least likely to own an electrical cooker. The likelihood of owning a plug in electric heater was more likely for profile groups P3 and P7 but in general the odds ratio for this appliance type was evenly spread across all profile groups. This is most likely down to the smaller percentage penetration for this particular appliance type. The likelihood of owning a stand-alone freezer was relatively high for profile groups P1 – P7.

Figure 8.19 shows percentage penetration for appliances electrical cooker, stand alone freezer and plug in heater for each profile. Penetration of electric cookers is notably less for profile groups P6, P8, P9 and P10, with these households more likely to use natural gas instead of electricity for cooking. Stand alone freezers have a relatively high penetration for profile groups P1 – P7. Plug in electric heaters used for space heating have a relatively low penetration across all groups, however, is slightly higher for profile groups P3, P7 and P10.

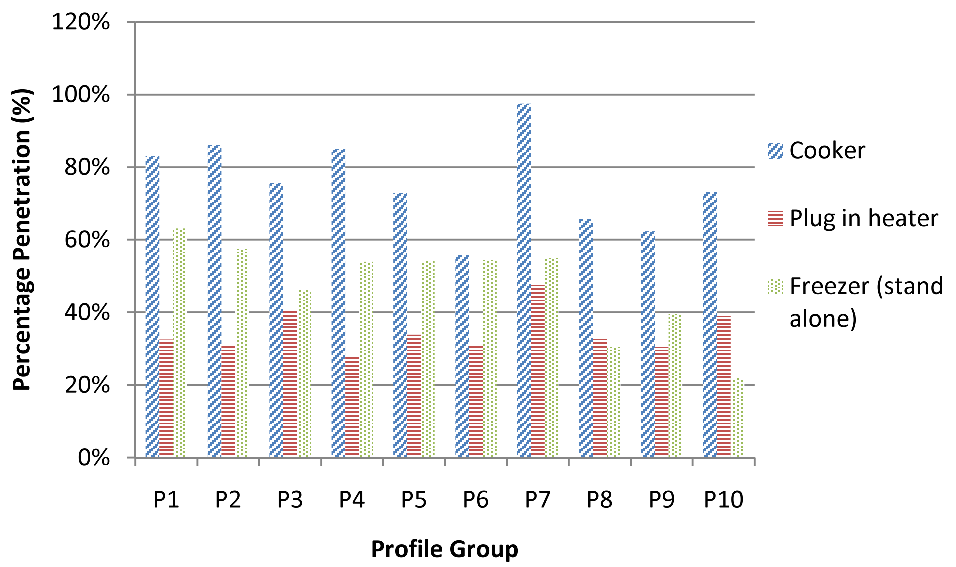


Figure 8.19: Percentage penetration of electric cooker, plug in heaters, and stand alone freezer by profile group

The likelihood of owning a television of less than 21 inches is higher for profile groups P1 and P2 compared to the other groups. However, households with televisions greater than 21 inches showed that they are more likely to use electricity in a similar manner to profile group P5. Households that owned desktop and laptop computers showed the greatest likelihood of using electricity in a similar manner to that of profile groups P1 and P5. Similarly, profile group P1 showed a high level of likelihood of households owning appliance type game console, with profile group P5 also likely but to a lesser extent.

Figure 8.20 shows percentage penetration of televisions greater and smaller than 21 inches, desktop and laptop computers and game console for each profile group. Large televisions (greater than 21 inches) have the highest penetration for profile group P5. Smaller televisions (less than 21 inches) have slightly less penetration for profile groups P8, P9 and P10 but are relatively constant across all groups. This trend is similar for computers (desktop and laptops) and game consoles. Profile group P5, has a comparatively high penetration of all appliances shown in Figure 8.20 whereas profile group P7 has the lowest penetration for laptop computers and game consoles between all groups.

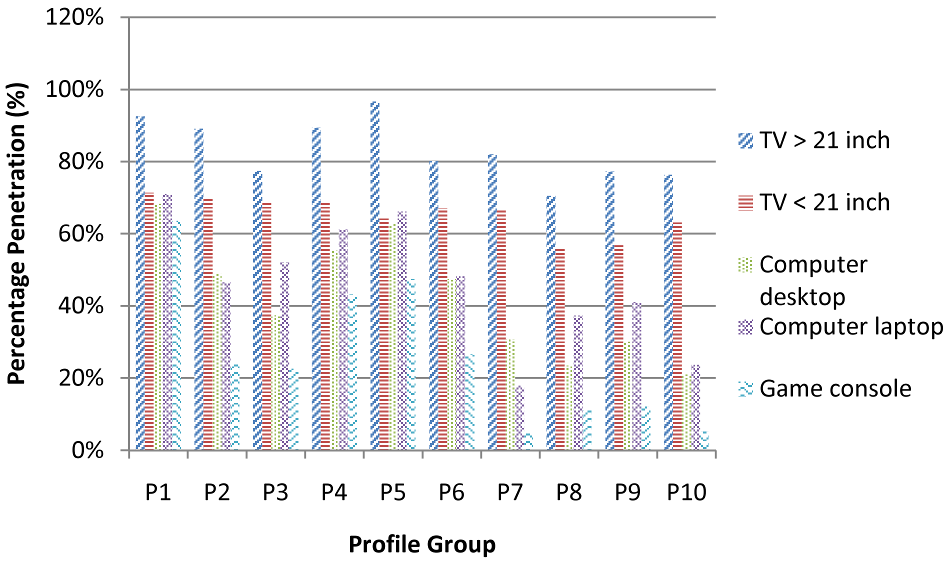


Figure 8.20: Percentage penetration of TV's (<21 inch, > 21 inch, desktop computer, laptop and game console by profile group

The results presented in Table 8.8 and Figure 8.17 to Figure 8.20 are summarised in Table 8.9. The table shows the most common electrical appliances associated with a particular profile group.

Overall, profile groups P1 – P5 are characteristic of households with a high percentage penetration of electrical appliances. Profile groups P6 and P7 are characteristic of households which have a medium level of penetration of electrical appliances. Finally, profile groups P8, P9 and P10 are characteristic of households with a low level of penetration of electrical appliances.

Table 8.9: Profile group descriptions by electrical appliance characteristics

	Profile Group									
Dwelling & Occupant Characteristics	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10
Washing machine	*	*	*	*	*	*	*			*
Tumble dryer	*	*		*	*					
Dishwasher	*	*	*	*	*					
Shower (instant)	*	*	*	*	*		*			*
Shower (pumped)	*	*	*	*	*	*	*		*	
Cooker	*	*	*	*			*			*
Heater (plug in convective)			*				*			
Freezer (stand alone)	*	*	*	*	*	*	*			
Water pump	*	*		*	*	*				
Immersion	*	*	*	*	*		*	*	*	*
TV's less 21 inches	*	*	*	*	*	*				*
TV's greater 21 inches	*	*		*	*		*			*
Desktop Computer	*	*	*	*	*	*				
Laptop computer	*		*	*	*	*				
Game console	*			*	*	*				

8.3 Conclusion

This chapter presented a methodology using clustering to segment electricity households into profile groups based on their pattern of electricity use across the day. A number of clustering techniques such as: k-mean, k-medoid and SOM were investigated and compared based on the DB validity index. SOM performed best, and was used to cluster *Dataset I* on a daily basis over the six month period.

In total, ten daily electricity load profiles were presented that represented common patterns of electricity use within the home in Ireland. Day type variations were also shown by weekday, Saturday and Sunday and as well as the seasonal variations between Summer–Autumn–Winter. On a daily basis, individual households tended to use different profile groups. Therefore, the mode (HMode) was calculated in order to determine the profile group that each household used for the majority of the time over the six month period. The APPT parameter then quantified the propensity for households to switch between profiles on a day-to-day basis. This was shown to be higher for some profile groups as opposed to others.

A number of electrical parameters, presented in Chapter 4, were calculated and compared against the original sample data. The results for mean E_{Total} matched the original sample dataset closely. However, descriptive statistics for Standard Deviation, Maximum and Minimum values for the same parameter were significantly different. This was a result of replacing the original 3,941 customer electricity demand load profiles with just ten in total on a single day and therefore resulted in a loss of variability. This was also true of parameters E_{MD} and E_{LF}. E_{ToU} tended to be later for the characterised profiles compared to the original data. This was largely due to the removal of the stochastic component of the original data and replacement with a more deterministic pattern of electricity use generated by the characterised profiles. The mean value for this parameter showed a E_{ToU} of 15:30 for original sample data and 18:00 for characterised profiles. However, when the mode was compared between original sample data and characterised profiles the results showed both peak electricity demands occurring at 18:00.

The paired sample t-tests showed that parameter E_{Total} was found not to be significantly different, however, the remaining three parameters were found to be significantly different at the 0.05 p-value level. The time series tests showed that small and large values of electricity demand were not fully represented by the characterised profiles and that this was also due to the loss in variability mentioned above. This was also visible in the time series plot across a weekly period. The autocorrelation function showed that for the majority of the time, electricity peaks were consistent with the original sample data. Similarly, the PSD periodgram showed similar frequencies for original sample and characterised time series. However, the characterised profiles did not show the same high frequency components again due to the loss of variability between original sample data and the newly characterised profiles.

The final part of the chapter presented results for associating dwelling and occupant characteristics to the ten electricity load profile groups. The results showed that in general, the ten electricity load profile groups can be broken down into three broad categories. The first category, consisting of profile groups P1, P2, P4 and P5 are characteristic of high electricity users. These profile groups correspond with large dwelling types (detached) with a greater number of bedrooms, with HoH age between 36 and 55 and families living with children and from a high social class. These households usually have a large stock of electrical appliances, with tumble dryers and dishwashers being more prevalent compared to the other profile groups.

The second category, consisting of profile groups P3, P6, P7 were generally associated with medium electricity users and are associated with mid-sized dwelling types such as small detached, semi-detached and terraced dwellings. The profile groups correspond with three and four bedrooms and usually have an older HoH from a middle social class of C and DE with no children living at home. These households have a medium stock of electrical appliances but differ from the larger electricity category in that tumble dryers and dishwashers are less prevalent within the homes. Lastly profile groups P8, P9 and P10 correspond with low electricity users and are associated with smaller dwelling types such as semi-detached, terraced and apartments and mainly comprise of one, two and three bedrooms. The dwelling is usually occupied by an older HoH age of 56 plus, usually lives alone and from a middle to lower social class of C and DE. These households are less likely to have a large stock of electrical appliances compared to the previous two categories.

Chapter 9

Conclusions and Recommendations

9 CONCLUSIONS AND RECOMMENDATIONS

9.1 Introduction

To date, research has focussed on either developing highly averaged demand load profiles or using probabilistic methods to characterise domestic electricity consumption. In terms of the former, these profiles do not reflect the true nature of electricity consumption patterns within the home, however, they generally can be considered to be representative. Conversely, probabilistic methods reflect more realistic patterns of electricity use within the home but often cannot be considered to be representative. This research addresses these issues and provides a series of domestic electricity demand load profiles that are both representative and reflect common patterns of daily electricity use in the home. The research also provides a method of linking domestic electricity load profiles to dwelling and occupant characteristics. This means that electricity load profiles can be assigned to particular households based on information relating to the dwelling, occupant or electrical appliance characteristics. In effect, no prior knowledge of a dwellings electricity demand is required to assign a particular electricity profile to a household; however, where this does exist it may also be used.

9.2 Conclusions

The relatively recent availability of smart metering data has meant that methods of characterising electricity consumption, which have traditionally been applied to other sectors, can now be applied to domestic use. The literature describing these methods as well as existing domestic electricity characterisation approaches were categorised based on the level at which the data was collected (i.e. at an aggregate or individual dwelling) and the interval period. A number of different characterisation approaches were evaluated throughout the research in order to meet the objectives outlined in Section 1.5, each one building upon the next.

Firstly a statistical regression approach was used to characterise electricity consumption within the home. Domestic electricity use was described as a function of key electrical parameters. This proved to be an effective approach to characterise key domestic electricity load profile features within a small number of parameters (E_{Total}, E_{MD}, E_{LF} and E_{ToU}) and therefore was subsequently used for validation purposes throughout the remaining part of the research. However, a disadvantage with this characterisation technique was that by describing electricity use within the home as a function of parameters a certain amount of information was lost in the process. As a result individual electricity load profiles cannot be extracted based solely on these parameter values alone. However, this did not affect the validation approach and the analysis also identified the main influential dwelling, occupant and appliance characteristics that influenced electricity consumption within the home which were also subsequently used throughout the research.

The next method used was an autoregressive approach where Markov chains characterised the probabilities of using electricity based on previous values within a household. Markov chains proved to be a very effective technique to characterise the magnitude component of electricity load profiles. Parameters: E_{Total}, E_{MD} and E_{LF} were all successfully reproduced within a reasonable degree of accuracy between original sample and characterised profiles. However, the Markov chain process was unable to characterise the temporal properties of the load profiles. This was most obvious when comparing profiles across a 24 hour period where unusual patterns of electricity use (normally during the night time period) were evident. This subsequently led to the introduction of a number of time series tests to interrogate the temporal properties of the load profiles more rigorously. Autoregression and spectral density functions were calculated for the original and characterised profiles and was useful for determining the performance of the characterisation process in the time and frequency domains.

The drawbacks with this technique included the computation time required to characterise each individual household and the temporal difficulties associated with characterising the magnitude component of electricity load profiles at appropriate times of the day. These disadvantages along with the fact that it would have been difficult to link dwelling and occupant characteristics to the transitional probability matrix meant that this approach was not pursued further. However, it still remains a good method for characterising domestic electricity use and it may be possible to divide the diurnal electricity load profile into segments and have three or four transitional probability matrices across a daily period.

Other time series approaches were considered next, particularly those that could specifically characterise the temporal properties of an electricity load profile. Two methods which showed the greatest potential were: Fourier transforms and Gaussian processes and therefore were evaluated further. Depending upon the electricity load profile shape each characterisation method had advantages and disadvantages and tended to be complementary to each other. Gaussian processes were better at characterising households which consumed a large amount of electricity relative to the rest of the day and for short time intervals. In contrast, Fourier transforms were better at characterising households that consumed electricity more evenly across the day. The two techniques are important as both patterns of electricity use are common within domestic households.

In addition Fourier transforms were shown to be very accurate when evaluated against parameter E_{Total} where as Gaussian processes performed better when evaluated against E_{MD}. The temporal properties were shown to remain for both autocorrelation and spectral density functions showing similar results between original and characterised time series. The application of multivariate linear regression between the characterised time series and dwelling and occupant characteristics proved unsuccessful, resulting in highly averaged electricity load profile shapes. This can be explained by the longitudinal averaging process applied to the dataset as well as the application of regression as a whole which tended to smooth out the characteristic shape of the individual domestic electricity demand load profile.

Finally, a clustering based approach was used to characterise domestic electricity load profiles. This technique was different to previous methods in so far as it was used to reduce the data first by segmenting similar patterns of electricity use into clusters. As each cluster represented a similar pattern of electricity use each profile was able to be combined without losing the characteristic shape for each profile. Ten profile groups were produced in total, representing common patterns of electricity use within the home in Ireland. Seasonality and intra-daily variations were accounted for by clustering each day separately. Intra-daily differences for day types: weekdays, Saturdays and Sundays were shown to exist and separated in order to illustrate different patterns of electricity use for different day types. Similarly, seasonality patterns were shown to exist on electricity consumption patterns which could be attributed to the effects of temperature (for cold appliances) and sunrise and sunset times (for lighting).

The application of a multi-nominal logistic regression was used to associate each profile group with household characteristics. As a result this meant that individual dwellings and the manner with which their occupants use electricity could be completely distinguished from their dwelling, occupant and appliance characteristics. Appendix A contains a complete library of electricity load profiles for each profile group (P1 to P10) representing diurnal, intra daily and seasonal patterns of electricity use.

The profiles can be broadly categorised into three main groups. The first group, high electricity users were characterised by profile groups P1, P2, P4 and P5. These were largely made up of detached dwellings of four and five bedrooms. They were mainly occupied by adults and children and with a mixture of middle aged to older HoH ages and from a higher social class. They also had a high penetration of household electrical appliances, particularly tumble dryers and dishwashers.

The second group were characterised by medium electricity users and correspond with profile groups P3, P6, and P7. These mainly comprised of a mixture of detached and semi detached and terraced dwellings of three and four bedrooms. The dwellings were mainly occupied by adults only and again had a mixture of middle aged to older HoH ages but from a lower social class than the last group. These households also owned what could be considered to be an average number of common household electrical appliances.

Finally profile groups P8, P9 and P10 represent the lowest electricity users. These were typically characterised by apartments and terraced dwelling types of one to three bedrooms. The dwellings were mainly occupied by people living alone and from a mixture of younger and older HoH ages and from a lower Social class than the previous two groups. They also had the lowest penetration of common household appliances.

9.3 Recommendations for Further Research

The research presented here used a number of different mathematical techniques to characterise domestic electricity demand. Each method applied built upon the strengths and weaknesses of the next with the purpose of achieving the objectives set out in Section 1.5.

The statistical method presented in Chapter 5 provided an analysis of the characteristics that were most influential in determining electricity demand within the home. Four electrical parameters were used to characterise diurnal domestic electricity demand, however, there is scope for investigating more. A multivariate linear regression was applied between the electrical parameters and dwelling and occupant characteristics. There is also further scope for carrying out a principal components analysis to include a larger number of characteristics and possibly improve the overall predictive power of using parameters to describe household electricity consumption. However, by combining dwelling and occupant characteristics the meaning of the characteristics will be lost.

The autoregressive Markov chain process provided a method of characterising the variable nature of domestic electricity load profiles. However, as discussed, it failed to capture the temporal components within the characterisation process. In addition, the ability to link transitional probability matrices and dwelling and occupant characteristics for a large number of households proved problematic. There is potential for further work in the area by splitting a diurnal electricity load profile up into a number of sections like that discussed in Chapter 1 (i.e. night time, morning, daily and evening periods). A method of linking dwelling and occupant characteristics to the transitional probability matrix would still need to be found but this could possibly be done through a combination of clustering and regression.

Fourier transforms and Gaussian processes provided a method of characterising the temporal components of domestic electricity load profiles within the descriptive coefficients. However, it would be interesting to investigate the performance of both these techniques at characterising at smaller time intervals of less than 30 minutes. A clustering based approach could also be applied where instead of clustering the actual data, Fourier transforms and Gaussian process coefficients could be segmented into groups.

Data mining, of which clustering is a part, is a dominant area of engineering and computer science in today's data-rich world. This research applied the most widely adopted clustering techniques and algorithms alongside a systematic engineering approach to produce a set of representative electricity load profiles. However, there are many different methods that can be used for data classification as well as various different algorithms within each technique to calculate cluster points. Therefore there is further scope to carry out additional work in the area, by applying different data mining techniques and algorithms as well as applying different methodologies to the area of domestic electricity load profiling.

REFERENCES

[1] European Smart Metering Industry Group (ESMIG), "A Guide to Smart Metering," Brussels.

[2] Electric Power Research Institute (EPRI), "Characterizing Residential Customer Preferences for Electric Service Plans," California, 2013.

[3] R. Hierzinger, M. Albu, H. van Elburg, A. J. Scott, A. Łazicki, L. Penttinen, F. Puente, and H. Sæle, "European Smart Metering Landscape Report 2012," Vienna, 2012.

[4] Sustainable Energy Authority of Ireland (SEAI), "Press Release - Full Data from National Smart Meter Trial Published." [Online]. Available: http://www.seai.ie/News_Events/Press_Releases/2012/National_Smart_Meter_Trial_Data_Release.pdf. [Accessed: 08-Aug-2013].

[5] Commission for Energy Regulation (CER) in Ireland, "Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report," Dublin, 2011.

[6] Irish Social Science Data Archive, "Data from the Commission for Energy Regulation (CER) - Smart Metering Project." [Online]. Available: http://www.ucd.ie/issda/data/commissionforenergyregulationcer/. [Accessed: 14-Aug-2013].

[7] J. V Ringwood, D. Bofelli, and F. T. Murray, "Forecasting Electricity Demand on Short , Medium and Long Time Scales Using Neural Networks," *Journal of Intelligent and Robotic Systems*, vol. 31, pp. 129–147, 2001.

[8] J. Lourenco and P. Santos, "Short Term Load Forecasting Using Gaussian Process Models," *Instituto de Engenharia de Sistemas e Comput adores de Coimbra*, pp. 1–12, 2010.

[9] L. Amaral, R. Souza, and M. Stevenson, "A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting," *International Journal of Forecasting*, vol. 24, no. 4, pp. 603–615, Oct. 2008.

[10] Eirgrid, "System Demand," 2013. [Online]. Available: http://www.eirgrid.com/operations/systemperformancedata/systemdemand/. [Accessed: 14-Aug-2013].

[11] Retail Market Design Service, "Standard Load Profiles." [Online]. Available: http://www.rmdservice.com/guidance/standard_load_profiles.htm. [Accessed: 14-Aug-2013].

[12] Elexon, "Load Profiles And Their Use In Electricity Settlement." [Online]. Available: http://www.elexon.co.uk/ELEXON Documents/load_profiles.pdf. [Accessed: 30-Nov-2011].

[13] Y. G. Yohanis, J. D. Mondol, A. Wright, and B. Norton, "Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use," *Energy and Buildings*, vol. 40, no. 6, pp. 1053–1059, Jan. 2008.

[14] S. Firth, K. Lomas, A. Wright, and R. Wall, "Identifying trends in the use of domestic appliances from household electricity consumption measurements," *Energy and Buildings*, vol. 40, no. 5, pp. 926–936, Jan. 2008.

[15] P. Conlon, "Development of Domestic and SME Time of Use Tariff Structures for a Smart Metering Program in Ireland," *ESBI*. [Online]. Available: http://www.esbi.ie/news/pdf/White-Paper-Time-of-Use-Tariff-Structures.pdf. [Accessed: 14-Aug-2013].

[16] European Commission DG Energy, "Green Paper - A 2030 framework for climate and energy policies," Brussels, 2013.

[17] Department of Communications Marine and Natural Resources, "Delivering A Sustainable Energy Future For Ireland," Dublin, 2007.

[18] L. G. Swan and V. I. Ugursal, "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 1819–1835, Oct. 2009.

[19] M. Aydinalp and V. Ugursal, "Comparison of neural network, conditional demand analysis, and engineering approaches for modeling end-use energy consumption in the residential sector," *Applied Energy*, vol. 85, no. 4, pp. 271–296, Apr. 2008.

[20] K. J. Baker and R. M. Rylatt, "Improving the prediction of UK domestic energy-demand using annual consumption-data," *Applied Energy*, vol. 85, no. 6, pp. 475–482, Jun. 2008.

[21] E. Gabreyohannes, "A nonlinear approach to modelling the residential electricity consumption in Ethiopia," *Energy Economics*, vol. 32, no. 3, pp. 515–523, May 2010.

[22] A. Capasso, W. Grattieri, R. Lamedica, and A. Prudenzi, "A bottom-up approach to residential load modeling," *IEEE Transactions on Power Systems*, vol. 9, no. 2, pp. 957–964, May 1994.

[23] R. Yao and K. Steemers, "A method of formulating energy load profile for domestic buildings in the UK," *Energy and Buildings*, vol. 37, no. 6, pp. 663–671, Jun. 2005.

[24] J. Widén and E. Wäckelgård, "A high-resolution stochastic model of domestic activity patterns and electricity demand," *Applied Energy*, vol. 87, no. 6, pp. 1880–1892, Jun. 2010.

[25] I. Richardson, M. Thomson, D. Infield, and C. Clifford, "Domestic electricity use: A high-resolution energy demand model," *Energy and Buildings*, vol. 42, no. 10, pp. 1878–1887, Oct. 2010.

[26] G. Wood and M. Newborough, "Dynamic energy-consumption indicators for domestic appliances: environment, behaviour and design," *Energy and Buildings*, vol. 35, no. 8, pp. 821–841, Sep. 2003.

[27] J. O'Doherty, S. Lyons, and R. S. J. Tol, "Energy-using appliances and energy-saving features: Determinants of ownership in Ireland," *Applied Energy*, vol. 85, no. 7, pp. 650–662, Jul. 2008.

[28] E. Leahy and S. Lyons, "Energy use and appliance ownership in Ireland," *Energy Policy*, vol. 38, no. 8, pp. 4265–4279, Aug. 2010.

[29] Martin Howley, E. Dennehy, B. Ó. Gallachóir, and M. Holland, "Energy in Ireland 1990 – 2011," Dublin, 2012.

[30] A. Field, *Discovering Statistics Using SPSS*. London: SAGE Publications, 2005.

[31] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, Jan. 2002.

[32] M. Aydinalp, V. I. Ugursal, and A. S. Fung, "Modeling of the appliance, lighting , and space-cooling energy consumptions in the residential sector using neural networks," *Applied Energy*, vol. 71, pp. 87–110, 2002.

[33] M. Aydinalp, "Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks," *Applied Energy*, vol. 79, no. 2, pp. 159–178, Oct. 2004.

[34] M. H. Beale, M. T. Hagan, and H. B. Demuth, "Neural Network Toolbox ™ User 's Guide," R2013a ed., Matlab.

[35] S. P. Moutter, P. S. Bodger, and P. T. Gough, "Spectral decomposition and extrapolation of variations in electricity loading," *IEEE Proceedings*, vol. 133, no. 5, pp. 247–255, 1986.

[36] P. S. Bodger, D. R. D. Brooks, and S. P. Moutter, "Spectral decomposition of variations in electricity loading using mixed radix Fast Fourier Transform," *IEE Proceedings C Generation, Transmission and Distribution*, vol. 134, no. 3, p. 197, 1987.

[37] E. Gonzalez-Romera, M. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting with neural networks and Fourier series," *Energy Conversion and Management*, vol. 49, no. 11, pp. 3135–3142, Nov. 2008.

[38] A. G. Riddell and K. Manson, "Parametrisation of domestic load profiles," *Applied Energy*, vol. 54, no. 3, pp. 199–210, Jul. 1996.

[39] L. Magnano and J. Boland, "Generation of synthetic sequences of electricity demand: Application in South Australia," *Energy*, vol. 32, no. 11, pp. 2230–2243, Nov. 2007.

[40] C. Chatfield, *The Analysis of Time Series: An Introduction*, Sixth Edit. Chapman & Hall/CRC, 2004.

[41] D. J. Leith, M. Heidl, and J. V Ringwood, "Gaussian Process Prior Models for Electrical Load Forecasting," *8th International Conference on Probabilistic Methods Applied to Power Systems*, 2004.

[42] MathWorks, *Curve Fitting Toolbox TM User 's Guide*, R2013a ed. .

[43] S. Pappas, L. Ekonomou, D. Karamousantas, G. Chatzarakis, S. Katsikas, and P. Liatsis, "Electricity demand loads modeling using AutoRegressive Moving Average (ARMA) models," *Energy*, vol. 33, no. 9, pp. 1353–1360, Sep. 2008.

[44] MathWorks, "Vector Autoregressive Models - MATLAB & Simulink." [Online]. Available: http://www.mathworks.co.uk/help/econ/var-models.html#btdp52f. [Accessed: 26-Jul-2013].

[45] G. E. P. Box and G. Jenkins, *Time Series Analysis, Forecasting and Control*. Holden-Day, Incorporated ©1990, 1990.

[46] O. Ardakanian, S. Keshav, and C. Rosenberg, "Markovian Models for Home Electricity Consumption," *GreenNet Conference 2011*, pp. 31–36, 2011.

[47] A. Shamshad, M. Bawadi, W. Wanhussin, T. Majid, and S. Sanusi, "First and second order Markov chain models for synthetic generation of wind speed time series," *Energy*, vol. 30, no. 5, pp. 693–708, Apr. 2005.

[48] Y. Y. Hsu and K. L. Ho, "Fuzzy expert systems: an application to short-term load forecasting," *IEEE Proceedings C - Generation, Transmission and Distribution*, vol. 139, no. 6, pp. 471 – 477, 1992.

[49] P. a. Mastorocostas, J. B. Theocharis, and a. G. Bakirtzis, "Fuzzy modeling for short term load forecasting using the orthogonal least squares method," *IEEE Transactions on Power Systems*, vol. 14, no. 1, pp. 29–36, 1999.

[50] H. Mori and N. Kobayashi, "Optimal Fuzzy Inference for Short-Term Load Forecasting," *Power*, vol. 11, no. 1, pp. 390–396, 1996.

[51] MathWorks, "Continuous Wavelet Transform - MATLAB & Simulink R2013b." [Online]. Available: http://www.mathworks.co.uk/help/wavelet/gs/continuous-wavelet-transform.html. [Accessed: 25-Jul-2013].

[52] V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, "An Electric Energy Consumer Characterization Framework Based on Data Mining Techniques," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 596–602, May 2005.

[53] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, and C. Toader, “Electric energy customer characterisation for developing dedicated market strategies,” *2001 IEEE Porto Power Tech Proceedings (Cat. No.01EX502)*, p. 6, 2001.

[54] S. V. Verdú, M. O. García, C. Senabre, A. G. Marín, and F. J. G. Franco, “Classification , Filtering , and Identification of Electrical Customer Load Patterns Through the Use of Self-Organizing Maps,” *IEEE Transactions on Power Systems*, vol. 21, no. 4, pp. 1672–1682, 2006.

[55] L. Hubert, H.-F. Köhn, and D. Steinley, “Cluster Analysis□: A Toolbox for MATLAB,” 2009, pp. 444–512.

[56] MathWorks, “Hierarchical Clustering - MATLAB & Simulink R2013b.” [Online]. Available: http://www.mathworks.co.uk/help/stats/hierarchical-clustering.html. [Accessed: 26-Jul-2013].

[57] E. Mooi and M. Sarstedt, *A Concise Guide to Market Research*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2011.

[58] E. W. Weisstein, “Distance -- from Wolfram MathWorld.” [Online]. Available: http://mathworld.wolfram.com/Distance.html. [Accessed: 30-Jul-2013].

[59] S. S. Singh and N. C. Chauhan, “K-means v/s K-medoids□: A Comparative Study,” *National Conference on Recent Trends in Engineering & Technology, B.V.M. Engineering College, V.V.Nagar, Gujarat, India*, no. May, 2011.

[60] M. Hynar, M. Burda, and J. Sarmanov, “Unsupervised clustering with growing self-organizing neural network – a comparison with non-neural approach,” *Department of Computer Science, VSB – Technical University of Ostrava*, vol. ISBN 80–01, pp. 58–68, 2005.

[61] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, “On Clustering Validation Techniques,” *Journal of Intelligent Information Systems*, pp. 107–145, 2001.

[62] E. W. Weisstein, “K-Means Clustering Algorithm -- from Wolfram MathWorld.” [Online]. Available: http://mathworld.wolfram.com/K-MeansClusteringAlgorithm.html. [Accessed: 30-Jul-2013].

[63] G. Chicco, R. Napoli, F. Piglione, P. Postolache, M. Scutariu, and C. Toader, “Load Pattern-Based Classification of Electricity Customers,” *IEEE Transactions on Power Systems*, vol. 19, no. 2, pp. 1232–1239, May 2004.

[64] MathWorks, “Neural Network Toolbox Users Guide,” no. Version 4, 2002.

[65] E. Leahy, S. Lyons, and S. Walsh, “Working Paper No. 421- Electrical Appliance Ownership and Usage in Ireland,” Economic and Social Research Institute - Dublin, 2012.

[66] M. Parti and C. Parti, “The Total and Appliance-Specific Conditional Demand for Electricity in the Household Sector,” *The Bell Journal of Economics*, vol. 11, no. 1, p. 309, 1980.

[67] A. Wright and S. Firth, "The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations," *Applied Energy*, vol. 84, no. 4, pp. 389–403, Apr. 2007.

[68] Commission for Energy Regulation (CER) in Ireland, "Decision on the National Rollout of Electricity and Gas Smart Metering," Dublin, 2012.

[69] M. Hart and R. de Dear, "Weather sensitivity in household appliance energy end-use," *Energy and Buildings*, vol. 36, no. 2, pp. 161–174, Feb. 2004.

[70] D. S. Parker, "Research highlights from a large scale residential monitoring study in a hot climate," *Energy and Buildings*, vol. 35, no. 9, pp. 863–876, Oct. 2003.

[71] I. C. Schick, P. B. Usoro, M. F. Ruane, and J. A. Hausman, "Residential end-use load shape estimation from whole house metered data," *IEEE Transactions on Power Systems*, vol. 3, no. 3, pp. 986–992, 1987.

[72] Aigner, C., Sorooshian, C., Kerwin, P., "Conditional demand analysis for estimating residential end-use profiles," *The Energy Journal*, vol. 5, pp. 81–97, 1984.

[73] R. Bartels, R., Fiebig, D.G., Garben, M., Lumsdaine, "An end-use electricity load simulation model (DELMOD)," *Utilities Policy*, vol. 2, no. 1, pp. 71–82, 1992.

[74] N. Cross and C. T. Gaunt, "Application of rural residential hourly load curves in energy modelling," *2003 IEEE Bologna Power Tech Conference Proceedings,*, vol. 3, pp. 853–856, 2003.

[75] B. M. Larsen and R. Nesbakken, "Household electricity end-use consumption: results from econometric and engineering models," *Energy Economics*, vol. 26, no. 2, pp. 179–200, Mar. 2004.

[76] A. Ljones, A., Nesbakken, R., Sandbakken, S., Aaheim, "Household Energy Consumption. The Energy Survey 1990 - Reports 92/2," Statistics Norway, 1992.

[77] K. H. Tiedemann, "Using conditional demand analysis to estimate residential energy use and energy savings," *European Council for Energy Efficient Economy (ECEEE) Summer Study*, pp. 1279–1283, 2007.

[78] J. Widén, M. Lundh, I. Vassileva, E. Dahlquist, K. Ellegård, and E. Wäckelgård, "Constructing load profiles for household electricity and hot water from time-use data—Modelling approach and validation," *Energy and Buildings*, vol. 41, no. 7, pp. 753–768, Jul. 2009.

[79] Y. Shimoda, T. Fujii, T. Morikawa, and M. Mizuno, "Residential end-use energy simulation at city scale," *Building and Environment*, vol. 39, no. 8, pp. 959–967, Aug. 2004.

[80] J. L. Walker, C F, Pokoski, "Residential Load Shape Modeling Based on Customer Behavior," *IEEE Transactions on Power Apparatus and Systems*, no. 7, pp. 1703–1711, 1985.

[81] C. N. Jardine, “Synthesis of high resolution domestic electricity load profiles,” *1st International Conference & Workshop on Micro-Cogeneration Technologies and Applications*, vol. 7, pp. 1–8, 2008.

[82] M. Stokes, “Removing barriers to embedded generation: a fine-grained load model to support low voltage network performance analysis,” Ph.D Thesis, De Montfort University, 2005.

[83] A. Prudenzi, “A Neuron Nets Based Procedure for Identifying Domestic Appliances Pattern-of-Use from Energy Recordings at Meter Panel .,” *IEEE Power Engineering Society Winter Meeting*, vol. 2, pp. 941 – 946, 2002.

[84] H. Chen, C. A. Canizares, and A. Singh, “ANN-based short-term load forecasting in electricity markets,” in *IEEE Power Engineering Society Winter Meeting. Conference Proceedings*, 2001, vol. 2, pp. 411–415.

[85] S. MohammadZadeh and A. A. Masoumi, “Modeling residential electricity demand using neural network and econometrics approaches,” in *The 40th International Conference on Computers & Indutrial Engineering*, 2010, pp. 1–6.

[86] R. Singh, B. C. Pal, and R. A. Jabr, “Statistical Representation of Distribution System Loads Using Gaussian Mixture Model,” *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 29–37, 2010.

[87] H. Xu and T. Niimura, “Short-Term Electricity Price Modeling and Forecasting Using Wavelets and Multivariate Time Series,” *Power Systems Conference and Exposition*, pp. 1–5, 2004.

[88] Y. Chen, P. Luh, C. Guan, Y. Zhao, L. Michel, M. Coolbeth, P. Friedland, and S. Rourke, “Short-Term Load Forecasting: Similar Day-Based Wavelet Neural Networks,” *IEEE Transactions on Power Systems*, vol. Vol. 25, no. No. 1, pp. 3353–3358, 2010.

[89] J. Pahasa and N. Theera-Umpon, “Short-Term Load Forecasting Using Wavelet Transform and Support Vector Machines,” *Power Engineering Conference, IPEC 2007. International*, pp. 47–52, 2007.

[90] H. T. Nguyen and I. T. Nabney, “Short-term electricity demand and gas price forecasts using wavelet transforms and adaptive models,” *Energy*, vol. 35, no. 9, pp. 3674–3685, 2010.

[91] P. E. McSharry, S. Bouwman, and G. Bloemhof, “Probabilistic Forecasts of the Magnitude and Timing of Peak Electricity Demand,” *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1166–1172, May 2005.

[92] S. W. Heunis and R. Herman, “A Probabilistic Model for Residential Consumer Loads,” *IEEE Transactions on Power Systems*, vol. 17, no. 3, pp. 621–625, 2002.

[93] A. Cagni, E. Carpaneto, G. Chicco, R. Napoli, I. Elettrica, and P. Torino, “Characterisation of the Aggregated Load Patterns for Extra-Urban Residential Customer Groups,” *IEEE Melecon 2004*, pp. 1–4, 2004.

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

[95] G. Chicco and E. Carpaneto, “Probabilistic characterisation of the aggregated residential load patterns,” *IET Generation, Transmission and Distribution*, vol. 2, no. 3, pp. 373 – 382, 2008.

[96] D. H. O. McQueen, P. R. Hyland, and S. J. Watson, “Monte Carlo Simulation of Residential Electricity Demand for Forecasting Maximum Demand on Distribution Networks,” *IEEE Transactions on Power Systems*, vol. 19, no. 3, pp. 1685–1689, Aug. 2004.

[97] J. A. Jardini, C. M. V Tahan, M. R. Gouvea, S. U. Ahn, and F. M. Figueiredo, “Daily Load Profiles for Residential, Commercial and Industrial Low Voltage Consumers,” *IEEE Transactions on Power Delivery*, vol. 15, no. 1, pp. 375–380, 2000.

[98] C. S. Chen, M. S. Kang, J. C. Hwang, and C. W. Huang, “Temperature Effect to Distribution System Load Profiles and Feeder Losses,” *IEEE Transactions on Power Systems*, vol. 16, no. 4, pp. 916–921, 2001.

[99] G. Chicco, R. Napoli, and F. Piglione, “Comparisons among clustering techniques for Electricity Customer Classification,” *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933–940, 2006.

[100] S. V. Verdú, M. O. García, F. J. G. Franco, N. Encinas, A. G. Marín, A. Molina, and E. G. Lázaro, “Characterization and Identification of Electrical Customers Through the Use of Self-Organizing Maps and Daily Load Parameters,” *IEEE PES Power Systems Conference and Exposition*, vol. 2, pp. 899 – 906, 2004.

[101] B. Pitt, “Applications of Data Mining Techniques to Electric Load Profiling,” Ph.D Thesis, University of Manchester Institute of Science and Technology, 2000.

[102] M. Gavrilas, V. C. Sfintes, and M. N. Filimon, “Identifying Typical Load Profiles Using Neural- Fuzzy Models,” *Transmission and Distribution Conference and Exposition, 2001 IEEE/PES*, vol. 2, pp. 421 – 426, 2001.

[103] G. J. Tsekouras, N. D. Hatziargyriou, and E. N. Dialynas, “Two-Stage Pattern Recognition of Load Curves for Classification of Electricity Customers,” *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1120–1128, 2007.

[104] E. Carpaneto, G. Chicco, R. Napoli, and M. Scutariu, “Electricity customer classification using frequency–domain load pattern data,” *International Journal of Electrical Power & Energy Systems*, vol. 28, no. 1, pp. 13–20, Jan. 2006.

[105] M. Espinoza, C. Joye, R. Belmans, and B. De Moor, “Short-Term Load Forecasting, Profile Identification, and Customer Segmentation: A Methodology Based on Periodic Time Series,” *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1622–1630, 2005.

[106] T. Zhang, G. Zhang, J. Lu, X. Feng, and W. Yang, “A New Index and Classification Approach for Load Pattern Analysis of Large Electricity Customers,” *IEEE Transactions on Power Systems*, vol. 27, no. 1, pp. 153–160, Feb. 2012.

[107] S. M. Bidoki, N. Mahmoudi-Kohan, M. H. Sadreddini, M. Z. Jahromi, and M. P. Moghaddam, “Evaluating Different Clustering Techniques for Electricity Customer Classification,” *Electrical Power Distribution Networks (EPDC)*, pp. 1–5, 2011.

[108] Commission for Energy Regulation (CER) in Ireland, “Reports and Publications,” *Smart Metering Documentation*. [Online]. Available: http://www.cer.ie/en/information-centre-reports-and-publications.aspx?article=1c6fdd02-da48-44b8-8703-7f0916c2de7a. [Accessed: 04-Jan-2013].

[109] Commission for Energy Regulation (CER) in Ireland, “Electricity Smart Metering Customer Behaviour Trials (CBT) Findings - Appendices,” Dublin, 2011.

[110] Central Statistics Office of Ireland, “Private Households in Permanent Housing Units by Type of Private Accommodation, Province County or City and CensusYear - StatBank - data and statistics,” *Census 2011*, 2012. [Online]. Available: http://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=CD511&PLanguage=0. [Accessed: 04-Jan-2013].

[111] Economic and Social Research Institute (ESRI), “Irish National Survey of Housing Quality (NSHQ),” Dublin, 2002.

[112] Sustainable Energy Authority of Ireland (SEAI), “Energy in the Residential Sector,” Dublin, 2008.

[113] Ipsos Mori, “Social Grade - A Classification Tool,” 2009. [Online]. Available: http://www.ipsos-mori.com/DownloadPublication/1285_MediaCT_thoughtpiece_Social_Grade_July09_V3_WEB.pdf. [Accessed: 08-Aug-2013].

[114] Central Statistics Office of Ireland, “Private Households by Composition of Private Household, Persons per Household and CensusYear - StatBank - Data

and Statistics.” [Online]. Available: http://www.cso.ie/px/pxeirestat/Statire/SelectVarVal/Define.asp?maintable=CD504&PLanguage=0. [Accessed: 20-Aug-2013].

[115] Central Statistics Office of Ireland, “Household Budget Survey 2009 - 2010,” 2012.

[116] E. W. Weisstein, “Autocorrelation -- from Wolfram MathWorld.” [Online]. Available: http://mathworld.wolfram.com/Autocorrelation.html. [Accessed: 27-Aug-2013].

[117] E. W. Weisstein, “Spectral Power Density -- from Wolfram MathWorld.” [Online]. Available: http://mathworld.wolfram.com/SpectralPowerDensity.html. [Accessed: 27-Aug-2013].

[118] SAGE, “Chapter 24 Logistic Regression,” *Extension Chapters on Advanced Techniques*. [Online]. Available: http://www.uk.sagepub.com/burns/website material/Chapter 24 - Logistic regression.pdf.

[119] D. Davies and D. Bouldin, “A cluster separation measure,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979.

[120] K. L. Lo, Z. Zakaria, and M. H. Sohod, “Determination of Consumers ’ Load Profiles based on Two-stage Fuzzy C-Means,” *Proceedings of the 5th WSEAS Int. Conf. on Power Systems and Electromagnetic Compatibility*, vol. 2005, pp. 212–217, 2005.

[121] A. Zainal, D. F. Samaon, M. A. Maarof, and S. M. Shamsuddin, “Fuzzy c-Means Sub-Clustering with Re-sampling in Network Intrusion Detection,” *2009 Fifth International Conference on Information Assurance and Security*, pp. 683–686, 2009.

[122] R. Srikanthan and T. a. McMahon, “Stochastic generation of annual, monthly and daily climate data: A review,” *Hydrology and Earth System Sciences*, vol. 5, no. 4, pp. 653–670, 2001.

[123] A. Graps, “An Introduction to Wavelets,” *IEEE Computational Science and Engineering, Summer 1995, vol. 2, num. 2*. [Online]. Available: http://www.amara.com/IEEEwave/IEEEwavelet.html. [Accessed: 30-Nov-2011].

[124] S. Zadeh and A. Masoumi, “Modelling Residential Electricity Demand Using Neural Network and Econometrics Approaches,” *Computers and Industrial Engineering (CIE), 2010 40th International Conference on*, pp. 1 – 6, 2010.

[125] F. McLoughlin, A. Duffy, and M. Conlon, “The Generation of Domestic Electricity Load Profiles through Markov Chain Modelling,” *Euro-Asian Journal of Sustainable Energy and Development Policy*, vol. 3, 2010.

LIST OF PUBLICATIONS

Journal Publications

Chapter 5

(1) 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, Vol. 48, pp. 240-248, May 2012.

Chapter 6

(2) McLoughlin, F., Duffy, A., Conlon, M., "The Generation of Domestic Electricity Load Profiles through Markov Chain Modelling," Euro-Asian Journal of Sustainable Energy Development Policy, Vol. 5, 2012.

Chapter 7

(3) McLoughlin F, Duffy, A., Conlon, M., Evaluation of time series techniques to characterise domestic electricity demand, Energy, Vol. 50, pp. 120-130, February 2013.

Conference Publications

Chapter 8

(1) McLoughlin, F., Duffy, A., Conlon, M., Analysing domestic electricity smart metering data using self-organising maps, CIRED Workshop, 29-30 May 2012, Lisbon, Portugal.

(2) McLoughlin, F., Duffy, A., Conlon, M., "A parametric analysis of domestic electricity consumption patterns in Ireland. Tenth International Conference on Environment and Electrical Engineering (EEEIC), 8-11 May 2011, Rome, Italy.

Appendix A: Domestic Electricity Demand Load Profile Groups

The following figures present each individual electricity load profile group over the six month period for weekdays, Saturdays and Sundays. Each load profile is discussed in terms of its pattern of electricity use and possible factors driving its use are discussed.

Profile Group 1

Figure A.1 to Figure A.3 show electricity use for profile group 1 for weekdays, Saturdays and Sundays across the six month period. Figure A.1 shows profile group 1 for weekdays where each individual line corresponds to mean electricity demand for a particular day over the six month period for that group. In total there are 132 diurnal periods shown, excluding Saturdays and Sundays which are shown in Figure A.2 and Figure A.3. A clear seasonality effect is observed from July to December, with the brighter colours representing mid/late summer through to the darker colours indicating mid/late winter. In terms of school and work holidays in Ireland, the beginning of July is associated with the beginning of the summer period. Similarly, the end of December is associated with the Christmas holiday period. Hence it is an appropriate time span as it tells us a lot about household occupant behaviour across the calendar season's summer, autumn and winter and significant holiday periods.

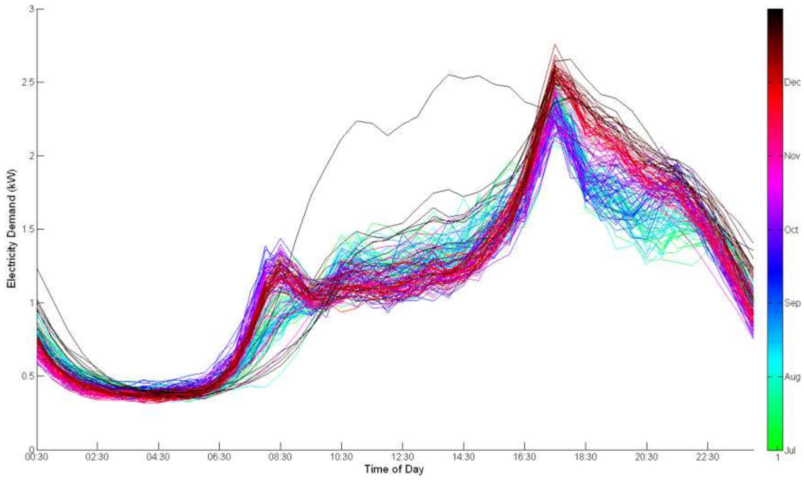


Figure A.1: Weekday profile group 1 over the six month period (July – Dec 2009)

Figure A.1 also shows an increase in electricity demand between the early morning hours of 00:30 to 06:30 during the summer months. The difference is only small, approximately 10-15 Watts, and is almost certainly a result of cold appliances such as fridges and freezers cycling more frequently and hence consuming more electricity during the summer. This effect is noticeable in the early morning period as there is little or no activity within the household at these times compared to other times of the day where the effect is lost to general electricity consumption throughout the household.

The occurrence of a morning peak in the winter time much earlier and more pronounced than in summer is also evident from Figure A.1. This would suggest that lighting is a significant contributor to the morning peak as electric central heating penetration in Irish households is reasonably small (<5%). Daytime electricity use is approximately the same between summer and winter months with possibly slightly more electricity being used during the former. This is most likely related to increased occupancy rates during the day over the summer period due to school holidays or vacation days from work. The evening peak changes by approximately 75 Watts between summer and winter. It is interesting to note the difference between the reduction in electricity demand after the evening peak for summer and winter periods. The variation is most likely attributed to lighting as indicated by each diurnal profile shape decreasing earlier during the summer months before returning to the overall trend of night time electricity use.

Figure A.2 and Figure A.3 show profile group 1 for Saturdays and Sundays respectively.

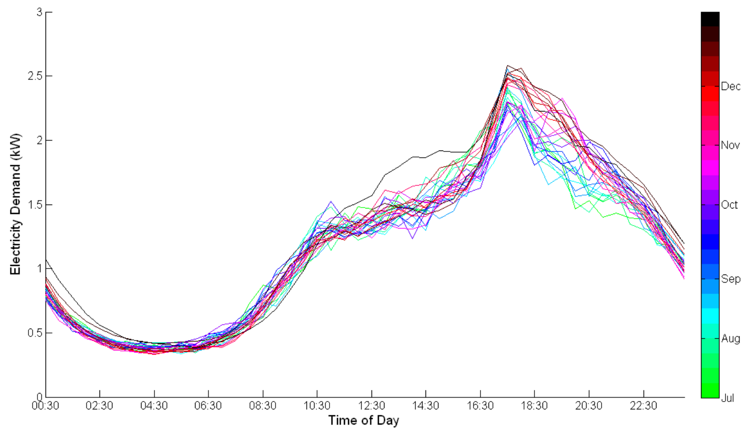


Figure A.2: Saturday profile group 1 over the six month period (July – Dec 2009)

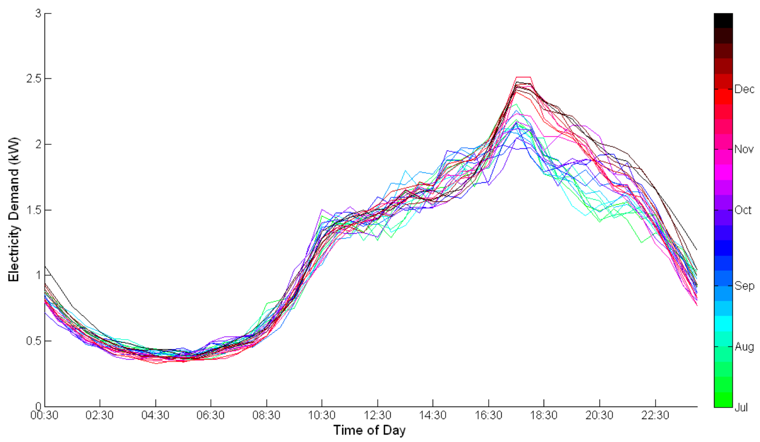


Figure A.3: Sunday profile group 1 over the six month period (July – Dec 2009)

As one might anticipate, there is not the same characteristic sharp morning peak shown in Figure A.2 and Figure A.3 for both Saturdays and Sundays. Electricity use in most households increases gradually as occupants get up at different times due to lower work or schooling commitments for these days. In addition, there is less of a seasonality effect to the profile shape for weekends as it may already be daylight when occupants are getting out of bed and are generally home during the day in both summer and winter. A similar amount of electricity is used at peak evening times to the same times on weekdays. A steeper evening peak is apparent for Saturday compared to Sunday, suggesting a more gradual increase in evening electricity consumption for the latter. The seasonality component is still evident in the evening time as occupants switch on lights.

Profile Group 2

Figure A.4 to Figure A.6 show electricity use for profile group 2 for weekdays, Saturdays and Sundays across the six month period. Figure A.4 shows a later and lower use of electricity demand in the morning time compared to the previous profile. The most significant characteristic of this profile is the large electricity peak centred at 1pm (lunch time). There is little change in the peak due to seasonality, with a similar amount of electricity being used in summer and winter time periods. This profile is characteristic of a household which uses electricity intensely at lunch time instead of the more common evening time peak. This could correspond to households which tend to have their main meal for the day at lunch time as opposed to the evening time. The seasonality component between summer and winter is clearly evident between the hours of 16:30 to 22:30, which most likely relate to the change in lighting up times in Ireland throughout the year.

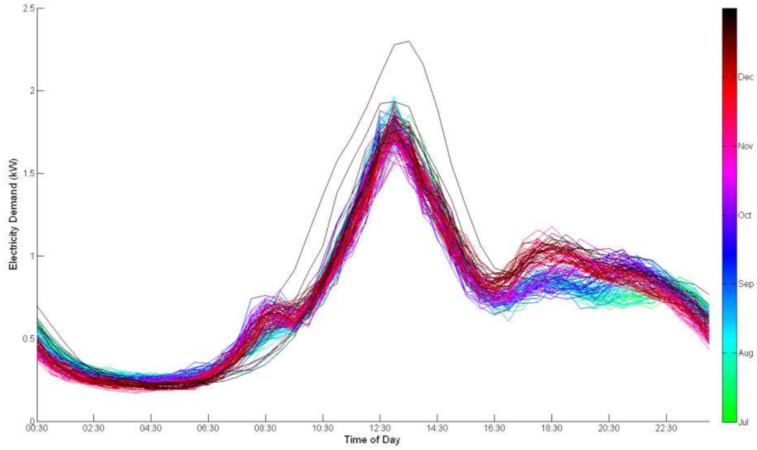


Figure A.4: Weekday profile group 2 over the six month period (July – Dec 2009)

Figure A.5 and Figure A.6 show profile group 2 for Saturdays and Sundays over the six month period. The main difference between weekday and weekend is the absence of the small morning peak.

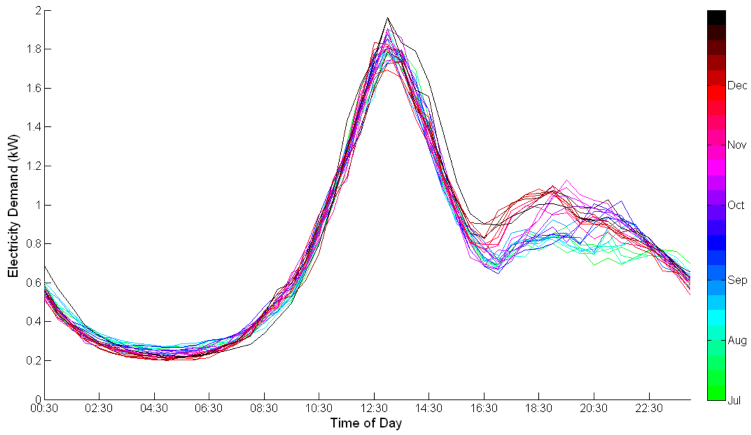


Figure A.5: Saturday profile group 2 over the six month period (July – Dec 2009)

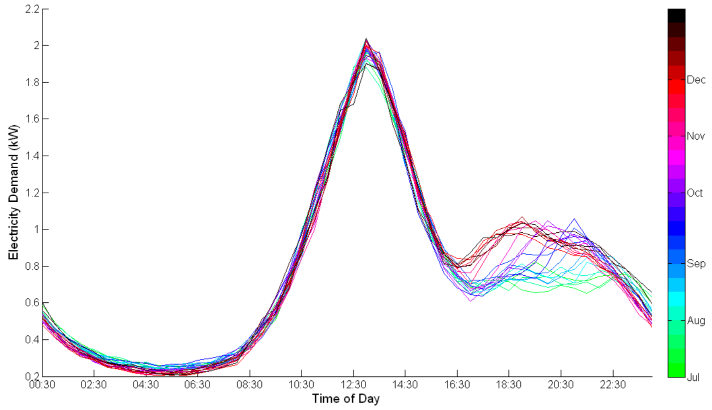


Figure A.6: Sunday profile group 2 over the six month period (July – Dec 2009)

Profile Group 3

Figure A.7 to Figure A.9 show profile group 3 for weekdays, Saturdays and Sundays over the six month period. Figure A.7 shows a large morning peak between the hours 07:00 and 10:30. There is very little electricity used across the later morning and afternoon periods before the evening peak starts at 16:30 which is mostly contributed to by lighting as evidenced by the strong seasonality effect. A significant amount of outliers are also apparent for this particular profile, all corresponding to the winter period. Some of these relate to holidays (such as Christmas and New Year's) and others may possibly be related to occupant vacation days.

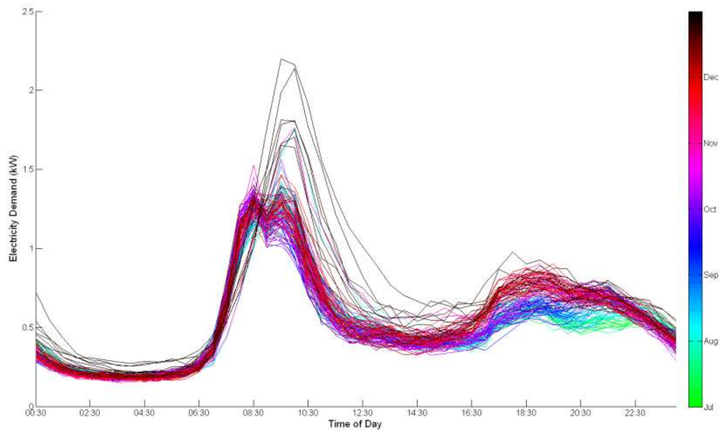


Figure A.7: Weekday profile group 3 over the six month period (July – Dec 2009)

Figure A.8 and Figure A.9 shows profile group 3 for Saturdays and Sundays over the six month period. Sundays tend to consume more electricity during the morning peak when compared against Saturday suggesting slightly more household activity at this time for the former.

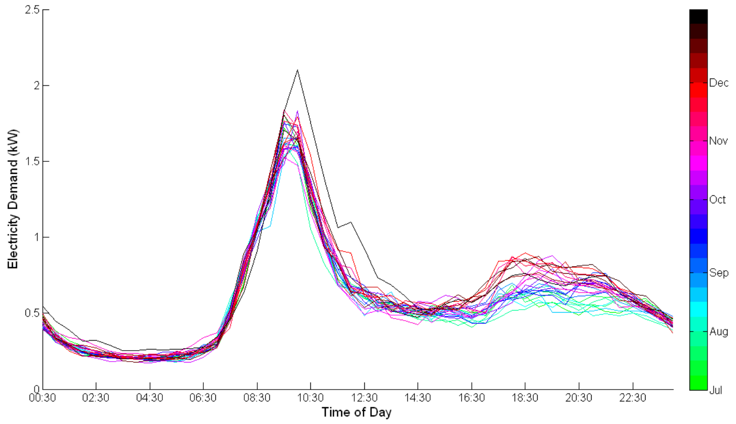


Figure A.8: Saturday profile group 3 over the six month period (July – Dec 2009)

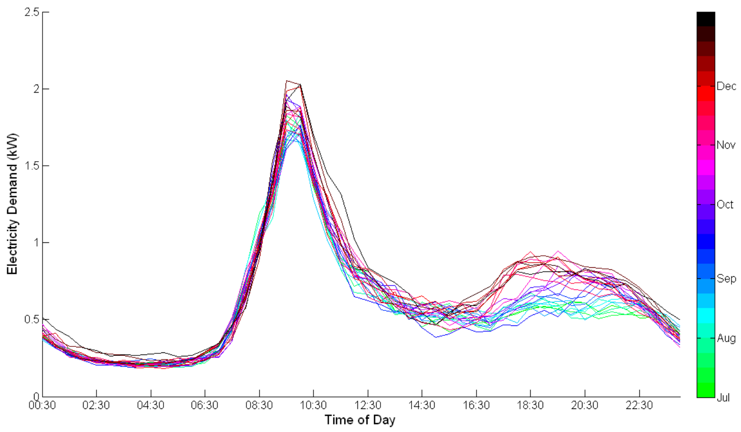


Figure A.9: Sunday profile group 3 over the six month period (July – Dec 2009)

Profile Group 4

Figure A.10 to Figure A.12 shows profile group 4 electricity use for weekdays, Saturdays and Sundays over the six month period. Figure A.10 shows the weekday profile which is similar in shape to profile group 1 presented earlier but significantly less in magnitude. Both a morning and evening peak are apparent, with a much smaller peak at lunch time. The morning peak starts at 06:30 and lasts until 09:00. It is likely that this is mainly composed of lighting as evidenced by the strong seasonality component. The evening peak starts at 16:00, with maximum occurring at 18:30. After the evening peak the difference between the decrease in electricity demand between summer and winter is clearly seen as lights do not need to be switched on till much later in the evening for the latter.

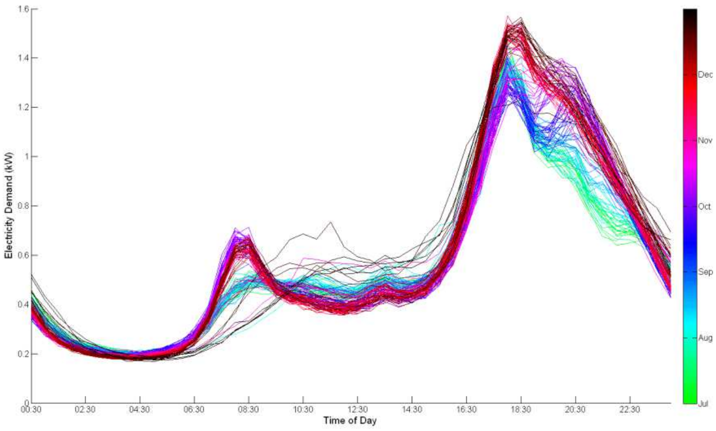


Figure A.10: Weekday profile group 4 over the six month period (July – Dec 2009)

Figure A.11 and Figure A.12 shows Saturdays and Sundays for profile group 4 over the six month period. The same morning peak shown in the weekday profile does not exist for Saturdays and Sundays. This again may suggest that this is most likely composed of lighting as occupants get up later at the weekends. More electricity is used over the late morning and early afternoon on Saturday as opposed to Sunday suggesting that occupants get out of bed a little earlier on the former compared to the latter.

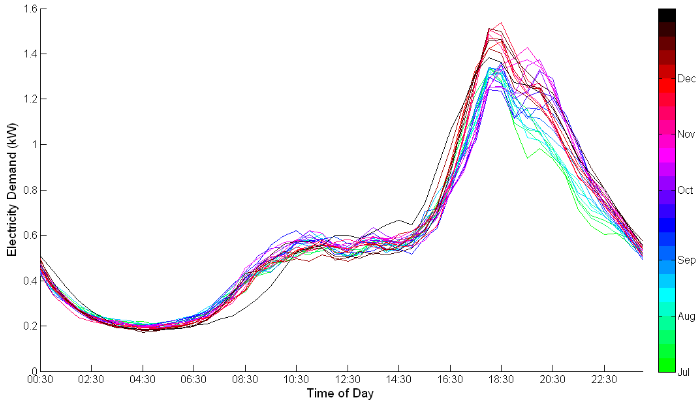
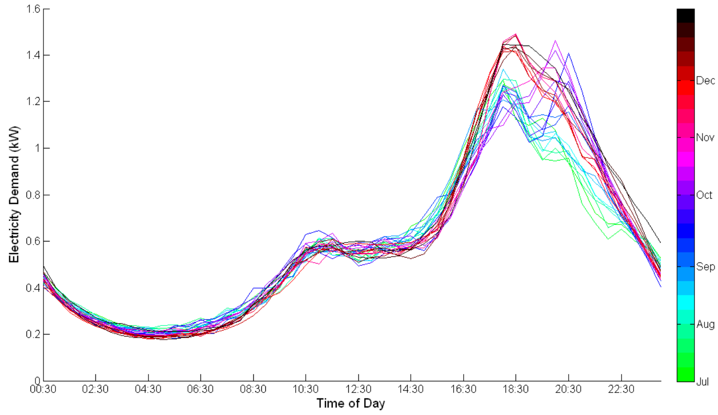


Figure A.11: Saturday profile group 4 over the six month period (July – Dec 2009)



Profile A.12: Sunday profile group 4 over the six month period (July – Dec 2009)

Profile Group 5

Figure A.13 to Figure A.15 shows electricity use for weekdays, Saturdays and Sundays for profile group 5. Figure A.13 shows a morning peak and a double evening peak with the first occurring at 16:30 and the second at 20:30. The difference between the two evening peaks is highly seasonal suggesting that a large part of the first evening peak is composed of lighting. The later evening peak does not have any seasonality component associated with it at all and is characteristic of households who tend to consume electricity late at night.

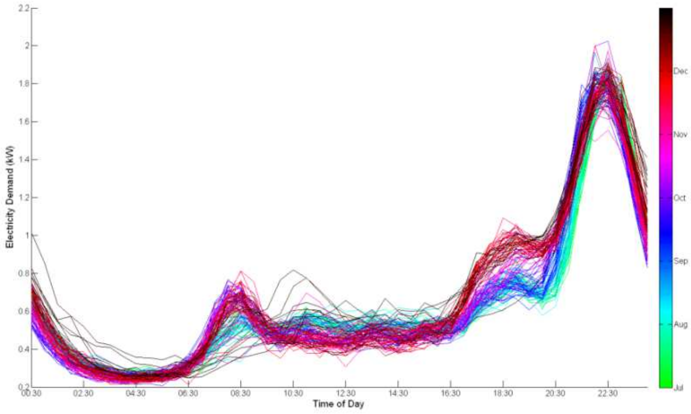


Figure A.13: Weekday profile group 5 over the six month period (July – Dec 2009)

Figure A.14 and Figure A.15 shows electricity profile group 5 for Saturdays and Sundays over the six month period. A seasonality component also exists in these figures, similar to that shown for profile group 4 indicating that it is not related to occupancy patterns and most likely a result of lighting as discussed above. The morning peak is smoother compared to the weekdays with no seasonal component suggesting later activity times in the household for the morning periods at the weekends.

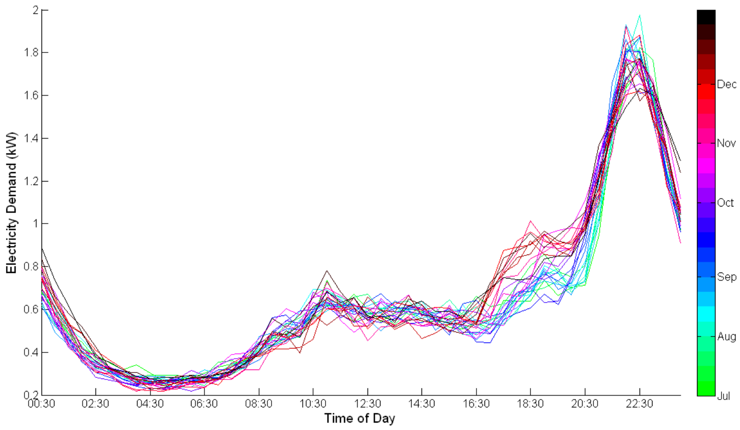


Figure A.14: Saturday profile group 5 over the six month period (July – Dec 2009)

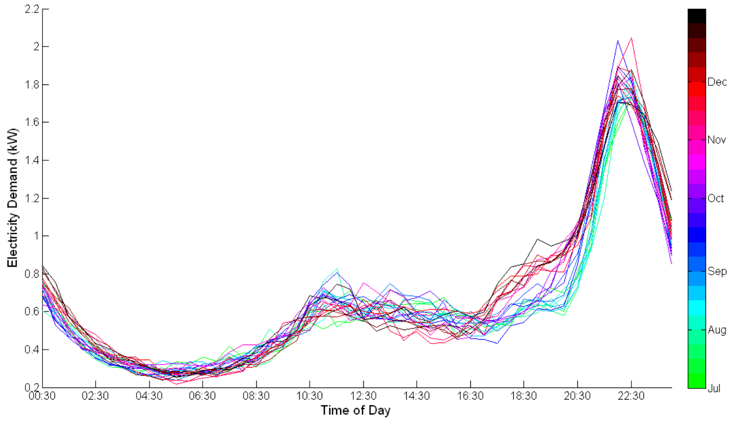


Figure A.15: Sunday profile group 5 over the six month period (July – Dec 2009)

Profile Group 6

Figure A.16 to Figure A.18 shows electricity use for profile group 6 for weekdays, Saturdays and Sundays over the six month period. Figure A.16 shows three distinct peaks indicating morning, lunch and evening time electricity use. The seasonal influence on the load profile shows more electricity being consumed over the late morning and early afternoon periods during the summer time. This is most likely related to increased occupancy over the day time period during the summer months but also could be due to increased cycling of cold appliances. As profile group 6 magnitude of electricity consumption is less compared to those presented previous, the contribution of cold appliances to overall electricity consumption (which are always left switched on) becomes more. Again the seasonality effect of lighting switching on and off in the evening time is apparent.

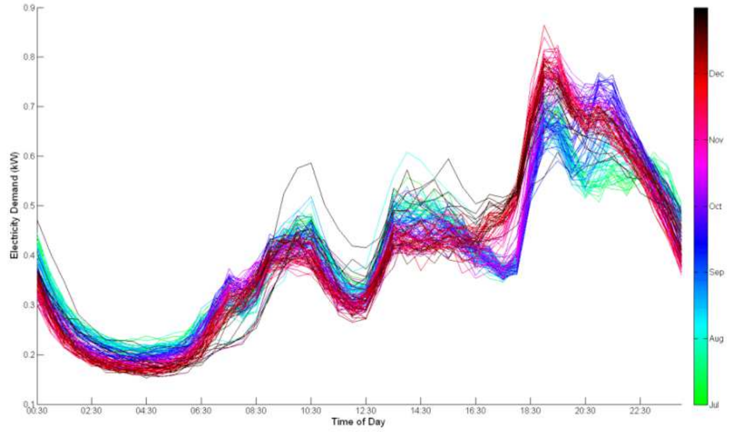


Figure A.16: Weekday profile group 6 over the six month period (July – Dec 2009)

Figure A.17 and Figure A.18 shows electricity profile group 6 for Saturdays and Sundays over the six month period. A greater use of electricity at the evening peak is apparent on the Saturday than on the Sunday which suggests a more gradual use of electricity on the latter.

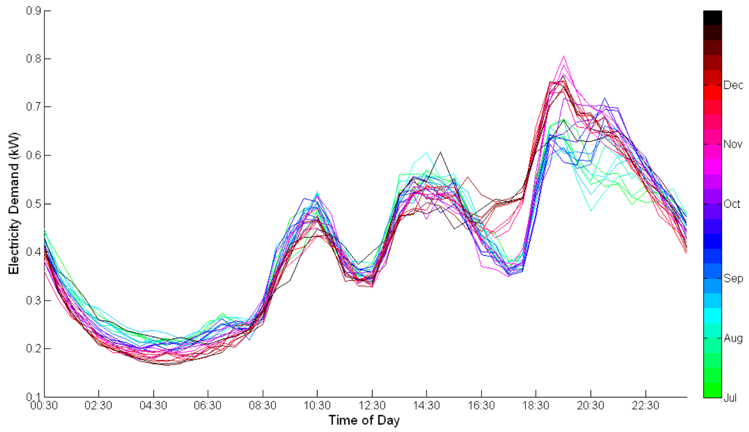


Figure A.17: Saturday profile group 6 over the six month period (July – Dec 2009)

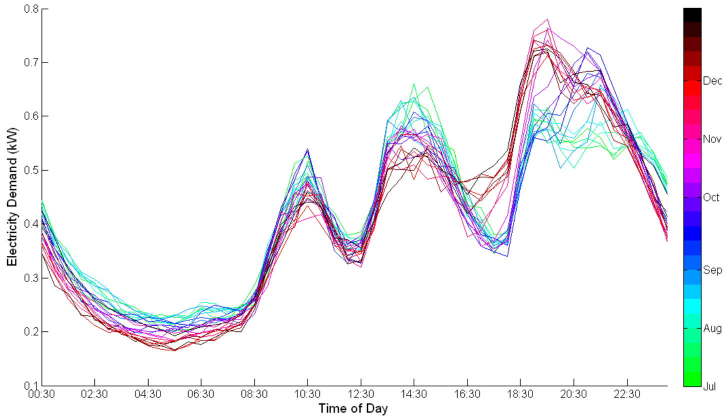


Figure A.18: Sunday profile group 6 over the six month period (July – Dec 2009)

Profile Group 7

Figure A.19 to Figure A.20 shows electricity profile group 7 for weekdays, Saturdays and Sundays over the six month period. Figure A.19 shows a weekday profile, similar to that shown for profile group 2, however, with two important differences. Profile group 7 shows an earlier lunch time peak at 12:00, as opposed to 13:00, and is significantly less in magnitude with a peak time electricity use of 1.5 kW as opposed to 2.4 kW. There is also less of a morning peak with profile group 7 compared to profile group 2. Again the seasonality component can be seen between the hours of 16:30 to 22:30 corresponding with changes in lighting up time as one goes toward the winter period.

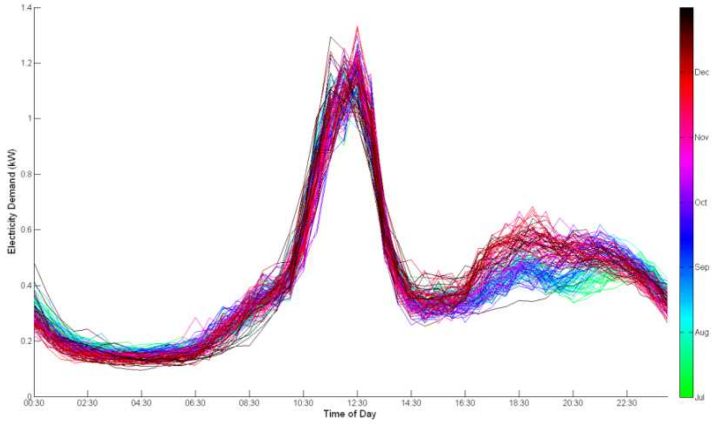


Figure A.19: Weekday profile group 7 over the six month period (July – Dec 2009)

Figure A.20 and Figure A.21 show electricity profile group 7 for Saturdays and Sundays. There is little difference between these two profiles except for a more prolonged use of electricity in the evening time on the Saturday compared to the Sunday for the winter months only.

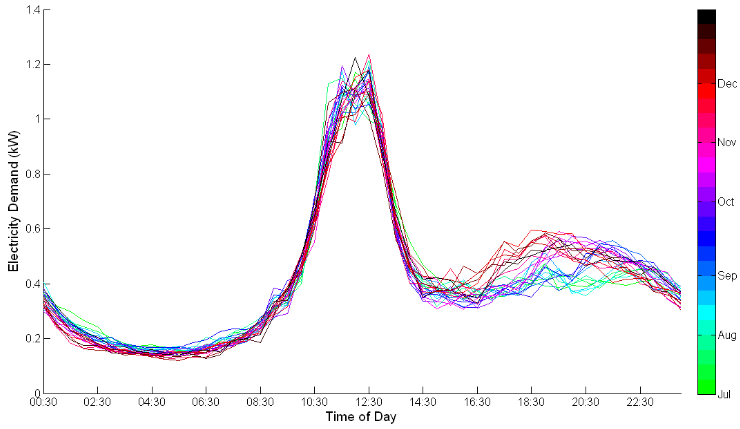


Figure A.20: Saturday profile group 7 over the six month period (July – Dec 2009)

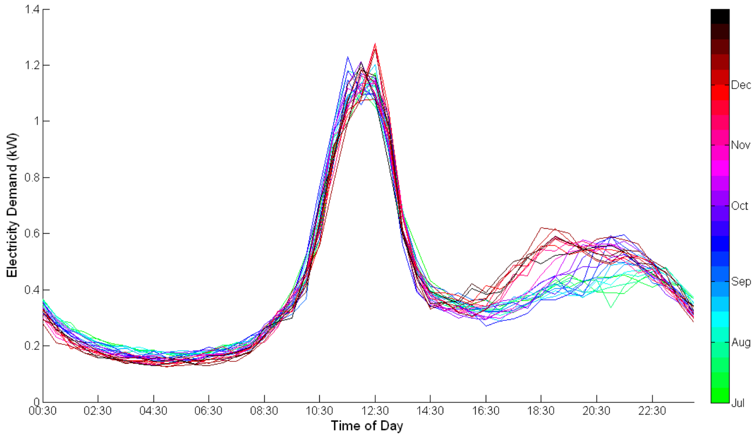


Figure A.21: Sunday profile group 7 over the six month period (July – Dec 2009)

Profile Group 8

Figure A.22 to Figure A.24 shows electricity profile group 8 for weekdays, Saturdays and Sundays over a six month period. Figure A.22 is different to any other profile group shown in previous figures, mainly because the magnitude component of the electricity load profile is very small, with a peak value of just 200 Watts. Therefore profile group 8 most probably represents a dwelling where there is little or no activity throughout the day and could possibly be classed as a vacant dwelling. The only significant contributors to electricity demand throughout the day almost undoubtedly come from one or more cold appliances and a possibly a small lighting component in the evening time. The strong seasonality component throughout the entire day highlights the influence of the increased cycling of cold appliances in the summertime due to an increase in external temperature.

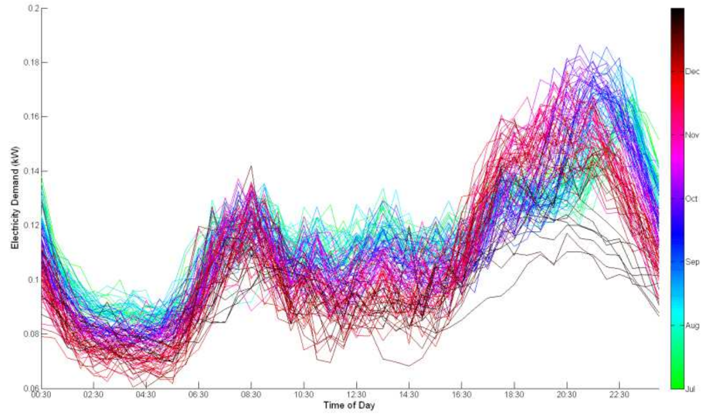


Figure A.22: Weekday profile group 8 over the six month period (July – Dec 2009)

Figure A.23 and Figure A.24 shows electricity profile group 8 for Saturdays and Sundays. Electricity consumption is marginally smaller at the weekend than during the week and tends to be larger in the evening time on Sunday compared to Saturday.

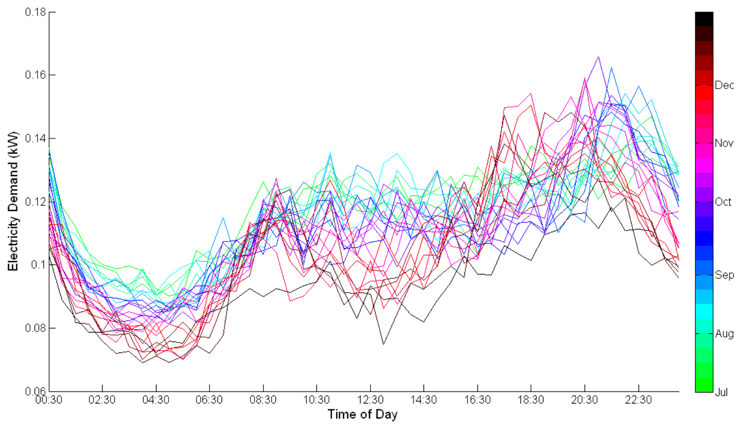


Figure A.23: Saturday profile group 8 over the six month period (July – Dec 2009)

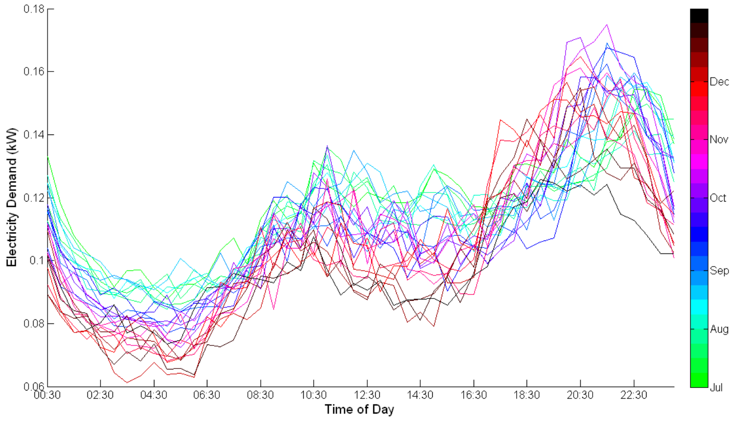


Figure A.24: Sunday profile group 8 over the six month period (July – Dec 2009)

Profile Group 9

Figure A.25 to Figure A.26 shows electricity profile group 9 for Saturdays and Sundays across the six month period. Profile group 9 is similar in shape to profile group 1 and 4 already shown but differs with a significantly less magnitude component to electricity consumption across the day. The increased activity of the cold appliances across the day is apparent as indicated by a strong seasonality component across the day and into the evening where lighting most likely becomes the dominant factor contributing to electricity consumption.

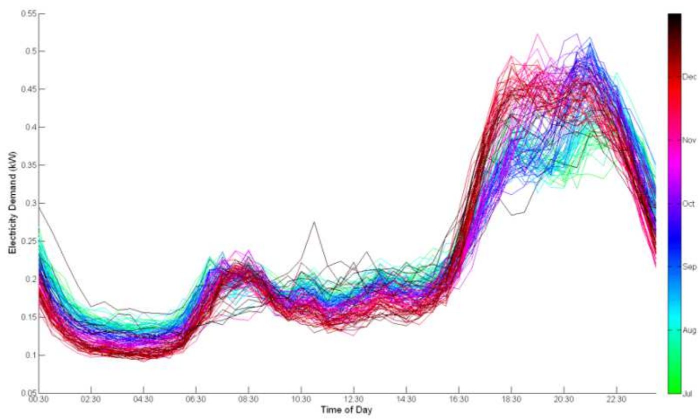


Figure A.25: Weekday profile group 9 over the six month period (July – Dec 2009)

Figure A.26 and Figure A.27 shows electricity profile group 9 for Saturdays and Sundays across the six month period. Both figures have the same seasonality component as the weekday with more electricity being consumed in the summer than in the winter time. Slightly more electricity is used over a longer time period on the Saturday compared to the Sunday suggesting more occupant activity in the home during these times on the former.

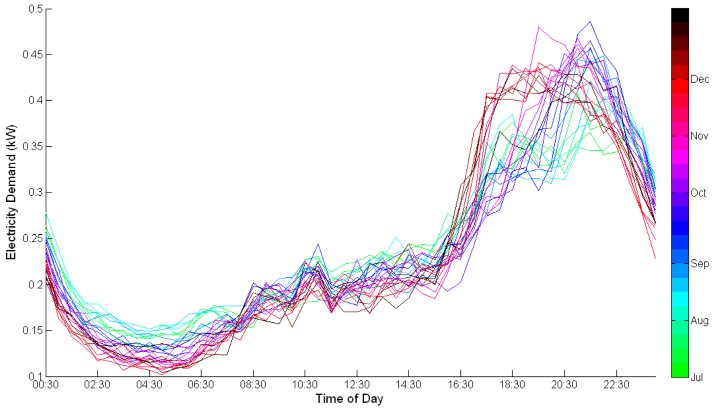


Figure A.26: Saturday profile group 9 over the six month period (July – Dec 2009)

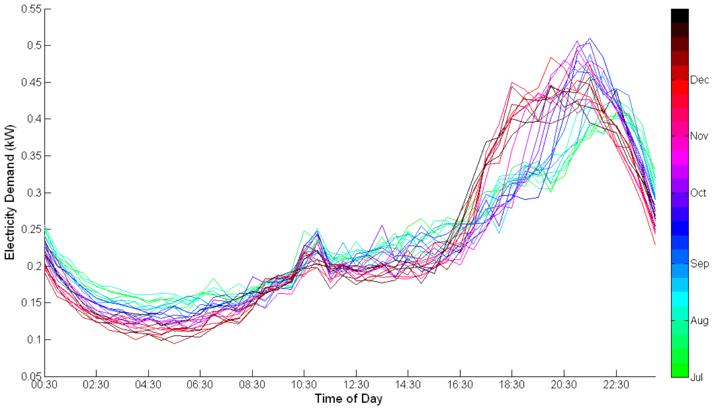


Figure A.27: Sunday profile group 9 over the six month period (July – Dec 2009)

Profile Group 10

Finally Figure A.28 to Figure A.30 shows electricity profile group 10 for weekdays, Saturdays and Sundays over the six month period. A late morning peak starting at 08:00 and ending at 10:30 with maximum electricity consumption of 700 Watts (excluding outlier) is apparent, in Figure A.28. This is followed by a period of smaller electricity consumption between the hours of 10:30 to 13:00. The seasonality component shows that largest amount of electricity is being used between the months July to October.

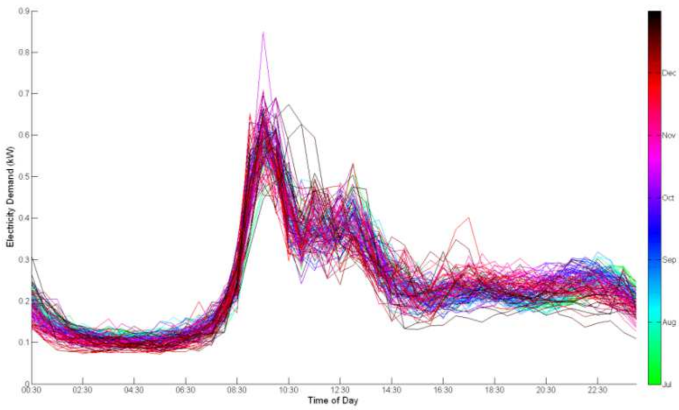


Figure A.28: Weekday profile group 10 over the six month period (July – Dec 2009)

Figure A.29 and Figure A.30 shows electricity profile group 10 for Saturdays and Sundays. The profiles show that electricity is used more continuously over the period of 14:30 to 23:30 for the Saturday whereas for the Sunday a small peak is apparent between the hours of 18:30 to 22:30.

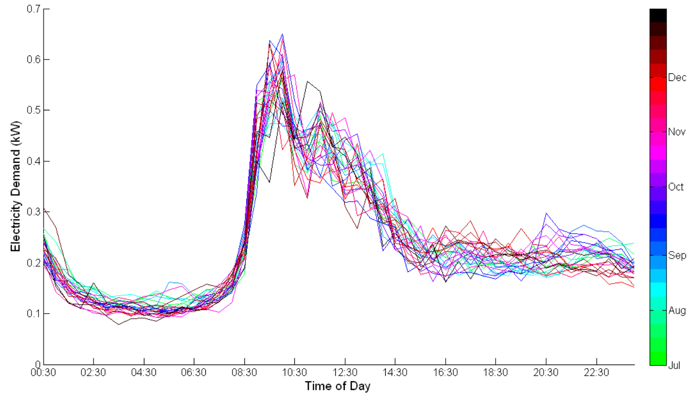


Figure A.29: Saturday profile group 10 over the six month period (July – Dec 2009)

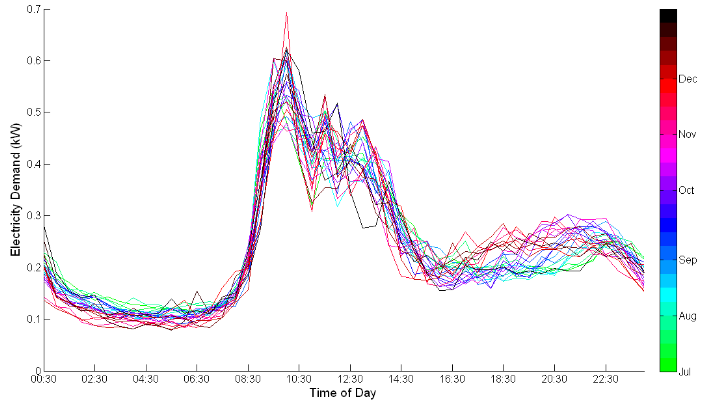


Figure A.30: Sunday profile group 10 over the six month period (July – Dec 2009)

Appendix B: Descriptive statistics for Average Percentage Profile Time (APPT)

Weekday

Profile Number	Profile 1 Mean	Profile 1 Std. Dev.	Profile 2 Mean	Profile 2 Std. Dev.	Profile 3 Mean	Profile 3 Std. Dev.	Profile 4 Mean	Profile 4 Std. Dev.	Profile 5 Mean	Profile 5 Std. Dev.
HMode (p1)	0.4834	0.1646	0.1102	0.0973	0.0334	0.0556	0.0713	0.0852	0.0979	0.0855
HMode (p2)	0.1099	0.0787	0.3786	0.1123	0.0690	0.0781	0.0728	0.0627	0.0593	0.0566
HMode (p3)	0.0666	0.0713	0.0827	0.0715	0.3852	0.1124	0.0690	0.0670	0.0450	0.0495
HMode (p4)	0.2014	0.0972	0.1792	0.0742	0.1730	0.0902	0.4699	0.1328	0.2444	0.0854
HMode (p5)	0.0803	0.0685	0.0556	0.0537	0.0382	0.0495	0.0733	0.0726	0.3835	0.0985
HMode (p6)	0.0275	0.0391	0.0871	0.0716	0.0983	0.0767	0.1203	0.0858	0.0894	0.0730
HMode (p7)	0.0060	0.0135	0.0562	0.0614	0.0406	0.0446	0.0306	0.0366	0.0150	0.0194
HMode (p8)	0.0132	0.0329	0.0223	0.0406	0.0616	0.0717	0.0306	0.0457	0.0276	0.0508
HMode (p9)	0.0096	0.0214	0.0189	0.0282	0.0679	0.0812	0.0497	0.0696	0.0303	0.0508
HMode (p10)	0.0022	0.0081	0.0092	0.0199	0.0329	0.0472	0.0126	0.0236	0.0076	0.0228

Profile Number	Profile 6 Mean	Profile 6 Std. Dev.	Profile 7 Mean	Profile 7 Std. Dev.	Profile 8 Mean	Profile 8 Std. Dev.	Profile 9 Mean	Profile 9 Std. Dev.	Profile 10 Mean	Profile 10 Std. Dev.
HMode (p1)	0.0143	0.0296	0.0133	0.0326	0.0061	0.0244	0.0034	0.0184	0.0047	0.0140
HMode (p2)	0.0427	0.0473	0.0896	0.0690	0.0119	0.0266	0.0089	0.0170	0.0236	0.0363
HMode (p3)	0.0485	0.0549	0.0466	0.0524	0.0317	0.0551	0.0364	0.0549	0.0751	0.0738
HMode (p4)	0.1932	0.0888	0.1229	0.0770	0.0578	0.0674	0.1099	0.0865	0.0653	0.0492
HMode (p5)	0.0532	0.0615	0.0093	0.0177	0.0104	0.0218	0.0139	0.0298	0.0066	0.0144
HMode (p6)	0.4077	0.1203	0.1284	0.0753	0.0412	0.0525	0.1321	0.0855	0.1074	0.0877
HMode (p7)	0.0675	0.0565	0.3920	0.1169	0.0182	0.0270	0.0299	0.0323	0.0870	0.0834
HMode (p8)	0.0327	0.0489	0.0453	0.0506	0.5973	0.2236	0.1584	0.1125	0.1371	0.1030
HMode (p9)	0.1087	0.0910	0.0733	0.0766	0.1652	0.1188	0.4387	0.1250	0.1414	0.0817
HMode (p10)	0.0316	0.0459	0.0794	0.0681	0.0604	0.0655	0.0684	0.0641	0.3521	0.0951

Saturday

Profile Number	Profile 1		Profile 2		Profile 3		Profile 4		Profile 5	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HMode (p1)	0.4579	0.1644	0.1067	0.0976	0.0445	0.0702	0.0653	0.0856	0.0890	0.0831
HMode (p2)	0.1630	0.0986	0.3776	0.1114	0.0989	0.0915	0.1009	0.0826	0.1045	0.0898
HMode (p3)	0.0630	0.0737	0.0758	0.0730	0.3761	0.1198	0.0555	0.0678	0.0316	0.0428
HMode (p4)	0.1712	0.0998	0.1915	0.0860	0.1670	0.0984	0.4341	0.1292	0.1808	0.0855
HMode (p5)	0.0651	0.0684	0.0523	0.0608	0.0239	0.0398	0.0548	0.0682	0.3645	0.1068
HMode (p6)	0.0356	0.0524	0.0828	0.0772	0.0964	0.0835	0.1293	0.0942	0.1177	0.1035
HMode (p7)	0.0103	0.0233	0.0506	0.0600	0.0503	0.0580	0.0497	0.0586	0.0333	0.0473
HMode (p8)	0.0166	0.0358	0.0290	0.0466	0.0580	0.0730	0.0406	0.0639	0.0281	0.0542
HMode (p9)	0.0130	0.0332	0.0208	0.0396	0.0508	0.0727	0.0463	0.0665	0.0419	0.0609
HMode (p10)	0.0044	0.0192	0.0130	0.0302	0.0341	0.0554	0.0235	0.0423	0.0086	0.0210

Profile Number	Profile 6		Profile 7		Profile 8		Profile 9		Profile 10	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HMode (p1)	0.0169	0.0368	0.0146	0.0347	0.0085	0.0329	0.0044	0.0235	0.0047	0.0172
HMode (p2)	0.0550	0.0658	0.1050	0.0743	0.0167	0.0399	0.0131	0.0286	0.0274	0.0451
HMode (p3)	0.0391	0.0568	0.0551	0.0745	0.0259	0.0507	0.0175	0.0380	0.0507	0.0626
HMode (p4)	0.1729	0.0897	0.1143	0.0787	0.0673	0.0799	0.0965	0.0886	0.0798	0.0702
HMode (p5)	0.0442	0.0597	0.0156	0.0324	0.0103	0.0266	0.0116	0.0325	0.0058	0.0194
HMode (p6)	0.4125	0.1309	0.1398	0.0879	0.0519	0.0632	0.1349	0.0931	0.1072	0.0791
HMode (p7)	0.0787	0.0726	0.3576	0.1286	0.0268	0.0459	0.0354	0.0489	0.0897	0.0782
HMode (p8)	0.0403	0.0593	0.0520	0.0503	0.5808	0.2271	0.1490	0.1164	0.1265	0.0944
HMode (p9)	0.1000	0.0920	0.0645	0.0709	0.1369	0.1136	0.4592	0.1409	0.1381	0.0892
HMode (p10)	0.0404	0.0580	0.0816	0.0818	0.0750	0.0804	0.0784	0.0800	0.3701	0.1060

Sunday

Profile Number	Profile 1		Profile 2		Profile 3		Profile 4		Profile 5	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HMode (p1)	0.4741	0.1661	0.0840	0.0915	0.0502	0.0825	0.0623	0.0837	0.0999	0.0898
HMode (p2)	0.1673	0.1070	0.4458	0.1509	0.1197	0.0982	0.1155	0.0934	0.1160	0.0879
HMode (p3)	0.0444	0.0585	0.0642	0.0777	0.4004	0.1353	0.0415	0.0634	0.0276	0.0513
HMode (p4)	0.1606	0.0974	0.1486	0.0944	0.1338	0.0913	0.4279	0.1309	0.1659	0.0884
HMode (p5)	0.0700	0.0767	0.0424	0.0600	0.0262	0.0470	0.0527	0.0673	0.3835	0.1102
HMode (p6)	0.0364	0.0521	0.0903	0.0834	0.1007	0.0890	0.1328	0.0957	0.1051	0.0960
HMode (p7)	0.0122	0.0281	0.0605	0.0713	0.0540	0.0702	0.0521	0.0589	0.0379	0.0503
HMode (p8)	0.0188	0.0427	0.0258	0.0435	0.0467	0.0712	0.0422	0.0617	0.0258	0.0516
HMode (p9)	0.0132	0.0277	0.0257	0.0435	0.0411	0.0584	0.0532	0.0699	0.0299	0.0526
HMode (p10)	0.0030	0.0151	0.0128	0.0310	0.0272	0.0549	0.0199	0.0390	0.0086	0.0220

Profile Number	Profile 6		Profile 7		Profile 8		Profile 9		Profile 10	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
HMode (p1)	0.0126	0.0365	0.0066	0.0204	0.0063	0.0259	0.0046	0.0243	0.0034	0.0140
HMode (p2)	0.0651	0.0741	0.0884	0.0877	0.0192	0.0438	0.0150	0.0345	0.0190	0.0444
HMode (p3)	0.0361	0.0605	0.0487	0.0732	0.0187	0.0419	0.0181	0.0447	0.0560	0.0718
HMode (p4)	0.1522	0.0866	0.0993	0.0904	0.0587	0.0749	0.0807	0.0791	0.0550	0.0665
HMode (p5)	0.0375	0.0612	0.0174	0.0362	0.0089	0.0278	0.0150	0.0397	0.0049	0.0155
HMode (p6)	0.4282	0.1377	0.1325	0.0891	0.0455	0.0641	0.1321	0.0955	0.0969	0.0903
HMode (p7)	0.0794	0.0725	0.4086	0.1350	0.0259	0.0448	0.0424	0.0600	0.0760	0.0699
HMode (p8)	0.0392	0.0577	0.0480	0.0618	0.5941	0.2326	0.1521	0.1140	0.1412	0.1086
HMode (p9)	0.1118	0.1014	0.0730	0.0768	0.1466	0.1172	0.4592	0.1406	0.1392	0.1033
HMode (p10)	0.0379	0.0560	0.0776	0.0909	0.0761	0.0899	0.0807	0.0890	0.4085	0.1022

Appendix C: Advantages and disadvantages of time series approaches to electricity load profile characterisation

Characterisation Type	Applied to Aggregate Demand	Applied to Individual Dwelling Demand	Time Resolution – High (≤ 1hr)	Time Resolution – Low (> 1hr)	Advantages	Disadvantages
Fourier Series	Yes [38][35][37]	No	Yes [38][35]	Yes [37]	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 [84][124][7]	Yes [32][33][19]	Yes [84][124][7]	Yes [32][33][19]	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 [86][41][8]	No	Yes [86][41][8]	No	Good at approximating small intervals of 'sharp' electricity demand.	Less good at approximating 'smother' average electricity demand profiles.
Autoregressive (incl. Markov chain)	Yes [9][21][39][43]	Yes [125][46]	Yes [9][39][125][46][43]	Yes [21]	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 a minimum of forty eight variables used (i.e. each half hourly period characterised separately).
Fuzzy Logic	Yes [48][49][50]	No	Yes [48][49][50]	No	Cause and effect clearly defined between input and output.	A minimum of forty eight variables required (i.e. each half hour period characterised separately).
Wavelets	Yes [87][88][89][90]	No	Yes [87][88][89][90]	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 [91][95]	Yes [12][92][93][94][96][97][98][66]	Yes [91][12][92][93][94][95][96]	Yes[66]	Widely used for generating standard load profiles (as shown in Figure 2)	Load profiles tend to be average rather than variable unless each half hourly period is characterised separately.

			[97][98]			

Appendix D: Smart metering survey questions

1	I would like to start by asking you a few questions about yourself. Are you the person in your home who is responsible or jointly responsible for paying the electricity bill?	
2	And are you the person who opted to sign up for the trial?	
3	May I ask your name please? NAME	
4	PLEASE RECORD SEX FROM VOICE	
5	May I ask what age you were on your last birthday? INT	IF NECCESSARY, PROMPT WITH AGE BANDS
6	What is the employment status of the chief income earner in your household, is he/she	
7	SOCIAL CLASS Interviewer, Respondent said that occupation of chief income earner was.... <CLASS> Please code	
8	Do you have internet access in your home?	
9	Do you have broadband in your home?	
10	Do you use the internet regularly yourself?	
11	Are there other people in your household that use the internet regularly?	
12	What best describes the people you live with? READ OUT	
13	How many people over 15 years of age live in your home?	
14	And how many of these are typically in the house during the day (for example for 5-6 hours during the day)	
15	How many people under 15 years of age live in your home?	
16	And how many of these are typically in the house during the day (for exanmple for 5-6 hours during the day)	
17	I/we am/are interested in changing the way I/we use electricity if it reduces the bill	
18	I/we am/are interested in changing the way I/we use electricity if it helps the environment	
19	I/we can reduce my electricity bill by changing the way the people I/we live with use electricity	
20	I/we have already done a lot to reduce the amount of electricity I/we use	
21	I/we have already made changes to the way I/we live my life in order to reduce the amount of electricity we use.	
22	I/we would like to do more to reduce electricity usage	
23	I/we know what I/we need to do in order to reduce electricity usage	
24	I/we have already done a lot to reduce the amount of electricity I/we use	
25	I/we have already made changes to the way I/we live my life in order to reduce the amount of electricity we use.	
26	I/we would like to do more to reduce electricity usage	
27	I/we know what I/we need to do in order to reduce electricity usage	

28	Thinking about the energy reduction activities undertaken by you or your family/household, in the last year, did your efforts reduce your bills?	
29	Approximately what % savings on average did you achieve on the average bill?	
30	It is too inconvenient to reduce our usage of electricity	
31	I do not know enough about how much electricity different appliances use in order to reduce my usage	
32	I am not be able to get the people I live with to reduce their electricity usage	
33	I do not have enough time to reduce my electricity usage	
34	I do not want to be told how much electricity I can use	
35	Reducing my usage would not make enough of a difference to my bill	
36	If you were to make changes to the way you and people you live with use electricity, how much do you believe you could reduce your usage by?	
37	I would now like to ask some questions about your home. Which best describes your home?	
38	Do you own or rent your home?	
39	What year was your house built INT ENTER FOR EXAMPLE	1981- CAPTURE THE FOUR DIGITS
40	Approximately how old is your home?	
41	What is the approximate floor area of your home?	
42	Is that	
43	How many bedrooms are there in your home	
44	Which of the following best describes how you heat your	Electricity (electric central heating/storage heating)
45	Which of the following best describes how you heat your	Electricity (plug in heaters)
46	Which of the following best describes how you heat your	Gas
47	Which of the following best describes how you heat your	Oil
48	Which of the following best describes how you heat your	Solid fuel
49	Which of the following best describes how you heat your	Renewable (e.g. solar)
50	Which of the following best describes how you heat your	Other
51	Do you have a timer to control when your heating comes on and goes off?	
52	Which of the following best describes how you heat water in	Central heating system
53	Which of the following best describes how you heat water in	Electric (immersion)
54	Which of the following best describes how you heat water in	Electric (instantaneous heater)
55	Which of the following best describes how you heat water in	Gas
56	Which of the following best describes how you heat water in	Oil
57	Which of the following best describes how you heat water in	Solid fuel boiler
58	Which of the following best describes how you heat water in	Renewable (e.g. solar)
59	Which of the following best describes how you heat water in	Other
60	Do you have a timer to control when your hot water/immersion heater comes on and goes off?	
61	Do you use your immersion when your heating is not switched on?	

62	Which of the following best describes how you cook in your home	
63	Returning to heating your home, in your opinion, is your home kept adequately warm?	
64	Do any of the following reasons apply?	I prefer cooler temperature
65	Do any of the following reasons apply?	I cannot afford to have the home as warm as I would like
66	Do any of the following reasons apply?	It is hard to keep the home warm because it is not well insulated
67	Do any of the following reasons apply?	None of these
68	Have you had to go without heating during the last 12 months through lack of money?	
69	Have any of the following ever applied to you?	I had to go without heat on a cold day
70	Have any of the following ever applied to you?	I had to go to bed to keep warm
71	Have any of the following ever applied to you?	I lit the fire late or switched on the heat late because I did not have enough fuel or money for fuel
72	Have any of the following ever applied to you?	None of these
73	Washing machine	
74	Tumble dryer	
75	Dishwasher	
76	Electric shower (instant)	
77	Electric shower (electric pumped from hot tank)	
78	Electric cooker	
79	Electric heater (plug-in convector heaters)	
80	Stand alone freezer	
81	A water pump or electric well pump or pressurised water system	
82	Immersion	
83	Washing machine	
84	Tumble dryer	
85	Dishwasher	
86	Electric shower (instant)	
87	Electric shower (electric pumped from hot tank)	
88	Electric cooker	
89	Electric heater (plug-in convector heaters)	
90	Stand alone freezer	
91	A water pump or electric well pump or pressurised water system	
92	Immersion	
93	TV’s less than 21 inch	
94	TV’s greater than 21 inch	
95	Desk-top computers	
96	Lap-top computers	
97	Games consoles, such as xbox, playstation or Wii	
98	TV’s less than 21 inch	

99	TV’s greater than 21 inch	
100	Desk-top computers	
101	Lap-top computers	
102	Games consoles, such as xbox, playstation or Wii	
103	Washing machine INT	
104	Tumble dryer INT	
105	Dishwasher INT	
106	Electric shower (instant) INT	
107	Electric shower (pumped from hot tank) INT	
108	Electric cooker INT	
109	Electric heater (plug-in) INT	
110	Water pump INT	
111	Immersion water INT	
112	Stand alone Freezer INT	
113	TV’s less than 21 inch INT	
114	TV’s greater than 21 inch INT	
115	Desk-top computers INT	
116	Lap-top computers INT	
117	Games consoles, such as xbox, playstation or Wii INT	
118	Does your home have a Building Energy Rating (BER) - a recently introduced scheme for rating the energy efficiency of your home?	
119	What rating did your house achieve?	
120	And now considering energy reduction in your home please indicate the approximate proportion of light bulbs which are energy saving (or CFL)? INT	READ OUT
121	Please indicate the approximate proportion of windows in your home which are double glazed? INT	READ OUT
122	Does your hot water tank have a lagging jacket?	
123	Is your attic insulated and if so when was the insulation fitted? INT	PROBE TO PRECODES
124	Are the external walls of your home insulated?	
125	I would now like to ask you about your expectations about	Learn how to reduce my energy usage
126	I would now like to ask you about your expectations about	Learn how to reduce my electricity bill
127	I would now like to ask you about your expectations about	Do my part to help the environment by my participation
128	I would now like to ask you about your expectations about	Do my part to make Ireland become more up to date
129	My household may decide to make minor changes to the way we use electricity	
130	My household may decide to make major changes to the way we use electricity	
131	My household may decide to be more aware of the amount of electricity used by appliances we own or buy.	
132	In future, when replacing an appliance, my household may decide to choose one with a better energy rating	

133	How do you think that your electricity bills will change as part of the trial?
134	By what amount?
135	By what amount?
136	Moving on to education, which of the following best describes the level of education of the chief income earner
137	And considering income, what is the approximate income of your household - this should be before tax, you should include the income of all adults in the household? Please note that this figure will remain completely confidential and will not
138	Can you state which of the following broad categories best represents the yearly household income BEFORE TAX?
139	Is that figure
140	Can I just double check is that figure..
141	The number of suppliers competing in the market
142	The percentage of electricity being generated from renewable sources
143	The overall cost of electricity
144	The number of estimated bills received by customers
145	The opportunity to sell back extra electricity you may generate (from solar panels etc) to your electricity supplier
146	The environmental damage associated with the amount of electricity used

Appendix E: Social Class Categories [113]

Social class of Chief Income Earner (CIE)

Social Class	Description
A	High managerial, administrative or professional
B	Intermediate managerial, administrative or professional
C1	Supervisory, clerical and junior managerial, administrative or professional
C2	Skilled manual workers
D	Semi and unskilled manual workers
E	State pensioners, casual or lowest grade workers, unemployed with state benefits only
F	Farmers