Residential electricity demand: A modelling analysis of the dynamics of household activity and energy consumption patterns



José Luis Ramírez Mendiola

St Peter's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary Term 2019

Acknowledgements

The research presented in this thesis would not have been possible without the support of the Consejo Nacional de Ciencia y Tecnologia (CONACyT), which provided me with a considerable portion of the funding needed to carry out my doctoral research.

I would like to thank my supervisors, Dr Phil Grünewald and Professor Nick Eyre, for their support throughout. I have benefited greatly from their advice, ideas and insights over past years, and I can say without a shadow of a doubt that this project would not be what it is without their valuable contributions.

Along this journey, I have had the honour of meeting so many wonderful people. And I am only grateful to have had the privilege of being a part of the warm community of the Environmental Change Institute and the School of Geography and the Environment, and having the pleasure of calling so many of you my friends. The list is long, but I would not forgive myself (and you probably would not forgive me either) if I didn't mention here those of you with whom I have shared the longest portion of the road:

A special mention goes to my gorgeous desk neighbours, Agne, Anabelle and Christine, who made all those days sitting in the DPhil room all the more interesting.

Huge shout out to Nico and Dina, the brother from another mother and the sister from another mister. I will be forever grateful for having been a member of The Gym and Tonic crew, and for all those days of days of hard training and all those nights of hard playing.

I would also like to thank you, my dearest Rachel, for being such an inspiration, and for your support throughout the final stages and the next steps. Gracias, amor.

A mis hermanos, Andrea y Pablo, a los que siempre he llevado y llevaré en mi corazón, y quienes me inspiran a ser la persona de quienes pueden estar orgullosos de llamar hermano.

Mamá y Papá, ustedes se merecen una mención particularmente especial. Nunca podré agradecerles lo suficiente for todo su apoyo a lo largo de todos estos años, pero har'e lo mejor que pueda, tal y como me han enseñado.

Statement of Originality

I hereby confirm that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Abstract

Understanding the diversity of energy requirements and energy-related behaviours in the residential sector is key to assessing the potential of this sector to contribute to the overall flexibility of future low-carbon power systems. Models offer the opportunity to explore the relationships between household activities and the associated electricity demand loads. However, if models are to assist in the research that aims to inform the design of effective interventions, understanding the role users play must be at the core of the modelling efforts. Conventional approaches to residential electricity demand modelling have been developed with a view to inform conventional power systems. The simplifying assumptions that rule these models mask the complexity of the links between user activities and demand for electricity. Therefore, assessing the potential of the implementation of energy efficiency measures targeting the residential sector calls for the development of enhanced modelling approaches. In this thesis the key shortcomings of conventional modelling approaches have been identified. In particular, the analysis has revealed that conventional models consistently fail to represent the diversity of residential users. Thus, the need to re-assess the adequacy of the simplifying assumptions used in the development of conventional modelling approaches has been highlighted. The diversity of residential energy requirements has been further investigated, and novel methods to characterise the socio-demographic heterogeneity of households have been presented. Furthermore, the thesis describes novel approaches that attempt to bridge the gap between energy research through social science lenses and the development of technical modelling tools. The findings of the validation analysis show that these approaches could indeed help improve our understanding of the current role of the residential sector and assess its potential to provide additional system flexibility.

Contents

List of Figures **xi**

List of Tables **xiv**

List of Acronyms and Abbreviations **xv**

Key definitions **xvii**

1 Introduction **1**

1.1 Background and Rationale 3

1.1.1 Conventional model for power generation 5

1.1.2 The power system of a low-carbon economy 7

1.1.3 Renewable power generation 10

1.1.4 Decentralised power systems 11

1.1.4.1 Advantages of decentralised systems 12

1.1.4.2 Drawbacks of decentralised systems 13

1.1.5 The demand-supply balancing challenge 14

1.1.6 Demand-Side Response mechanisms 17

1.1.7 The residential sector: a key target for system efficiency measures 21

1.1.8 Electricity demand modelling 23

1.1.8.1 Residential electricity demand modelling 25

1.1.8.2 The case for improving residential electricity demand modelling 28

1.2 Summary of aims and objectives 40

1.2.1 Research questions 40

1.3 Structure of the thesis 41

1.4 Notes on thesis format 44

2 The diversity of residential electricity demand: A comparative analysis of metered and simulated data 47
2.1 Introduction . . . 51
2.2 Residential electricity demand modelling . . . 53
2.3 The CREST model . . . 56
2.3.1 Structure . . . 57
2.3.2 Validation . . . 59
2.4 The UK Household Electricity Survey . . . 60
2.5 Comparison of CREST model's simulations against HES data . . . 62
2.5.1 Methodology . . . 63
2.5.1.1 Data processing . . . 63
2.5.1.2 Comparative analysis . . . 67
2.6 Analysis results . . . 70
2.6.1 Variations in annual demand levels . . . 70
2.6.2 Empirical cumulative distributions and significance tests . . . 73
2.6.3 Individual appliances: Total annual consumptions . . . 79
2.6.4 Mean daily load profile general features . . . 81
2.7 Discussion and Future work . . . 83
2.8 Conclusions . . . 88

3 Linking intra-day variations in residential electricity demand loads to consumers' activities: what's missing? 91
3.1 Introduction . . . 95
3.2 Activity-based load modelling background . . . 97
3.3 Data and methods . . . 100
3.3.1 Activity data . . . 100
3.3.2 Appliance usage data . . . 102
3.3.3 Extraction of daily activity and appliance usage patterns . . . 105
3.4 Results and Discussion . . . 108
3.4.1 Activity profiles and appliance usage . . . 109
3.4.1.1 Food preparation/Cooking . . . 109
3.4.1.2 Laundering . . . 111
3.4.1.3 Dish washing . . . 114
3.4.1.4 TV watching & ICT related . . . 115

3.4.2 Dependence of simulated power consumption on active occupancy . . . 117
3.4.3 Activity and activity-related consumption patterns in the literature . . . 122
3.4.4 Limitations and further research . . . 124
3.5 Conclusions . . . 125

4 Residential activity pattern modelling through stochastic chains of variable memory length 129
4.1 Introduction . . . 133
4.2 Activity-based demand modelling: a brief overview . . . 136
4.2.1 Time-use data . . . 137
4.2.2 Markov chain technique . . . 137
4.2.2.1 Higher-order Markov models . . . 139
4.2.3 General approach to activity-to-demand conversion . . . 141
4.2.4 Insights from social practice theory . . . 142
4.2.5 Towards a more robust approach to activity modelling . . . 143
4.3 Data and methods . . . 145
4.3.1 The 2015 UK Time-Use survey . . . 145
4.3.2 Variable memory length models . . . 146
4.3.3 Algorithm . . . 149
4.3.4 Implementation . . . 154
4.4 Validation analysis results . . . 158
4.4.1 Distribution of relevant past states . . . 159
4.4.2 State probabilities profiles . . . 162
4.4.3 State durations distributions . . . 169
4.4.4 Transition Occurrence Frequency error profiles . . . 171
4.5 Discussion . . . 174
4.5.1 Potential for improvement . . . 179
4.6 Conclusions . . . 180

5 Simulating residential electricity demand: A modelling framework for activity-driven appliance usage of socio-demographically diverse households 183
5.1 Introduction . . . 187
5.1.1 Drivers of residential electricity demand . . . 188

5.1.2 Residential electricity demand modelling 190
5.1.3 Motivation 191
5.2 Modelling Framework: Data and Methods 195
5.2.1 Stage 1: Household characterisation 196
5.2.2 Stage 2: Simulation of household activity patterns 204
5.2.3 Stage 3: Simulation of daily power consumption 206
5.2.3.1 Characterisation of appliances 207
5.2.3.2 Simulating activity-driven power usage 211
5.2.3.3 Simulating non-activity-related power usage . . . 218
5.3 Validation analysis 218
5.3.1 Simulated power consumption and average daily load profile 220
5.3.1.1 Comparison with previous approaches 223
5.3.2 Load curve metrics 225
5.3.3 Monthly demand estimates 231
5.3.4 Distribution of households with respect to total annual electricity consumption 232
5.4 Discussion 234
5.4.1 Advantages of the proposed approach 235
5.4.1.1 Characterisation of simulated households 235
5.4.1.2 Approach to the simulation of household activity 236
5.4.1.3 Approach to the simulation of appliance-level power consumption 237
5.4.2 Potential for improvement 238
5.5 Conclusions 240

6 Conclusions **243**
6.1 Introduction 245
6.2 Summary of findings 246
6.2.1 Chapter 2: Characterising the diversity of energy requirements 247
6.2.1.1 Rationale and aim 247
6.2.1.2 Findings and discussion 248
6.2.2 Chapter 3: Characterising activity and appliance usage patterns 249
6.2.2.1 Rationale and aim 249
6.2.2.2 Findings and discussion 250

6.2.3 Chapter 4: Improving the approach to residential activity modelling . . . 253
6.2.3.1 Rationale and aim . . . 253
6.2.3.2 Findings and discussion . . . 254
6.2.4 Chapter 5: Improving the approach to appliance usage modelling . . . 256
6.2.4.1 Rationale and aim . . . 256
6.2.4.2 Findings and discussion . . . 258
6.3 Future outlook . . . 260
6.3.1 Contributions of this thesis . . . 261
6.3.1.1 Relevant insights for Demand-Side Management . 263
6.3.1.2 Final overview of core chapters . . . 267
6.3.2 Outlook for future research . . . 268

Appendices

A Supplementary information – Paper 1 [Chapter 2] 275
A.1 Empirical probability density estimation . . . 275
A.2 Trimodal Gaussian Mixture model . . . 277
A.3 Limitations of Household Electricity Survey data . . . 280

B Supplementary information – Paper 2 [Chapter 3] 283
B.1 Misrepresentation of appliance usage . . . 283
B.2 Activity and appliance usage data collection . . . 284
B.3 Further limitations of the comparative analysis . . . 286
B.4 Individual appliance usage patterns and theories of social practice 289

C Supplementary information – Paper 3 [Chapter 4] 291
C.1 Transition Occurrence Frequency Error profiles . . . 291
C.2 Activity sequences and social science . . . 294

D Supplementary information – Paper 4 [Chapter 5] 297
D.1 Household characterisation variables . . . 297
D.2 Monthly consumption estimates . . . 300

References **303**

List of Figures

1.1 Graphic illustration of the conventional model for power systems. Source: (UK National Infrastructure Commission 2016) 6
1.2 Distribution of the generation plants in the UK by fuel source (as of 2015). The size of the pointers is indicative of the generation capacity of each plant. Source: Interactive Map of UK power sources - Carbon Brief (Pearce 2019). 8
1.3 Graphic representation of the model for the power systems of the future. Source: (UK National Infrastructure Commission 2016) . . 10
1.4 Structure of the thesis. 43

2.1 Annual electricity demands by household, ranked by magnitude: A) Annual demands of both metered and simulated households in CREST22 dataset (adapted from (Richardson et al. 2010, Fig. 6)). B) Annual demands of both metered and simulated households in HES24 dataset. C) Annual demands of both metered and simulated households in HES243 dataset. Households were grouped in sets of five. The levels observed correspond to the average consumption of each set. 72
2.2 Distribution of total annual consumptions: A) Probability Density Functions of metered and simulated HES243 datasets. B) Empirical Cumulative Distribution of metered and simulated data. The dotted lines around the ECDF of the metered HES243 dataset are the 95% confidence bounds. 74

2.3 Total annual consumptions distributions: A) Probability Density Function of metered HES243 dataset and trimodal Gaussian mixture fit. Each vertical bar at the bottom of the plot represents a data point. B) Empirical Cumulative Distributions of metered annual consumptions and trimodal Gaussian mixture fit. The dotted lines around the ECDF of the metered HES243 dataset are the 95% confidence bounds. 77
2.4 Distribution of households according to size and annual consumption range: A) EPDFs corresponding to households grouped by size. The vertical line in each distribution indicates the mean value. B) Percentage composition of household sizes for each cluster. 78
2.5 Annual consumptions ECDFs of individual appliances: A) Main cooking appliances: cooker, or equivalently, hob+oven. Present in 70% of households. B) ECDFs of simulated and monitored kettles. Present in 94% of households. C) ECDFs of simulated and monitored washing machines. Present in 83% of households. D) ECDFs of simulated and monitored main TVs. Present in 98% of households. 80
2.6 Mean daily load profiles corresponding to metered and simulated HES243 datasets. Typical UK profile corresponds to the load profile of the average class 1 (unrestricted domestic tariff) customer. 82

3.1 Average daily load profiles. Adapted from (Ramírez-Mendiola et al. 2017). Based on data from (UKERC Energy Data Centre et al. 2013, Zimmermann et al. 2012). Class 1 customers refers to ELEXON's customer classification (see (UKERC Energy Data Centre et al. 2013)). 98
3.2 Characterisation of activities according to their relationship with energy use (example). 101
3.3 Daily food preparation and cooking activity profiles. **A)** General cooking appliances' usage profile. **B)** Individual cooking appliances' usage profiles. 110
3.4 Daily laundering activity profile. **A)** General laundering appliances' usage profile. **B)** Individual laundering appliances' usage profiles. 113

3.5 Daily dish washing activity and metered dishwasher usage profiles. 114
3.6 Daily TV watching and ICT related activity profiles. **A)** TV watching activity and metered overall TV usage profiles. **B)** ICT related activity and metered overall IT appliances' usage profiles. 116
3.7 Average occupancy: **A)** Weekday average occupancy profile. **B)** Weekend average occupancy profile. The different shaded areas represent, at any given 10 min interval, the overall probability that an individual is in one of three occupancy states: Active at home, sleeping or absent. 118
3.8 **A)** Average occupancy profile. **B)** Comparison between overall active occupancy profile and average daily load profiles. 120

4.1 Graphical representation of a daily activity sequence. $S1 \cdots Sn$ =states in the sequence 149
4.2 Distribution of past states with respect to their associated gain function value. Sub-figures correspond to sets of states grouped by length (ℓ = length of state). 161
4.3 Daily evolution of the distribution of past states with respect to their length (ℓ). The shading in each band indicates the proportion of states, measured from 0 to 1, as indicated by the scale atop this figure. 162
4.4 Empirical activity state probability profiles extracted from the UK 2015 time-use data. The order of the elements listed in the legend is the same as the order of the stacked state probability profiles. . 163
4.5 Root-mean-square deviation of the daily state probability estimates produced by the first-order and variable memory length models. Dotted line - Variable memory length model; Continuous line - first-order model. 165
4.6 Comparison between state probability profiles as extracted from datasets with 100 and 1.5×10^4 simulated activity sequences. Dotted line - Empirical state probability profiles; Continuous line - Model estimates. 168
4.7 Distribution of state durations. Dotted line - empirical distribution; Red line - Variable length model; Blue line - first-order model. 170

4.8 Transition Occurrence Frequency (TOF) Error profiles: Blue line - first-order model error; Red line - Variable memory length error; the shaded regions along the profiles represent the standard deviation ranges. 173

5.1 Modelling framework architecture. 196
5.2 Probability distributions of households with respect to total annual consumption. The households have been grouped by size to show the differences in variability within same-size groups. 199
5.3 Example of the daily switch-on probability profiles for the different types of day considered. Case: Microwave oven. 209
5.4 Example of the joint probability distributions of duration of cycle of operation with respect to time of day for the different types of day considered. The shading intensity indicates the probability in accordance with the scale at the bottom of this figure. Case: Microwave oven. 210
5.5 Sample appliance power consumption traces for different durations. Case: Microwave oven. 211
5.6 Top: Sample weekly metered power consumption from a 2-person household. Bottom: Sample weekly simulated power consumption from the simulated counterpart. 221
5.7 Top: Average daily load profile for metered and simulated households. Bottom: Daily profile of load variability with respect to time of day. 223
5.8 Comparison of metered and simulated average daily load profiles. The diagonal dashed line indicates the optimal warping path. The continuous red line and dotted blue line associated with the simulated profiles correspond to their respective best warping path. 225
5.9 Top: Change in Load Factor with respect to level of aggregation. Bottom: Hourly Coincidence factor profiles for metered and simulated loads. 229
5.10 Average load duration curves for a 1-year demand period. 230
5.11 Seasonal trend of the average total monthly demand for electricity. 231
5.12 Distribution of households with respect to total annual consumption. 233

A.1 Comparison of Density Estimation methods. A sample of $N = 1000$ observations drawn from a Standard Normal Distribution is used to test the effectiveness of KDE and k-NN density estimation methods. In the KDE estimator, a Gaussian kernel is used with a bandwidth chosen to minimise the mean squared error. In the k-NN estimator, the conventional value $k = N^{1/2}$ is used. 277

C.1 Transition Occurrence Frequency (TOF) Error profiles (Figure 4.8 in Chapter 4). Blue line - first-order model error; Red line - Variable memory length error; the shaded regions along the profiles represent the standard deviation ranges. 293

D.1 Comparison of the seasonal trend of the average total monthly demand for electricity between HES sample households and the national figures for the years 2009-2012. 301

List of Tables

1.1 Methods employed in the core chapters. 44

2.1 Categories of appliances and the appliances contained in each one, according to the appliances present in both metered and simulated datasets. 67
2.2 Mean values and standard deviations of annual electricity consumptions datasets. 71
2.3 Trimodal Gaussian mixture model's parameters. 76
2.4 Summary of parameters of the Weibull distributions fitted to the annual consumptions ECDFs of individual appliances. 81

3.1 Activities and list of monitored appliances associated with them. . 103

4.1 Model dimension stratification as a function of the order. 141
4.2 Proportion of states by gain range by length group and maximum gain value. 160
4.3 Relative change in RMS deviation by activity by dataset size. . . 166
4.4 Relative difference in RMS deviation of distributions of state durations by activity. 170

5.1 Breakdown of households by composition type. 201
5.2 Breakdown of family types by number of children. 201
5.3 Set of appliance ownership statistics for one-person households, for people aged less than 65 years. 203
5.4 Percentage breakdown of households by number of TV sets owned. 203
5.5 Activities and associated appliance types. 213

A.1 Ranking of models based on their Akaike Information Criterion (AIC) score. 279

D.1 Review of predictor variables and modelling methods employed in 10 models recently developed. 299

List of Acronyms and Abbreviations

AC	Alternating Current
BEIS	UK Department for Business Energy and Industrial Strategy
CCC	UK Committee on Climate Change
CDF	Cumulative Distribution Function
CHP	Combined Heat and Power
CO_2	Carbon dioxide
CREST	Centre for Renewable Energy Systems Technology
DC	Direct Current
DECC	UK Department of Energy and Climate Change
Defra	UK Department of Environment, Food and Rural Affairs
DSM	Demand-side Management
DSR	Demand-side Response
DTW_d	Dynamic Time Warping distance
ECDF	Empirical Cumulative Distribution Function
EPDF	Empirical Probability Density Function
EU	European Union
GHG	Greenhouse Gases
GW	Gigawatt

HES	UK's Household Electricity Survey
HETUS	Harmonised European Time-use Survey
ICT	Information and Communication Technologies
IEA	International Energy Agency
IPCC	Intergovernmental Panel on Climate Change
KDE	Kernel Density Estimation
kW	Kilowatt
kWh	Kilowatt-hour
MW	Megawatt
MWh	Megawatt-hour
NILM	Non-intrusive load monitoring
Ofgem	Office of Gas and Electricity Markets
ONS	UK Office for National Statistics
PDF	Probability Density Function
PV	Photo-voltaic
RMSD	Root-mean Square Deviation
TUS	Time-use survey
UK	United Kingdom
UKERC	UK Energy Research Centre

Key definitions

Active occupancy: describes the time periods during which one or more members of a given household are found awake and active within their home, and therefore, likely to engage in energy-related activities.

Coincidence factor: is defined as the ratio of the maximum of an aggregate load registered during a given period of time to the sum of the maxima of the individual loads aggregated during the same period. The magnitude of this factor gives an idea of how likely it is that the maximum power consumption of individual users/devices would occur at the same time.

Cumulative Density Function: is a function whose value is the probability that a corresponding continuous random variable has a value less than or equal to the argument of the function.

Decentralised generation: also referred to as distributed or on-site generation, is the activity of power generation by small-scale power generation units. When connected to the main grid, these are typically integrated at the distribution network level. Distributed generation systems typically use renewable sources such as small hydro, bio-fuels, solar PV, wind and geothermal power, and play an increasingly important role in power distribution systems.

Demand load profile: is a term used in electrical engineering to refer to a graph that shows the variation in the electrical load in a system with respect to time. Load profiles vary across customer types (e.g. residential, commercial, industrial), and seasons of the year. The information contained in load profiles is used by power providers to plan how much electricity they will have to supply at any given time.

Demand load shifting: is the activity of shifting energy demand in order to match it with supply and to assist in the integration of variable supply resources.

Demand-side management: refers to a group of initiatives or actions designed to efficiently manage consumer demand for energy. A number of methods such as financial incentives and behavioural change through education can be used to encourage customers to optimise their energy use.

Demand-side response: can be broadly defined as the set of mechanisms by which demand loads can be shifted in time in response to certain control signals and with varying degrees of end-user engagement. This sits alongside Demand Reduction under the umbrella term Demand-side management.

Dispatchable power plant: refers to the power generation assets that can be brought on line, or ramped up or down at will, in order to meet the demand for electric power.

Distribution: refers to the transportation of electric energy and power by means of medium to low voltage lines.

Dynamic Time Warping distance: is a time-series analysis metric commonly used to measure the similarity between two temporal sequences sampled at different speeds or resolutions.

Gaussian Mixture Model: is a probabilistic model for representing the presence of normally (Gaussian) distributed sub-populations within an overall population. These models do not require previous knowledge about the sub-populations present, as the model learns the sub-populations automatically. This type of models constitute a form of unsupervised learning.

Generation assets: are the sites, buildings and plant used for the purposes of power generation.

Household: is defined as one person living alone or a group of people (not necessarily related) living at the same address and sharing living spaces and common facilities of a housing unit.

Household appliances: are devices, machines or pieces of equipment, especially electric ones, that are used in the house such as cooker or washing machines.

Intermittent power source: refers to any source of energy that is not continuously available for the generation of electric power. This type of sources may be predictable but they are outside of direct control and cannot be dispatched to meet the demand loads in a power system.

Load factor: is defined as the ratio of the average load during a given period of time to the maximum load registered during the same period. In electrical engineering the load factor is a measure of the efficiency of electric energy usage. The higher the load factor, the smoother the load profile observed, and the more efficiently the generation capacity is being used.

Load-duration curve: is a graph that illustrates the amount of time during which the electricity demand is greater than or equal to a given load. Load-duration curves are also used to illustrate the relationship between generation capacity requirements and capacity utilisation.

Markov process: is a stochastic process that satisfies the Markov property. This property states that previously observed states in the process are not relevant to the evolution of the process. That is, predictions of subsequent event occurrence are made based on the current state only. For this reason, this kind of processes are characterised by what is referred to in the statistical literature as memorylessness.

Operating reserve: is surplus generating capacity that can be made available to the power system operators to instantly respond to a sudden increase in the demand loads or sudden decreases in intermittent renewable power output.

Power generation: refers to the activities associated with the generation of electricity in a power plant.

Power system: refers to the network of components deployed to supply, transport, distribute and use electric power. The system can be broadly divided into generation, transmission and distribution systems. Unless otherwise stated, the term is used to refer to national-scale systems.

Power system flexibility: refers to the extent to which the system can modify electricity production or consumption in response to variability, expected or otherwise.

Python: is a high-level programming language developed with a design philosophy that emphasises code readability.

Residential sector: refers to the group of users that dwell in either owned or rented housing units.

Transmission: refers to the transportation of electrical energy and power by means of high- and very-high-voltage lines.

1 | Introduction

1.1 Background and Rationale

The debate about climate change no longer focuses on whether anthropogenic emissions of greenhouse gases are part of the problem. It has long been established that they are one of the main drivers of the observed climate change trends (IPCC 2007), and the focus has now shifted to the actions that need to be taken to stabilise global warming to 1.5°C and the potential implications of the corresponding mitigation measures (IPCC 2018).

Delivering any acceptable target for long-term warming requires a complete and rapid reworking of our energy systems as they are the largest single contributor to anthropogenic greenhouse gases (GHG) emissions (Eyre et al. 2018, IPCC 2014). In particular, achieving a near-complete decarbonisation of the power sector is critical, and this needs to happen sooner rather than later (Eyre et al. 2018, Pfeiffer et al. 2016). On a global scale, power generation presently accounts for approximately 25% of the total anthropogenic GHG emissions (IPCC 2014). In the UK, this sector alone currently contributes approximately 30% of the total carbon dioxide (CO_2) emissions (Alderson et al. 2012, Hammond & Pearson 2013)

There is, however, huge potential for substantial decarbonisation in this sector. Therefore, a global transition to low-carbon power generation systems is anticipated and necessary to reduce climate change impacts (Pfeiffer et al. 2016, Solomon & Krishna 2011). To this end, a number of countries have committed to taking substantive measures to reduce their emissions. The UK was one of the first countries to set legally binding agreements to curtail GHG emissions. The 2008 UK Climate Change Act (Parliament of the United Kingdom 2008)

set out a target for an 80% reduction of Kyoto Protocol GHG emissions relative to the 1990 baseline.

The first comprehensive plan for a transition towards a low-carbon economy in the UK was set out in the 2009 Low Carbon Transition Plan (DECC 2009). The emission reduction targets and the strategies for meeting them have been further detailed in the 2011 UK Carbon Plan (DECC 2011). This report emphasizes the need for the transition towards low-carbon energy systems as, according to the strategies for meeting the overall emission reduction targets, it is expected that the net emissions from this sector will have to be close to zero by 2050 (DECC 2011).

In the UK, the power sector has historically been the largest contributor to GHG emissions. In the year of the publication of the Carbon Plan (2011) this sector was responsible for about 27% of total emissions by source (DECC 2011). According to the last reported figures, as of 2016 the sector accounts for over a quarter (25.7%) of total emissions (BEIS 2018*a*). Between 1990 and 2016 an overall reduction in greenhouse gas emissions of 57% has been observed (BEIS 2018*a*).

The main source of emissions from this sector is the combustion of natural gas and coal for electricity generation. The observed decrease in emissions from this sector has resulted mainly from an increased uptake of renewables and changes in the mix of fuels used for electricity generation (BEIS 2018*a*). The transition towards low-carbon energy systems will entail even more substantial changes to the electricity generating mix. There are still major uncertainties over the most cost effective mix of technologies and the pace of the transition (DECC 2011, Strbac et al. 2016). The overall direction seems to be quite clear, however, but it

entails making a radical move away from the current model for power generation and supply (Chmutina & Goodier 2014, DECC 2009, 2011, Eyre et al. 2018, Strbac et al. 2016).

1.1.1 Conventional model for power generation

The origins of the conventional model for power systems can be traced back to the last decade of the 19th century; this model has remained largely unchanged since (Altmann et al. 2010, Hammond et al. 2013).

After the end of the "war of the currents", power systems across the world have been built with a view to accommodate large AC thermal power plants. The motivations for this have been primarily economic; larger plants can reach higher energy conversion efficiencies, and take advantage of economies of scale while making use of the cheapest fuels available (Alderson et al. 2012, Hammond et al. 2013, Hughes 1983). As a result, power systems based on the conventional model are systems where generation is highly centralised and largely based on the use of fossil fuels.

In most developed countries – the UK included – power systems are characterised by a 'lock-in' to centralised power generation, and carbon-intensive forms of generation account for large proportions of the total generation capacity (Bergman & Eyre 2011, Chmutina & Goodier 2014, Unruh 2000). Over two thirds of the electricity consumed in the UK is still generated by large – fossil fuel-fired and nuclear – thermal power plants (BEIS 2018*b*). As of 2017, 40.4% of the electricity was generated by burning natural gas and 9.0% burning coal. A very small amount is produced from other fuels (2.9%) (BEIS 2018*b*). The

volume of electricity generated by coal and gas-fired power stations changes each year, as some stations switch between the two depending on fuel prices (BEIS 2018*b*).

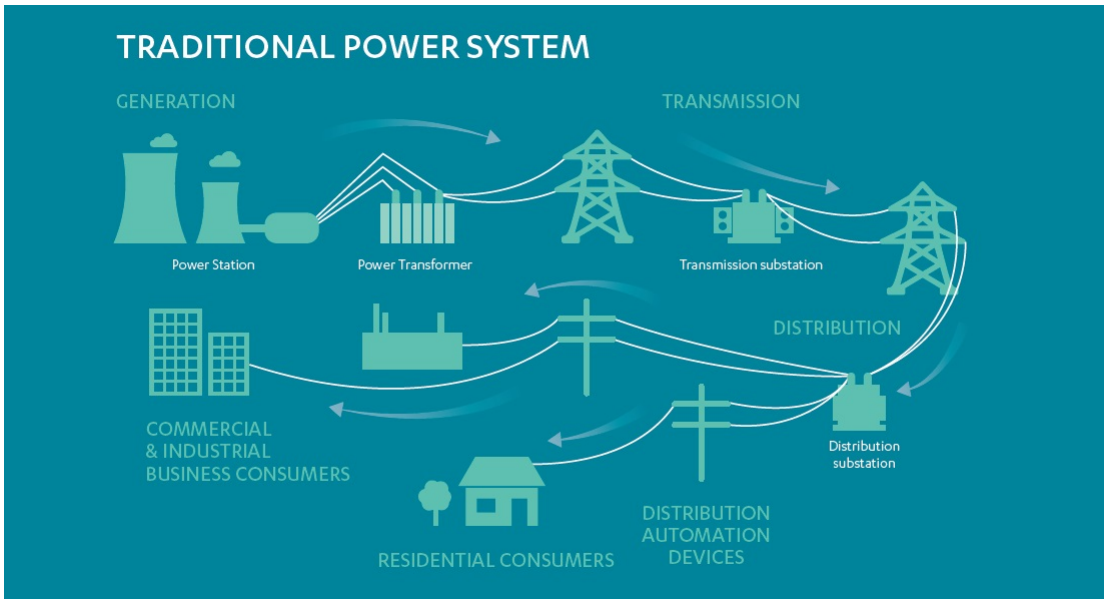


Figure 1.1: *Graphic illustration of the conventional model for power systems. Source: (UK National Infrastructure Commission 2016)*

The centralised model of power generation has certainly offered a number of economic benefits.However, since carbon-intensive forms of power generation account for large shares of the generation capacity, it also has significant drawbacks in terms of the environmental impact of power generation.

As the generated power flows from source to end-user, energy is lost at each stage of the generation and distribution processes. When fossil fuel-fired generation is one of the main forms of power generation, this translates into excess GHG emissions.

In systems predominantly based on thermal power generation, overall energy system losses amount to approximately 65% of the primary energy input

(Allen, Hammond & McManus 2008, Hammond et al. 2013); energy conversion losses associated with burning fuels for power generation are about 58% (Allen, Hammond & McManus 2008); the losses associated with the transmission and distribution of the power generated amount to approximately 1.5% and 5%, respectively (Hammond et al. 2013).

The increasing pressure to meet the carbon reduction targets has driven researchers and policy-makers to devise new pathways for the systems that will power our future in a secure, efficient, cost-effective and environmentally friendly way.

1.1.2 The power system of a low-carbon economy

The imperative to move away from fossil fuel-based power generation has spurred the integration of increasing shares of cleaner power generation technologies into the systems (Anderson & Torriti 2018, Ringkjøb et al. 2018).

Moreover, the need to further reduce overall GHG emissions while ensuring a cost-effective power generation model is also driving a transition towards systems based on a configuration that allows for a more efficient use of the energy resources available close to where energy is needed (Altmann et al. 2010, Chmutina & Goodier 2014, Rezaee Jordehi 2016).

Thus, the old power systems where the generation of electricity was dependant on a few centralised power stations are being gradually replaced by systems where thousands of smaller plants distributed all over the country are providing the required energy (Altmann et al. 2010). In the case of the UK, this trend can be better observed overleaf in Figure 1.2.

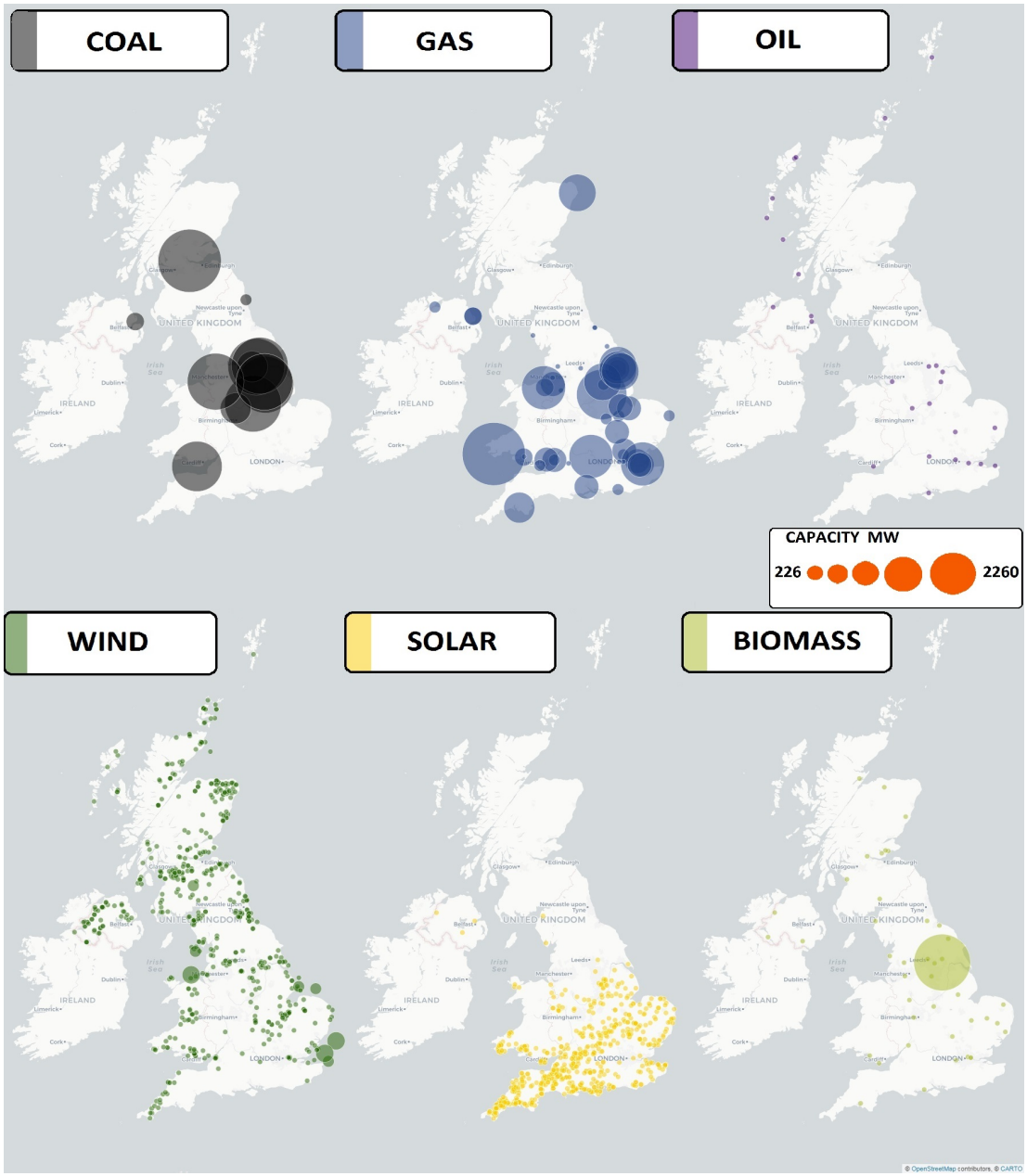


Figure 1.2: *Distribution of the generation plants in the UK by fuel source (as of 2015). The size of the pointers is indicative of the generation capacity of each plant. Source: Interactive Map of UK power sources - Carbon Brief (Pearce 2019).*

In addition to the technical factors such as the potential for energy efficiency increases, other factors such as the rising demand for electricity and increasing concerns over security of supply also play an important role in encouraging the

decentralisation of the power systems (Altmann et al. 2010, Chmutina & Goodier 2014).

Thus, over the last decade, ensuring a secure, low-carbon supply of electricity drawing on locally available renewable energy resources has become a key policy objective of a number of countries, including the member states of the European Union (EU) (Altmann et al. 2010, Welsch et al. 2014).

The large-scale integration of renewable power generation technologies across EU member states has been largely driven by the commitment to comply with the European Parliament directive on the promotion of the use of renewable energy sources (Altmann et al. 2010, European Parliament & Council of the European Union 2009). In the case of the UK in particular, a number of policies introduced in recent years are likely to result in an increased drive towards decentralised energy systems (Chmutina & Goodier 2014). However, decentralised generation is yet to play a significant role in the UK power system (Bergman & Eyre 2011).

It is still not clear what combination of generation technologies and balancing mechanisms would yield the best results. Moreover, there is likely to be a wide gap between the theoretical estimates of the technical potential of the different configurations and what is actually technically and economically feasible. However, it is generally envisioned that the model of a power system based on centralised power generation plants sending electricity to the end-users via a one-way grid (Figure 1.1) will be replaced by a model where the power system is connected by a two-way grid and where resources such as physical storage, active demand management and the output of hundreds (if not thousands) of small power generation units will be balanced in real time with the aid of complex intercommunication links (Figure 1.3).

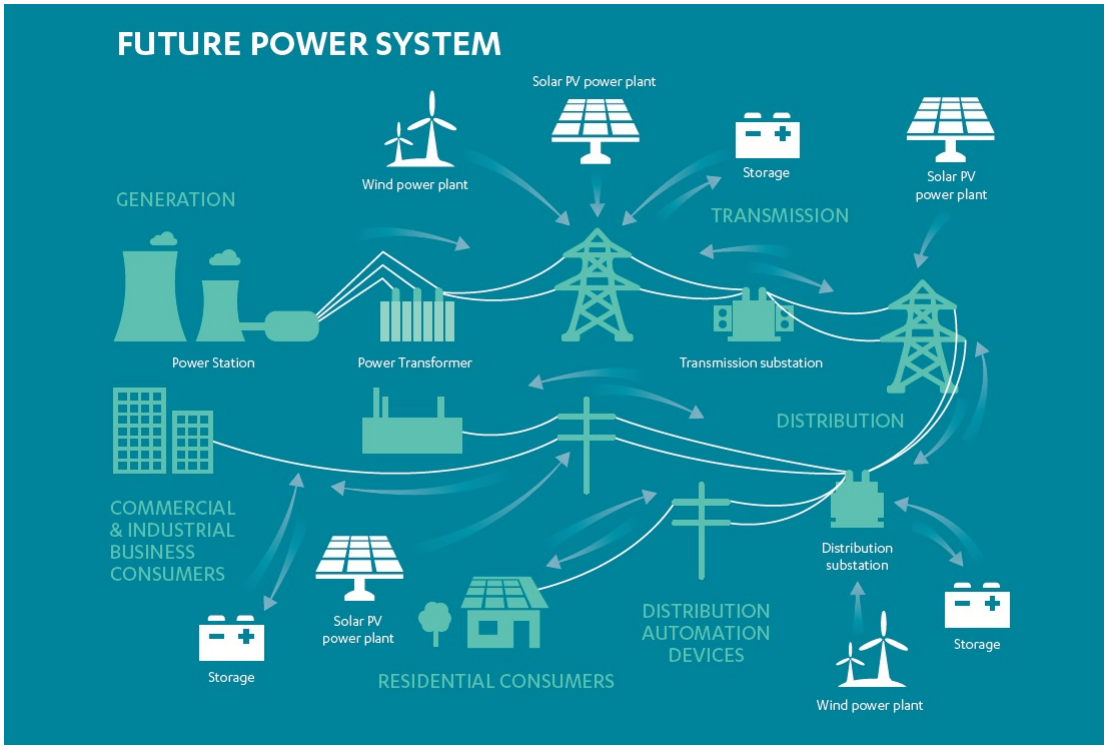


Figure 1.3: *Graphic representation of the model for the power systems of the future. Source: (UK National Infrastructure Commission 2016)*

1.1.3 Renewable power generation

The term 'renewable energy sources', or 'renewables' for short, is used to identify those primary energy resources which are naturally replenished on human timescales. With the exception of tidal and geothermal energy, renewable energy sources are ultimately solar-derived.

Renewables such as solar photo-voltaic (PV) and wind power fall within another category commonly discussed in the literature: intermittent or variable renewables. These forms of power generation depend upon resources that fluctuate over different time scales: throughout the day, day-to-day and season-to-season.

Since the GHG emissions associated with the production of power through the use of these technologies is virtually zero, these technologies are also classified as clean forms of power generation. While renewable energy sources alleviate some of the concerns about GHG emissions, their increased integration into power systems poses new complex challenges. There are other sources of energy that are classified as renewables by some of the academic and policy literature (e.g. geothermal and hydro-power). However, these are typically left out of the debate on the implications of the integration of renewables as they do not present any greater challenges than conventional power generation technologies in terms of their integration into the systems (IEA 2011).

1.1.4 Decentralised power systems

The definition, or what is meant by the term 'decentralised generation' often varies depending on the context (Allen, Hammond & McManus 2008, Altmann et al. 2010, Chmutina & Goodier 2014). In the broader sense, decentralised or distributed energy refers to the generation of energy close to the point of use (Allen, Hammond & McManus 2008).

In the context of EU legislation, decentralised generation is defined as plants connected to the distribution system (European Parliament 2009), where the term 'distribution system' refers to the medium- and low-voltage network (110kV or lower), which excludes the transmission system lines that operate at high- and extra high-voltages (Altmann et al. 2010).

An example of a somewhat more specific definition can be found in the 2004 UK Energy Act (Parliament of the United Kingdom 2004) where distributed

generation is defined as small-scale electricity generation technologies installations which are renewable or low-carbon, and have a generation capacity of up to 50 kW. In the context of UK legislation, these are also sometimes referred to as micro-generation. The term micro-generation, initially coined in the UK, has been used to refer to very small power generation installations based on solar PV cells, solar thermal heating, small-scale wind power, biomass-fuelled boilers and micro combined heat and power (CHP) units (Bergman & Eyre 2011).

1.1.4.1 Advantages of decentralised systems

There are a number of advantages to decentralised power systems that suggest they might play an important role in the transition to low-carbon economies (Allen, Hammond & McManus 2008). Studies emphasising the potential benefits of systems based on a configuration with more distributed power generation assets, which are able to make better use of local energy resources, have been around for over two decades (Walker et al. 2007). More recent studies on the potential gains derived from adopting a model of decentralised power generation frequently claim that systems based on this model would offer greater resilience, reliability and efficiency. In addition, they would also appear to be more cost-effective and accessible alternatives with potential for offering enhanced security of supply (Chmutina & Goodier 2014, Coaffee 2008). Generating power locally can offset some of the cost of transmission and distribution system upgrades, and the additional generation resources can contribute to grid reliability.

1.1.4.2 Drawbacks of decentralised systems

Decentralised systems, however, also have a number of drawbacks, and there are a number of barriers to the transition towards this kind of systems. In order to reap the full benefits of such systems, these barriers will have to be overcome.

The barriers to the transition towards decentralised power systems can be broadly categorised as technical, economic and information-related (Allen, Hammond & McManus 2008). The research presented in this thesis aims to help overcoming some of the technical barriers. Thus, the discussion presented in this section will focus on these.

From the technical perspective, one of the main barriers to the integration of generation resources at distribution network level is the complexity of the interconnection process (Allen, Hammond & McManus 2008, Altmann et al. 2010, Howard et al. 2014). The interconnection of distributed generation resources requires an assessment of the potential effects on the existing grid as the addition of generation assets may risk the existing electric infrastructure. While the installation of distributed generation resources might seem technically feasible there might be some limitations such as differences in fault current limits at different points in the grid.

From a network management standpoint, increasing the amount of distributed generation resources is often considered with reluctance as it brings the complexity of transmission networks to distribution network level (Allen, Hammond & McManus 2008, Altmann et al. 2010). Some of the complications are caused by current methods for managing the distribution networks, but the

features of different active resource components themselves also play a role, as they might not be sufficiently developed to enable easy interconnection.

The complications associated with the interconnection of distributed generation are due to a large extent to the *status quo* of power networks and distribution network managing methods and practices, and to a lesser extent, to the characteristics of the generation assets themselves. Existing power networks have been optimised for a one-way flow of power originating from centralised power stations. Therefore, major changes to the power networks will be necessary if distributed generation assets are to play a significant role (Allen, Hammond & McManus 2008, Hammond et al. 2013). Nevertheless, the transition towards decentralised power systems is feasible. It has been estimated that by 2050 between 30 and 40 per cent of the electricity consumed in the UK could be produced by distributed power generation units (Allen, Hammond, Harajli, Jones, Manus & Winnett 2008, Energy Saving Trust et al. 2005, Hammond et al. 2013).

Increasingly large shares of distributed generation would come in hand with proportionally large shares of intermittent renewables (Altmann et al. 2010). The variable nature of these generation technologies entails additional efforts in terms of their integration into existing power systems (IEA 2011).

1.1.5 The demand-supply balancing challenge

The ability to balance a fluctuating demand load has been a fundamental feature of all power systems. In this sense, the need to deal with variability and uncertainty in the system is not a new challenge as this has long been taken into account.

The operators of current power systems have considerable experience of responding to the variability observed in highly aggregated demand loads. However, the inherent steadiness of power systems based on a centralised configuration has made this task relatively more simple compared to the challenges that would be posed by systems with large shares of decentralised and variable renewable power generation units.

The stability and reliability of current power systems has been a rather fortuitous consequence of centralised power generation (Grünewald & Diakonova 2018). The large synchronous generators that have historically dominated the generation of power add a very valuable property to the systems. This property is the so-called 'system inertia' and, quite literally, refers to the mechanical inertia possessed by the massive spinning turbines that drive the large centralised power generators (Tielens & Van Hertem 2016, Vogler-Finck & Früh 2015). These generators produce an alternating current (AC) with a frequency of oscillation proportional to the frequency with which the spinning turbine of the generator rotates. The voltage and frequency of this alternating current need to be kept relatively constant for the power system to operate adequately.

Demand-supply imbalances manifest themselves as changes in the frequency of the current flowing through the power system. However, a change in the frequency of the current can only occur if the rotational frequency of the generator has changed. The larger the amount of inertia present in the system, the harder it is to change the frequency of the currents and therefore, the better the system is able to handle demand fluctuations. Thus, in centralised power systems, the enormous amounts of inertia contained in the spinning generators provide the immediate response to minor demand-supply imbalances (Tielens & Van Hertem 2016, Vogler-Finck & Früh 2015). The re-stabilisation mechanism provided by

the system inertia is a rather passive response to mismatches between supply and demand. That is, no active response was required.

When more abrupt and significant changes to the demand load are experienced and quick active responses are required, operators call upon the most flexible resources. These are commonly referred to as dispatchable resources as they can be dispatched upon request of the grid operators. Dispatchable generators have been traditionally open-cycle gas plants or, whenever available, hydro plants (Altmann et al. 2010, IEA 2011). The fast response that this kind of plants offer adds substantially to the power system flexibility[1] and contributes to the overall system stability.

From the power system's perspective, the behaviour of distributed renewable generation units is substantially different to that of centralised generation facilities (Bevrani & Raisch 2017, Ringkjøb et al. 2018, Tielens & Van Hertem 2016). Thus, the integration of increased volumes of variable power generation units generally entails a reduction of the system's flexibility needed to offer a reliable and secure power supply (Bevrani & Raisch 2017, Ringkjøb et al. 2018, Strbac et al. 2012, Tielens & Van Hertem 2016).

In addition to the issues associated with the intermittent nature of their output, variable renewable sources do not contribute (on their own) to the system inertia (Bevrani & Raisch 2017, Ringkjøb et al. 2018, Tielens & Van Hertem 2016). Moreover, the generation output of these sources cannot be fully controlled by network operators. That is, they are not dispatchable. Therefore, as the number of fossil fuel-fired generators displaced by intermittent renewables

[1] In the context of power systems, the term 'flexibility' refers to the extent to which the system can modify electricity production or consumption in response to variability, expected or otherwise(IEA 2011, Welsch et al. 2014).

increases, the systems' ability to provide flexibility (based on the traditional approaches) decreases.

From the discussion of the traditional approaches to demand-supply balancing it becomes evident that flexibility provision has historically been a responsibility of the supply side. Thus, it is also evident that a successful transition to a low-carbon energy system will depend not only on the decarbonisation of power generation through an increased uptake of intermittent renewables, but also – and to a great extent – on our ability to find alternative ways of providing the flexibility needed to maintain the system balance (Grünewald & Diakonova 2018, Ringkjøb et al. 2018, Torriti et al. 2015).

1.1.6 Demand-Side Response mechanisms

The loss of system flexibility associated with the increased uptake of intermittent renewables means that alternative forms of operating reserve will have to be available to the system (Welsch et al. 2014).

There are four key mechanisms that can enhance the overall efficiency and balancing capabilities of future low-carbon systems: Installation of highly-efficient and more flexible conventional generation plant as back-up capacity; Reinforcement of transmission and distribution networks; Deployment of energy storage technologies, and; Demand-Side Response mechanisms.

Over the last few years, the increasing recognition of the importance of energy consumption behaviours along with the declining costs of enabling technologies have brought the development of Demand-Side Response (DSR) mechanisms

closer to the focus of flexibility enhancement efforts (Hammond & Pearson 2013, Torriti et al. 2010).

If effectively implemented, DSR mechanisms could help in the provision of primary and secondary frequency response, short-term operating reserve, network congestion management services, and overall security of supply (Ofgem 2013, Strbac et al. 2016). For this reason, several studies consider DSR as an essential component of future low-carbon systems (Aghaei & Alizadeh 2013, Altmann et al. 2010, Bradley et al. 2013, Grünewald & Torriti 2013, McKenna et al. 2017, Meyabadi & Deihimi 2017, Ringkjøb et al. 2018, Torriti et al. 2015, 2010).

DSR can be broadly defined as the set of mechanisms by which demand loads can be shifted in time or curtailed completely in response to certain control signals. In the more general sense, effective DSR mechanisms enhance the systems' ability to absorb increasing shares of variable renewable generation (Bradley et al. 2013, Finn & Fitzpatrick 2014, Moghaddam et al. 2011, Moura & de Almeida 2010, Strbac et al. 2012, Zheng et al. 2014), which is a key advantage in the context of current long-term decarbonisation ambitions. In the coming decades leading up to 2050, the EU intends to cut on GHG emissions by 80-95% below 1990 levels (European Commision 2012, Welsch et al. 2014). Therefore, the value of the system flexibility enhancements resulting from effective DSR will increase even further (Bradley et al. 2013, Strbac et al. 2012, Welsch et al. 2014).

DSR offers both technical and economic value to future power systems. From the technical point of view, the effective implementation of DSR mechanisms would allow for a quicker system response by adjusting both generation and demand if needed (Ela et al. 2010, Welsch et al. 2015). Therefore, power systems assisted by DSR programmes would become more efficient. In the context of DSR

implementation, the efficiency of a system does not refer to the thermodynamic efficy associated with energy conversion processes. Rather, it defines an enhanced ability to cope with fluctuations in overall supply capacity, primarily through the ability to shift certain demand loads to the points in time where there is surplus capacity. Thus, systems with the ability to deploy such balancing mechanisms would consequently be able to make a much more efficient use of the output of variable renewable generating assets (Bradley et al. 2013, Grünewald & Diakonova 2018, Welsch et al. 2014).

From the economic point of view, DSR mechanisms offer cost-effective alternatives to other means of dealing with an increased integration of low-carbon generation, such as additional backup capacity (based on conventional thermal generation), network reinforcements or physical storage systems (Barton et al. 2013, Grünewald & Torriti 2013). It is estimated that the gross benefits of flexibility provision by non-generation means could reach £10 billion per year (Strbac et al. 2016).

Effective DSR implementation would, in general, offset the need for additional investment (capital expenditures) required for the installation of additional backup capacity (Barton et al. 2013, Bradley et al. 2013, Strbac et al. 2012). Moreover, the kind of generating units typically used as backup capacity tend to have very low capacity factors. This means that they have very low energy efficiency and are, therefore, more expensive to run. Thus, by avoiding having to install more of these, the resulting increases in operational expenditure are avoided. At the same time, the potential for a more efficient use of the capacity already in place would result in further reductions in operational expenditures.

When comparing the advantages of DSR to those of the alternatives, the case of storage is particularly relevant as the potential of DSR can be directly compared to the potential of storage resources. From the power system's perspective, the capacity of shifting loads to different times of day has the same effect as using physical electric stores (Grünewald & Diakonova 2018, Strbac et al. 2012). Moreover, they bring about the same benefits to the system; for instance, both storage and DSR alleviate the need for network reinforcements. However, studies looking at how the potential of DSR compares to the potential of physical storage systems have consistently found DSR to be a more cost-effective alternative (Hammond & Pearson 2013, Strbac et al. 2012). The capital expenditure associated with DSR is generally lower than that associated with distributed storage resources (Strbac et al. 2012). In addition, the losses associated with re-distributing demand across time are generally lower than the losses associated with re-distributing power across space (Strbac et al. 2012).

In spite of the substantial benefits of DSR, the development and implementation of DSR policies has been slow (Hammond & Pearson 2013, Torriti 2015, Torriti et al. 2010). This is partly a consequence of the limited spread of the knowledge about the potential of effective DSR and partly the overestimation of the costs associated with the implementation of DSR programmes (Bradley et al. 2013, Hammond & Pearson 2013).

The look for solutions to the flexibility problem of low-carbon systems has largely focused on technology options. However, as it has been argued in the literature, a disproportionate emphasis on the technical aspects of the problem is in fact a barrier to a wider understanding (Hammond & Pearson 2013, Verbong et al. 2013).

A number of studies have shown that social aspects such as people's daily routines, individual preferences and the interactions within households are key to achieving energy demand reductions (Hammond & Pearson 2013, Hargreaves et al. 2010, McKenna et al. 2017, Torriti 2015, Torriti & Grünewald 2014). However, the development of DSR policies targeting the residential sector has been particularly stagnant. This can be largely attributed to the fact that studies looking in detail at how the dynamics of electricity demand patterns vary in relation to households' activities are scarce (McKenna et al. 2017, Torriti & Grünewald 2014).

It is, however, increasingly recognised that paying a closer look to the social dimensions to the flexibility problem would allow for better informed transition strategies, and thus help in preventing a premature lock-in into particular transition pathways (Hammond & Pearson 2013, Verbong et al. 2013). As a result, the interest in the effects of people's daily routines and their social context (e.g. household or communities) on the observed energy demand patterns has surged, as it has become clear that this is essential to the development of effective DSR portfolios (Hammond & Pearson 2013, McKenna et al. 2017, Torriti 2015, Verbong et al. 2013).

1.1.7 The residential sector: a key target for system efficiency measures

The residential sector is a major contributor to the final energy consumption across countries. This also extends to electricity consumption. Among the 28 EU member states, this sector accounted on average for about 30% of the national total (European Environment Agency 2017).

Efforts to reduce residential energy consumption (and the associated carbon emissions) have traditionally focused on improving thermal insulation of homes and increasing the thermal efficiency of space heating systems (Bergman & Eyre 2011, Haines et al. 2010). More recently, with the prospect of future energy systems with large shares of renewable power generation, the focus has somewhat shifted to the electrification of energy-intensive services such as heating and transport. While this is expected to deliver an overall reduction in annual energy consumption, it is also expected that this could increase annual electricity consumption and peak demand load by as much as 35% and 93%, respectively (Baruah et al. 2014).

As the transition to a low-carbon future progresses, and homes become more energy efficient and electricity supplies a much greater proportion of residential energy demands, having a clear understanding of how energy, or electricity in particular, is used in this sector will prove even more relevant as it is critical to identifying future opportunities for energy use reduction and sustainable utilisation (Harris et al. 2015).

Based on the way electricity has been traditionally produced, it has been widely assumed that an overall reduction in energy consumption would also entail an overall reduction of the carbon emissions from this sector. This remains largely true because, although the mix of power generating technologies has begun to see significant changes, the bulk of the electricity is still generated by fossil-fuelled power plants. The assumption that the carbon intensity associated with different end-uses would remain the same as current estimates has made estimating the potential carbon emission reductions associated with DSM mechanisms seem relatively straightforward (Haines et al. 2010, Harris et al. 2015). However, in future low-carbon energy systems, where it is expected that even larger shares

of the electricity flowing through the grid will be generated by intermittent renewables, the carbon intensity of particular end-uses could vary substantially from one hour to the next. Thus, as we move towards lower-carbon systems, the increasing time sensitivity means that assessing the potential for carbon emissions reductions will become increasingly difficult.

An improved understanding of how the different residential end-uses give rise to the observed electricity demand loads would allow for the identification of further opportunities for cost-effective carbon mitigation. For instance, in the case of the UK, studies have shown that reductions in appliance energy use is one of, and perhaps even the most, cost-effective methods for reducing carbon emissions (CCC 2008, Haines et al. 2010). Nevertheless, if substantial reductions are to be achieved – and sustained – it is essential that we learn more about how the demand loads associated with different end-uses behave, and more importantly, the ways in which we could control them.

1.1.8 Electricity demand modelling

Estimates of likely demand for energy are crucial for the planning and operation of energy systems (Baetens et al. 2010, Engle et al. 1992, Filik et al. 2011, Howard et al. 2014, Piwko et al. 2005, Ringkjøb et al. 2018). They are essential for the development of generator management strategies, maintenance scheduling, and even to coordinate purchases and sales of electricity across utilities.

The steadiness that has characterised power systems up to this point has made the task of producing the required estimates relatively straightforward. In the past, these estimates have often been produced based on "historical rules of

thumb" developed by skilled technicians with many years of experience (Engle et al. 1992).

There are many factors at play that determine how, when and how much energy is used. Models allow us to explore how these factors relate to each other and to the observed energy consumption patterns. For this reason, the development of models to produce accurate estimates of the demand loads in a system has become a key part of energy research.

Up until the first half of the twentieth century, few – and largely unsuccessful – attempts had been made to model the demand for electricity or energy for that matter (Gates 1942). Nowadays, however, models play a crucial role in both energy planning and decision-making, and are being continuously developed in both industry and academia (Pfenninger et al. 2018, Ringkjøb et al. 2018). With the complexity of energy systems set to increase as we move towards low-carbon economies, the role of modelling will only increase in importance.

The approaches to electricity demand modelling are numerous and varied. The nature of these is determined by both the input data and the intended purposes of the model. Perhaps the most common example of the application of energy demand models is the production of total demand forecasts at different time and space scales (Kannan & Strachan 2009). However, more complex issues call for more temporally and spatially resolved analyses.

In addition to total energy consumption estimates, load profiles also play an important role in the design, planning and control of power systems as they contain information regarding the variation of the level of power demand throughout the day. The ranges of the temporal variation of the loads experienced by the system are critical for determining system capacity

requirements and developing adequate operation plans (Baetens et al. 2010, Howard et al. 2014, Piwko et al. 2005).

Demand profiles vary across economic sectors. However, the standardisation of most industrial and commercial practices, along with typically high levels of regulation, impose some major constrains on the demand profiles associated with these sectors. Although there is considerable diversity of loads in these sectors, the constrains imposed on them have facilitated the analysis of their demand loads to a certain extent (Swan & Ugursal 2009). The residential sector, however, has a much greater diversity of users who are largely unconstrained, and the huge differences in their lifestyle choices and behavioural patterns result in equally diverse demand profiles.

1.1.8.1 Residential electricity demand modelling

Over the last couple of decades, the interest in residential electricity demand modelling has soared. This is partly due to the significant contribution from the residential sector towards overall total and peak demands for electricity. However, another reason for the increasing interest in this sector in particular is the potential for the implementation of energy saving measures targeting residential users as a way of easing the burden posed on the supply side to keep the system in balance and the pressure to ensure a secure and reliable power supply.

The total demand for electricity for a given dwelling is the aggregate demand for energy required to support all electricity end-uses, which includes the losses due to appliance and system (in)efficiencies. Since the demands associated with

the different end-uses are complex and typically inter-related, comprehensive models are needed to assess the impact of the adoption of energy efficiency measures and technologies suitable for residential applications (Swan & Ugursal 2009).

The approaches to residential electricity demand modelling have been broadly grouped into two categories: "top-down" and "bottom-up". The terminology refers to the "hierarchical position of data inputs as compared to the housing sector as a whole" (Swan & Ugursal 2009). Swan & Ugursal (2009) carried out a detailed review of the modelling techniques used up until the point of their study, and classified them accordingly.

The key difference between top-down and bottom up approaches is that models pertaining to the top-down category part from estimates of the total demand from the residential sector and attribute a certain demand to the individual elements (i.e. households). Conversely, bottom-up models typically account for the demands associated with individual end-uses and calculate the total demand of individual households or groups of households and then extrapolate the results to produce estimates for a given region or nation.

Both approaches have proved useful for different applications. Depending on the intended application, one approach might be preferred over the other. However, attempts to benefit from the strengths of both approaches have driven the development of hybrid approaches as well.

Among the strengths of the top-down approach is the fact that only aggregate data – typically widely available – is needed. In addition, this type of models tends to rely heavily on historic trends. Since drastic paradigm shifts rarely occur, a well-calibrated top-down model provides enough prediction capability to model

small deviations from the status quo (at the aggregate level) (Swan & Ugursal 2009). This heavy reliance on historical trends, however, is also a drawback as this, by definition, implies that top-down models lack the ability to model, for instance, sudden technological changes.

Bottom-up models, on the other hand, use highly-resolved data to calculate electricity consumption associated with specific end-uses. This high level of detail gives bottom-up models the ability to model new technological options, which is perhaps their greatest strength. In addition, from an energy efficiency perspective, this kind of models also prove useful in identifying areas with potential for improvement.

Another key feature of bottom-up models is that they are very well suited to study the effect of user behaviour on the observed demand patterns.

In recent years, the relevance of the role user behaviour plays in determining the observed demand has been increasingly recognised (Grandjean et al. 2012, Torriti 2014, 2017). Several attempts have been made to incorporate user behaviour into the simulation of residential electricity demand, particularly in recent years.

From these attempts, a new approach has emerged driven by the use of time-use data to underpin the simulation of household activity patterns. While the use of inputs such as time-use data are definitively a positive development in the quest for a behaviour-based residential electricity demand model, there is still considerable room for improvement in the development of the techniques used to simulate user behaviour based on this kind of data.

1.1.8.2 The case for improving residential electricity demand modelling

Residential electricity demand modelling has found applications in supporting a variety of studies on a wide range of issues such as (1) building performance analysis of low-carbon buildings (Hoes et al. 2009, Yan et al. 2015), (2) planning and performance analysis of local energy systems (Howard & Modi 2017, Yamaguchi & Shimoda 2017, Yao & Steemers 2005), (3) the impacts the integration of micro-generation (Howard et al. 2014, Shimoda et al. 2017), (4) analysis of long-term decarbonisation scenarios (Boßmann & Staffell 2015, Kannan & Strachan 2009), (5) the impacts of electric vehicles' charging and discharging on demand load profiles (Grahn et al. 2013, Kelly et al. 2015), and (6) to achieve a better prediction of the time dependences and variability of the demand loads to analyse the impacts of energy efficiency measures and demand response mechanisms (Bajada et al. 2013, Campillo et al. 2012, Fujimoto et al. 2017, Gottwalt et al. 2011, Paatero & Lund 2006). However, the increasing complexity of the energy systems, as well as the increasing need for a better account of the inter-dependencies of the various system components calls for better tools to support this kind of analyses.

Despite the multitude of modelling approaches developed up until this day, some prevailing issues have been consistently identified. These can be broadly summarised as follows:

- misrepresentation of the diversity of behaviour/energy demands across households
- low time resolution of simulated daily demand loads

- loose representation of the time variability of demand loads
- lack of support for the analysis of potential demand response mechanisms

In what remains of this section, these issues will be discussed in more detail.

1.1.8.2.1 Misrepresentation of the diversity of behaviour/energy demands across households

It is rather well known that there is considerable variation in the total demands for electricity across different types of households. However, most current residential electricity demand models are blind to such variability (Haines et al. 2010, Yamaguchi et al. 2018).

When looking at the effects of household behaviour on the observed demand for electricity, or energy in general, it is useful to make a distinction between intra- and inter-household behavioural differences. Yamaguchi et al. (2018) define intra-household variation as the variation resulting from differences in occupant behaviour within the household, whereas inter-household variation is defined as the difference in behaviour among households.

Some modelling approaches have tried to integrate the behavioural aspects that give rise to the great diversity of energy demands. However, the distinction between intra- and inter-household variations is often ignored (O'Brien et al. 2017, Yamaguchi et al. 2018). In this section, the discussion will focus on the effects of inter-household behaviour. The relevance of intra-household behavioural differences will be discussed in the following sections.

A number of studies have focused their efforts on studying these differences through the use of clustering approaches (e.g. Aerts et al. (2014)). In the context of the research presented in this thesis, clustering approaches may enable the identification of groups of users with statistically significant differences in their activity and energy consumption patterns. Some of these differences might give some indications on which groups might be more or less susceptible to certain intervention mechanisms. Thus, the use of clustering approaches would assist in the development of a user ontology which would, in turn, allow for the development of more targeted approaches that focus on a particular type of user.

However, while the identifications of clusters or similar types of users is advantageous and desirable, the representation of the heterogeneity of users within and between different groups remains important.

Actively managing energy demand effectively results in altering the diversity of loads on the system. However, artificially altering the diversity of the loads associated with different end-users might have undesirable consequences such as "payback" or "rebound" effects, which increase or pose new challenges to the system. Models of residential electricity demand that produce a good representation of the diversity of end-users offer the opportunity to explore the effects of changes in such diversity. Therefore, they allow for a more effective assessment of the effectiveness of interventions that might induce such changes, and assist in the identification of potential shortcomings that might not necessarily had been taken into account in previous efforts.

Several studies have shown that the great diversity in the socio-demographic factors that characterise households results in an equally great diversity of energy demands (Delzendeh et al. 2017, Fischer et al. 2015, Haldi & Robinson 2011,

Jones et al. 2015, López-Rodríguez et al. 2013, Matsumoto 2016, Santiago et al. 2014, Sekar et al. 2016). However, the influence of socio-demographic conditions is rarely taken into account in residential electricity demand modelling approaches.

The incentives to take the influence of socio-demographic factors into account are strong and numerous. At the high levels of aggregation that current power systems are used to deal with, the influence of individual households or particular household types is negligible (Flett & Kelly 2015). However, in the context of systems with small, highly distributed generation units, the reductions in the number of consumers whose main supply are these local resources comes with a corresponding increase in the influence of the different household behaviours (Flett & Kelly 2015).

Previous research by Huebner et al. (2015, 2016) has focused on analysing how much of the variability of residential energy demand can be explained by different categories of predictors such as building-related variables, socio-demographics and, behaviours and attitudes towards energy. Based on data collected from a representative sample of the English housing stock, Huebner et al. were able to contrast the explanatory power of such variables. The socio-demographic factors considered were: Household size and type (e.g. family with/without children), number of dependent children (if any), equivalised income range, age range of adult members of the household, dwelling tenure and length of residency, and employment status of the household. This research found that socio-demographic factors on their own account for about 21% of the variability observed in the demand for energy (Huebner et al. 2016). In spite of the evidence in favour of the relevance of the influence of socio-demographic factors, further research on the effects of different community compositions, i.e. the type of households present, is limited (Flett & Kelly 2015).

When looking at the distribution of households with respect to total demand for electricity, it has been observed that households in the upper quartile consume much more than those in the lower quartile (Haines et al. 2010, Jones & Lomas 2016). In the UK, for instance, the upper quartile accounts for nearly half (49%) of the electricity consumed by the residential sector (Jones & Lomas 2016). Therefore, accounting for such variability is necessary for producing accurate estimates of overall energy demands.

In addition, the projected trends for the change in the overall socio-demographic composition further incentivise the analysis of the effects of such factors. For instance, in the UK the proportion of people aged 65 and over is expected to increase from about 18% – as of the last census – to about 20% by 2025 and about 23% by 2035 (ONS 2012). The foreseen trends would also appear to indicate that there will be larger proportions of one-person households (ONS 2012). It is expected that the increase in one-person households will account for two thirds of the overall increase in the number of households (Haines et al. 2010). If we consider that each household is equipped with a basic set of appliances, this increase in number of households has significant implications in terms of demand for energy and consequent carbon emissions.

1.1.8.2.2 Low time resolution of simulated daily demand loads

In the history of the analysis of energy demand loads, the attention paid to the time resolution of such loads has been negligible. This is largely due to the following two reasons. Firstly, sufficiently time-resolved energy demand data has been historically scarce, or even unavailable. In addition, the computational cost of the analysis of such data and its use for model development has been typically

prohibitively large. Thus, in order to limit the computational costs associated with the simulation of demand loads, the approaches to energy demand modelling have focused on producing a stylised representation of the temporal characteristics of the demand loads.

These simplifications, however, have been proven to have significant impacts (Poncelet et al. 2016). Research has shown that the effects of analysing the demand for energy at low time resolutions are large enough to qualitatively affect the conclusions of such analyses (Pfenninger 2017, Pfenninger et al. 2018, Poncelet et al. 2016).

The technological advances over the last decades have made not only possible, but also increasingly feasible and cost-effective, both the collection of large volumes of high-resolution energy demand data and the analysis of such data for the purpose of energy demand modelling. The proliferation of smart meters is a clear proof that there are increasing opportunities for collecting larger and better datasets that could be used to underpin residential energy demand modelling. The data collected by smart meters, however, might be more difficult to gain access to due to privacy concerns. Nevertheless, the interest in studying the temporal dynamics of residential electricity consumption has sparked the development of alternative metering technologies and control devices that target the residential sector (e.g. (Sense 2019)).

Given the pressing need to address the issues associated with the transition to low-carbon energy systems, taking advantage of such advances is more of a necessity rather than an option. Therefore, the development of modelling approaches that improve the temporal representation of the simulated demand loads should be prioritised (Poncelet et al. 2016).

1.1.8.2.3 Loose representation of the time variability of demand loads

Increasingly time-resolved data analysis and demand modelling should come hand in hand with an improved representation of the time variability of the observed demand loads. Up until now, however, residential electricity demand modelling has focused on the ability to reproduce the smoothed out daily load profiles that result from the aggregation of demand loads from a large number of end-users.

Historically, this (low) level of detail has proven sufficient for the purposes of conventional network planning (Palmintier & Webster 2011, Pfenninger et al. 2014). And the approaches commonly used in widely implemented models have proven to be good enough for producing such stylised representations.

Partly due to this, the efforts to try and learn more about the intra-day variability observed at lower levels of aggregation and its potential implications in terms of performance requirements of local distribution networks have been limited. However, this lack of understanding of the variability of demand loads has been identified as a key barrier to the successful deployment of decentralised low-carbon generation and the implementation of demand response programmes (DECC 2014, Flett & Kelly 2015).

Recently it has become clear that the conventional approaches used in planning models fall short when dealing with power systems with large shares of intermittent renewable power sources (Poncelet et al. 2016).

Enabling the successful transition to low-carbon energy systems means that network operators will have to implement relevant strategies to accommodate for changes to the generation resources on the evolving networks (UK Power

Networks 2014). In essence, more suitable balancing mechanisms will need to be developed. To this end, however, detailed studies on system load profiles will need to be conducted to uncover the nature of the different system loads (Elombo et al. 2017).

The ability to model the time variability of energy demand is relevant at many, if not all, scales of energy system planning and operation (Flett & Kelly 2015). However, the ideal levels of granularity are determined by the intended applications. For the purpose of local networks and micro-grids planning, considerations of the intra-day variability of the demand loads are essential (Shimoda et al. 2017). In addition, in systems where residential users account for a substantial portion of the user base, considerations of intra-household variability are also relevant (Cao & Sirén 2016, Yamaguchi et al. 2018). Therefore, understanding the way the observed variability in the daily demand loads depends on the number and type of end-users is becoming increasingly necessary.

Among the literature on residential electricity demand modelling, it is increasingly recognised that user behaviour is one of the key drivers of the observed variability. Over the last few years a number modelling approaches that seek to incorporate the effects of user behaviour have been developed. However, the highly stochastic nature of residential user behaviour and the associated demand for electricity has made the task of developing accurate models considerably difficult.

The lack of data on residential user behaviour means that the approaches developed up until now are typically supported by simplifying assumptions about the nature of the stochasticity of the observed user activity patterns. However, it

is increasingly acknowledged that more needs to be known in order to unravel the relationship between the daily residential activity patterns and the corresponding demand for electricity.

Recent studies have shown that the time dependencies of the observed demand loads depend heavily on the residential activity patterns. New approaches to residential electricity demand modelling have emerged that attempt to look into this issue. However, the evidence provided by these studies would appear to indicate that more needs to be known about the nature of residential activity scheduling. In order to capture the minute-by-minute time variability of residential demand loads, a detailed analysis of residential activity patterns and the links with demand for electricity need to be carried out.

The knowledge provided by such analyses would allow for the development of more suitable approaches to residential electricity demand modelling that accurately capture the complexity of the relationships between residential activity patterns and electricity demand loads.

1.1.8.2.4 Lack of support for the analysis of potential DSR mechanisms

The issues that future low-carbon systems are expected to have to deal with entail that the overall system flexibility will have to increase. In this context, the term flexibility refers to the ability of the system to maintain the supply-demand balance at all times, thus ensuring an uninterrupted service provision.

Some of the measures to increase systems' flexibility concern the supply side. However, providing all the required flexibility from the supply side would require

considerable investment and in some cases might not even be economically or technically feasible. For this reasons, flexibility measures targeting the demand side are increasingly seen as a key element of future low-carbon systems.

The concept of demand response mechanisms is not new to power systems. However, it has traditionally involved sectors such as the industrial or commercial, where the demand loads are better understood and it is easier to implement.

Since the residential sector is one of the main contributors to total demand for electricity and the main contributor to peak demand, there is substantial interest to implement demand response mechanisms that might reduce the impact of this sector in particular. However, the development of demand response measures for the residential sector has been relatively slow, at least across European states (Torriti & Grünewald 2014).

One of the main reasons for such a slow progress is the fact that the relationship between people's everyday-life doings and the resulting demand for energy or electricity is not very well understood yet (Torriti & Grünewald 2014, Torriti et al. 2015).

Currently, the predominant approach to demand side response is based on the idea of using time-dependent differential pricing schemes to incentivise demand load shifting to more 'convenient' times (Anderson 2016). However, research has shown that the price elasticity of residential demand for electricity is generally too low to significantly influence the demand profiles (Anderson 2016, Fan & Hyndman 2011, Fouquet 2011, Reiss & White 2005). Therefore, it is clear that alternative, more effective approaches are needed.

Until recently, the nature of the activities that give rise to the observed demand had been largely ignored (Anderson 2016, Shove & Walker 2014).

However, it is increasingly recognised that residential electricity consumption patterns are strongly linked to the activities of the households. A result of this has been the emergence of research that aims at developing residential demand side response mechanisms based on the reconfiguration of the households' activity patterns (Strengers 2011, 2012).

Despite the differences between the approaches to residential demand response, it is clear that in order to understand the practical value and estimate the potential of demand response in the residential sector it is necessary to understand what different people do at different times of day and how that results in the observed demand loads (Anderson 2016, Torriti 2017).

Demand response programmes aiming at shifting residential energy demand loads in time assume, be it directly or indirectly, that the activities that give rise to such demand loads can be straightforwardly shifted to different times of day (Anderson 2016). However, as of yet very little is known as to what extent and for what kind of users this might actually be true (Powells et al. 2014, Walker 2014).

The need to improve our understanding of the dynamics of residential energy demand patterns has driven the development of energy demand modelling approaches that attempt to take account of the timing of certain activities (McKenna et al. 2015, Palm & Ellegård 2011, Richardson et al. 2010, Torriti 2014, Widén et al. 2009, 2012, Wilke et al. 2013). These are data-driven approaches that are commonly based on time-use data which is, to date, the best source of empirical household activity data.

The approaches to activity-based energy demand modelling developed up until now have proven effective in representing the average activity profiles

observed in empirical time-use data. However, little has been done in terms of looking at the sequencing of activities. Learning more about the sequencing of activities would allow us to better understand the temporal variation of demand that derives from the differences between household activity sequences (Anderson 2016, Palm & Ellegård 2011). This calls for the development of approaches to residential activity demand modelling that focus on understanding the household activity patterns rather than on the reproduction of average aggregate behaviour.

A better understanding of the sequencing of the households' daily activities would offer the opportunity to group households not only by their socio-demographic characteristics, but also by the type of activity patterns they typically follow. Among other things, this would provide the information necessary to focus the efforts of specific intervention programmes aiming at modifying such observed patterns (Anderson 2016, Palm & Ellegård 2011, Strengers 2012).

If demand response programmes are to provide an effective way of aiding in the supply-demand balancing challenge, any load shifting intervention targeting the residential sector will have to be informed not only by energy demand load profiles, but also by our knowledge about the flexibility of the daily activity patterns (Torriti et al. 2015).

1.2 Summary of aims and objectives

The primary objective of this thesis is to develop a modelling framework that better captures the diversity of electric energy demands in the residential sector. If models of demand for electricity are to inform the development of energy efficiency measures and aid in the planning of future low-carbon energy systems, it is essential that the differences in energy demands across different groups of the end-user base are well represented. This concerns not only the total estimates, but also the way energy demands vary in time.

1.2.1 Research questions

This thesis will seek to contribute towards improving our understanding of where residential demand for electricity originates from by exploring the following questions:

1) How does the diversity of socio-demographic characteristics map to the diversity of total energy demands?

2) How do household activity patterns relate to domestic appliances usage patterns?

3) How can the influence of habitual behaviours on daily household activity scheduling be incorporated into household activity modelling?

4) How can the simulation of user-appliance interactions and activity-to-power conversion be improved to enhance load modelling performance?

As well as identifying the shortcomings of current approaches to (residential) activity and energy demand modelling, the thesis provides general insights into the ways in which these can be improved. These findings are then used as the basis for the development of the modelling framework proposed in this thesis.

1.3 Structure of the thesis

Chapter 1 has provided a summary of the background literature and the aims and objectives of the research presented in this thesis. The review of the more specific relevant literature is presented at the beginning of each of the core research chapters. In particular, reviews of the state of the art in activity and electricity demand modelling are presented in Chapters 4 and 5, respectively.

The core of the thesis can be divided into two as follows:

Part 1: Characterisation of the residential demand for electricity

Chapter 2 (Paper 1) looks at the diversity of households in terms of their total demand for electricity.

Chapter 3 (Paper 2) analyses the relationship between household activity patterns and the usage patterns of a number of domestic appliances.

Part 2: Development of proposed modelling framework

Chapter 4 (Paper 3) presents a novel methodology for the analysis of empirical household activity data and proposes an approach to activity modelling based on the identification and use of characteristic behavioural patterns.

Chapter 5 (Paper 4) integrates the findings of the analysis that comprises the first part of this thesis with the proposed model for household activity and adds the final elements of the proposed electricity demand modelling framework.

Finally, Chapter 6 draws out the overall conclusions and discusses how the overall aims and objectives were addressed. The contributions of the findings of the individual papers to the broader literature are further discussed, and potential agendas for future research in this field outlined.

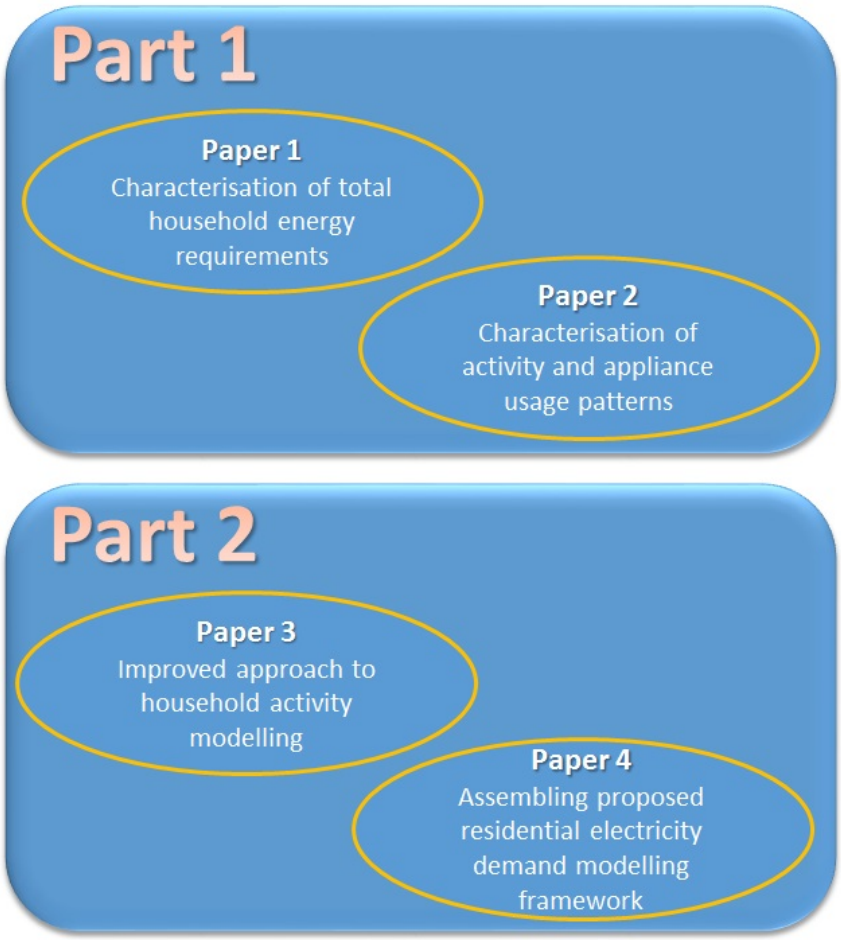


Figure 1.4: *Structure of the thesis.*

Table 1.1: *Methods employed in the core chapters.*

	Aim	Key methods
Chapter 2	Characterising the diversity of total household electricity consumption	• Statistical modelling • Parametric and non-parametric probability density estimation • Analysis of probability distribution functions
Chapter 3	Characterising activity and appliance usage patterns	• Temporal profiling of intra-day variability of activity engagement • Temporal profiling of intra-day variability of appliance usage • Qualitative comparison of associated probability profiles
Chapter 4	Development of a household activity pattern model	• Implementation of pattern identification algorithm • Stochastic modelling of daily household activity schedules • Development of new performance metrics
Chapter 5	Development of a residential electricity demand modelling framework	• Identification of appliance power consumption signatures • Characterisation of conditional probability distributions of the duration of appliance operation cycles • Agent-based modelling of user-user and user-appliance interactions

1.4 Notes on thesis format

The format of this thesis is based upon the guidelines for the article-based thesis, whereby the substantive research chapters are (a minimum of three) academic papers, provided as published in, or submitted to, peer-reviewed journals. Four first-author articles are provided (chapters 2 to 5) which together present a coherent body of research investigating the diversity of electric energy demands in the residential sector, and ways in which this can be better represented in modelling approaches.

At the time of submission of this thesis, three of these papers have been published and one submitted, as follows:

Published

- Ramírez-Mendiola, J. L., Grünewald, P., Eyre, N. (2017) The diversity of residential electricity demand – A comparative analysis of metered and simulated data. Energy and Buildings 151, 121 - 131
  DOI: 10.1016/j.enbuild.2017.06.006

- Ramírez-Mendiola, J. L., Grünewald, P., Eyre, N. (2018) Linking intra-day variations in residential electricity demand loads to consumers' activities: What's missing? Energy and Buildings 161, 63 - 71
  DOI: 10.1016/j.enbuild.2017.12.012

- Ramírez-Mendiola, J. L., Grünewald, P., Eyre, N. (2019) Residential activity pattern modelling through stochastic chains of variable memory length. Applied Energy 237, 417 - 430
  DOI: 10.1016/j.apenergy.2019.01.019

Submitted

- Ramírez-Mendiola, J. L., Grünewald, P., Eyre, N. (2019) Simulating residential electricity demand: A modelling framework for activity-driven appliance usage of socio-demographically diverse households
  Submitted to Applied Energy journal.

The papers are presented in a standardised chapter format, but the contents of each chapter are the same as those of the published or submitted papers. In order to provide a coherent format throughout the thesis, referencing to sections,

figures, tables, as well as the citations, has been standardised. The references to external sources in all chapters have been gathered into a single list at the end of the thesis.

Supplementary material and information that was considered relevant to each of the core chapters but was not included in the published articles is provided in the form of appendices.

2 | The diversity of residential electricity demand: A comparative analysis of metered and simulated data

Abstract

A comparative study between simulated residential electricity demand data and metered data from the UK Household Electricity Survey is presented. For this study, a high resolution probabilistic model was used to test whether this increasingly widely used modelling approach provides an adequate representation of the statistical characteristics of the most comprehensive dataset of metered electricity demand available in the UK. Both the empirical and simulated electricity consumption data have been analysed on an aggregated level, paying special attention to the mean daily load profiles, the distribution of households with respect to the total annual demands, and the distributions of the annual demands of particular appliances. A thorough comparison making use of both qualitative and quantitative methods was made between simulated datasets and their metered counterparts. Significant discrepancies were found in the distribution of households with respect to both overall electricity consumption and consumption of individual appliances. Parametric estimates of the distributions of metered data were obtained, and the analytic expressions for both the density function and cumulative distribution are given. These can be incorporated into new and existent modelling frameworks, as well as used as tools for further analysis.

2.1 Introduction

It is a well-known fact that the residential sector is a major contributor to overall electricity consumption in most countries. Residential electricity consumption in the UK accounts for over a third of the national total (DECC 2015). Moreover, it is also well known that the domestic sector contributes significantly to peak demand, especially during winter. Over the last decade, however, the potential of this sector to contribute towards reductions in energy consumption and CO_2 emissions has been increasingly recognised. This, in turn, has sparked the interest in this research area.

In particular, there is an increasing interest in understanding the consumption patterns of the residential sector, and how these might change in response to changes in climate and energy prices, and the implementation of new supply-demand balancing strategies and other low-carbon measures.

To this end, several efforts have been made to model electricity demand, seeking to quantify the energy requirements at different levels of analysis. Two fundamentally different approaches divide these modelling efforts: top-down and bottom-up. This terminology refers to the hierarchical level of the data inputs. Top-down models focus on describing the overall trends observed in historical records, with little or no interest in individual end-uses. In contrast, bottom-up models focus on representing the contribution of each end-use towards the aggregate consumption and overall trend. A more detailed description of these two approaches can be found in (Swan & Ugursal 2009).

Electricity demand modelling has proved very important in both academia and industry. In particular, bottom-up models' ability to incorporate many

different factors into the modelling has been key to identifying and understanding the essential elements associated with the production of demand loads. As the transition towards low-carbon energy systems progresses, robust modelling tools become even more relevant. The effectiveness of low-carbon measures such as the implementation of Demand-Side Management (DSM) strategies relies heavily on the extent to which we understand residential electricity consumption. A better understanding of the electricity consumption patterns, and the potential changes these may undergo, would allow us to devise the most suitable measures.

As new, relevant data become available, opportunities for further improvement in our models arise. The lack of information about consumer behaviour and appliance usage during the development of models such as the one used in this study (see Section 2.3) may lead to problems which could not have been revealed with the data available at the time. However, by applying these models in a different context to the one in which they were validated, these issues can now be revealed and explored. Moreover, the new sources of data currently available provide us with an opportunity to re-assess the models' performance. In particular, it is worth verifying that the simulated electricity consumption data produced by the models aggregate such that the statistical characteristics observed in the metered data are well represented, as this is key to a robust bottom-up model.

In the following section brief descriptions of the evolution of residential energy demand modelling and the state of the art are given. The next two sections concern a description of the key elements of the analysis presented, namely the probabilistic demand model used and the sub-metered electricity consumption data. The methodology for the comparative analysis, which encompasses both qualitative and quantitative methods, is discussed in Section

2.5.1. This is followed by a discussion of the analysis results, where the critical shortcomings identified are highlighted. Finally, a summary of concluding remarks based on the results previously described is given, as well as further details about potential directions of future work.

2.2 Residential electricity demand modelling

In terms of demand for electricity, the variability across households is high. In the UK, for instance, according to data from the Office of Gas and Electricity Markets (Ofgem), 'Typical Domestic Consumption Values' for electricity range from 2000 to 7200 kWh/year (Ofgem 2015). The results of the most detailed study (to date) of electricity use in English households would appear to indicate that there is an even higher variability, with values ranging from 560 to 14580 kWh/year (Palmer et al. 2013).

There are many factors associated with this variability. The variations in energy consumption associated with differences in dwelling characteristics and the impact of ambient temperature/climate conditions have been extensively studied. However, the impacts of the differences in dwellings' appliance content, and the usage patterns of such appliances have not been studied as much.

Accounting for the impact different combinations of these factors have is critical in representing the range of consumptions currently observed. Consequently, the bottom-up approach to demand modelling has been favoured for this purpose. Methodologies that follow this approach offer the advantage of modelling demand with greater level of detail. Demand estimates can be calculated at the household or even end-use level.

A branch of bottom-up modelling is characterised by a focus on user activity-based simulations of demand patterns. Models falling within this group can be further divided into deterministic and probabilistic models. Deterministic models are based on assumptions about seemingly direct causal links between chosen drivers and expected outcomes. In contrast, probabilistic models make use of stochastic methods for the simulation of electricity consumption patterns. This kind of models make predictions about different outcomes based on the stochasticity of the input data. These models have, therefore, the ability to represent the diversity of energy demand across time and populations.

The model developed by Yao & Steemers (2005) is a good example of a deterministic model. The simulation of demand loads is based on fixed, pre-determined household activity patterns characterised mainly by the periods of absence. Appliances are allocated based on national ownership statistics. Their associated demand is calculated based on average ratings and frequency of use estimates. The model is able to produce rough estimates of half-hourly daily load profiles. Another example of a deterministic model is presented in (Widén et al. 2012).

Recognising the limitations of the deterministic approach, several attempts to capture the variability observed in real households were made. An emerging approach, which continues to attract interest, is based on the use of data from Time-Use Surveys (TUS) for the extraction of characteristic activity patterns. These, in turn, are used in the development of probabilistic models. Page et al. (2008) developed a general methodology for the use of Markov chain modelling techniques for the simulation of activity sequences. Several variations of this methodology, supported by the use of country-specific TUS data were implemented. Some of the most commonly cited examples are the

models developed by Widén & Wäckelgård (2010) and Richardson et al. (2010), which are based on Swedish and UK data, respectively. The reader is referred to (Torriti 2014) for a more comprehensive review of models based on this approach.

A few attempts at improving this widely adopted modelling approach have been made. These, however, have focused on the refinement of the simulation of occupancy profiles. To this end, different methods have been applied.

One of these approaches has focused on improving the representation of the duration of the occupancy states through the use of semi-Markov models. In the model presented in (Wilke et al. 2013) the transition probabilities are based on the French TUS, and Weibull distributions are used to model the duration of the occupancy states. In the model presented in (Aerts et al. 2014) the transition probabilities are based on the Belgian TUS and non-parametric estimates are used to model the duration of the occupancy states.

Another approach has focused on the identification of characteristic activity patterns which exhibit statistically significant differences. In the model presented in (Bondu & Dachraoui 2015) a Bayesian clustering technique is used for the identification of household-level activity profiles. In the model presented in (Aerts et al. 2014), a hierarchical agglomerative clustering approach is used for identifying characteristic individual occupancy patterns. In both models, the clustering approach is complemented by the development of a Markov model for the identified clusters.

The model presented in (McKenna & Thomson 2016) is an extension of the model developed by Richardson et al. (2010). In addition to the extension of the model to include thermal demand loads, the simulation of occupancy

sequences was refined. The original model was based on two-state occupancy patterns simulations. The refined version is based on the simulation of four-state occupancy patterns. Both models, however, make use of the first-order Markov chain technique for the production of the occupancy sequences. The electricity demand model was left largely unaltered. The only modification being the removal of end-uses associated with thermal demand loads. This original model for electricity demand has found widespread applications in academia and industry, for instance (Ballanti & Ochoa 2015, Ballanti et al. 2013, EA Technology 2012, Good et al. 2015, Navarro et al. 2013).

The structure of the model developed by Richardson et al. (2010), also known as the CREST model, shows great similarity to that of other, more recent models. There are, however, differences between some of the key elements of the structure. The main difference concerns the approach to simulating occupancy patterns, as discussed in the paragraphs above. A lesser difference concerns the use of country-specific TUS data and statistics on appliance ownership and average ratings. Despite these differences, the approach to converting activity patterns into electricity demand loads has remained essentially the same. Therefore, although this model is based on UK-specific data, the results of the analysis presented in this paper are relevant to any model using a similar approach. In the following section, a more detailed description of the CREST model is presented.

2.3 The CREST model

In 2010 a new model for residential electricity demand was developed at CREST (Centre for Renewable Energy Systems Technology, Loughborough

University). This is a probabilistic model that provides an estimate of the electricity consumption of individual households based on the number of residents, occupancy patterns and dwelling appliance content. Based on the approach developed by Page et al. (2008) for the implementation of Markov chain models based on time-use data, Richardson et al. developed a residential building occupancy model (Richardson et al. 2008). Later on, this occupancy model was combined with the approach developed by Yao & Steemers (2005) for generating switch-on events of electric appliances to create a new probabilistic model that has found widespread application in the literature (Richardson et al. 2010). The model is capable of generating electricity load profiles of individual households as derived from the simulated use of individual appliances.

2.3.1 Structure

The model is divided into two modules. The first one is responsible for the simulation of dwelling occupancy patterns. The second one is responsible for the conversion of the simulated activity patterns into electricity demand loads. Occupancy profiles are simulated for each household, for each day of the simulated period. These profiles are based upon data from the UK 2000 Time-Use survey (Ipsos-RSL & ONS 2003). The data from this survey is representative of national time budgets (ONS 2003*a*). The version of the model used for the simulations described in this paper features a two-state occupancy model. Individual occupancy profiles are essentially binary sequences of states corresponding to periods of active and inactive occupancy. The active occupancy states also take account of the number of household members currently active.

Each simulated household is allocated a particular set of appliances.

Appliances are taken from a pre-defined set of 33 common appliances and randomly allocated based on national ownership statistics. The simulated appliances are configured using statistics that include their mean total annual electricity consumption, average power rating, and average cycle length. This configuration is meant to ensure that the appliances represent some general behaviour or pattern rather than emulate some specific real-life appliance (a specific make, or any other specific characteristics). In this sense, CREST's 'appliances' are rather categories of sorts into which many real-life appliances are expected to fit.

The activation of some of these 'appliances' during the simulations depends on the presence of active occupants in the dwelling. Therefore, the variation in appliance usage is based upon the activity of the members of the household and their number. A demand load profile is generated for each allocated appliance for each day of the simulated period. The loads generated depend on the occupancy profiles previously generated. The total daily load profile is the result of aggregating all of the appliances' daily demand loads.

The variation in electricity consumption between different households is achieved through changes in the household composition, set of appliances, and occupancy patterns of each simulated dwelling. Moreover, a calibration factor provides a means to align the overall average household consumption of the simulated sample to the relevant average annual consumption (e.g. sample group, regional or national average). For the simulations used in the original validation analysis this value was set between 4100 and 4300 kWh; this in order to represent dwellings from a specific region in the UK. The calibration of the model is meant to ensure that the overall average annual consumption of the simulated set of households is around a particular value, which is one of the

input parameters for the model.

2.3.2 Validation

For the purpose of validating the CREST model, household-level electricity consumption data was metered from 22 households in the East Midlands (Richardson et al. 2010). The households' electricity consumption was monitored for a full year. Thus, the resulting metered dataset contained around eight thousand 1 min resolution daily load profiles. The model was then calibrated against the collected data, and used for simulating the electricity consumption from the 22 monitored households. The simulations runs covered the equivalent of a full year period, and generated data with the same resolution as the metered data.

The simulation output was compared against the original metered data. The validation analysis was based on comparisons of different aspects, including the variation of annual demand between dwellings. One of the aims was to formally assess the extent to which these datasets differ from one another in terms of this variation. To this end, a statistical significance test, namely the Mann-Whitney U test, was used (Richardson et al. 2010). The purpose of this test is to determine how likely it is that the simulated dataset corresponds to a sample drawn from the same population as the reference sample (metered dataset). One of the advantages of this test is that it can be applied even when the distributions of the data are not known. The resulting measured and simulated annual consumptions are shown below in Figure 2.1(A).

In summary, the comparison between metered and simulated datasets appeared to show that there was no significant statistical difference between

them, thus validating the model. The CREST model is a good example of the attempts to link activity profiles to electricity consumption via the simulation of appliance usage. For this reason, we were particularly interested in looking at how well represented is this link by the simulations when compared with a larger, highly disaggregated metered dataset. A more detailed description of the data used for this purpose is given below in Section 2.4.

2.4 The UK Household Electricity Survey

The UK's Household Electricity Survey (HES) was the result of a study jointly commissioned by Defra, DECC and the Energy Saving Trust (Energy Saving Trust et al. 2012). The study had four main objectives:

- To identify the range and quantity of electrical appliances found in the typical home.
- To understand their patterns of use and their impact on peak electricity demand.
- To monitor total electricity consumption of the homes as well as that of individual major appliances in the household.
- To collect user habit data when using the range of appliances present in the households.

For this study 250 owner-occupier households were monitored over the period May 2010 to July 2011. The final report on this study observes that the monitored households were chosen such that the statistical characteristics of the sample would match closely those of the typical socio-economic mix (Zimmermann et al.

2012). Only owner-occupiers were recruited for the study. The households, however, were still fairly typical in terms of socio-demographic factors. The average annual electricity consumption across the sample was 4093 kWh/year, which compares very well to the 4115 kWh/year average (as of 2015) across all UK homes (DECC 2015).

Appliance-level electricity consumption was recorded for each of the 250 households for each day of the monitoring period. For some households total demand load, as read from the mains, was also recorded. The overall length of the monitoring period varies across households. 26 of them were monitored for a full year. The rest were monitored for "one-month-long" periods with effective lengths varying between 20 to 30 days, and covering different parts of the trial period.

The resolution of the metered electricity consumption profiles also varies. For households monitored for a full year, electricity consumption was metered every 10 min. For the rest of the households, it was metered every 2 min. The resulting dataset is a collection of highly detailed electricity consumption profiles, which already has provided very important insights into the way electricity is used in the UK's residential sector. However, the opportunities for exploiting such a rich source are far from exhausted. Further analysis would allow us to have a better understanding of some of the issues associated with residential electricity consumption, such as the scope for demand shifting, estimation of stand-by consumption, the effects of changes in the size and efficiency of appliances, and differences in the way different socio-economic groups use electricity.

The characteristics of this dataset make it ideal for testing current approaches to modelling of residential electricity use. In particular, we are interested in

testing whether a model such as CREST could provide us with simulated data that adequately represents the statistical characteristics observed in HES data.

2.5 Comparison of CREST model's simulations against HES data

The information about the relationship between consumer behaviour and appliance usage was more limited at the time models such as CREST's were developed. The issues caused by such lack of information could not be revealed using the metered electricity consumption data available at the time. However, the arrival of richer datasets allows us to explore these issues now by applying these models in a different context to the one in which they were validated. Now that more comprehensive datasets are available, such as the one provided by HES, this paper re-assesses CREST model's assumptions on appliance usage and their corresponding impacts on daily electricity load profiles and the variability of annual consumption across households.

The interest in testing this model against the HES data resides in great measure in the fact that the group of monitored households is much more diverse than any previous group. Moreover, the appliance-level electricity consumption profiles provide us with a further point of comparison between simulated and metered data.

2.5.1 Methodology

We wanted to reproduce, to some extent, the original model validation. For the original validation a relatively small sample consisting of 22 households was used (see Section 2.3.2). Total daily demand for electricity was monitored for a full year. These metered data allowed for comparisons of total annual demands and aggregated load profiles between metered and simulated households. In contrast, HES data offers the opportunity to compare total annual consumption levels, individual appliances' annual consumption and aggregated load profiles of a group of households over 10 times larger.

2.5.1.1 Data processing

Preliminary analysis of HES data revealed that some of the households monitored during the study are composed of 6+ members. Based on its present configuration, the CREST model can only simulate households of up to 5 members. For this reason and the sake of consistency, the original HES sample was "filtered" so as to be able to make a direct comparison between metered and simulated samples. The final sample resulting from excluding 6+ member households consisted of a total of 243 households, of which 24 had annual records. We will refer to this dataset as HES243. The extent to which this filtering affects the analysis will be discussed in Section 2.7.

In order to test the simulation outputs against the whole HES dataset, the comparative analysis between simulated and metered data was divided into two parts. Firstly, we focused on the 24 HES households that were monitored for a full year. The corresponding dataset will be referred to as HES24 throughout the

rest of the paper. The existence of this sub-set of households with annual records proves important, as it gives us the opportunity to make a closer comparison between the results of the analysis of this HES24 sub-set and the set of 22 households used for the original model validation (See Section 2.3.2). We will refer to the latter dataset as CREST22.

Secondly, we then extended the analysis to the rest of the HES sample so as to be able to assess the model's ability to represent the statistical characteristics observed in a larger, more comprehensive metered dataset. However, the sub-set of 219 households for which only monthly records were present had to be processed in a different way to the set of households with full year records.

Total household annual consumptions are not explicitly included in the HES data. This information had to be extracted from the records available for each household. Most households had monthly records only, so these had to be used for producing an estimate of the household's total annual consumption. For each household with monthly records only the available data was used for calculating total daily consumptions for each day in the monitoring period. Then, based on the distribution of total daily consumptions, the total daily consumptions for the missing days needed to complete a full year were generated stochastically to add some variation. The generated daily consumptions were then grouped into the months corresponding to the same monitoring period as that of households with annual records, and the corresponding monthly seasonal factor was applied. Finally, all daily loads were added up thus providing an estimate of the total annual consumption of that particular household. The case of the 24 households with annual records was more straightforward as we only had to add up the consumptions recorded throughout the year. In addition to the total annual consumption, a mean daily load profile for the whole HES sample was calculated

based on all available data.

For the simulation of electricity consumption data the CREST model was used. A full version of the CREST model is publicly available in the form of an Excel spreadsheet model (See Richardson et al. (2010)). However, for the purpose of simulating the electricity consumption of the same number of households as in the HES sample, the CREST model was re-implemented in Python in order to allow for faster, multiple automated runs. Features, functionality and results of Python and Excel versions of the model are equivalent. We used this Python implementation for generating a year's worth of artificial electricity consumption data for 243 households. Each simulated household profile was matched to a metered counterpart in the HES243 dataset. Since it is possible to specify both the day and the month when running simulations, the simulated period was also matched to the monitoring period of the corresponding metered households. Total annual consumptions corresponding to each household were calculated, as well as the mean daily load profile for the whole simulated HES243 sample. For the first part of the analysis, which focused on the HES24 dataset, simulations were run based on the specifications of this sub-set of households only.

As the electricity consumption of individual appliances is available from both metered and simulated data, there was an interest in looking at how the estimates of the annual consumptions of the simulated appliances are distributed and how these distributions compare to those of the corresponding metered appliances. To this end, we also extracted the total annual consumptions of individual appliances for the households in the HES243 dataset.

As observed in Section 2.3.1, the CREST model generates artificial electricity consumption data based on the simulation of appliance usage for a pre-defined set

of appliances. Among the model's pre-defined set, there are 'appliances' which are meant to represent the consumption of the appliances listed in Table 2.1. This pre-defined set includes some other 'appliances' which represent the electricity consumption of some other devices with more predictable consumption patterns (e.g. cold appliances). However, we focus the analysis on the appliances listed in Table 2.1, as the electricity loads generated by the use of these appliances are the ones with direct links to households' activities. We compare the simulated demand of these appliances with the data from their corresponding metered counterparts.

The appliances of interest were grouped based on the activities they are associated with: cooking, washing, TV watching, and ICT related activities. Both CREST and HES appliances are grouped in these more general categories, as shown in Table 2.1. In most cases it is possible to make a one-to-one comparison between HES and CREST appliances in spite of the differences that may exist between them.

In order to obtain the distributions of the simulated total annual consumption of individual appliances, it was necessary to run a new set of simulations. Based on the original model configuration, appliances are allocated to the simulated households on a random basis. Therefore, the original model implementation had to be modified so as to ensure that the appliances of interest were allocated to the relevant households. The goal was to match the appliance configuration of the simulated households to that of the metered counterparts.

HES records show which appliances are owned by which households. Based on this information, households were grouped accordingly. Using the modified model implementation, simulation runs corresponding to one year long periods were carried out.

Table 2.1: *Categories of appliances and the appliances contained in each one, according to the appliances present in both metered and simulated datasets.*

Categories	Appliances contained	
	CREST	HES
Cooking	hob, oven, microwave, kettle	cooker1, cooker2, hob1, hob2, oven1, oven2, hob+oven, microwave1, microwave2, kettle1, kettle2
Washing	Washing machine, Tumble dryer, Dish washer	Washing machine, washer dryer, Tumble dryer, Dish washer
TV watching	TV1, TV2, TV3	TV1, TV2, TV3, CRT TV, LCD TV, plasma TV, Audio-visual site
ICT related	PC	Desktop PC 1, Desktop PC 2, Computer site, Laptop 1, Laptop 2

2.5.1.2 Comparative analysis

In order to assess whether the simulated data is consistent with the metered one, both qualitative and quantitative methods were used.

The qualitative comparison between datasets focused on the Empirical Cumulative Distribution Function (ECDF) of the different datasets. The ECDF is a formal direct estimate of the Cumulative Distribution Function[1] from which simple statistical properties can be derived. Moreover, the ECDF properties form the basis of various statistical significance tests.

[1]Cumulative distribution function (CDF) - It is a function whose value is the probability that a corresponding continuous random variable has a value less than or equal to the argument of the function.

The Probability Density Function (PDF) of the total annual consumptions was obtained as well, as it provides a better graphical representation of some of the statistical characteristics of the metered HES243 dataset. In order to obtain a reliable representation of the PDFs, the Kernel Density Estimation (KDE) method was used. Kernel Density estimation is often used as a data smoothing technique. However, KDE is not simply a smoothing method. This method offers a way of obtaining an accurate non-parametric estimate of the PDF of a given dataset. This estimate is in fact the best possible estimate of the distribution of the original population from which the sample was drawn. It can be shown that the kernel density estimate converges to the real distribution as the sample size increases, and its convergence rate is the highest among all estimates (Wahba 1975).

As observed above, some statistical significance tests are based on the properties of ECDFs. This leads naturally to the quantitative part of this comparative analysis, which focuses on the use of such tests. For the quantitative comparison we will use two tests, namely, the Kolmogorov-Smirnov test and the Mann-Whitney U test. Both tests are non-parametric, which means they do not make assumptions about the distribution of the data. Both tests can be used to compare two unpaired datasets, both focus on comparing the statistical characteristics of the distributions of the two datasets being compared, and both are robust to the presence of outliers. However, the two tests work in a very different way.

The Kolmogorov-Smirnov (KS) test compares the cumulative distribution of the two data sets in question, and calculates a p-value that depends on the largest difference between the two distributions. The test is sensitive to any differences in the two distributions. Significant differences in shape, spread or median values

will result in a small p-value. The KS test can also be used to test whether an empirical distribution is consistent with an ideal distribution, and therefore, it is commonly used as a test of goodness of fit, or for testing for normality. In essence, this test answers the question: If the two samples were randomly drawn from different parent populations, how likely it is that the two parent populations have the same statistical distribution?

The Mann-Whitney (MW) test compares two datasets by ranking the elements in the dataset from smallest to largest, and calculating the average of the rank scores of the elements in the two datasets. A p-value is calculated which depends on the difference between the average ranks of the two datasets. Compared to the KS test, the MW test is mostly sensitive to changes in the median value. The MW test is closely related to the t-test. However, the MW test is more widely applicable than other more popular tests such as the t-test, as it does not require the assumption of normally distributed data, and it is much more robust to the presence of outliers. In essence, the MW test answers the question: If both samples come from the same population, how likely it is that random sampling would result in the differences observed in this particular case?

As we are interested in determining whether the simulated data generated by the model accurately capture the statistical characteristics of the metered data, we use these tests to obtain a quantitative measure of how closely related the two datasets are. In other words, we use these tests to assess whether both datasets are consistent enough with statistical samples drawn from the same population, or whether the two datasets are consistent with samples drawn from populations which follow the same statistical distribution. The MW test provides us with a quantitative measure of how likely the former situation is, whereas the KS test

provides us with a measure of how likely it is the latter.

The MW test was also used in the original validation of the CREST model (see Section 2.3.2). The positive results of the test were one of the arguments in favour of the validation of the model. We take this as an opportunity to show the effect that sample size may have when using this kind of test by applying the test to the datasets HES24 and HES243 separately and comparing the results.

2.6 Analysis results

2.6.1 Variations in annual demand levels

The high variability observed in annual demand levels across the monitored households clearly reflects the fact that different households mean different needs, lifestyles and consumption patterns. In this context, when we talk about variability, we refer to the natural variation that is observed in the demand levels as a result of the differences in the underlying causes of said demand.

Figure 2.1 shows both simulated and metered annual demands, obtained from the different datasets. Both the variability of annual demand among households and the differences between the metered and simulated levels can be appreciated, as well as how they compare across datasets. This figure starts to show that the empirical data has greater diversity than the model simulations would suggest. The red horizontal lines indicate the overall average annual consumption of the corresponding metered dataset. Only one line per section is needed because, as observed in section 2.3, an adequate calibration of the CREST model ensures

that the overall average annual consumption of the simulated sample is around a specified value. The corresponding values are listed in Table 2.2 which shows, in numbers, how the variability of both simulated and metered data compares across the different datasets.

The mean annual electricity demands per dwelling calculated from metered and simulated data for each of the analysed datasets differ by around 1%. The fact that this difference is small only shows that the model was appropriately calibrated. Despite this, it is observed that the differences in standard deviations are considerably larger. For both HES24 and HES243 datasets the standard deviation of the metered annual consumptions is about 40% higher than the standard deviation of the simulated ones.

The results of the first part of the analysis, concerning the HES24 dataset, already appeared to suggest that the variation in annual demand levels was being under-represented. However, the simulated HES24 dataset used for this study appears to be consistent with the simulated CREST22 dataset used in the original validation analysis[2] (See Table 2.2).

Table 2.2: *Mean values and standard deviations of annual electricity consumptions datasets.*

	CREST22		HES24		HES243	
	metered	simulated	metered	simulated	metered	simulated
Mean annual demand (kWh/y)	4172	4124	4624.6	4629.7	3944.5	3942.9
Standard Deviation (kWh/y)	1943	1372	2510	1545	2348	1405
Coefficient of variation	0.47	0.33	0.54	0.33	0.60	0.36

[2]CREST22 data was taken from Richardson et al. (2010))

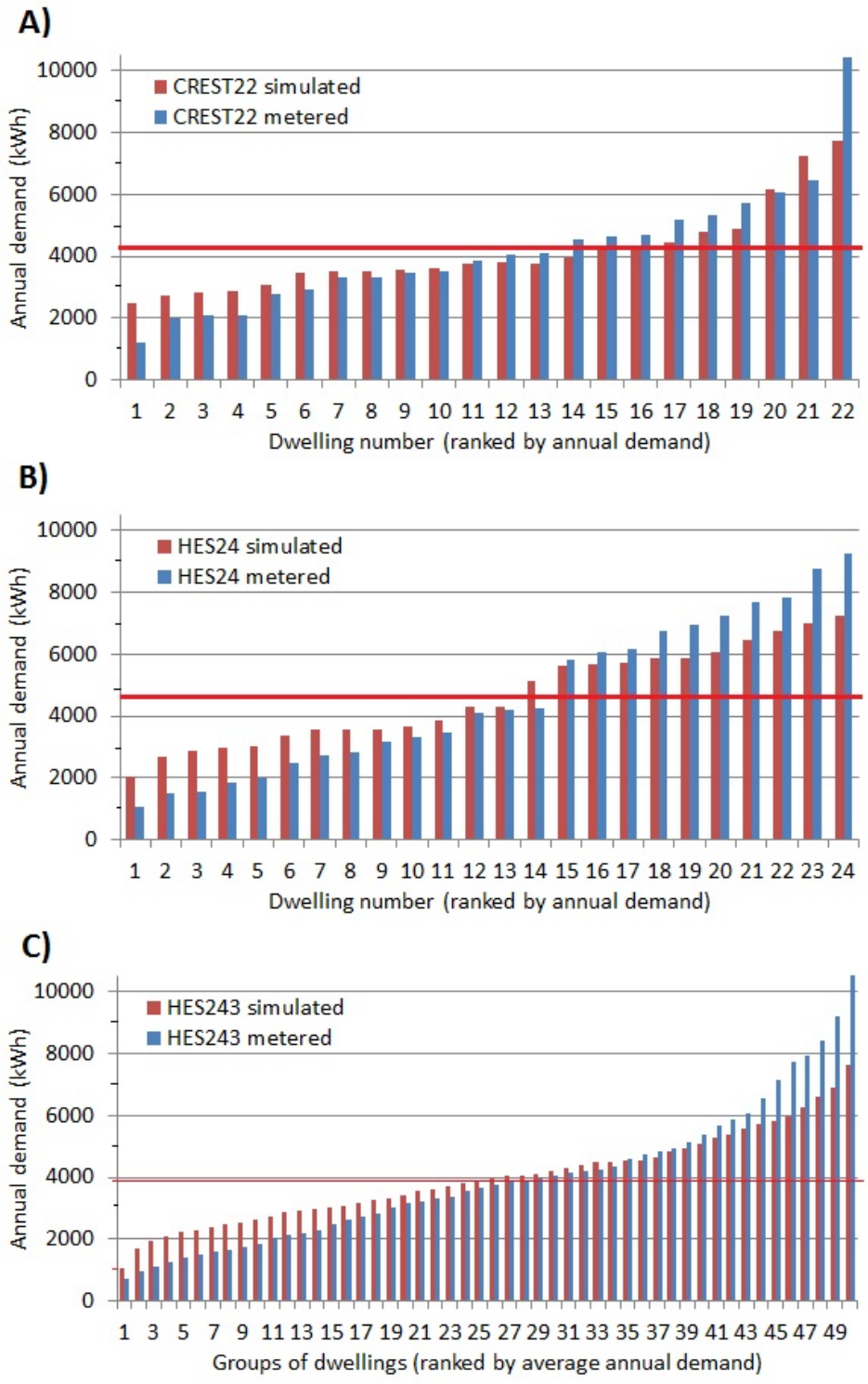


Figure 2.1: *Annual electricity demands by household, ranked by magnitude: A) Annual demands of both metered and simulated households in CREST22 dataset (adapted from (Richardson et al. 2010, Fig. 6)). B) Annual demands of both metered and simulated households in HES24 dataset. C) Annual demands of both metered and simulated households in HES243 dataset. Households were grouped in sets of five. The levels observed correspond to the average consumption of each set.*

The analysis of CREST22 datasets showed that the standard deviation of the metered annual consumptions was about 30% higher than the standard deviation of the simulated ones. Despite this seemingly large difference, the result of applying a statistical significance test would appear to indicate that the datasets were not significantly different (See Section 2.3.2).

2.6.2 Empirical cumulative distributions and significance tests

The HES24 dataset is very similar in size to the CREST22 dataset used for the original validation. As part of the original validation analysis, a Mann-Whitney (MW) test with a 5% level of significance was used to compare metered and simulated CREST22 datasets. Based on the test results, it was concluded that there was no significant difference between them (Section 2.3.2). In addition to verifying this, and in order to further test how well represented was the metered data by the simulations in the original analysis, we extracted the annual consumption data from the original validation results and performed a Kolmogorov-Smirnov (KS) test on these datasets. The results of this test also supported the hypothesis that the two datasets are consistent with each other.

The same tests were then applied to metered and simulated HES24 datasets. With a p-value of 0.38, the MW test with 5% significance level suggests that it is likely that these two samples were drawn from the same population. Moreover, with a p-value of 0.44, the KS test with 5% significance level suggests that it is likely that these two samples come from populations equally distributed. The results of this part of the analysis seem to be well in agreement with those of the

original validation, thus confirming that the model produces reasonable estimates when dealing with small samples ($\sim$ 20 households).

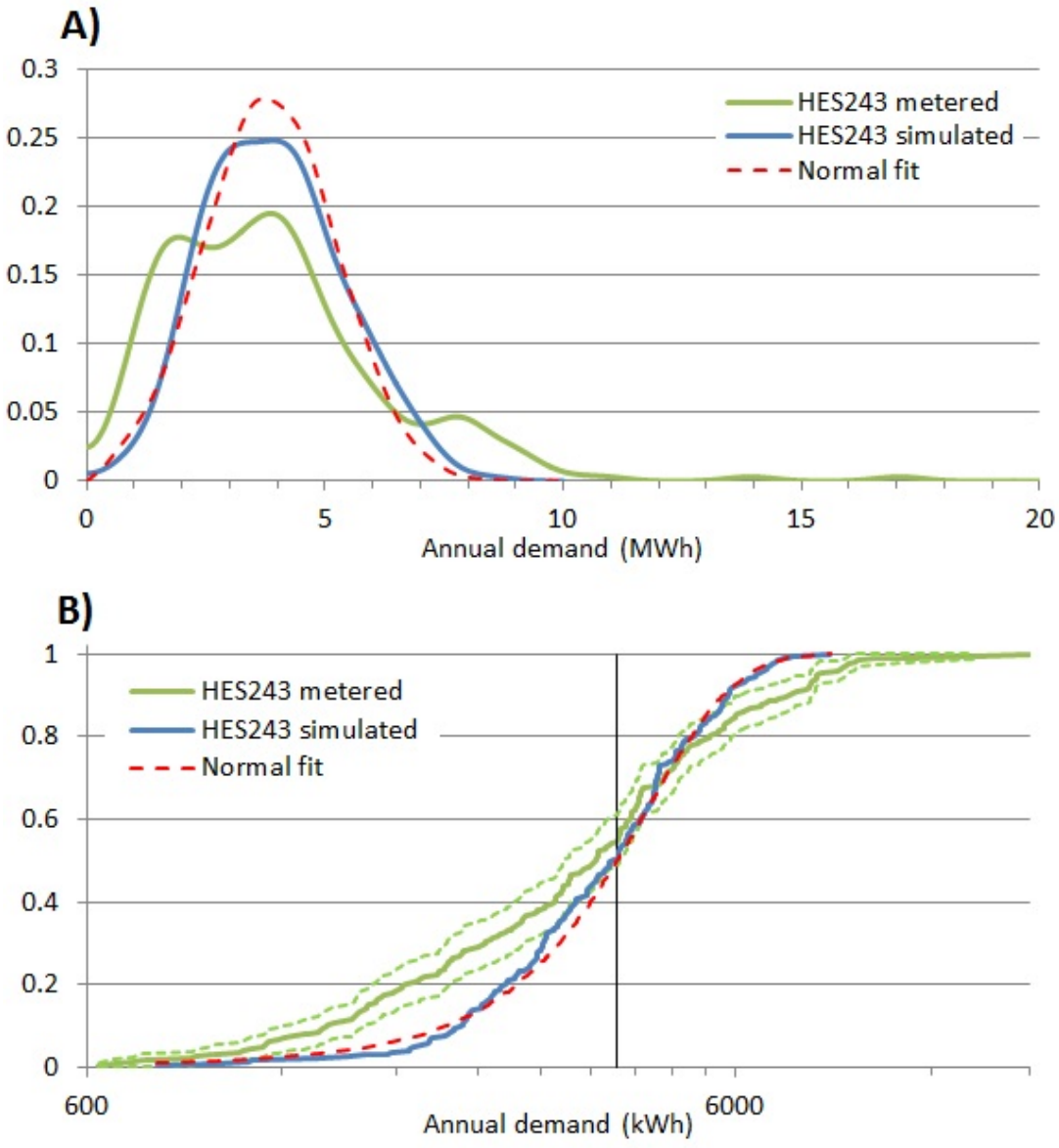


Figure 2.2: *Distribution of total annual consumptions: A) Probability Density Functions of metered and simulated HES243 datasets. B) Empirical Cumulative Distribution of metered and simulated data. The dotted lines around the ECDF of the metered HES243 dataset are the 95% confidence bounds.*

The second part of the analysis, which focused on the HES243 datasets, was aimed at determining whether the model simulations produce a reasonable representation of the statistical characteristics of a larger, more diverse set of households. Both MW and KS test were used again to provide a quantitative measure of the differences between metered and simulated datasets.

On this occasion, however, the tests results revealed the existence of important differences in the statistical characteristics of both datasets. With a p-value of 0.049, the MW test with 5% significance level now suggests that it is unlikely that the metered and simulated HES243 samples come from the same population. In a similar manner, with a p-value of 0.001, the KS test with 5% significance level suggests that it is very unlikely that these two samples come from populations with the same statistical distribution. Moreover, by using the KS test to test for normality, we found that the simulated HES243 dataset appears to be consistent with normally distributed data (p-value of 0.81). The metered HES243 sample is not. These results are best appreciated graphically. As Figure 2.2 shows, the simulated HES243 dataset is well in agreement with its corresponding normal fit (dashed line).

Probability density functions corresponding to each set of annual consumptions can also be appreciated in Figure 2.2. The Kernel Density Estimate of the distribution of metered annual consumptions revealed the existence of a more complex underlying distribution. Based on this estimate, three clusters can be readily identified. What this shows is that the distribution of households with respect to total annual consumption is concentrated in three clusters, centred around the values (modes): 1.58 MWh, 3.92 MWh and 7.85 MWh. Based on this information, and on the same principles behind the Kernel Density Estimation, we produced a parametric estimate for the empirical PDF

observed in Figure 2.2. This estimate can be expressed by the following trimodal Gaussian mixture:

$$f(x) \sim C_1\mathcal{N}_1(x \mid \mu_1, \sigma_1) + C_2\mathcal{N}_2(x \mid \mu_2, \sigma_2) + C_3\mathcal{N}_3(x \mid \mu_3, \sigma_3) \tag{2.1}$$

where $\mathcal{N}_i(x \mid \mu_i, \sigma_i) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_i}{\sigma_i}\right)^2}$ is the general normal distribution with mean μ_i and standard deviation σ_i, and C_i is a proportionality constant.

A summary of the parameters of this estimate and their corresponding values can be found in Table 2.3. Figure 2.3(A) shows how this trimodal Gaussian model compares to the empirical PDF.

The ability to produce a parametric estimate for the PDF means that it is also possible to find an expression ($F(x)$) for the CDF, as these two functions are related analytically by the equation: $F(x) = \int_{-\infty}^{x} f(t)dt$. Therefore, by integrating $f(x)$ we obtain that the CDF is given by:

$$F(x) = \sum_{i=1}^{3} \frac{C_i}{2}\left[1 + \text{erf}\left(\frac{x-\mu_i}{\sigma_i\sqrt{2}}\right)\right] \tag{2.2}$$

where $\text{erf}(x)$ is the error function, and C_i, μ_i and σ_i are the same parameters used in equation (2.1) (see Table 2.3). The way this estimate compares to the ECDF is shown in Figure 2.3(B).

Table 2.3: *Trimodal Gaussian mixture model's parameters.*

	$\mathcal{N}_1(\mu_1, \sigma_1)$	$\mathcal{N}_2(\mu_2, \sigma_2)$	$\mathcal{N}_3(\mu_3, \sigma_3)$
C_i	0.261	0.622	0.115
μ_i	1.575	3.92	7.85
σ_i	0.78	1.28	1.0

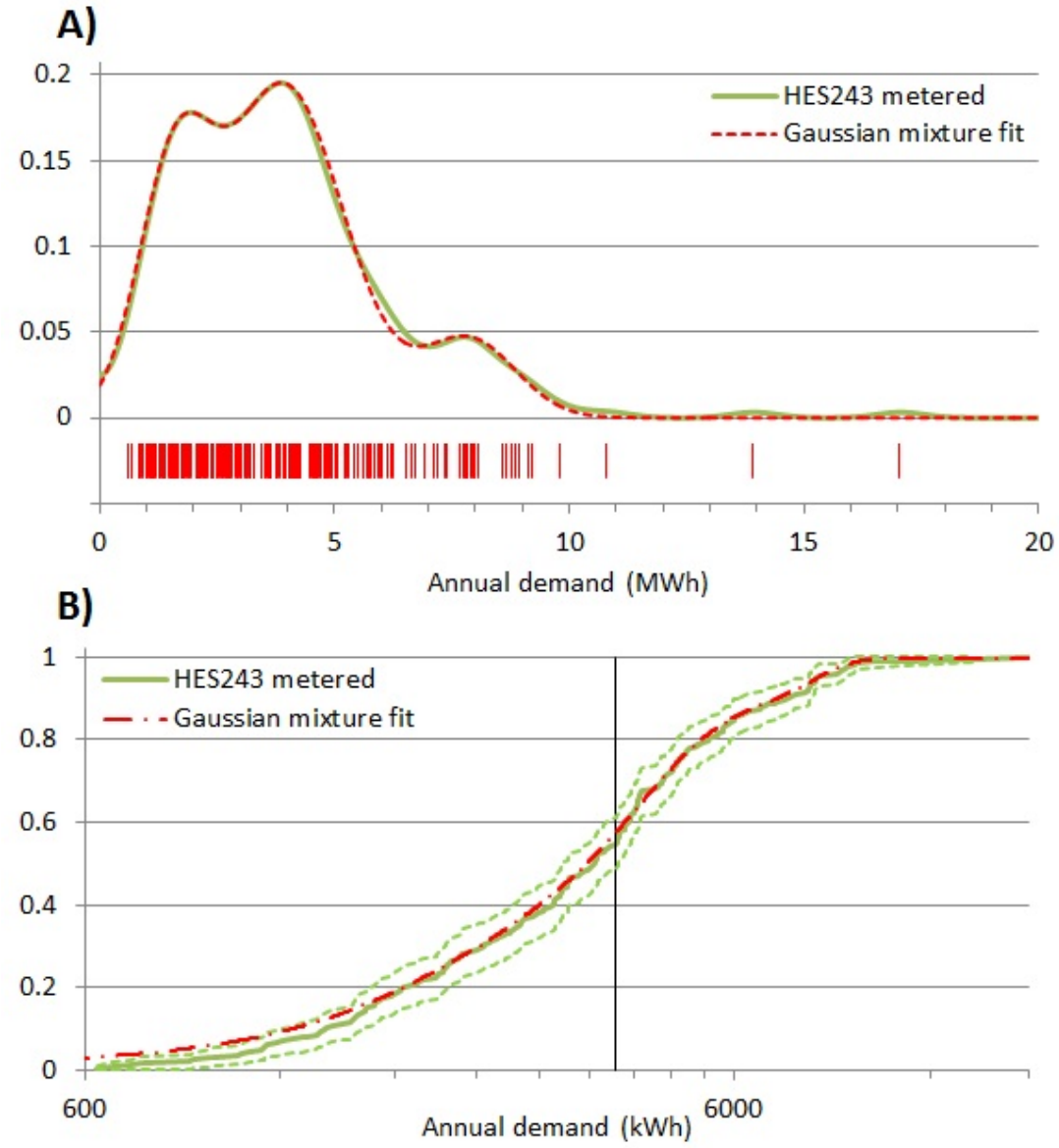


Figure 2.3: *Total annual consumptions distributions: A) Probability Density Function of metered HES243 dataset and trimodal Gaussian mixture fit. Each vertical bar at the bottom of the plot represents a data point. B) Empirical Cumulative Distributions of metered annual consumptions and trimodal Gaussian mixture fit. The dotted lines around the ECDF of the metered HES243 dataset are the 95% confidence bounds.*

In order to gain a better insight into the relationship between household size and total annual consumption, we looked at the sub-groups of households characterised by these features in two different ways.

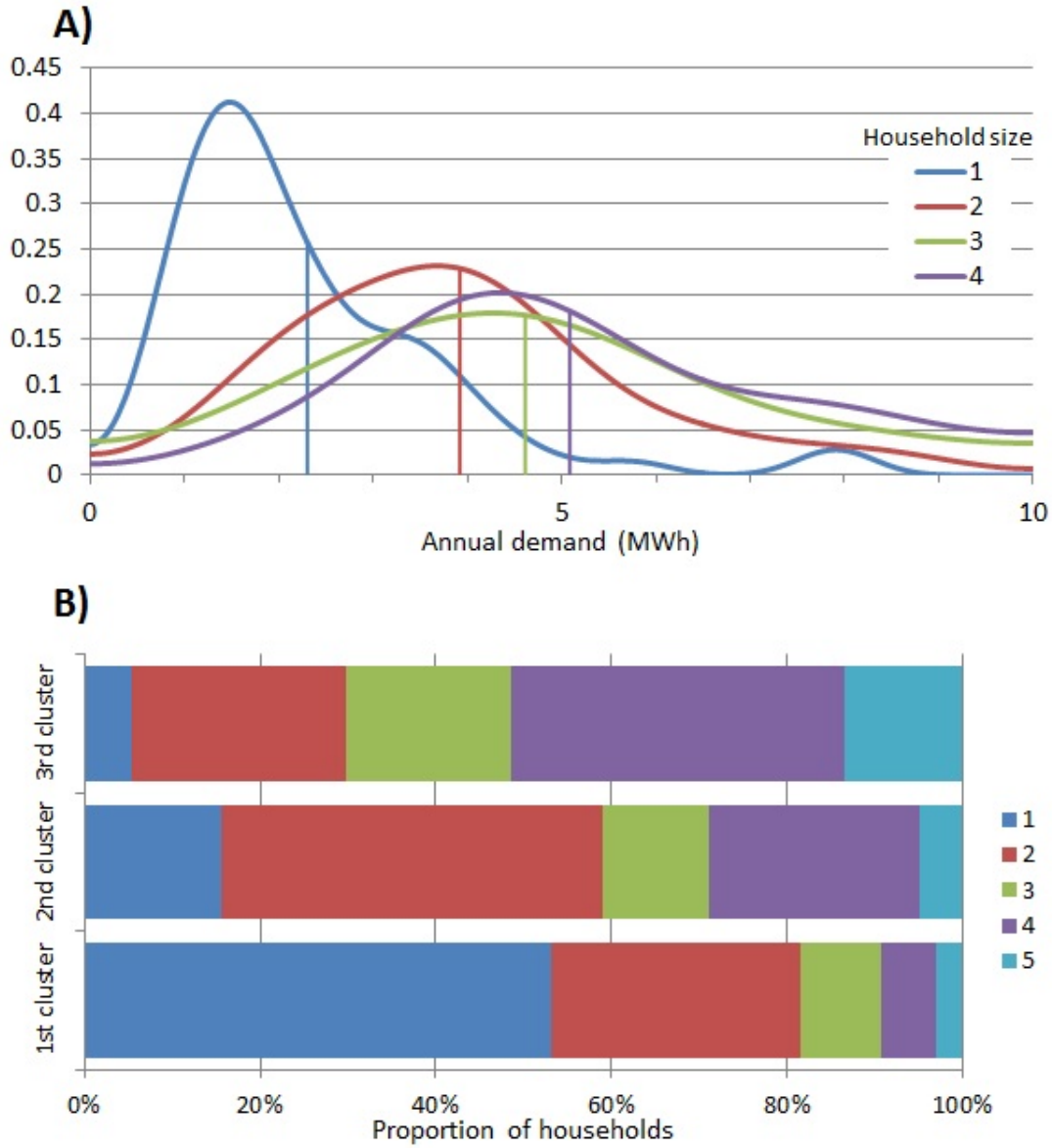


Figure 2.4: *Distribution of households according to size and annual consumption range: A) EPDFs corresponding to households grouped by size. The vertical line in each distribution indicates the mean value. B) Percentage composition of household sizes for each cluster.*

Firstly, we grouped the households according to size, and used the KDE method to obtain an estimate of the empirical PDF corresponding to each group. These distributions can be observed in Figure 2.4(A). The number of households with 5 members is not large enough as to provide a reliable estimate of the

overall distribution of annual consumptions for this sub-group. Therefore, this distribution was omitted.

Secondly, based on the parametric estimate we obtained for the overall distribution, we grouped the households according to the three observed clusters. The criteria for the allocation of households to the different clusters was whether their annual consumption falls within the interquartile range of the cluster in question. Households in each cluster were further grouped according to size. The percentage composition of each cluster is shown in Figure 2.4(B).

2.6.3 Individual appliances: Total annual consumptions

We looked at the appliance content of the HES243 households and, based on the list of appliances of interest presented in Section 2.5.1, we grouped the households that contained each one of those appliances. We then calculated the ECDF of both simulated and metered data for each one of those groups. Examples of the obtained distributions are shown in Figure 2.5.

The results of this analysis made evident that the variability observed in the annual consumption of individual appliances is completely misrepresented by the simulated data. In some cases (see Figure 2.5(A) and (D)) the average annual consumption of both datasets is well in agreement. In some others (Figure 2.5(B) and (C)), the differences between the average annual consumptions are evident as well.

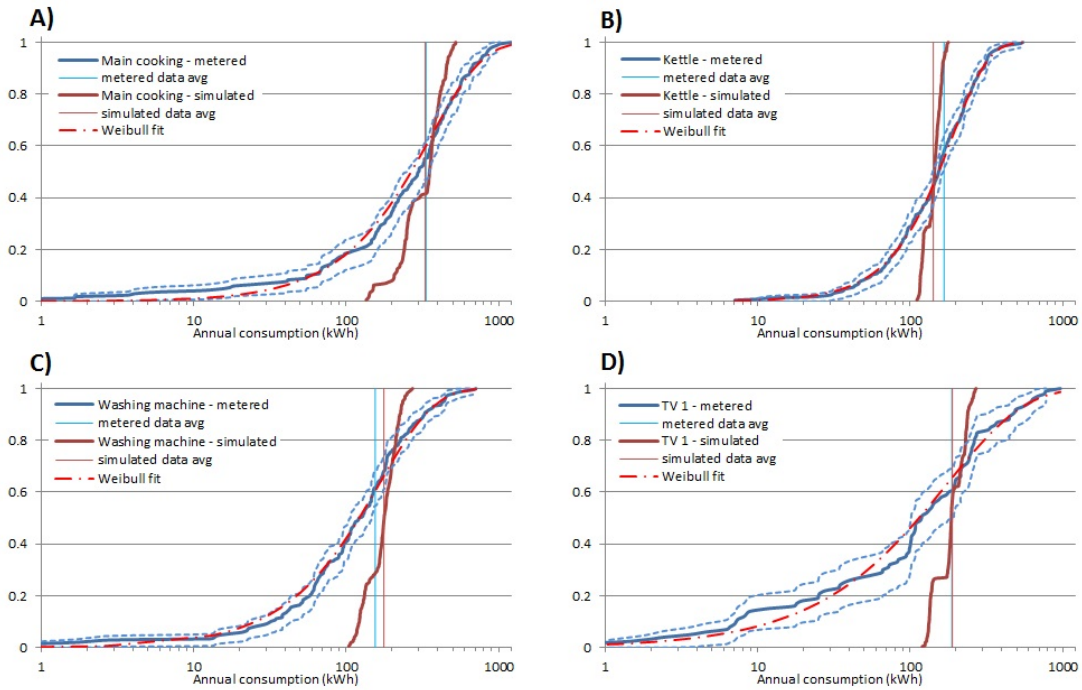


Figure 2.5: *Annual consumptions ECDFs of individual appliances: A) Main cooking appliances: cooker, or equivalently, hob+oven. Present in 70% of households. B) ECDFs of simulated and monitored kettles. Present in 94% of households. C) ECDFs of simulated and monitored washing machines. Present in 83% of households. D) ECDFs of simulated and monitored main TVs. Present in 98% of households.*

In general, it was observed that the empirical distributions of individual appliances' annual consumptions were skewed. Log-normal and Weibull models were tested for possible fits, as these models are particularly good at representing skewed data.[3] The analysis revealed that the best fits were provided by Weibull models. The summary of parameters of the distributions fitted to the ECDFs of the annual consumptions of individual appliances is presented in Table 2.4.

[3]A more empirical argument for the use of these models is rooted in the fact that they are generally applicable for modelling size/magnitude distributions (e.g. particle size, wind speed).

Table 2.4: *Summary of parameters of the Weibull distributions fitted to the annual consumptions ECDFs of individual appliances.*

	Weibull fit parameters	
	Scale	Shape
Main cooking	356.54	1.26
cooker	344.72	1.28
oven	248.34	1.04
microwave	57.66	1.19
toaster	21.90	0.94
kettle	188.20	1.82
washing machine	164.85	1.20
tumble dryer	367.87	0.96
washer dryer	262.81	1.06
dishwasher	301.57	1.47
TV1	174.90	0.86
TV2	109.22	0.69
laptop	24.34	0.87
desktop	156.37	0.89
computer site	213.01	0.98

2.6.4 Mean daily load profile general features

Mean daily load profiles were obtained from both metered and simulated HES243 datasets. These can be observed in Figure 2.6, along with the typical UK profile. This typical UK load profile corresponds to the half hour load profile for the average Profile Class 1 customers[4](UKERC Energy Data Centre et al. 2013), and is presented for comparison purposes only.

[4]Profile Class 1 customers are domestic customers supplied on unrestricted tariffs, with a maximum demand below 100 kW as measured by metering systems containing only one meter register. There are eight profile classes, and for each profile class a sample group of electricity supply customers is randomly drawn from the population of electricity supply market customers. These samples are designed to provide an accurate estimate of the load pattern for each class of customers for use in electricity settlement. Consumption data is obtained by either installing half-hourly meters at the sites or getting half-hourly consumption data directly from suppliers.

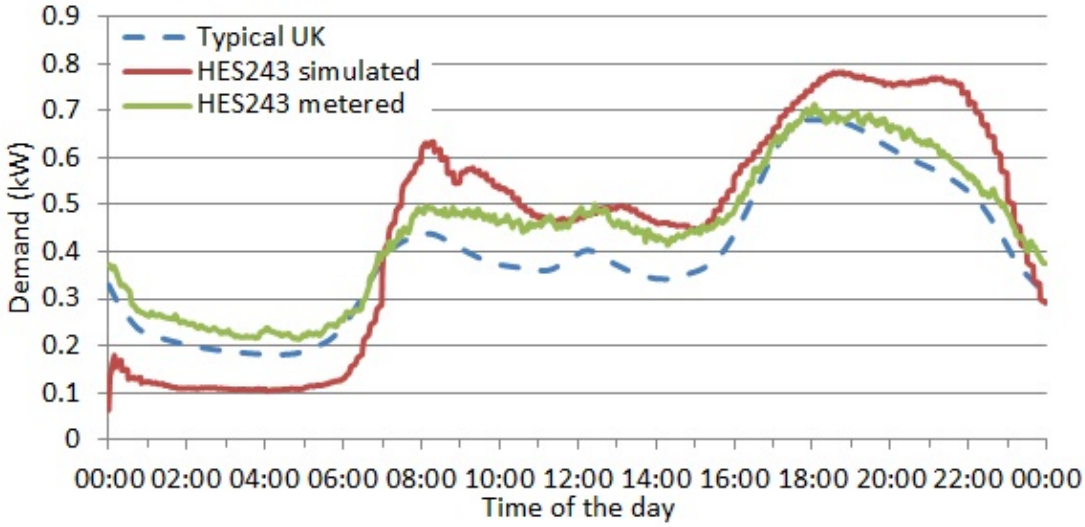


Figure 2.6: *Mean daily load profiles corresponding to metered and simulated HES243 datasets. Typical UK profile corresponds to the load profile of the average class 1 (unrestricted domestic tariff) customer.*

Both the metered and simulated mean daily load profiles have, generally speaking, similar shapes to that of the typical profile. However, when the profiles are compared in more detail, marked contrasts are observed:

- There are significant differences between the values of the lowest and highest demands observed in the two profiles.
- The simulated profile appears to overestimate the duration of the peak period.
- The variations observed in the CREST profile, such as the transitions between periods of low/high demand and periods of high/low demand, correspond to more abrupt changes than those observed in the other two profiles.

Further analysis of the characteristics and differences of these profiles may be desirable, and indeed, necessary. In particular, it might be interesting to pay a

closer look at how the activity profiles and the use of appliances are linked, and how well represented the duration of certain activities is. However, a more detailed analysis of these profiles falls beyond the scope of this paper. The prospects for a more detailed study on the intra-day variability of demand are left to be discussed in the next section.

2.7 Discussion and Future work

The primary aim of the analysis presented in this paper was to determine whether a probabilistic model based on currently used approaches is capable of providing a good representation of the variability observed in the electricity consumption levels of a highly detailed dataset. The choice of the CREST model for this study was due in great measure to the fact that this model seeks to emulate the variability of household electricity use (Richardson et al. 2010). According to the way in which the model was constructed, the simulated appliances are configured such that their annual consumptions result in figures typical of an average household. Moreover, the overall average annual consumption of the simulated households was also expected to be in agreement with the national figures. The model calibration involves the scaling of the simulated sample's annual consumptions. This is meant to ensure that the simulated average annual consumption matches the average annual consumption of the real-life counterpart (see Section 2.3).

The high level of detail offered by HES data justifies its use for the analysis. The appliance-level metered electricity consumption records make it ideal for a comparison against the simulated appliance consumption data. Due to the

technical limitations of the model discussed in Section 2.6, the original HES dataset had to be filtered so as to allow for a consistent analysis. The filtering consisted in removing the households with more than five members from the sample. However, we believe this filtering had little impact on the overall statistical characteristics of the dataset and the analysis results. The proportions of households in the HES sample corresponding to household sizes ranging from 1 to 4 are well in agreement with the proportions observed at the national level (ONS 2015). If we extrapolate this, by assuming that the proportion of households with five and 6+ residents are equally well represented in the sample, this would mean that households with 6 or more members account for just over 2% of households in the UK. A sense of the scale of the effect of removing said households can be given by comparing the overall average annual consumptions; the average annual consumption across the whole 250 households HES sample is just over 3% higher than the average across the HES243 dataset.

Great part of the analysis presented in this paper focused on the distribution of households with respect to their annual consumptions. The analysis of the annual consumption data extracted from the HES dataset revealed the existence of a complex trimodal distribution (see Section 2.6.2). Further analysis revealed that the underlying distribution appears to be very well estimated by a trimodal Gaussian mixture with analytical expression given by equation (2.1). Each of the three clusters identified in the empirical distribution is consistent with a normal distribution. The resultant composite distribution, however, does not exhibit normal-like features. A breakdown into sub-groups of households characterised by size revealed that: 1) Over 50% of the households concentrated around the first cluster are one-person households. The first cluster is also the largest one, which means that the majority of the one-person households are found in this consumption range. 2) Over 40% of the households concentrated around

the second cluster correspond to two-people households, making this sub-group the predominant one. 3) Over 33% of the households concentrated around the third cluster are 4-people households. The second biggest group corresponds to two-people households (see Figure 2.4). It appears to be a linear correlation between the household size of the predominant group and the mean value of the corresponding cluster. However, further analysis is needed in order to determine whether this is a general trend or just an interesting coincidence.

The distributions of the simulated annual consumptions were found to be consistently normal. It would appear that the current simplifying modelling assumptions are causing simulation outputs consistent with normally distributed data. These assumptions would appear to lead to reasonable first-order estimates as misrepresentation issues appear less severe when simulations are restricted to small sets of households. However, the differences between simulated and metered data become evident when simulating larger samples.

Modelling assumptions concerning the conversion of activity profiles into electricity consumption patterns have remained essentially the same over the last decade. It is therefore necessary to reassess these assumptions. Assumptions leading to normally distributed data will need to be modified or replaced entirely. If replacing these assumptions in their entirety proves overly complicated, then the implementation of mixture models based on normal components might be a good alternative.

In addition to the differences in the overall shape of the distributions, it was observed that the variability in the levels of annual consumption is consistently misrepresented in the simulated data. Regardless of sample size, it was found that the level of variability observed in the metered data is consistent across

independent datasets. Moreover, the variability observed in the metered data is about 20% higher than in the simulated data (see Table 2.2). The HES sample is still arguably small, statistically speaking. However, we presume it provides a better insight into the kind of distribution and variability that would be observed in larger samples. Given the method used for estimating the overall distribution, it is also very likely that the statistical characteristics observed in the estimate are close to those of the true population distribution. It is important that a refined representation of the characteristics of the distributions observed in Figure 2.3 is achieved, as this is key to a robust approach to residential electricity demand modelling.

The analysis of the distribution of metered annual consumptions generated many interesting results, which include analytic expressions for the probability density function (PDF) and the cumulative distribution function (CDF) (see Section 2.6.2). Knowing the analytical expression for the cumulative distribution allows for the use of methods such as the inverse transform sampling. Given a cumulative distribution function, this method allows for the generation of samples consistent with data drawn from said distribution. A better idea of the statistical characteristics of the distribution of annual consumptions gives us the opportunity to fine tune demand models. The improvements derived from this would be reflected in simulated datasets that are more in agreement with the kind of distribution and levels of variability observed across real-life households.

Total annual consumptions from individual appliances were used as a further point of comparison between simulated and metered datasets. Based on each dataset, distributions of total annual consumptions were obtained for the appliances listed in Table 2.1. The choice of the appliances of interest was based on the reasonable expectation that their demand loads would be linked to

households activities. In the CREST model's simulations, a calibration factor is calculated for each appliance. According to the model documentation, this factor is adjusted so that over a large number of runs the average annual consumption of the appliance would match typical levels (Richardson et al. 2010). In some cases it was found that the average annual consumption of the simulated appliances was well in agreement with that of their metered counterparts (see Figure 2.5(A) and (D)). In most cases, however, discrepancies were found. In general, the simulated values are consistently higher than those obtained from the metered data. In terms of the distribution, similar problems to those observed in the distributions of total household annual consumptions were found. The variability observed in the metered data is consistently under-represented by the simulations' output. Moreover, the annual consumptions of individual simulated appliances present skewed distributions. These distributions were found to be consistent with Weibull distributions. A summary of the parameters for the fitted models is presented in Table 2.4.

A comparison was also made between metered and simulated mean daily load profiles extracted from the corresponding HES243 datasets. Important differences in the overall features of both profiles were identified. Further in-depth analysis is needed in order identify the underlying causes of these issues. However, a more detailed characterisation of said differences falls beyond the scope of this paper. In particular, it would be interesting to investigate in more detail the relationship between household activity patterns, appliance usage and the intra-day variations in demand levels. This will be the subject of future work.

2.8 Conclusions

In this paper, a comparative analysis of the statistical characteristics of metered and stochastically simulated electricity consumption datasets was presented. Metered household- and appliance-level electricity consumptions were extracted from the UK's Household Electricity Survey dataset (see Section 2.4). Corresponding electricity consumptions were simulated using a widely implemented probabilistic model based on the UK residential sector (see Section 2.3). The model was configured such that the simulated households matched the relevant features of the metered counterparts. The analysis of the differences between datasets allowed us to provide a measure of how well represented are the observed features by probabilistic models based on currently used approaches and to identify key shortcomings.

In particular, we compared the way simulated and metered households and individual appliances are distributed with respect to total annual consumptions. An element of the assessment of the model's performance concerns how well represented are these distributions by the simulated output. Based on the analysis presented in this paper we can conclude that current probabilistic models fail to capture some of the key characteristics observed in the metered datasets (see Section 2.6.2). Significant discrepancies were found when the corresponding distributions were compared (see Figures 2.2 & 2.5). In particular, it was found that metered household annual consumptions follow a complex distribution from which three clusters can be identified. The distributions observed in the simulated data are much simpler. Current underlying assumptions about electricity use appear to lead to normally distributed data with consistently limited variability. Therefore, these assumptions need to be rectified so as

to prevent the misrepresentation of the statistical characteristics observed in metered data.

The diversity of real-life households is reflected by the complexity of the distribution of total annual consumptions. A more complex distribution does not necessarily mean over-complicated assumptions need to be used in the models. We propose a trimodal Gaussian mixture model which provides a very good estimate of the observed distribution. Analytical expressions for both the density function and cumulative distribution are given. In a similar manner, the distributions of annual consumptions of individual appliances were found to be well represented by Weibull models. As this paper demonstrates, it is possible to use estimates that capture reasonably well the complexity of the observed distribution while preserving models' simplicity.

The analysis presented shows that residential demand for electricity is more diverse than hitherto assumed. A better idea of how households and individual appliances are distributed with respect to annual consumption provides us with opportunities for further refinement of current probabilistic models. In addition, the results further support the fact that current and future modelling approaches would benefit greatly from the use of larger, highly resolved metered datasets.

3 | Linking intra-day variations in residential electricity demand loads to consumers' activities: what's missing?

Abstract

An increasing interest in the representation of the intra-day variability observed in domestic electricity consumption patterns has driven the development of various modelling frameworks that incorporate consumer behavioural patterns as a key element for the simulation of electricity consumption. Some of the existing models produce reasonable representations of the broader characteristics of user activity patterns, typically rendered as dwelling occupancy patterns. However, when these activity patterns are used to produce estimates of the electricity demand loads associated with such activities, little attention is paid to the links between activities and the actual use of the electric equipment responsible for the production of the actual demand loads. Instead, the simulation of demand loads from activity patterns is ruled by simplifying assumptions that mask the reality behind those links. This paper therefore seeks to unpack the relationship between activity and electricity demand profiles by focussing on the underlying activity patterns in more detail, and how these relate to the usage patterns of the associated appliances. These relationships are studied based on currently available datasets. The analysis of the activities associated with the use of more than one appliance revealed the differences between the likelihood of each appliance being activated throughout the day relative to user engagement in the activity the appliance is associated with. In practice, what this shows is how each appliance's share of the demand load associated with the activity varies throughout the day. As a result, the power consumption associated with a particular activity is subject to the same kind of variability; activity-related demand. The results of this analysis can be used in conjunction with current or new modelling approaches with a view to linking the user activity patterns with the simulation of more realistic electricity demand loads.

3.1 Introduction

Daily demand load profiles are becoming increasingly relevant in the planning and design of energy systems. System flexibility is more costly to deliver in low-carbon systems and more detailed information about the intra-day variations in energy demand could therefore afford more cost-effective infrastructure developments.

In countries like the UK, the transition towards a low-carbon economy may lead to the creation of new major demand loads, such as those derived from the electrification of transport and heating (CCC 2008, DECC 2011). After years of declining electricity demand, which eased capacity constraints, these new loads could pose serious new challenges to capacity provision.

A better understanding of the relationship between user behaviour and electricity demand loads would provide us with an opportunity for improving energy systems' performance. With a more detailed knowledge of the underlying processes that give rise to the observed electricity demand profiles we might be able develop adequate Demand-Side Management (DSM) strategies, which seek to mitigate the need for costly storage or other surplus capacity by shifting the timing and reducing the magnitude of demand (Haney et al. 2011).

DSM is seen by many as a key component of a smart grid. Robust models, capable of producing accurate representations of energy demand loads and their flexibility in relation to the societal functions they fulfil, are crucial to assess the potential of DSM strategies (Torriti 2017, Yao & Steemers 2005).

The branch of energy demand modelling based on the simulation of user activity is an area of energy research that has recently attracted considerable

attention. Several models of residential demand based on occupancy data have been developed over the last decade (Andersen et al. 2014, Lee & Malkawi 2014, Page et al. 2008, Richardson et al. 2008, Robinson et al. 2011, Wilke et al. 2013, Yao & Steemers 2005).

The earliest examples have taken a deterministic approach when it comes to linking the occupancy profiles to the generation of simulated demand loads (Yao & Steemers 2005). The simplicity of this approach yields satisfactory results at the aggregate level, but fails to capture the diversity of demand at household level (Richardson et al. 2010). Probabilistic methods are capable of producing more realistic representations of the variability of metered consumption patterns as demonstrated by (Widén & Wäckelgård 2010) and (Ramírez-Mendiola et al. 2017).

As Figure 3.8(B) illustrates, daily average occupancy would make for a good first-order predictor of demand. However, this figure also shows that at some of the most critical times the two measures diverge. Simulations overestimate the morning ramp up and also the duration of the evening peak. Conversely, baseload demand is significantly underestimated.

As has been pointed out elsewhere, any model seeking to represent electricity demand needs to take a detailed account of the timing of energy-related activities (Flett & Kelly 2017, Torriti 2017). In addition, accurate representations of the timing of the associated demand loads, which result from the use of the associated appliances, plays a vital role in the analysis of scenarios such as the implementation of DSM. The potential future changes in demand make it much more relevant to understand the specific appliance use in more detail, as strategies such as load shifting are highly appliance specific (Flett & Kelly 2017).

This paper thus aims to help further our understanding of the activity-to-demand dynamics by delving into the links between people's engagement in activities and the (potential) appliance usage that results from it. Section 3.3 discusses the key elements of the datasets used for the analysis and the data processing methods. A methodology for comparing activity profiles with electrical profiles was developed and is discussed in Section 3.3.3. We then present activity-specific results and associated appliance use in Section 3.4.1. The emerging discrepancies and their implications for demand models are discussed in our conclusions section.

3.2 Activity-based load modelling background

In the residential load modelling literature there is an emerging effort to link electricity demand with the timing of user activities. This approach focuses on simulating household activity patterns, usually through probabilistic methods. These models tend to be based on data from national time-use surveys. A review of models developed based on this approach can be found elsewhere (Torriti 2014).

The release of the data from the UK time-use survey of 2001 (Ipsos-RSL & ONS 2003) motivated the development of one of the most widely implemented models for residential electricity demand based on activity data (Richardson et al. 2010). This model uses time-use data to simulate daily occupancy profiles which are then used to determine whether certain simulated appliances are activated throughout the day. During states of active occupancy appliance activation is determined based on the probability that members of the household are engaged in the activity the appliance is associated with. The model has been used

elsewhere (Ramírez-Mendiola et al. 2017) to create a simulated counterpart of an empirical dataset (DECC & Intertek Testing and Certification Ltd. 2016). As the analysis presented in Ramírez-Mendiola et al. (2017) [Chapter 2] points out, the simulated households were configured such that each one of them had a counterpart in the empirical dataset. That is, parameters such as household size and average household annual consumption were assigned on a one-to-one basis and the appliance contents of the simulated households were matched as closely as possible with those of the corresponding empirical counterpart. As Figure 3.1 shows, the model provides a reasonable estimate of the overall average daily load profile. However, as the analysis presented in Ramírez-Mendiola et al. (2017) [Chapter 2] reveals, the assumptions about appliance usage in relation to engagement in the activities the appliances are associated with lead to a misrepresentation of the associated demand for electricity.

Models based on the same approach – that is, the conversion of active occupancy into demand for electricity – have been developed for other countries,

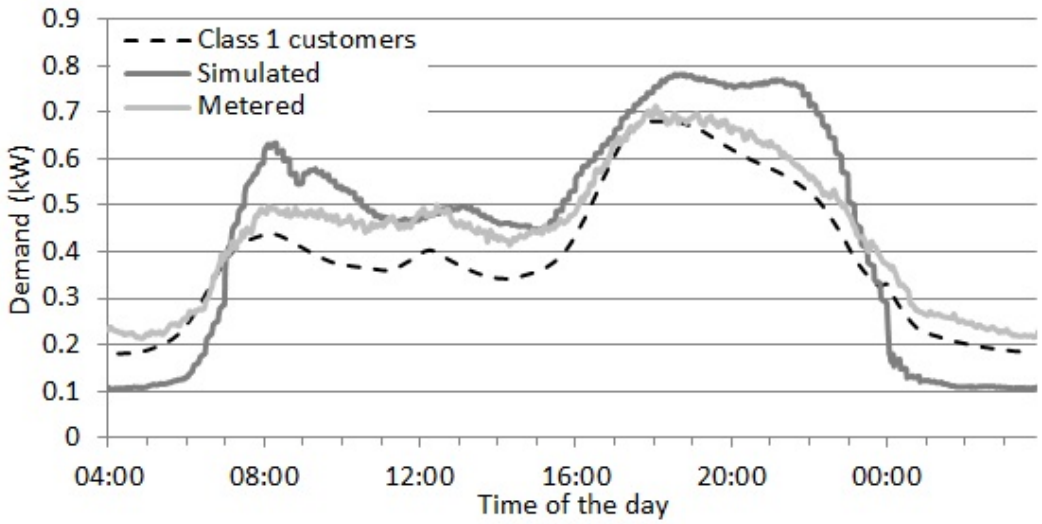


Figure 3.1: *Average daily load profiles. Adapted from (Ramírez-Mendiola et al. 2017). Based on data from (UKERC Energy Data Centre et al. 2013, Zimmermann et al. 2012). Class 1 customers refers to ELEXON's customer classification (see (UKERC Energy Data Centre et al. 2013)).*[1]

namely Belgium (Aerts et al. 2014), Spain (López-Rodríguez et al. 2013) and Sweden (Widén & Wäckelgård 2010). While there are some differences in terms of the way activity patterns are generated, the assumptions about appliance usage are essentially the same; activity engagement is the only proxy for appliance usage. A more detailed classification of models based on this occupancy-to-demand approach can be found elsewhere (Flett & Kelly 2017).

People's engagement in activities is not in itself the cause of the demand loads associated with those activities; rather, it is the use of the electrical appliances associated with or needed in order to perform such activities (see Figure 3.2). Since the model developed by Richardson et al. (2010) was published, detailed datasets on appliance electricity consumption have become available (DECC & Intertek Testing and Certification Ltd. 2016). These new datasets can now be used in conjunction with time-use data to provide new insights into the links between activity patterns and the associated appliance usage.

As has been pointed out in the load modelling literature, small-scale demand modelling needs to account for the influence of household characteristics and the stochasticity of individual behaviours, but it also requires that an accurate representation of the linkage between activity and activity-initiated energy use is achieved (Flett & Kelly 2017, Torriti 2017). To that end, this paper seeks to help provide a better picture of the activity engagement-appliance usage links so that better representations of them can be realised.

[1]The load profile corresponding to Class 1 customers is depicted to serve as a point of reference to show not only how the simulated and metered profiles compare to one another, but also how they compare to the reference profile. Class 1 comprises residential customers supplied on unrestricted tariffs, with a maximum demand below 100 kW as measured by metering systems containing only one meter register. There are eight profile classes, and for each profile class a sample group of electricity supply customers is randomly drawn from the population of electricity supply market customers. These samples are designed to provide an accurate estimate of the load pattern for each class of customers for use in electricity settlement processes.

3.3 Data and methods

Brief introductions to data on activities and appliance usage are given in Sections 3.3.1 and 3.3.2, respectively. This is followed by a description of the methodology implemented for the extraction of the relevant patterns from the source datasets (Section 3.3.3).

3.3.1 Activity data

Time-use surveys are studies concerned with how people spend their time. Their primary aim is to provide a direct measure of the amount of time spent by the population on various activities; these statistics are collected at a national level. Most EU member states have undertaken this kind of studies since the late 1990s following the guidelines of the Harmonised European Time-Use Study. In 2001, the UK Time-Use Survey (TUS) (ONS 2003*b*) which has been the base of the most recent time-use studies and activity-based models for the UK was conducted on behalf of several government departments. The surveyed sample comprises 6,414 households from across the UK. All survey respondents were asked to record their daily activities in detailed 24-hour diaries, producing over 18,000 (adult) diary entries. Each respondent was asked to complete diaries for two days: one weekday and one weekend day. Respondents' activities were logged every 10 minutes starting at 4:00 AM on the day in question. The analysis presented in this paper will focus on three key elements of this dataset: main and secondary activity, and location diaries.

The activity information reported in both main and secondary activity diaries is recorded in the form of a pre-defined set of codes which cover a wide range of

everyday activities (ONS 2003*b*). The two activity diaries combined allow us to get a much better picture of how people spend their time at home; in particular, when they engage in activities that may result in demand for electricity.

Among the surveyed activities, a distinction can be made depending on whether they can be linked to electricity consumption through the use of domestic appliances. For some of those activities it is possible to make a clear distinction. For some others, however, the distinction is somewhat ambiguous. People engaging in these kinds of activities may not necessarily decide/need to make use of appliances. The diagram in Figure 3.2 illustrates this point, taking some of the activities surveyed by TUS as examples.

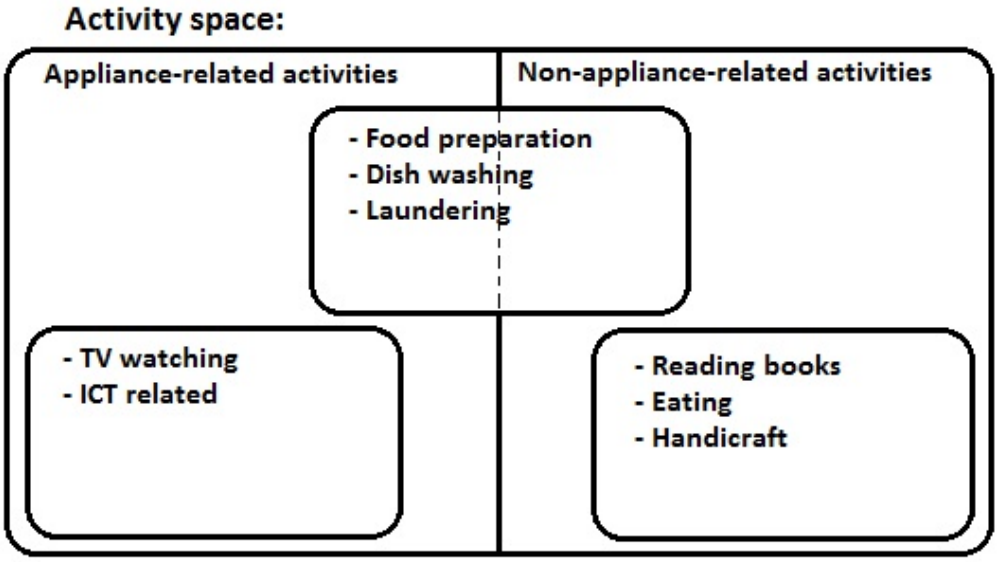


Figure 3.2: *Characterisation of activities according to their relationship with energy use (example).*

The activities on which the present analysis focuses are listed in Table 3.1, along with the monitored appliances that can be (unambiguously) associated with them. A similar mapping between activities and appliances was used by Stankovic et al. (2016). In (Stankovic et al. 2016) the authors make use of

these associations to draw inferences about the household activity profiles based on metered electricity use. They define an 'activities ontology' that maps a set of 16 activities to the different energy-using devices that can be associated with the corresponding activities. The set of activities is grouped into four broad categories, which include activities corresponding to 'daily routines' and 'computing and leisure'. Within the 'daily routines' categories activities such as cooking, laundering, eating, washing, and sleeping are found. Within the 'computing and leisure' category activities such as watching TV, gaming or using the computer are included.

3.3.2 Appliance usage data

Over the period May 2010 to July 2011, daily electricity consumption levels of 250 UK households were monitored and recorded as part of a national study (Energy Saving Trust et al. 2012). The Household Electricity Survey (HES) was jointly commissioned by Defra, DECC and the Energy Saving Trust and its main aim was to monitor total electricity consumption of the surveyed homes as well as that of individual major appliances (Zimmermann et al. 2012). The final report on this study observed that the monitored households were chosen such that the whole sample matched the typical socio-economic mix (Zimmermann et al. 2012). Only owner-occupiers were recruited for the study. The households, however, were still fairly typical in terms of socio-demographic factors (Zimmermann et al. 2012).

The data from this study (DECC & Intertek Testing and Certification Ltd. 2016, dataset) allow us to identify overall daily appliance usage patterns. The dataset is a collection of daily load profiles corresponding to the appliances monitored within the surveyed households. The overall length of the monitoring

period of each appliance varies across households; some households were monitored for a full year, whereas others were monitored for "one-month-long" periods with effective lengths varying between 20 and 30 days.

The monitored appliances can be grouped into categories corresponding to the groups of appliances typically used when performing certain activities (Table 3.1). There are also some appliances which could be broadly categorised as non-activity-related energy uses (e.g. cold appliances). Despite the fact that these appliances remain important in terms of their energy use, they are not significantly affected by the activities of the members of the household. As this study focuses on exploring the linkages between user activities and appliance usage, this kind of appliances was excluded from the analysis.

Table 3.1 summarises the areas in which there are reasonable expectations that a mapping between user activities and monitored energy use exists. It lists the activities that are studied in this paper, as well as the monitored appliances that are being associated with them. The names of the appliances listed correspond to the labels used in HES records.

Table 3.1: *Activities and list of monitored appliances associated with them.*

Activities (from TUS diaries)	Associated appliances (monitored by HES)
Food preparation	cooker, oven, microwave, kettle
Laundering	Washing machine, washer dryer, Tumble dryer
Dish washing	Dish washer
TV watching	TV1, TV2, TV3, CRT TV, LCD TV, plasma TV, Audio-visual site
ICT related	Desktop PC, Computer site, Laptop

As we observed in the previous section, the mapping of appliances onto activities defined in this paper is consistent with the one used in (Stankovic et al. 2016). However, we focus on those activities that are meaningful in terms of households' daily routines and likely to have a significant share of total electricity consumption. The appliances associated with the activities analysed in this paper account for over 40% of the total household annual consumption (BEIS 2016). Therefore, there is great potential for the use of these appliances to reduce both daily and annual household electricity consumption. However, in order to provide a good estimate of such potential, and more importantly to be able to exploit it, it is necessary to have a better understanding of the way these appliances are used and how their use relates to the reported activity patterns.

There are many other activities that may result in demand for electricity. However, these activities may not necessarily be a determinant in the everyday lives of households[2]. Therefore, identifying their corresponding electricity demand loads based on currently available data would prove considerably more complicated. In this context, we refer to an activity as a 'determinant' in the daily routines when there is a reasonable expectation that this kind of activity is likely to be of high priority. That is, some 'less important' activities are likely to be planned around those high-priority activities, and this in turn would determine the overall daily routine/activity pattern. There is always a possibility that this kind of activities may in fact be determinant of the overall daily activity patterns of certain individuals or groups of people. However, based on what is observed in the empirical data, this is rarely the case and therefore

[2]The UK Time Use Survey of 2015 codes the activities of the survey respondents based on a detailed list of 270 activity categories, many of which can be broadly considered energy-related activities. For instance, the category 'cleaning dwelling' could be associated with the use of vacuum cleaners, and the categories 'handicraft', 'woodcraft', 'metalcraft', 'sculpture and pottery' and 'repairs of dwelling' could all be associated with the use of electricity-powered tools. However, this kind of activities are much more likely to be carried out in the spare time of the household rather than being highly prioritised.

neglected.

3.3.3 Extraction of daily activity and appliance usage patterns

Activity and appliance usage data stem from different studies undertaken at different times. However, meaningful comparisons can be made between them focusing on the generic links. In order to do this, two assumptions are made. First, it is assumed that the patterns observed in these two samples are indeed representative of the overall patterns across the broader population; the second assumption is that the potential changes in the timing of activities and appliance usage during the time gap that separates both studies were small enough so as to have the overall patterns preserved.

Daily activity profiles were extracted from the time-use data for each of the activities listed in Table 3.1. At any given 10 min interval of the day, the rate of engagement in a particular activity is given by the number of times TUS respondents reported engaging in that activity at that time over the total number of diary slots in TUS records for the corresponding interval. In other words, it is the probability that a random person would engage in this activity at this time of the day. To produce the activity profiles, a combination of main and secondary activity diaries was used. The need for the use of the two diaries is due to the fact that some activities with a very direct link to electricity consumption, such as TV watching, are often deemed to be secondary activities. In order to account for this possibility, both diaries are analysed simultaneously. In this way, it is ensured that the most accurate representation of people's engagement in certain activities throughout the day is achieved. The activities recorded in the activity

diaries are by no means restricted to a particular location. Therefore, we made use of the location diaries to filter for "in home" activities.

In addition, based on the activities and locations reported in these diaries, overall occupancy profiles were also extracted. These occupancy profiles essentially show the probability of finding a random person engaging in any of three occupancy states considered: active occupancy, passive occupancy (sleeping) and absence. Based on each individual adult diary, an occupancy profile featuring the three occupancy states previously mentioned was created. Then, in a similar manner to the calculation of the activity profiles, the overall occupancy profile was calculated by aggregating all the individual profiles.

Daily appliance usage profiles were extracted from the HES data for each of the appliance groups listed in Table 3.1. The HES dataset is basically a collection of daily appliance-level electricity consumption profiles. Therefore, a direct comparison against the TUS data is not possible. The methodology for the extraction of the appliance usage profiles is different to the one used for TUS data processing, as this involved a more complex treatment of the relevant data. The main goal was to determine when an appliance was being actively used; hence, any baseload-like contributions to the load profile had to be discarded.

For the treatment of HES data two things had to be taken into consideration. Firstly, the length of the monitoring period of each appliance varies, due to the issues discussed in Section 3.3.2. What this means in terms of the analysis is that the number of days each kind of appliance (e.g. "kettle") was monitored across households is different. For this reason, the monitoring period of each individual appliance had to be treated separately.

Secondly, appliances found in different households have different consumption

characteristics, in spite of sharing a common label in HES records. That is, appliances with the same label may have different power requirements, different duty cycles, stand-by consumption and so on. Based on all the available data on each of the monitored appliances, the baseload consumption of each appliance (i.e. stand-by or losses) was estimated. This was then used as a threshold value for determining the moments of the day in which the appliance was actually in use. An appliance is considered to be "on" when its stand-by consumption threshold is exceeded by 50%. This is an extra precaution to account for the fluctuations observed in the power drawn during the stand-by phases of the different appliances. It is not possible to determine whether these fluctuations are metering errors or actual variations in the power drawn by the appliance. However, it is still necessary to account for them in order to avoid false positives.

A daily usage profile was thus generated based on this information for each day of the monitoring period of that particular appliance. These usage profiles are binary sequences, where 1 indicates an active state and 0 a state of inactivity. All the daily usage profiles corresponding to appliances sharing the same label are then aggregated. At any given 10 min interval, the rate of use of a particular appliance is given by the number of times the appliance was being actively used at that particular time over the total length of the monitoring period of that particular appliance in HES records. In other words, it is the probability of finding that type of appliance in use at that time of the day.

The purpose of these profiles is to provide a graphic summary of the likelihood of finding a given appliance in use at a given time of day. Although based on metered appliance electricity consumption data, the profiles should not be taken as probabilistic models for appliance usage as such. The data used for the analysis could also be used to create a simple probabilistic model of appliance usage by

assuming, for instance, a first-order Markov process. To this end, transition probability matrices would need to be extracted for each transition between the time intervals into which the day is divided. For the purpose of the analysis presented in this paper, however, the likelihood profiles presented are sufficient.

For the activities associated with more than one type of appliance, in addition to individual appliance usage profiles, general usage profiles based on the use of all the appliances associated with the activity were extracted. These profiles are meant to provide an estimate of users' engagement in these activities based on the use of the associated appliances. The data processing for this part of the analysis was carried out in a similar manner to that of the individual appliance usage profiles; this time, however, taking all the relevant appliances into account. For example, to extract the general cooking appliances' usage profile, daily usage profiles were obtained for every day in the monitoring periods of every cooking appliance. Then, the general appliance usage profile was derived from the overall aggregate of daily usage profiles and the total number of days in the combined monitoring periods.

By using these profiles we are able to compare the two datasets in terms of activity and appliance use probabilities. In the following section we present the results of such comparisons.

3.4 Results and Discussion

Activity and appliance usage profiles were extracted from the TUS and HES data respectively in order to explore the links between activity and activity-related energy use. First, we present the results of the analysis of the five activities

of interest and their corresponding appliances' usage profiles, and discuss the observed relationships. Next, we use the average occupancy profile extracted from the time-use data to show how the oversimplification of these linkages and the influence that active occupancy has over electricity demand loads may lead to severe misrepresentations when simulating electricity consumption.

3.4.1 Activity profiles and appliance usage

As we discussed in Section 3.3, the activities analysed as part of this study were chosen based on: (1) their general importance in terms of the representation of residential users' behavioural patterns, and (2) the reasonable expectation that such activities could results in electricity use. Table 3.1 summarises the activities relevant to this study, as well as the monitored appliances which were associated with them for the purpose of the analysis.

In the sections below activity-specific results are presented and discussed.

3.4.1.1 Food preparation/Cooking

A common general pattern is observed in the food preparation activity profile and the general cooking appliances' usage profile derived from all the relevant appliances in HES data; as Figure 3.3(A) shows, three peaks approximately centred around mealtimes can be identified. Two main differences between these two profiles can be pointed out. In the general appliances' usage profile the evening peak occurs about 40 minutes later than in the activity profile. Moreover, this general appliances' usage profile would appear to indicate that engagement

in cooking activities throughout the day is more likely than the activity profile would suggest.

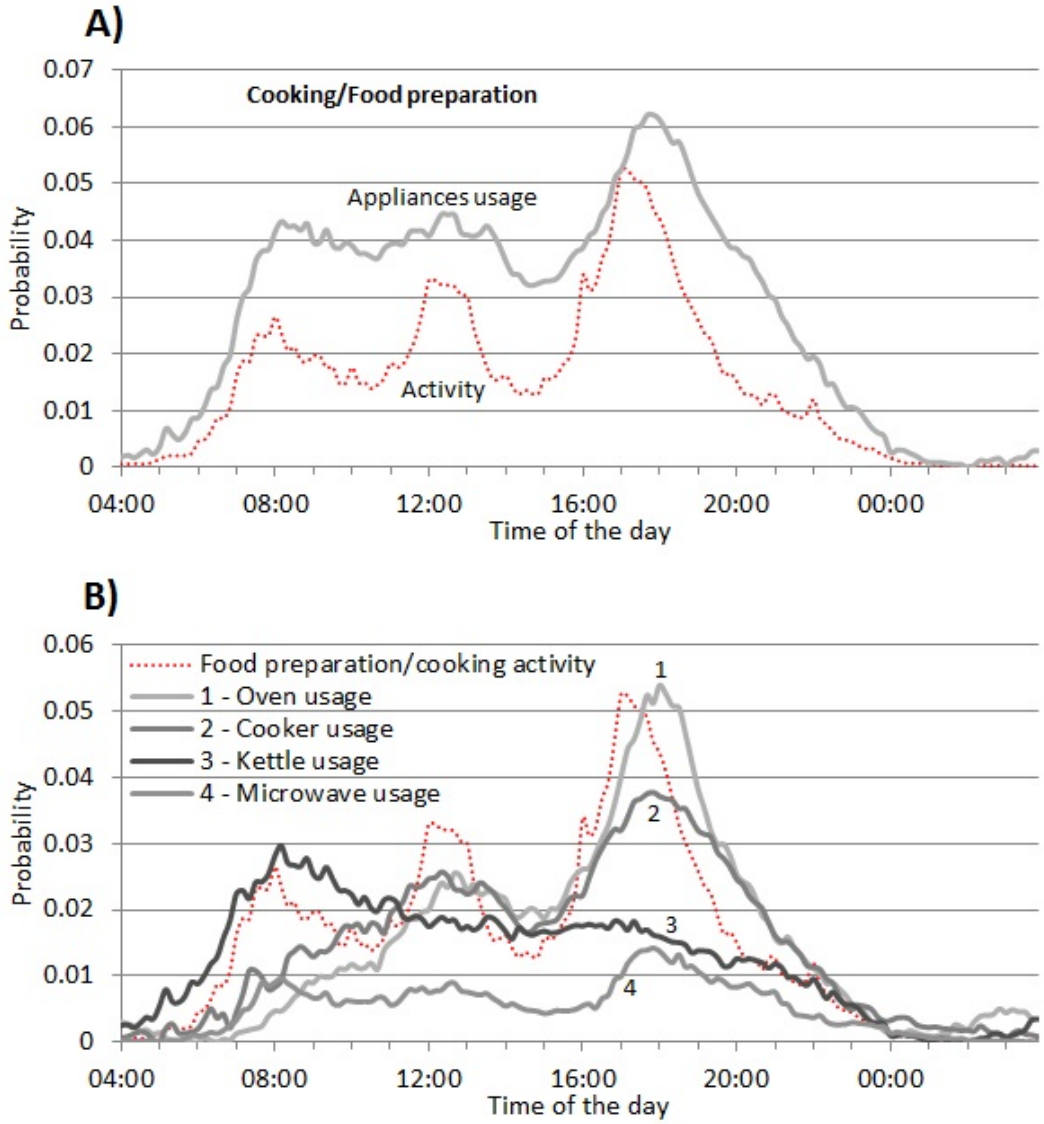


Figure 3.3: *Daily food preparation and cooking activity profiles.* **A)** *General cooking appliances' usage profile.* **B)** *Individual cooking appliances' usage profiles.*

The differences in the likelihood levels are particularly noticeable in the period between 7:00 and 15:00 hr. The higher levels, however, are not necessarily an indication that the amount of people engaging in cooking activities during this time is larger. These discrepancies are rather due to the fact that simultaneous use

of cooking appliances appears to be more common during daytime. This could potentially be attributed to people making preparations for meals in advance. Figure 3.3(B) shows that as dinner time approaches and the final stages of the day's cooking or food preparation activities get under way the use of ovens and cookers ramps up, whereas the use of the other smaller appliances (e.g. kettles) continues to decline.

The usage profiles corresponding to the four main cooking appliances in HES records are plotted alongside in Figure 3.3(B). These are also the appliances primarily associated with cooking activities in a more general context. Their corresponding usage profiles clearly show that the likelihood of each one of these appliances being used throughout the day varies at different rates. The comparison of individual usage profiles makes evident that the contribution to the overall demand load of each one of the cooking appliances varies independently throughout the day. What this could mean in practice is that the 'relative usefulness' of each appliance changes depending on the time of the day. For instance, it is observed that food preparation activities in the morning appear to be strongly correlated with the use of kettles; the use of the other appliances appears to be much less relevant in comparison. In contrast, in the evening, the correlation between food preparation activities and the use of ovens appears to be the strongest.

3.4.1.2 Laundering

Laundering activities appear to be more prominent in the morning hours. Figure 3.4(A) shows that there is a major peak in both laundering activity and appliance usage centred around 10:30 hr, followed by a steady decline towards the end of

the day. As in the case of food preparation activities, a general common pattern is observed in both activity and appliance usage profiles. Two main differences between activity and appliance usage profiles stand out. Both profiles follow each other very closely up until the peak occurs. However, after the peak the general appliances' usage profile would appear to indicate that engagement in laundering activities is more likely than the activity profile would suggest. In addition, the general appliances' usage profile would appear to indicate that there is a second peak, approximately centred around 19:00 hr. The higher likelihood levels observed in the general appliances' usage profile could potentially be attributed to the differences between the perceived duration of laundering activities from a user perspective and the actual duration of an appliance's cycle of operation. Users would only report being engaged in laundering activities when they are actively taking part in the process; loading and activating a washing machine, for instance. From the appliance's perspective, however, laundering activity would be registered from the moment the appliance is activated and until its duty cycle is completed.

As Figure 3.4(B) shows, the individual usage profiles corresponding to washing machines and tumble dryers are essentially different. The peak in washing machines' usage is consistent with the peak observed in the laundering activity profile; both peaks occur roughly at the same time of day. Washing machine usage follows the same general pattern observed in the activity profile. However, the appliance usage profile shows consistently higher levels of likelihood. The observed differences can be attributed to the same phenomenon observed in Figure 3.4(A). Tumble dryers' usage profile shows a different trend to the one observed in the laundering activity profile. Rather than presenting a well-defined peak in the morning, this profile looks more like a plateau with a maximum likelihood period extending approximately from 13:00 to 17:00

hr. This period is followed by a gradual decline towards the end of the day, consistent with what is observed in both the washing machine usage and activity profiles.

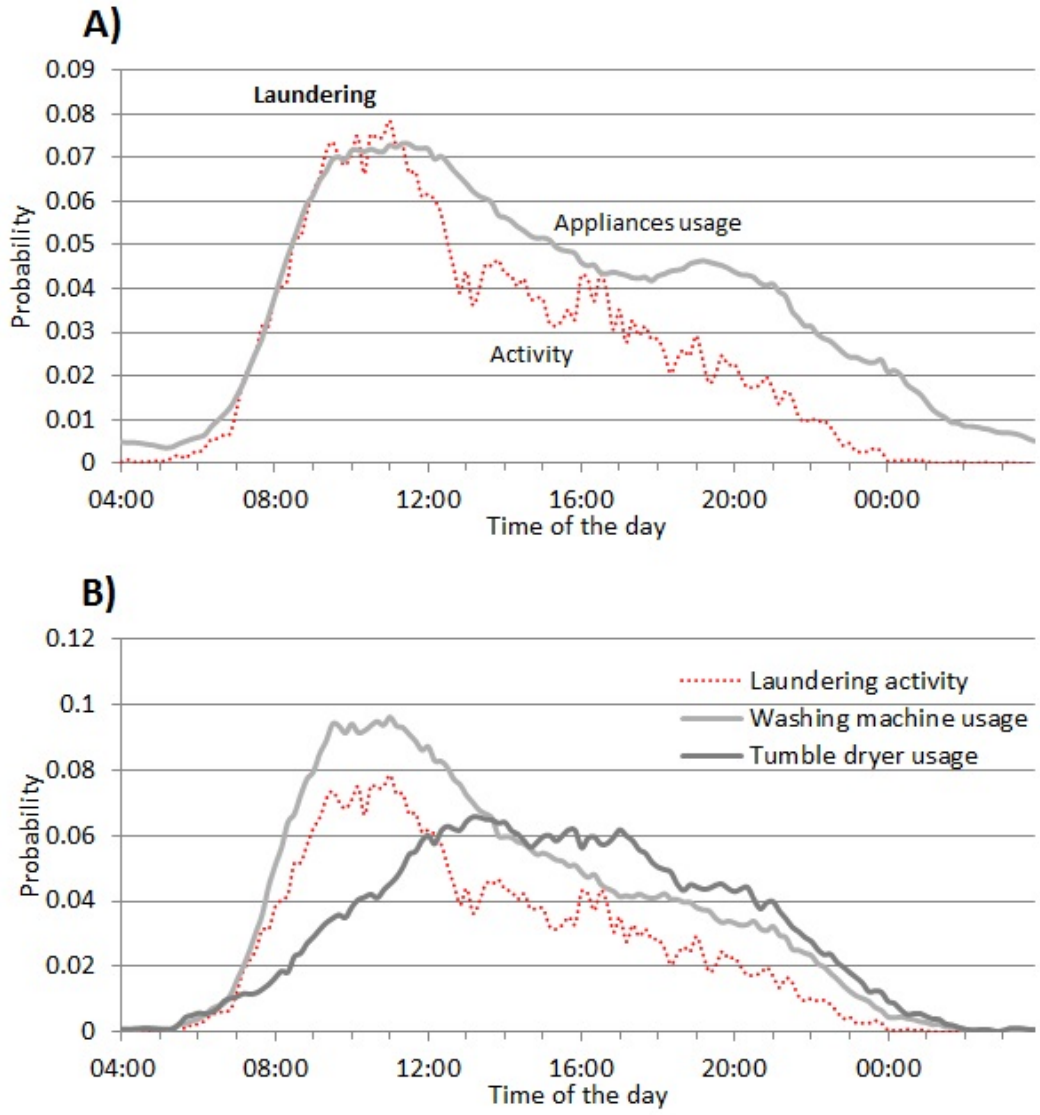


Figure 3.4: *Daily laundering activity profile.* **A)** *General laundering appliances' usage profile.* **B)** *Individual laundering appliances' usage profiles.*

The differences between the washing machine and tumble dryer usage profiles appear to be consistent with what is observed in practice. Firstly, they reflect the fact that when tumble dryers are used they are almost invariably run after

washing machines. Secondly, the generally lower rates of use reflected by the tumble dryer usage profile can also be attributed to the fact that tumble dryers are less ubiquitous than washing machines. In addition, tumble dryers are less commonly used due to the use of other drying methods (e.g. line drying).

3.4.1.3 Dish washing

The dish washing activity profile shows three well defined peaks, clearly centred around typical postprandial times. The dishwasher usage profile, however, as Figure 3.5 shows, appears to be divided into two broader regions: before and after 16:00 hr. The first region exhibits a single peak that is consistent with the morning peak observed in the activity profile. However, the peak in the appliance usage profile occurs about 40 minutes later than in the activity profile. During the period when the midday peak occurs in the activity profile, the usage profile exhibits a more levelled region. As for other appliances, this could be attributed to the effect of the duration of the appliance's cycle of operation.

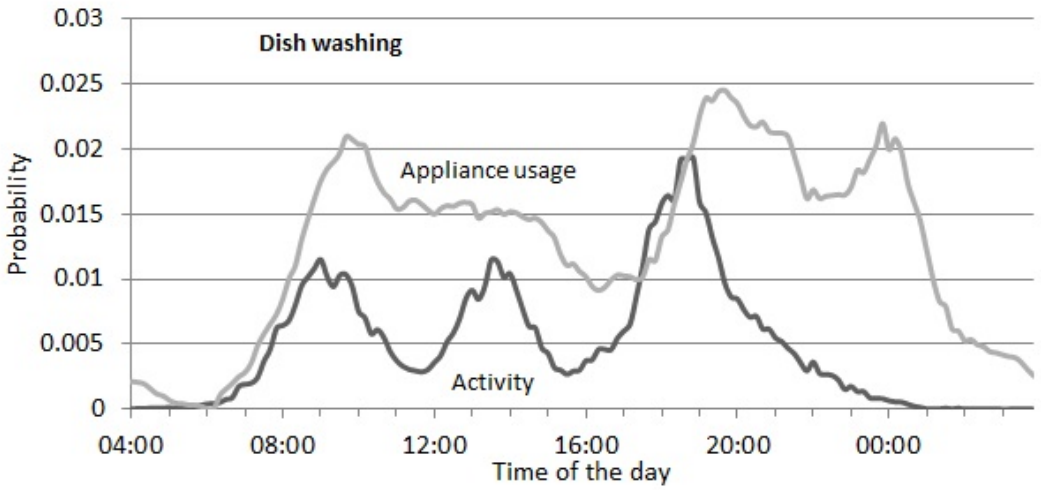


Figure 3.5: *Daily dish washing activity and metered dishwasher usage profiles.*

The second region of the dishwasher usage profile exhibits two peaks. The first one is consistent with the evening peak observed in the activity profile. As the morning peak, this peak in appliance usage occurs about 1 hr later than in the activity profile. Figure 3.5 also suggests that there is a strong tendency among dishwasher owners to run the appliances overnight, as the peak occurring around midnight indicates.

3.4.1.4 TV watching & ICT related

Perhaps not surprisingly, TV watching is most prominent in the evening hours. However, Figure 3.6(A) shows that reported TV watching is significantly lower than the measured consumption from TV sets would suggest. In the first instance, this can be attributed to the systematic under-reporting of the time spent watching TV. It may, however, also be related to the definition of the activity itself.

Participants in time-use surveys are asked to report on what they do, not which appliances are in use. A TV set, which is permanently running in the background while people go about their lives, will not be reported as an activity. Figure 3.6(A) suggests that this under-reporting effect is greatest during the daytime (up to 70%), whereas the evening peak at around 10 pm is captured fairly accurately in time and also better captured in magnitude. The lower under-reporting at peak (25%) could be a result of more people 'actively' watching TV.

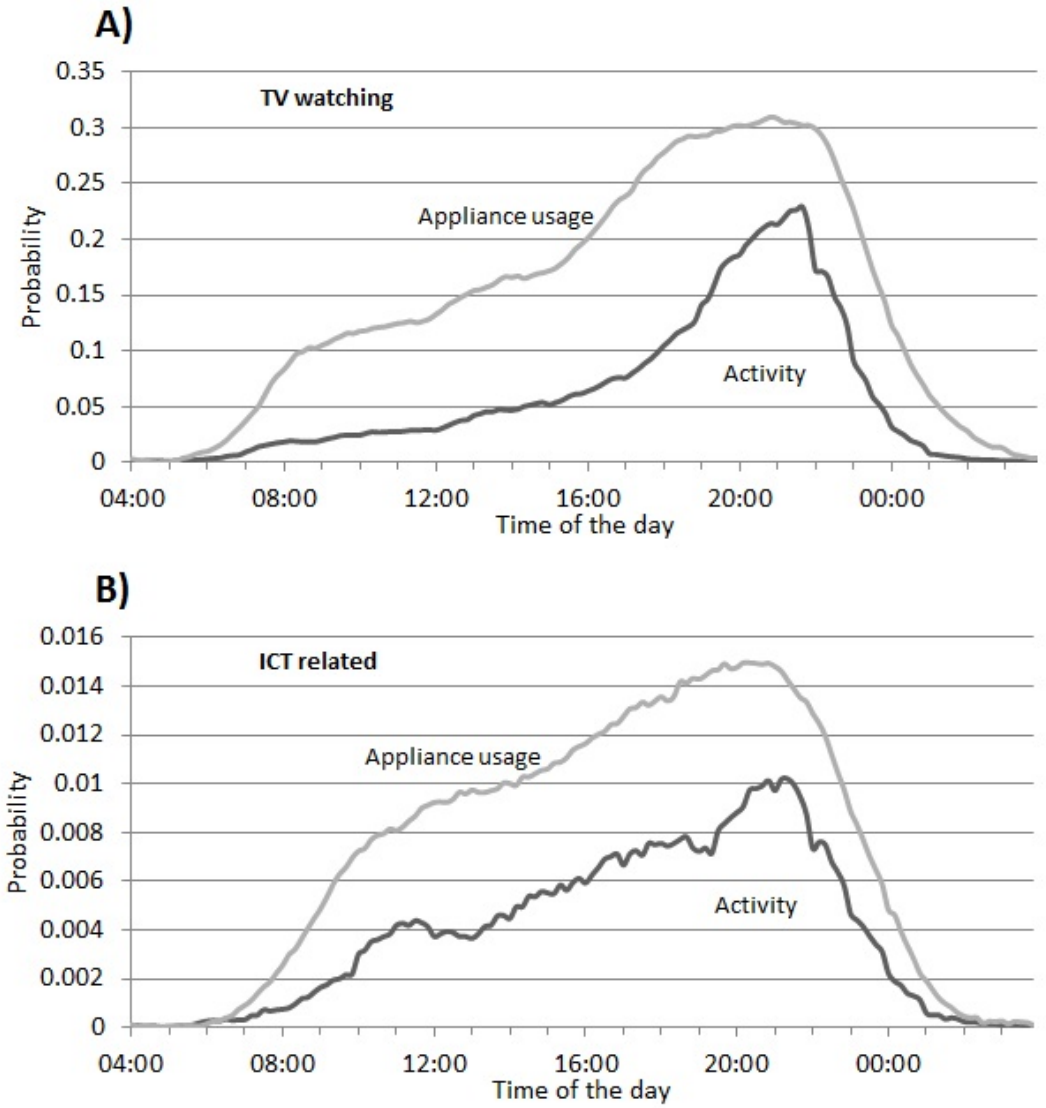


Figure 3.6: *Daily TV watching and ICT related activity profiles.* **A)** *TV watching activity and metered overall TV usage profiles.* **B)** *ICT related activity and metered overall IT appliances' usage profiles.*

3.4.2 Dependence of simulated power consumption on active occupancy

Previous studies have pointed out that the daily electricity load profile is strongly affected by changes in the occupancy of dwellings throughout the day. A combination of the activity and location diaries from TUS data allows us to determine the proportion of people likely to be engaged in a certain occupancy state at each 10 min interval of the day. The resultant average daily residential occupancy profile has a distinctive shape that translates into clear indications about people's preferences when it comes to where they spend their time at different times of day.

These preferences, of course, are subject to some change; the differences in these preferences manifest themselves, for instance, when the average occupancy profiles corresponding to weekdays and weekend days are compared (see Figure 3.7). In each of these profiles, three shaded regions can be observed. Each one of those regions corresponds to one of three basic occupancy states: active occupancy, passive occupancy (sleeping) and absence. The vertical axis indicates the probability of finding a random person engaging in any of the three states considered.

The differences in the typical everyday activities for the two types of days are evident from these profiles. As Figure 3.7 shows, despite the daily occupancy profiles having roughly the same shape, the probabilities of finding a random individual at home at a given time of the day differ. For instance, on average, it is about 20% more likely to find a random person at home between 9:00 and 15:00 hr on weekend days than on weekdays. These differences also provide a measure

of people's tendency to get up later on weekends and stay in longer. Figure 3.7(A) shows that on weekdays there is a short, sharp morning peak centred around 8:00 hr.

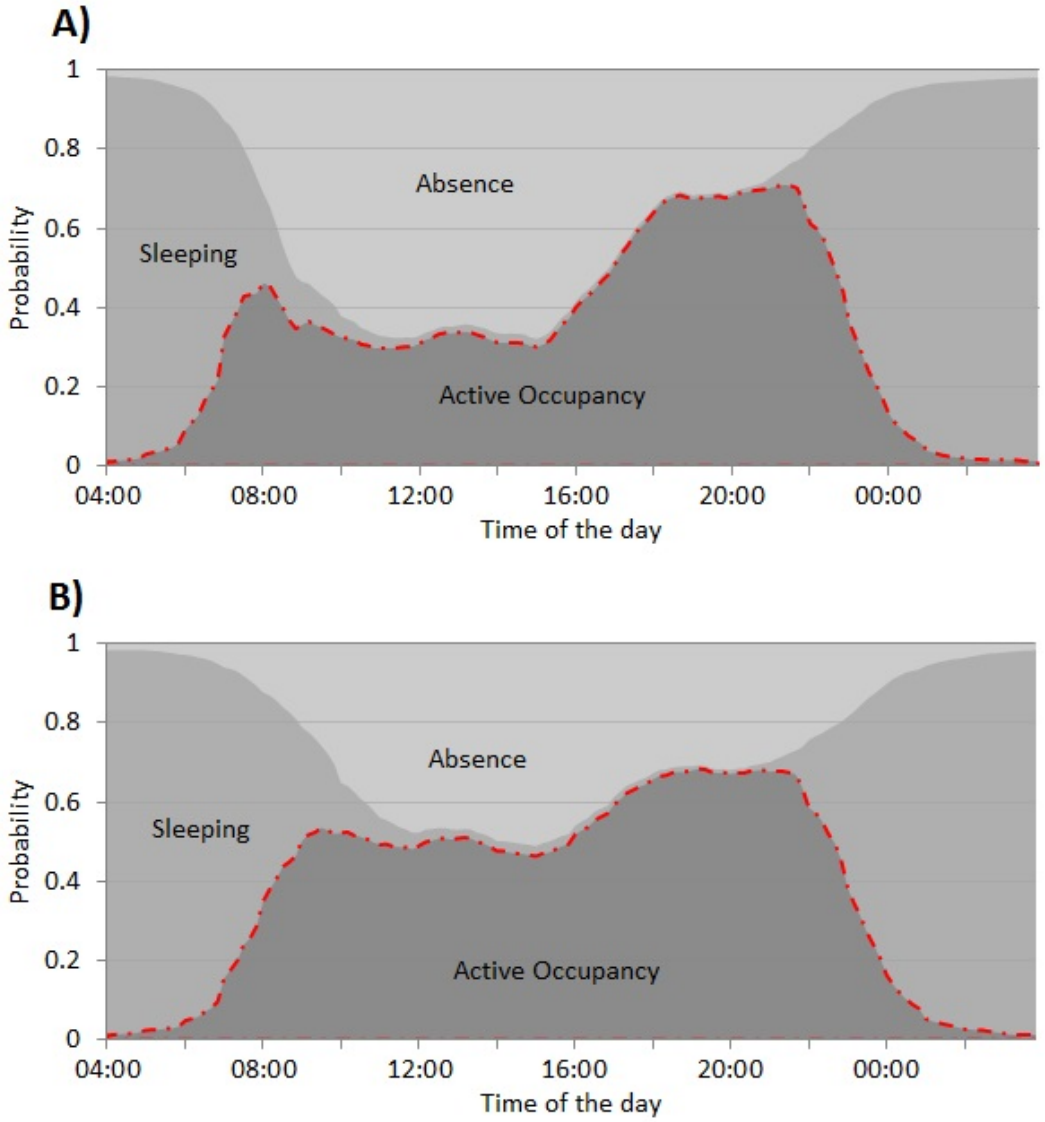


Figure 3.7: *Average occupancy:* **A)** *Weekday average occupancy profile.* **B)** *Weekend average occupancy profile. The different shaded areas represent, at any given 10 min interval, the overall probability that an individual is in one of three occupancy states: Active at home, sleeping or absent.*

The period leading to the observed peak is consistent with the typical transition period from being asleep to being active, preparing for the start of

the day. The peak is followed by a sharp decline, corresponding to the sudden change in occupancy that occurs when people leave home for work, school, or any other activities taking place at a different location. In contrast, Figure 3.7(B) shows that on weekend days the transition from being asleep to being active appears to be less sharp, with the peak in active occupancy showing up about 90 minutes later than in weekdays. Moreover, the decline in the likelihood of finding a random person asleep as the morning progresses is more gradual during weekends.

The times at which the evening peak is observed are also different. In both types of days, the likelihood of finding a random individual active at home gradually increases as the evening approaches, arriving at a peak of about 70%. On weekdays the peak is observed from 18:00 to 21:30 hr, whereas on weekend days it extends from 19:00 to 22:00 hr. Towards the end of the day, the decline in the likelihood of finding a random individual active is accompanied by a considerable increase in the probability of finding them asleep in both profiles.

The differences observed between the average occupancy profiles corresponding to the two types of days are relevant for a more detailed analysis of the daily occupancy patterns or for the refinement of occupancy models. These differences also have effects in terms of the observed electricity consumption, of course. For the purpose of illustrating and discussing the relationship between occupancy and the daily electricity demand load profile, however, a comparison between the overall average occupancy profile (Figure 3.8(A)) and the average electricity load profile (Figure 3.8(B)) will suffice.

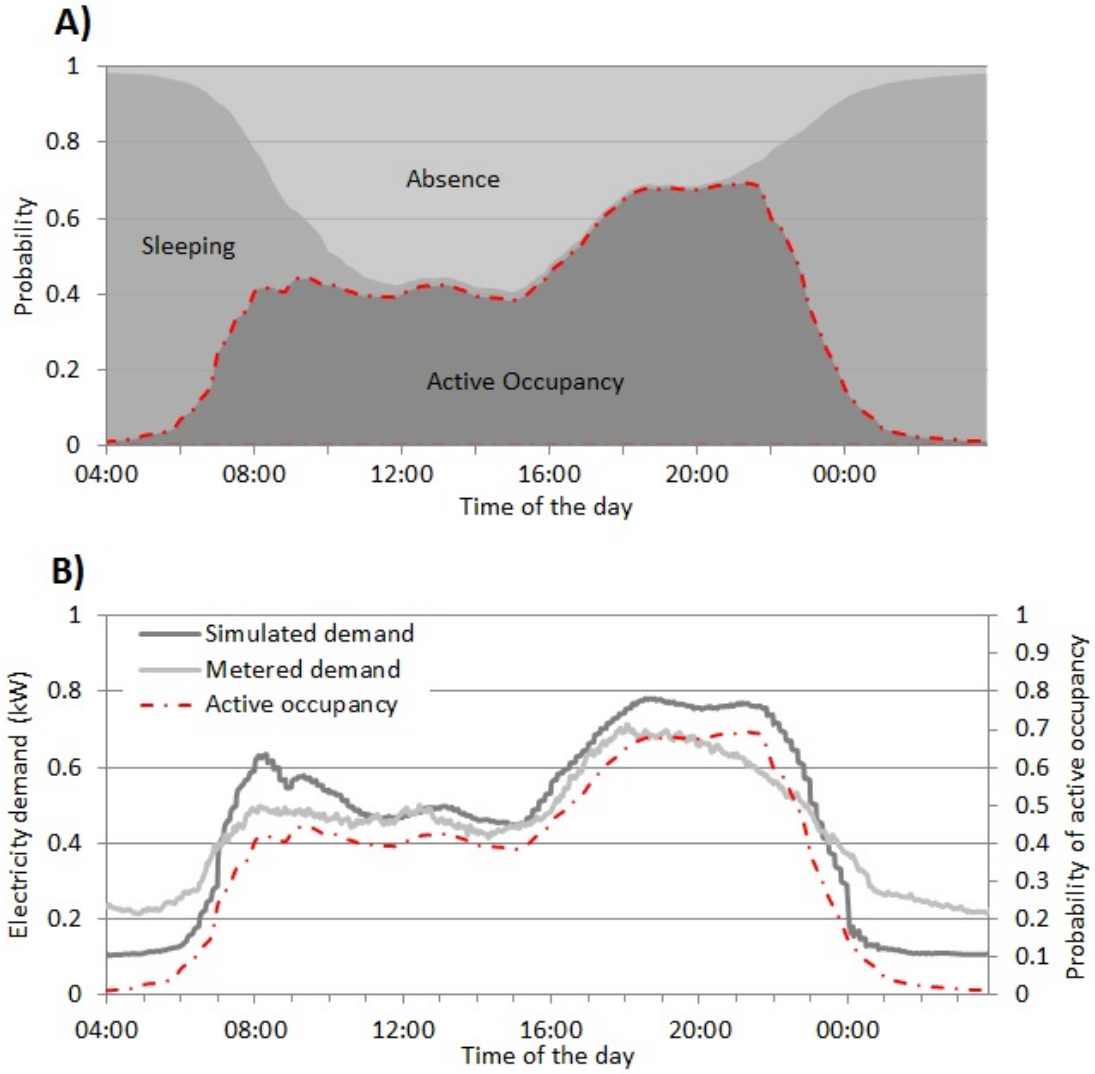


Figure 3.8: **A)** *Average occupancy profile.* **B)** *Comparison between overall active occupancy profile and average daily load profiles.*

In Figure 3.8(B) metered and simulated load profiles, and average occupancy profile are superposed to allow for an easier contrast. In this figure the strong correlation between active occupancy and demand for electricity can be clearly appreciated. Clearly, daily average occupancy would make for a good first-order predictor of demand. That is, a reasonable estimate of the overall daily load profile could be obtained by simply multiplying the average occupancy (a dimensionless quantity) by a linear (first-order) power consumption factor.

However, it can also be observed that the strength of this correlation varies throughout the day.

To further discuss this, let us divide the load profiles into three regions characterised by the level of demand: lowest, medium, and highest demand. If the duration of these periods is compared pairwise, it is observed that the durations of the first two regions are very similar in both metered and simulated load profiles. Incidentally, this also seems to be in agreement with the duration of the equivalent periods in the active occupancy profile. However, looking in more detail at the region corresponding to the evening peak, discrepancies can be observed in terms of the duration of this period in the different profiles. As Figure 3.8(B) shows, the duration of the peak period of the simulated demand profile is virtually the same as the duration of the active occupancy peak. In contrast, in the metered demand profile, the duration of the period during which the maximum demand is held is roughly half as long as in the simulated profile.

A related issue has to do with the over-estimation of the influence that active occupancy has over demand loads. A potential indicator of this can be identified by noting that the shape of the simulated load profile follows more closely the shape of the average daily occupancy profile than that of the metered profile. Furthermore, demand levels, especially during peak periods, are over-estimated in the simulated load profile.

Special attention should be paid to the most energy intensive appliances, such as those typically used for cooking. The over-representation of the use of one or more of these appliances can result in the kind of effects observed in the simulated profile used as example here. For instance, the simulated profile in Figure 3.8(B) shows a morning peak considerable higher than the one observed in the metered

profile. This would appear to be consistent with an over-representation of the appliances that, as we showed in the previous section, are not as heavily used as in other times of the day.

3.4.3 Activity and activity-related consumption patterns in the literature

Over the last decades, several studies on time-use data have provided us with some insight into the ways people go about their lives. As discussed in Section 3.2, this kind of studies has also provided opportunities for the development of models that seek to capture the stochasticity observed in the empirical behavioural data. While the simulation of activity patterns appears to have become the focus of attention for this kind of studies, the link between engagement in activities and actual generation of electric loads, that is, appliance usage, seems to be largely neglected.

Few attempts have been made to better understand the activity-related electricity consumption (Anderson 2016, Durand-Daubin 2013, Sekar et al. 2016, Torriti & Santiago 2016, Vassileva et al. 2012). However, these studies have focused on the analysis of electricity consumption aggregates over certain periods of time (e.g. day, month, year), rather than on the intra-day variations. This has probably been largely due to the lack of high-resolution appliance-level consumption datasets. In addition, these studies have also focused on single activities associated with single appliances.

For instance, the analysis presented by Torriti & Santiago (2016) focuses on an activity in particular: TV watching. The authors used a rather simple model

to produce estimates of aggregated electricity demand based on the number of households potentially engaged in TV watching at any given time of day. As in the other modelling examples (Section 3.2), the model used in (Torriti & Santiago 2016) assumes that activity data on TV watching, or any other single-appliance activity for that matter, is good enough a proxy for the activity-related electricity demand so that it can be used to estimate demand with a simple linear model. As the results of the analysis presented in this paper clearly show (See Section 3.4.1.4), this is not necessarily the case. There are effects that simplistic linear models cannot account for, such as the observed background TV usage which can be inferred from the metered TV usage. As we discussed in section 3.4.1.4, both the activity engagement and appliance (TV) usage profiles resemble each other in terms of overall shape. However, by comparing the two profiles it becomes evident that the data on TV watching activities is not enough in itself to produce an accurate estimate of the metered TV electricity consumption.

In the study carried out by Anderson (2016), the daily laundering activity patterns extracted from time-use data from different years (1985 and 2005) were compared. The aim of this study was to analyse the evolution of the laundering activities throughout the years, hoping that this would reveal ways in which practices in the residential sector have changed, and more importantly, ways in which these could be modified to achieve specific outcomes, which is one of the main aims of DSM strategies.

The analysis revealed that there have been significant changes in terms of the days of the week during which laundering activities take place. However, as the author himself points out, whilst the changing distribution of laundry across the week offers insights into the evolving nature of laundry overall, it is the nature of its timing within the day which is likely to be of greater concern in the DSM

context.

In addition to the relevance of the timing of engagement in laundering activities, the results for the associated appliance usage presented in this paper (See Section 3.4.1.2) point out the need to account for the fundamentally different usage patterns observed for washing and drying machines in relation to engagement in laundering activities.

3.4.4 Limitations and further research

Due to the nature of the kind of studies from which the data used for the analysis stems, the results of the analysis are by definition country-specific. While there are some clear limitations to the data used for the analysis presented in this paper, the results provide a number of insights into the relationship between engagement in activity and actual appliance usage.

Activity patterns are, in general, different in different countries and therefore, it only makes sense to compare them to consumption patterns from the same country. However, it has been pointed out elsewhere that there is a high level of similarity between the active occupancy profiles of the 15 European countries that have taken part in the Harmonised European Time-Use Survey (HETUS)[3] (Torriti & Santiago 2016). It is therefore natural to expect that there might be some similarities in the activity patterns, and perhaps in the associated appliance usage as well.

Given the relevant data, i.e. large enough appliance-level electricity consumption datasets, similar analyses could be carried out for different

[3]HETUS - Participant countries: Italy, Spain, Estonia, Latvia, Lithuania, Poland, Bulgaria, Slovenia, UK, France, Germany, Belgium, Finland, Norway and Sweden.

countries, which would provide opportunities for further discussion of country-specific results in terms of the discrepancies or similarities observed across countries, and what these entail.

What this study tries to point out is the need to refine our assumptions about activity-related appliance usage in the context of electricity demand modelling in the residential sector. As the results of the analysis show, even appliances associated with the same activity may have completely different daily usage profiles, which is something that models developed to date have failed to take into account. Whatever the reasons might be for the overlooking of these dynamic links between engagement in activities and appliance usage, the results of the analysis presented in this paper show that we are now in a position to better represent these links, and therefore, generate better estimates.

3.5 Conclusions

In this paper, the links between residential consumers' behavioural patterns and appliance usage patterns were explored to improve our understanding of the intra-day variations in residential electricity demand loads.

We presented a methodology for the extraction and comparison of probability profiles derived from time-use data and appliance-level electricity consumption.

The analysis of the discrepancies between these two profiles allows us to challenge commonly held assumptions about the relationship between activities and electricity consumption.

Four types of activities with high relevance to energy consumption are presented in detail. These are food related, laundering, dish washing and TV watching.

The energy intensity of food preparation varies throughout the day. Both the number of appliances involved in food preparation activities and the level of simultaneous use change depending on the time of the day. One therefore cannot use the activity 'food preparation' as the only proxy for electricity consumption, irrespective of the time of day. The analysis further shows that the type of appliances predominantly in use varies throughout the day. The use of high power/low energy demanding kettles is predominant in the morning, whereas evenings see a predominant use of ovens, which entails significantly higher energy consumption.

Dish washing is a good example for time shifts between activity and appliance usage. The reported activity follows an expected pattern featuring peaks after morning, midday and evening meals, with evening dish-washing being twice as common as daytime. Dishwasher usage does not follow this pattern. The peaks in appliance usage occur later than those in the reported activity, which is to be expected, since 'starting the dish-washer' will lead to consumption later on. Yet, at midday this sequence is not preserved. Instead the dish-washer cycle appears to be delayed until after the evening meal, and some of this is delayed even further to operate at the end of the day.

Laundering appliances' usage has the same 'delay effect' as dish-washers', but the intra-day activity pattern is a more reliable indicator of the link between the two. The strong sequencing of washing machine use and tumble drying is evident from the data.

TV watching revealed potentially systematic shortcomings in the link between reported activities and appliance usage. We conclude that a combination of under-reporting of TV watching and inactive use of TVs in the background both lead to a profound gap between activity and appliance usage data during daytime.

The analysis presented leads to conclude that the relationship between activities and load profiles is not as simple as current models assume, and more importantly, provides the elements for the development of new models that are able to provide a better representation. Using activity data as the basis for the simulation of load profiles requires careful additional calibration, or entirely new methodologies, to account for the variety of effects observed here. In some cases, like for food preparation activities, we have shown that considering the time of day as an additional factor could improve the predictive power of such models. In other cases, such as dish-washing, the relationship is more complex, due to the storage function of a dishwasher to accumulate dishes before being run.

In summary, this paper provides an example of the kind of analysis that is required in order to devise robust linkages between household activity patterns and activity-initiated electricity consumption, and provides further evidence for the need to further research these linkages.

4 | Residential activity pattern modelling through stochastic chains of variable memory length

Abstract

Residential activity modelling has attracted considerable attention over the last years. This is particularly due to the fact that residential energy demand loads are highly dependent on the activity patterns of the household. Therefore, activity models are being increasingly used to underpin high-resolution energy demand models. This paper details the implementation of a new methodology for the analysis of empirical activity data that allows for the identification of characteristic behavioural patterns within them. The identified patterns are then used as the basis for the construction of a high-resolution residential user activity model. The model attempts to capture the statistical characteristics of the empirical data in the form of a stochastic process with memory of variable length. The proposed model is compared to a model based on the predominant first-order Markov chain approach. In addition to the modelling approach, a new metric for assessing the quality of activity sequences simulations is proposed. Given the amount of empirical data contained in any of the individual time-use datasets currently available, it would appear that the performance improvement over the predominant first-order Markov chain approach is modest. However, the validation results show that the proposed approach has the potential for broadening our understanding of the scheduling of activities in people's day-to-day lives and how this relates to the observed variability in both activity and energy consumption patterns.

4.1 Introduction

The past couple of decades have seen a substantial increase in the interest in the branch of energy research that looks at the dynamics of energy demand. In particular, considerable efforts have been made to include the associated behavioural aspects into the modelling of residential demand loads through the simulation of people's activity patterns (Torriti 2014, 2017).

In the context of a transition towards a more decentralised electricity generation system, in which generation technologies with highly variable outputs are expected to play a major role, it is clear that a complete reworking of the relationship between supply and demand will be needed. Rather than solely expect the supply to adjust to the variations in demand loads, the active management of the dynamics of such demand loads is increasingly seen as a key component of future low-carbon systems (Torriti et al. 2010). In addition, it is acknowledged that experimenting with this active demand management in ways beyond conventional approaches is necessary in order to explore the potential for energy efficiency enhancements and to better assess the feasibility and problems of the proposed management strategies (Darby & McKenna 2012, Peacock & Owens 2014, Strengers 2012). However, in order to effectively explore these issues, more needs to be known about the dynamics and temporalities of the underlying behavioural patterns that give rise to the observed demand loads (Walker 2014).

Central to the approach that places activities at the base of our understanding of the dynamics of the observed residential demand loads is the idea that the way activities are ordered in time determines the timing of demand (Shove 2004,

Torriti 2017). Therefore, a better understanding of the scheduling of activities in people's day-to-day lives and how this relates to the observed variability is of the essence. However, up until now, activity modelling has largely focused on reproducing the overall features of mean activity profiles which are averaged over large numbers of individual daily activity schedules. While attempts have been made to further unravel the complexity of activity scheduling (Aerts et al. 2014, McKenna et al. 2015, Torriti 2014, Wilke et al. 2013), current approaches to the simulation of activity patterns still pose some serious limitations when it comes to learning more about daily activity scheduling.

In particular, the probabilistic methods used in these approaches produce an oversimplified representation of the dynamics observed in empirical activity data. Empirical activity schedules are observed to evolve in numerous ways. Moreover, the evolution of these daily activity schedules is arguably ruled by a self-adaptation process in which the parts of the schedule that are already in place, or a subset of them, play an important role in determining the subsequent steps in the evolution process. These complexities manifest as a wide behavioural diversity, and a probabilistic model aiming to produce a close representation of this diversity would have to attempt to capture such complexity.

However, the stochastic processes represented by current approaches are first-order approximations to the behavioural diversity observed in empirical activity data. Thus, the diversity in the range of behaviours that is observed in the simulated output of one of these first-order models results naturally in a narrow normal distribution around a mean activity pattern. Therefore, the behavioural diversity that results from simulating the observed dynamics by means of such simple processes is equally narrow. The simplicity of the formulation of these first-order models might increase their appeal. However, their simplifying assumptions

mask the complexity of the dynamics of the observed activity patterns.

If the ultimate goal is to develop a better understanding of the scheduling of activities in people's day-to-day lives, an approach to the simulation of daily activity patterns that takes account of the observed complexities would be greatly advantageous and indeed necessary.

In this paper we introduce a novel approach to the analysis and simulation of activity data based on variable memory length stochastic models. The idea behind the development of this approach is addressing the issues associated with the use of the predominant Markov chain technique. Particularly, its inability to take into account the influence of more than one state in the chain of events to determine the subsequent steps in the evolution of daily activity schedules. At the heart of the proposed approach is the analysis of the empirical activity data with a view to identifying characteristic patterns present in the empirical activity sequences so that they can be exploited in the simulation of daily residential activity scheduling. The primary intended purpose of the approach is the simulation of activity schedules to underpin the simulation of residential electricity consumption patterns. Therefore, the activities considered are restricted to those in-home activities that can be reasonably associated with the use (or non-use) of electric domestic appliances. The simulated activity sequences are constructed as the result of a series of transitions in which the likelihood of transitioning into the different activity states considered depends on the sequence of activity states observed prior to the transition in question.

The paper is structured as follows. Section 4.2 presents a brief and general review of current approaches to activity-based electricity demand modelling. In Section 4.3.2 we introduce the concepts relevant to the modelling approach

developed in this paper. The algorithmic process that allows for the identification of activity sub-sequences relevant to the transitions into subsequent states is detailed in Section 4.3.3, and the implementation for the analysis of time-use data is detailed in Section 4.3.4. The validation of the proposed model is presented in Section 4.4 where the activity data simulated by the model are compared with: 1) the data produced by a model based on the commonly employed first-order Markov chain technique, and 2) the empirical activity data contained in the 2015 UK Time-Use Survey. Finally, we conclude by discussing the potential of the approach introduced in this paper, its shortcomings and potential for improvement.

4.2 Activity-based demand modelling: a brief overview

In the residential electricity demand modelling literature there is a number of examples of models that attempt to simulate activity patterns with a view to incorporating them into the demand modelling process. Some of the first examples of this take a deterministic approach like the one used by Yao & Steemers (2005). Modelling approaches have evolved since and the most recent examples are generally based on the use of empirical activity datasets such as those provided by time-use surveys and make use primarily of the Markov chain technique. A detailed review of the different models and modelling approaches based on time-use data can be found elsewhere (Torriti 2014).

4.2.1 Time-use data

Time-use datasets are the best currently available sources of empirical data on the activities carried out by people throughout the day. These datasets come from large-scale surveys that look specifically at how people spend their time. The survey samples are typically based on households and the sample selection procedures aim at gathering the most statistically representative sample of a given population. Central to this kind of studies is a time-diary instrument in which respondents record their activities at regular time intervals. These time-diaries record the activity sequences for prescribed periods, which usually correspond to a single day. These surveys based on the collection of time-diaries are an effective means of gathering comprehensive datasets on how people spend their time, where do they spend it, and who they spend it with. Respondents are typically asked to complete time-diaries for one weekday and one weekend day in order to account for differences in time-use for the different types of day.

4.2.2 Markov chain technique

A commonly adopted approach to the simulation of stochastic sequences of potential events is based on the representation of these processes as Markov chains. Every step in the evolution of the sequence will result in one of the potential states considered. The set of different potential states is also referred to as the state space of the stochastic process. A Markov chain model consists of this state space and the probabilities associated with the transitions from each of the states into the others.

An abstraction of the way people go about their lives is thinking of them as if they were transitioning between the different elements of a set of potential states of activity. Based on this idea, it is thus possible to simulate the evolution of daily activity sequences as a Markov chain where the state space of the Markov process corresponds to the set of potential states of activity considered. The simplest example of this kind of model consists of a Markov process associated with a binary state space where the only potential states of activity considered are 'active' and 'inactive'. An example of such a model is the occupancy model developed by Richardson et al. (2008). An extension of these model was later introduced by McKenna et al. (2015). The proposed extension included one more activity state, but the technique used for the simulations remained the same.

The stochastic processes that can be adequately described by a first-order Markov chain model are said to satisfy the Markov property. Thus, they are characterised by what is referred to in the statistical literature as memorylessness. In this context, the concept of 'memory' refers to the number of previously observed states at a given moment in the evolution of the stochastic process. Thus, term memorylessness refers to the fact that previously observed states are not relevant to the evolution of the process. That is, predictions for subsequent events are made based on the current state only. When the stochastic process of interest shows no potential for the presence of complex patterns or particularly distinctive features, a well calibrated first-order Markov chain model can provide a good description of it.

In the case of residential daily activity scheduling, it would be reasonable to assume that the process that describes the observed dynamics is somewhat more complex.

Recognising the limitations of the first-order Markov chain approach, new approaches that aim to capture more of this complexity have been developed. Examples of these are the approaches developed by Wilke et al. (2013) and Aerts et al. (2014). The proposed models are variations of the first-order Markov chain models that are known in the statistical literature as semi-Markov models. The key difference between these and simple first-order models is the fact that in a simple first-order model the length of the blocks corresponding to the states in the chain are fixed, whereas in a semi-Markov model these are allowed to vary as a means of altering the 'duration' of the state. Although this class of models do represent an improvement over the conventional first-order Markov models, they still fail to address the memorylessness issue that characterises the simple first-order models. Therefore, the development of approaches that aim to address this issue would prove greatly advantageous, and is indeed necessary if we are to develop a better understanding of residential activity patterns.

4.2.2.1 Higher-order Markov models

In terms of improving the approach to the simulation of activity patterns, a natural step in this direction would be the development of higher-order Markov models. This kind of models are interesting for the purposes of activity modelling as they attempt to incorporate an element of 'memory' into the simulation of the stochastic process in question.

There are, however, serious issues associated with the use and development of this kind of models. Firstly, the number of free parameters in the model increases exponentially as a function of their order. The higher the order of the model the larger the amount of data required to construct the model. Consequently,

as the number of parameters that need to be estimated and the amount of data that needs to be processed increases, so does the computational intensity associated with both constructing and using the model. Secondly, and more importantly, the collection of all possible full high-order Markov chain models is limited and completely stratified. That is, there are no models with a number of free parameters in between those defined by the order of the models. This is illustrated in Table 4.1, where the columns correspond to the levels (strata) represented by the orders of the model.

Every element of the set of models that represent a given stochastic process has an associated error. The sources of this error are the model bias, the variance of the simulated output, and an irreducible component associated with the noise of the data used to construct de model (Hastie et al. 2009). The bias is associated with the order of the model and the variance is associated with the number of parameters present in the model. A model with high bias will result in a poor representation of the relevant statistical characteristics of the empirical data; a condition known in the modelling literature as underfitting. A model with high variance will result in simulations overly sensitive to small fluctuations in the empirical data; a condition known as overfitting. Thus, in order to achieve an adequate representation of the stochastic process in question, the model has to strike a good trade-off between bias and variance.

The model bias decreases as the order of the model increases. The variance increases as the dimension (number of parameters) of the model increases. A consequence of such discontinuous and drastic increase in the dimension of the high-order Markov chain models is that, given a real process for which a given set of states is observed, it is not possible to obtain a parsimonious representation of the sate space. That is, the lack of models 'in-between' the exponentially

increasing dimensions of the full high-order Markov chain models does not allow for an adequate balancing of the bias-variance trade-off.

The class of full higher-order Markov models offers good alternatives for the representation of theoretically consistent stochastic processes. From the empirical estimation point of view, however, the issues discussed above render this kind of models rather ill-suited for the study of many 'real-life' processes, including the daily scheduling of activities. Therefore, alternative approaches to the simulation of activity patterns still need to be sought.

Table 4.1: *Model dimension stratification as a function of the order.*

Order (n)	1	2	3	4	5
Dimension	56	448	3,584	28,672	229,376

**The model dimension is given by* $Dim = (|S| - 1)\,|S|^n$ *, where* $|S|$ *is the size of the state space and* n *is the order of the model. For the example presented in this paper (see Section 4.3.4), the size of the state space is* $|S| = 8$ *(8 different activity states).*

4.2.3 General approach to activity-to-demand conversion

Activity modelling for the purpose of electricity demand modelling has focused on the simulation of home occupancy patterns. Periods of active occupancy are then used as prompters for estimating the engagement in the different activities that may result in or affect the observed electricity consumption patterns. The reader is referred to Torriti (2014) for reviews of concrete examples of this approach. The implicit assumption underlying this approach is that the engagement in the different types of activities is a consequence of being in a state of active home occupancy (i.e. not asleep).

4.2.4 Insights from social practice theory

The literature on social practice theory points out that in order to effectively analyse people's daily schedules these should be treated as a whole (Shove et al. 2012, Torriti 2017). Therefore, for the sake of consistency, the study of the sequences of activities corresponding to these schedules should be carried out in the same manner.

The allocation of time to the different activities depends on different factors, but there are certain factors that prevent certain activities from happening at certain times due to the inconvenience they may cause, or the actual impossibility of the action. The scheduling of activities is restricted to some extent by the social constructions of time. For instance, the vast majority of people leave home for work or school at roughly the same time during workdays, lunch takes place at around midday and so on. These restrictions in turn result in a certain degree of synchronisation of the societal activities, which gives rise to the development of characteristic habitual sequences (behavioural patterns) (Walker 2014).

These characteristic patterns determine to a great extent the way energy demand loads arise. As Torriti (2017) points out, energy demand in households is shaped by the fundamental temporal characteristics of social practices, which include time dependence, duration and sequence. Moreover, as the kind of activity or activities being performed by the members of the household dictates the amount of energy required, the rate at which it is consumed (intensity), and the rate at which it varies in time, a more thorough analysis of empirical activity data is becoming increasingly relevant for the power sector. A better understanding of the way demand loads arise and vary would be most helpful for the purpose

of future network planning and tackling peak demand issues through effective implementation of Demand-Side Response mechanisms.

4.2.5 Towards a more robust approach to activity modelling

The simplicity of the Markov chain technique, as well as its ability to provide a reasonable representation of the overall features observed in aggregate daily activity profiles, have led it to become the current activity modelling paradigm. However, as we discussed in Section 4.2.2, this approach by definition disregards the influence characteristic patterns might have on the evolution of the stochastic process being modelled. When it is suspected that this kind of structural complexity might exist, as it is arguably the case of daily human activity scheduling, seems appropriate to try and learn more about the structural dependencies that might be present in the empirical data available.

Time-use datasets are, to date, the best source of empirical activity data (see Section 4.2.1). Therefore, it is reasonable to expect that the kind of characteristic patterns described in the social practice theory literature might be present in the reported activity sequences. In fact, the quality of the data is such that, if the sole purpose were to get a perfect snapshot of the current behaviour patterns, a very large set of activity sequences could essentially work as a model. This might seem like a feasible solution to the complications associated with activity modelling. However, such an approach presents major drawbacks. One of them is that the studies from which the empirical activity datasets stem are very expensive and they are difficult to carry out. But a more important drawback is the fact that in taking such an approach we lose the opportunity to learn more about the

behaviours that these activity sequences represent. The development of activity modelling approaches allows us to devise ways in which the characteristics of the behaviours observed in these activity sequences can be quantified. This in turn allows us to develop methods to study how these might change. There is, therefore, great incentive to develop a robust approach to the simulation of activity patterns.

An activity modelling approach that focuses on explicit engagement in activities rather than the representation of periods of home occupancy would give us the opportunity to learn more about the scheduling of daily activity and identify characteristic patterns. Moreover, an approach that seeks taking into account the influence observed characteristic patterns have on the development of daily activity sequences might help: 1) produce better representations of the different behaviours observed, and 2) gain a better understanding of how associated energy demand loads arise.

In the context of activity modelling, the approach to the analysis of the data used as a basis for the model is of great importance. Therefore, if we want to be able to exploit the information provided by characteristic activity patterns we need to integrate this goal into both the approach to data analysis and the development of activity modelling techniques. New approaches to data analysis that focus on the study of the underlying activity patterns observed in empirical data would allow us to build on the strengths and insights provided by previously developed activity modelling approaches and assist in the development of new ones that take into account the influence of such characteristic activity patterns. The approach presented in this paper attempts to do this.

4.3 Data and methods

4.3.1 The 2015 UK Time-Use survey

The latest major time-use dataset for the UK became available in 2017. It comes from the national survey carried out from April 2014 to December 2015 (Gershuny, J. & Sullivan, O. 2017). The survey was designed such that it would follow the guidelines set out by Eurostat on Harmonised European Time Use Studies (HETUS) (EUROSTAT 2009), and the data collected would be compatible with the previous major time-use survey carried out by the Office for National Statistics in 2000 (Ipsos-RSL & ONS 2003).

This study provides information about how individuals aged 8 years and over in the UK use their time on a given weekday and a weekend day. The collected time-diaries provide information about activities, location, co-presence, the use of computers and mobile devices, and even the level of enjoyment of time throughout the day. Questionnaire data provides information about characteristics of individuals and households such as employment and education, and demographic information such as age, gender, marital status, citizenship status and housing.

Respondents were asked to log their activities every 10 minutes. The responses were then re-coded based on a list of over 250 pre-defined activity categories. In addition to the main activity at any given time of day, they are able to specify secondary activities as well as location. Time-diaries are grouped into two broader categories: adult and children diaries. Filtering for adult diaries only we end up with $\sim 1.5 \times 10^4$ individual diary entries. All of this data

was used to construct the activity models used for the analysis presented in this paper.

4.3.2 Variable memory length models

Incorporating an element of 'memory' into the simulation of activity patterns is a non-trivial issue. Taking first-order Markov models as the starting point, a natural way of achieving this would be through the development of higher-order Markov models. However, as we have pointed out, the use of these models poses two main problems: the number of parameters increases exponentially with the order of the models, and the stratification of the dimension of the models does not allow for an adequate balancing of the model fitting to the empirical data (see Section 4.2.2.1).

The two problems discussed above can be addressed by a conceptually simple idea: allowing the order of the model to vary as the stochastic process evolves.

In the context of activity modelling, this varying order (memory length) proves particularly useful as it allows us to take account of the self-adaptation of the scheduling of activities throughout the day. Self-adaptation refers to the fact that, as the daily sequence of activities evolves, the probability of observing a given state further down the line might change. Thinking about a particular activity, say laundering, this could be thought of as the effect of having fulfilled the laundering duties for the day. While, on average (fixed-order model), the likelihood of engaging in laundering activities might still be relatively high, if the day's laundering duties have already been 'ticked-off' then the probability of having to do laundry again might not be the same now (variable memory model).

The implementation, however, presents some challenges. The way the order varies is determined as a function of the 'past'. That is, for a given transition in the course of the evolution of the process, the order of the model that would best describe such transition is determined by the sequence of states observed prior to the transition in question. The number of relevant prior states, that is, the length of the relevant 'past', depends in turn on the transition.

When the order of the model is allowed to vary in this way, we effectively end up with a hierarchy of embedded models of different order with a well-defined structure of variable memory length. In practice, this means that a number of the transitions that would be described by the models embedded in the hierarchy are collapsed into transitions associated with a lower-order model that are deemed equivalent. The result of this is an overall reduction of the size of the state space. Consequently, the state space associated with the variable memory length model can be considerably smaller than the space associated with the full higher-order model corresponding to the highest order considered in the variable memory length model. The extent to which the size of the state space can be reduced depends on the nature of the stochastic process that wants to be modelled.

As Bühlmann & Wyner (1999) point out, if the data for the stochastic process that wants to be modelled allows for the construction of such variable memory length models, there is nothing to lose but only to gain in comparison with the class of ordinary Markov chain models.

Variable memory length models originated in the field of information theory (Rissanen 1983). However, the concepts and methodologies developed for the purposes particular to this field have been expanded and adapted for a variety of purposes across different fields (Bühlmann & Wyner 1999, Ron et al. 1997, Willems et al. 1996).

Such variable memory length models apply generally well to problems involving categorical data. It is therefore sensible to assume that these models can be applied to the study of the categorical time series that correspond to the observed activity sequences in time-use data. Moreover, a variable memory length model has the potential to be used as an exploratory tool for the dynamics of the simulated processes. The variable memory structure allows for capturing the structural dependencies observed in the sequences of states resulting from the process, which in turn allows for a better representation of them. In the context of daily activity sequences, this would allow for a better representation of the underlying behavioural patterns observed in the empirical data.

The key problem when constructing a variable memory length model resides in finding an effective method for identifying the relevant portions of the 'past' based on their influence on the outcomes of the transitions to subsequent states. This can only be done based on local, pairwise comparisons between the states present in the hierarchy of nested models. To this end, a variety of algorithms have been developed that allow for the comparison of the statistical characteristics observed for the transition from states associated with models based at different levels of the hierarchy. Various examples of these algorithms can be found in the literature where they feature in greatly varied applications ranging from information theory (MacKay 2003, Rissanen 1983, Ron et al. 1997, Willems et al. 1996) to bioinformatics (Bejerano & Yona 2001, Leonardi 2006, Miele et al. 2005) to statistical physics (Anderson 2007, Lampoudi et al. 2009) to language pattern analysis (Brown et al. 1992, Charniak & Eugene 2000).

In this paper we derive a modified version of one of such algorithms and implement it to construct a variable memory length model of daily activity patterns. In the implemented algorithm the criterion for discriminating between

relevant past states is based on the Kullback-Leibler divergence and the log-likelihood ratio test. The process involved is detailed in the following section.

4.3.3 Algorithm

The aim of the algorithm is to adaptively estimate the lengths of the past sequences of activity states that are relevant to determining the transitions into the next state. As explained in the previous section, these relevant past sequences can be of variable length ranging from 1, which would correspond to the first-order model, to a certain maximum length ℓ_{max}.

A conditional transition probability distribution is associated with each one of these relevant pasts. This distribution determines how likely it is that the next element in the sequence corresponds to each one of the activity states considered. The probabilities are estimated based on the number of times the relevant past sequences are observed followed by each one of the activity states.

Let us denote by t the index of the corresponding time interval of the day. The resolution of the time-use diaries is 10 min. This means that the day is partitioned into 144 intervals. Therefore, in this particular case, $t = 1, 2, ..., 144$.

Let $\mathcal{S} = \{s_\alpha, s_\beta, s_\gamma, ...\}$ be the set of activity states that make up the daily activity sequences.

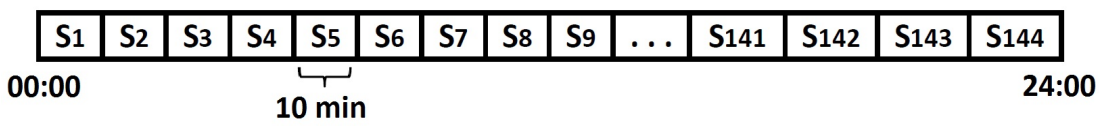


Figure 4.1: *Graphical representation of a daily activity sequence. $S1 \cdots Sn$ =states in the sequence*

Let $A_d = s_1 s_2 s_3 \cdots s_t \cdots s_{144}$ denote a given daily activity sequence, where $s_a = s_\alpha, s_\beta, s_\gamma, ...$ for any s_a in A_d, and let $s_{t-\ell}^{t-1} = s_{t-\ell} \cdots s_{t-2} s_{t-1}$ denote a given subsequence of A_d preceding the state s_t. This sequence corresponds to a 'past' of length ℓ for the state s_t.

We denote by $N_t\left(s_{t-\ell}^{t-1}\right)$ the number of occurrences of the 'past' denoted by $s_{t-\ell}^{t-1}$.

Assuming that the 'past' corresponding to the sequence $s_{t-\ell}^{t-1}$ exists, that is, it has been observed in the reported activity sequences, we have that $\sum_{s_a \in \mathcal{S}} N_t\left(s_{t-\ell}^{t-1} s_a\right) > 0$, where $s_{t-\ell}^{t-1} s_a$ denotes the activity sequence that results from concatenating the past $s_{t-\ell}^{t-1}$ and the subsequent state s_a.

Therefore, the estimates of the probabilities that make up the transition probability distributions associated with the past in question are given by

$$p_t\left(s_t \mid s_{t-\ell}^{t-1}\right) = \frac{N_t\left(s_{t-\ell}^{t-1} s_t\right)}{\sum_{s_a \in \mathcal{S}} N_t\left(s_{t-\ell}^{t-1} s_a\right)}\,. \tag{4.1}$$

The set of transition probabilities obtained for each one of the past sequences considered provide us with probability measures that can then be used to determine how relevant a particular 'past' is for determining the transitions into subsequent states compared to the past states associated with models at lower levels in the hierarchy.

Given two probability distributions defined over the same state space, the Kullback-Leibler divergence provides us with a measure of the difference between them (MacKay 2003). In the context of information theory, this metric is used

to estimate the amount of information lost when one distribution is used to approximate the outcomes resulting from using the other.

If $P_1(x)$ and $P_2(x)$ are the two probability distributions in question, the Kullback-Leibler divergence is given by

$$D\left(P_1 \mid\mid P_2\right) = \sum_x P_1(x) \log\left(\frac{P_1(x)}{P_2(x)}\right) . \tag{4.2}$$

Based on this, we can then estimate the relevance of a given past state by multiplying this measure by the occurrence count of said state (Bühlmann & Wyner 1999). That is, the function for assessing the potential of a model with extended memory $\ell \geq 1$ is defined as

$$\Delta_t\left(s_b s_{t-\ell}^{t-1}\right) = \sum_{s_a \in \mathcal{S}} N_t\left(s_b s_{t-\ell}^{t-1}\right) p_t\left(s_a \mid s_b s_{t-\ell}^{t-1}\right) \log\left(\frac{p_t\left(s_a \mid s_b s_{t-\ell}^{t-1}\right)}{p_t\left(s_a \mid s_{t-\ell}^{t-1}\right)}\right) , \tag{4.3}$$

where $s_b s_{t-\ell}^{t-1} = s_b s_{t-\ell} \cdots s_{t-2} s_{t-1}$, and

$$p_t\left(s_a \mid s_b s_{t-\ell}^{t-1}\right) = \frac{N_t\left(s_b s_{t-\ell}^{t-1} s_a\right)}{\sum\limits_{s_n \in \mathcal{S}} N_t\left(s_b s_{t-\ell}^{t-1} s_n\right)} \tag{4.4}$$

Given a complete activity sequence $A = s_1 s_2 s_3 \cdots s_t \cdots s_m$, we denote by $\mathcal{P}_\ell\left(A\right)$ the likelihood function estimate of that sequence based on the model of fixed memory length ℓ, where

$$\mathcal{P}_\ell\left(A\right) = \prod_{t=\ell+1}^{m} \mathcal{P}\left(s_t \mid s_{t-\ell}^{t-1}\right) . \tag{4.5}$$

Therefore, when individual states associated with models of fixed memory lengths $\ell < \ell'$ are being compared, many terms cancel each other out due to the multiplicative structure of the likelihood function estimate; the only remaining terms are the ones concerning the transition in question. If it were feasible to construct the full models of fixed memory length ℓ and ℓ', and obtain the likelihood function estimates for each one of their associated states, we could then re-state the discrimination criterion as

$$\Delta_t \left(s_b s_{t-\ell}^{t-1} \right) = \log \left(\frac{\mathcal{P}_{\ell'}(A)}{\mathcal{P}_{\ell}(A)} \right) . \tag{4.6}$$

What this shows is that the discrimination criterion is essentially a log-likelihood ratio test, only that the acceptance region is extended from [0,1] to [0,K]. The upper bound of the acceptance region for the discrimination criterion is chosen by asymptotic considerations. If n denotes the size of the empirical dataset (number of empirical activity sequences), and $|S|$ denotes the size of the set of activity states that make up the observed activity sequences, this upper bound can be estimated by

$$K \sim C \log(n) , \tag{4.7}$$

where C has to satisfy the condition $C > 2|S| + 4$. The derivation of this bound is provided elsewhere (Bühlmann 2000).

All the past states for which the value of the discrimination criterion falls within this region represent an improvement over lower-order models. The closer this associated value is to zero, the higher the relevance of the past state considered.

For the purposes of the algorithm, the extended log-likelihood ratio test (Eq. 4.3) acts as a gain function that will determine whether it is worth looking (further) back in the past in order to determine the probability distributions for the transitions into subsequent activity states. That is, whether it is worth extending the model with initial memory length $\ell = 1$ (first-order model) into a model with memory length $\ell' > \ell$ (higher-order model) for the observed set of transitions in question.

The log-likelihood ratio is a statistic, as it depends on the data. The test based on this statistic consists in comparing the value of this statistic to a critical value, or calculating the corresponding p-value. The null hypothesis of the test is rejected if the p-value of the statistic is too small, or the critical value is exceeded. The threshold is determined by the statistical significance level chosen for the test.

For the log-likelihood ratio test to be meaningful it is required that the models that are being compared are nested. That is, that the more complex model(s) can be transformed into the simpler one by imposing constrains on the parameters. As explained above, the variable length model is essentially an approximation of a full higher-order Markov chain model. The 'order' of the approximation obtained corresponds to the maximum length of the memory sequences ('pasts') deemed relevant. By restricting the length of the memory sequences considered for the generation of the model, we can downgrade the 'higher-order' nested models into the base first-order model.

As Eq. 4.6 shows, in the log-likelihood ratio the numerator corresponds to the likelihood of the observations being consistent with the extended memory (higher-order) model. This is the null hypothesis for the test. The denominator corresponds to the maximum likelihood of the observations for the different activity states.

It has been proven elsewhere that algorithms based on the log-likelihood assessment criteria converge asymptotically to the most suited variable memory length model (Bühlmann & Wyner 1999).

4.3.4 Implementation

Focusing on the applications to the analysis of time-use data, we then start by re-arranging the data contained in the time-use diaries such that it is consistent with the formalism presented in the previous section. As observed in Section 4.3.1, the activity data in time-use diaries is coded into a fairly large number of categories. The set of pre-defined codes often correspond to very specific cases of the same kind of activity. Therefore, these can be regrouped into more general categories. We can thus summarise the information in the time-diaries by converting them into sequences of more generally defined activities performed throughout a given day. These sequences are thus composed of a series of 144 random categorical variables[1] with a common activity state space. Namely, we consider the following eight activity states: absence, sleep, generic active occupancy, food preparation, dishwashing, laundering, TV watching and ICT related activities. We take the first letter of each of these categories as the corresponding label. That is, $S = \{a, s, g, f, d, l, t, i\}$ is the state space of the individual transition probability distributions[2].

As we observed in Section 4.1, we intend to apply the simulation of activity sequences to the simulation of residential electricity consumption patterns.

[1] A random categorical variable is a type of statistical variable that can take on one of a limited, and usually fixed, number of possible values, assigning each observation to a particular category on the basis of some qualitative property, i.e. a property that cannot be measured or assigned a numerical value.

[2] The elements of the state space are mutually exclusive. That is, it is assumed that only one activity is taking place at each 10 min interval.

Therefore, the activity states relevant to the analysis are those which correspond to activities that can be reasonably associated with the use (or non-use) of electric domestic appliances. A characterisation of the activity categories and the domestic appliances associated with them can be found in Ramírez-Mendiola et al. (2018) [Chapter 3].

The starting point is the construction of the first-order model which will be the base of the variable memory length model. Based on the empirical activity sequences contained in the time-use dataset, we extract the transition probability distributions corresponding to this first-order model for each of the state transitions in the activity sequences.

Once the first order estimate is available, the algorithm recursively looks for the observed pasts of lengths larger than one for each of the transitions in the activity sequences. At the same time, estimates for the transition probability distributions associated with the corresponding past are calculated. The maximum length the algorithm is allowed to look back into the past can be estimated based on the characteristics of the empirical dataset used to construct the model. Given the sample size, that is, the number of empirical activity records (n), and the number of activity categories considered ($|S|$), the estimate of the suitable maximum length is given by

$$\ell_{max} = \frac{\log(n)}{\log\left(|S|\right)} \ . \tag{4.8}$$

A derivation of this estimate can be found elsewhere (Weinberger et al. 1995). Using the values corresponding to the 2015 UK time-use data, $n \sim 10^4$ and $|S| = 8$, we have that this maximum length estimate is $\ell_{max} \sim 5$.

In order to prevent model overfitting we take the following two measures.

Firstly, we impose a restriction over the identified pasts in terms of their occurrence count. That is, we only take into account those pasts for which the condition $N_t\left(s_b s_{t-\ell}^{t-1}\right) \geq N_{min}$ is satisfied. The value of the lower bound for the occurrence count is in fact arbitrary. However, we choose this value to be $N_{min} = |S| = 8$ so that for any given past there is at least one possibility of observing a transition into each one of the eight activity states considered.

Secondly, we implement a probability smoothing step in the algorithm. Generally speaking, no single transition is absolutely impossible. Therefore, if we estimate the transition probability distributions based solely in terms of occurrence counts the risk of overfitting increases[3]. The transition probability distribution smoothing process occurring during this step helps ensure that no activity state is predicted to have a null transition probability, regardless of the past observed before it.

We start by defining a minimum transition probability which we will denote by p_{min}. If when the transition probabilities are estimated it is determined that the value of the estimate is zero, we replace this null probabilities by p_{min} and then re-normalize the non-null estimates. In analytical terms, if we denote the new transition probability estimates by $p'_t(s \mid \cdot)$, the new set of transition probabilities has to satisfy the condition

$$\sum_{s \in S} p'_t(s \mid \cdot) = 1\,. \tag{4.9}$$

That is, the new set of transition probabilities has to be an actual probability distribution.

[3]Given a particular empirical dataset, it is possible to find past states for which the transitions to subsequent states are restricted to a subset of the activity states considered.

If we denote the re-normalization constant by α, we can then express the new transition probability estimates in terms of the previous ones as

$$p'_t(s \mid \cdot) = \alpha p_t(s \mid \cdot) + p_{min}\,, \tag{4.10}$$

so that when $p_t(s \mid \cdot) = 0$, the value of the new estimate is p_{min}. Therefore, when we impose the condition stated above in Eq. 4.10, we find that the value of this constant is $\alpha = 1 - |S|\, p_{min}$.

Since the value of α has to be greater than zero, this in turn imposes a restriction over the value of p_{min}, which is $p_{min} < \frac{1}{|S|}$. That is, p_{min} can take any value within the range $\left(0, \frac{1}{|S|}\right)$.

Thus far, all model parameters have been estimated based on the characteristics of the dataset used to build the model. For the sake of consistency, we determine the value of this parameter based on these same elements. Therefore, during the smoothing of the transition probability distributions all null probabilities will be replaced by

$$p_{min} = \frac{N_{min}}{n} \sim 5 \times 10^{-4}\,, \tag{4.11}$$

and the non-null ones scaled accordingly. By estimating the value of p_{min} in this way we ensure that 1) the value is within the appropriate range, and 2) the value does not exceed the probabilities estimated directly from occurrence counts.

When a past has been identified, and its corresponding transition probability distribution has been estimated, the discrimination criterion described in Section 4.3.3 is used to determine whether the model would benefit from including said past in the state space of the transition in question. The process is repeated for

all possible pasts of different lengths and for all the time intervals of the activity sequences.

The algorithm was implemented in the Python programming language[4] (Ramírez-Mendiola 2018). The resulting set of parameters for the model was then exported to a file, and arranged in a particular format so that they could be reimported later for the purpose of simulating the sets of activity sequences used for the validation analysis.

4.4 Validation analysis results

The aim of the validation of this kind of models is to verify that the data generated by the stochastic process represented by the model possesses similar statistical characteristics as the empirical data used to construct the model.

Among the examples of occupancy and activity modelling found in the literature, a statistical measure commonly used for the validation of such models is the overall likelihood of finding a given person engaged in one of the occupancy/activity states considered throughout the day. These are commonly called state probabilities, and they vary with respect of time of day.

As we observed in Section 4.1, the motivation for developing the model presented is to produce a better representation of the characteristics of time-use data than what can be achieved by a first-order Markov model. The primary

[4]The original implementation makes use of standard modules only. As the number of calculations that need to be performed is large, the implementation makes use of Python's multiprocessing capabilities. This is by no means necessary, but it certainly speeds up the data processing.

aim of the proposed approach is to exploit the inherent patterns observed in the empirical activity data.

The validation of this variable memory length model will therefore focus on assessing the improvements over the first-order estimates. This will be assessed in terms of the statistical characteristics of the representations compared to those of the empirical time-use data. In particular, the validation analysis will look at 1) the daily state probability profiles, and 2) the distribution of the state durations. As an additional quality measure, we will also look at how the past states considered by the model are distributed with respect to their relevance.

4.4.1 Distribution of relevant past states

The relevance of the potential past states considered by the model is assessed by the gain function defined in Eq. 4.3.

The name 'gain function' can be slightly misleading, as one might intuitively expect higher values of this function to indicate higher gain, and therefore higher relevance. However, as the gain function is based on the Kullback-Leibler divergence, lower values are in fact an indicator of higher gain. The lower the divergence the higher the relevance of the past state considered. The relevance of past states is independent of their length. That is, pasts of different lengths can be equally relevant in terms of the transition into subsequent states at a different or the same time of day.

Figure 4.2 shows the distribution of past states with respect to their associated gain. The gain function value corresponding to each of the past states considered depends on the transition the state is associated with. However, once this value

has been estimated, we can simply regroup the past states by length and calculate their distribution with respect to their associated gain.

As we can observe in the different sub-figures in Figure 4.2, the distributions of the past states considered by the variable memory length model are heavily skewed towards zero. In each of the length groups over a third of the states is concentrated in the vicinity of zero.

As a measure of model quality, what these distributions show is that the benefit of adding the past states in question to the model is substantial. From an information theory point of view, the smaller the gain value of the associated states the larger the amount of additional information gained about the stochastic process. Therefore, the fact that large proportions of the past states considered are concentrated in the vicinity of 0 (Table 4.2) means that the overall amount of information gained is proportionally large.

Proportions of over 50 % of the states in each length group associated with a gain of less than 1 are clear indicators that the past states considered by the model have significant influence on the outcomes of transitions to subsequent activity states relative to the estimates associated with the first-order model. A summary of the maximum gain and percentage of states with gain values of less than 1 is presented in Table 4.2.

Table 4.2: *Proportion of states by gain range by length group and maximum gain value.*

Length group	Maximum gain	Proportion of states with gain ≤ 1	Proportion of states with gain ≥ 10
2	15.939	71.37 %	0.214 %
3	45.334	57.27 %	1.949 %
4	51.995	61.38 %	1.185 %
5	48.281	69.45 %	0.584 %

In terms of the distribution of the past states considered by the model, another aspect of interest is the way the distribution of states with respect to their length changes throughout the day. The evolution of said distributions can be observed below in Figure 4.3.

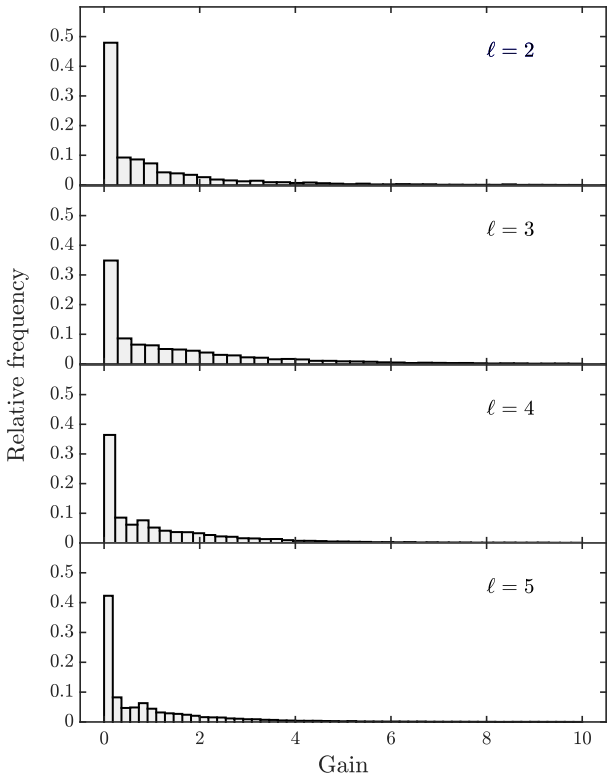


Figure 4.2: *Distribution of past states with respect to their associated gain function value. Sub-figures correspond to sets of states grouped by length (ℓ = length of state).*

The distribution in Figure 4.3 provides us with graphical evidence of the effects of the self-adaptation of the daily activity schedules. A more detailed analysis of the changes in these distributions might prove useful. For instance, it could potentially allow for the development of a method for assessing the potential for modifying activity schedules, or quantifying the effects of it. Although this is an interesting line of research, this falls beyond the scope of this paper.

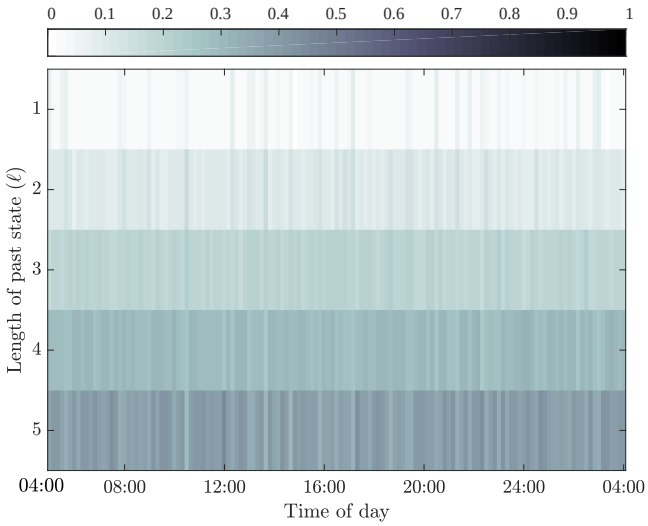


Figure 4.3: *Daily evolution of the distribution of past states with respect to their length (ℓ). The shading in each band indicates the proportion of states, measured from 0 to 1, as indicated by the scale atop this figure.*

4.4.2 State probabilities profiles

Based on a set of daily individual activity sequences it is possible to calculate the proportion of people engaging in the different activity states at any given time of day. These proportions correspond to the overall probabilities of finding a given individual engaging in any of the activity states considered and the evolution of these probabilities throughout the day is what is called the state probability profile.

Figure 4.4 shows the empirical state probability profiles from the UK 2015 time-use data. The eight shaded areas correspond to the different activities considered, as specified in the legend. For a given time of day the shaded area between the delimiting curves correspond to the state probability of the activity in question. These profiles exhibit some expected features such as a considerably low likelihood of finding people engaging in activities such as food preparation or

laundering during night-time, or the increase in the likelihood of people engaging in food preparation around mealtimes.

Based on the simulated activity data the corresponding state probability profiles are extracted, so that the estimates obtained from the simulated data can be compared to the empirical values. The discrepancies between these state probability profiles are quantified as root-mean-square deviation.

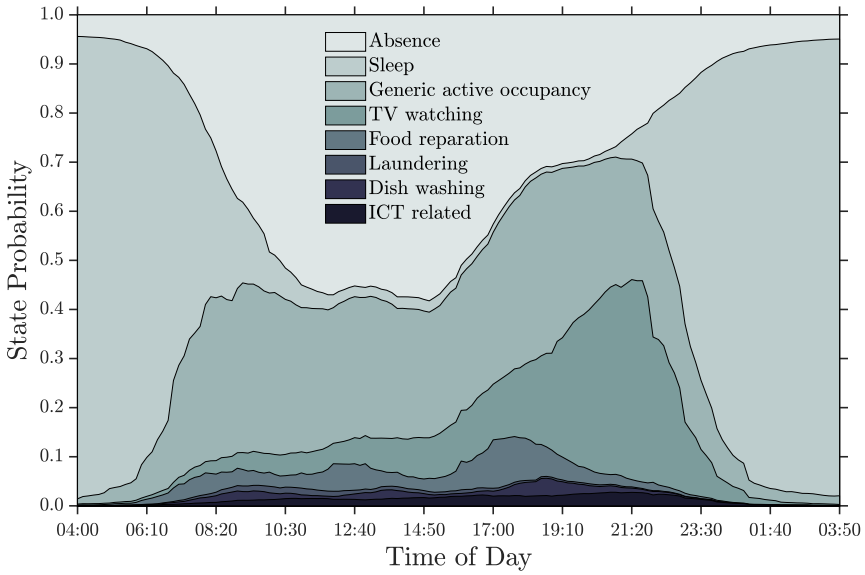


Figure 4.4: *Empirical activity state probability profiles extracted from the UK 2015 time-use data. The order of the elements listed in the legend is the same as the order of the stacked state probability profiles.*

Figure 4.5 shows the comparison between the activity state profiles obtained from empirical activity data and the estimates produced by the first-order Markov model and the variable memory length model. Both the variable memory length model and the first-order Markov model were used to produce sets of simulated activity sequences, with the number of simulated sequences increasing from 10 to 2×10^4. This is done so that the models' performance can be assessed and compared in terms of the rate of convergence of the root-mean-square deviation;

as the size of the simulated datasets approaches the size of the empirical dataset, the root-mean-square deviation converges to its minimum value.

As Figure 4.5 shows, the root-mean-square deviations of the state probability estimates produced by both models converge exponentially to an asymptotic lower bound in the range 0.1 - 0.3 % for all the activity states considered. However, we can also observe that the deviation of the estimates produced by the variable length model are generally lower than those of the first-order estimates, resulting in a faster convergence of the root-mean-square (RMS) deviation.

For the simulated sets containing 10 activity sequences only, the deviations of the first-order estimates range between 12.4 and 2.1%, whereas for the estimates produced by the variable memory length model the deviations range between 9.4 and 1.5%.

Table 4.3 presents a summary comparison of the RMS deviations of the two different models in terms of absolute and relative percentage differences. These differences correspond to the improvements over the estimates produced by the first-order Markov chain model. The overall relative percentage difference averaged over the different dataset sizes is 28%.

Figure 4.5 and Table 4.3 give us a good idea of how the estimates produced by the first-order model compare with the estimates produced by the variable memory length model with respect to the size of the simulated dataset. In general, we can observe that the variable length model performs better than the first-order model, and that the level of improvement also appears to increase as the dataset size does. The clearest example is perhaps the case of Food preparation activities, where the improvement over the first-order model starts at a mere 13%

and reaches nearly 50% when the size of the simulated dataset is comparable to the size of the empirical dataset.

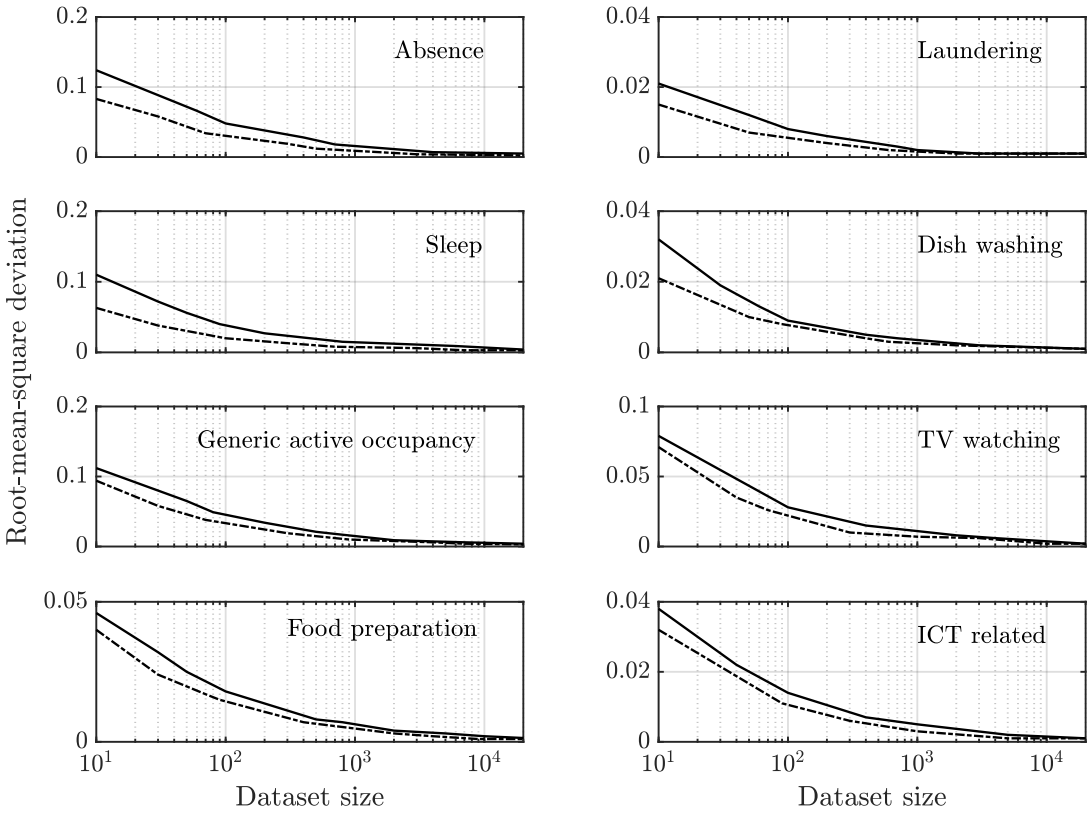


Figure 4.5: *Root-mean-square deviation of the daily state probability estimates produced by the first-order and variable memory length models. Dotted line - Variable memory length model; Continuous line - first-order model.*

Table 4.3: *Relative change in RMS deviation by activity by dataset size.*

Absence				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.1241	0.0468	0.0161	0.0051
Absolute difference*	0.0412	0.0159	0.0075	0.0023
Relative percentage difference*	33%	34%	46%	45%
Sleep				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.1111	0.0388	0.0159	0.0038
Absolute difference*	0.0472	0.0182	0.0079	0.0007
Relative percentage difference*	43%	46%	49%	19%
Generic active occupancy				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.1122	0.0461	0.0154	0.0039
Absolute difference*	0.0181	0.0121	0.0051	0.0009
Relative percentage difference*	16%	26%	33%	23%
Food preparation				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.0462	0.0178	0.0058	0.0019
Absolute difference*	0.0061	0.0036	0.0015	0.0009
Relative percentage difference*	13%	20%	26%	47%
Laundering				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.0211	0.0079	0.0019	0.0009
Absolute difference*	0.0062	0.0025	0.0005	0.0001
Relative percentage difference*	29%	32%	26%	11%
Dish washing				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.0322	0.0089	0.0034	0.0012
Absolute difference*	0.0111	0.0012	0.0011	0.0003
Relative percentage difference*	34%	13%	32%	25%
TV watching				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.0792	0.0283	0.0114	0.0024
Absolute difference*	0.0081	0.0059	0.0042	0.0001
Relative percentage difference*	10%	21%	37%	4%
ICT related				
Dataset size	10	10^2	10^3	10^4
RMS deviation - 1^{st} order model	0.0383	0.0141	0.0051	0.0014
Absolute difference*	0.0061	0.0034	0.0021	0.0003
Relative percentage difference*	16%	24%	41%	21%
Average percentage difference	24.5%	27%	36.3%	24.4%

** Difference with respect to the RMS deviation of variable memory length model estimates.*

In Figure 4.6 we can now observe how the estimates for small and large datasets produced by the variable length model compare with one another. For each of the activities considered the estimates for daily state probability profiles produced by the model are compared with the empirical state probabilities. On the left column we have the estimates corresponding to a set of 100 simulated daily activity sequences. On the right column we have the estimates corresponding to a set of simulated activity sequences of the same size as the empirical dataset.

The estimates extracted from the sets of 100 daily activity sequences only are noisier than those obtained from larger simulated datasets. However, as Figure 4.5 shows, these state probability estimates provide a closer representation (lower RMS deviation) of what is observed in the empirical data than the estimates produced by the first-order model.

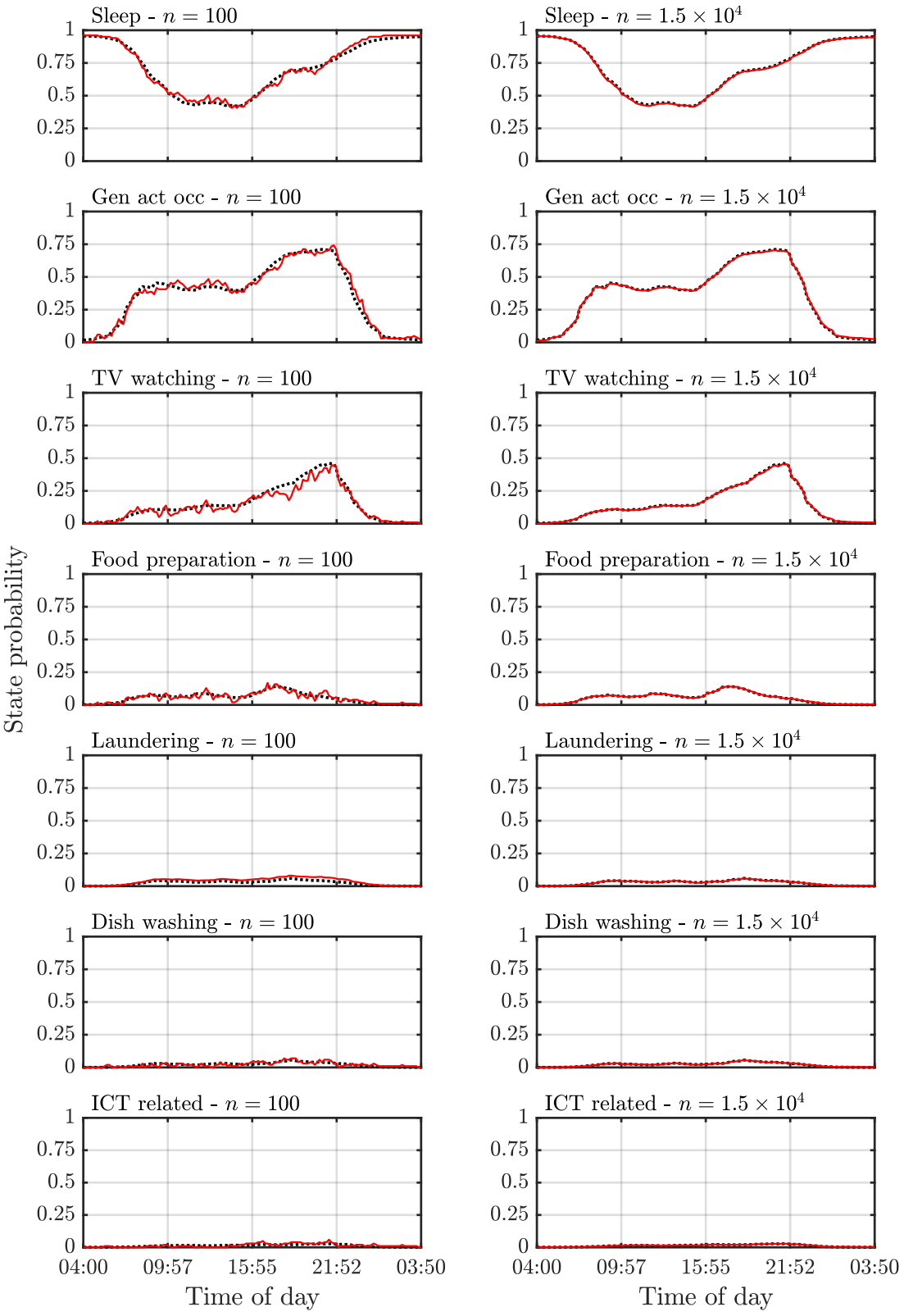


Figure 4.6: *Comparison between state probability profiles as extracted from datasets with 100 and* 1.5×10^4 *simulated activity sequences. Dotted line - Empirical state probability profiles; Continuous line - Model estimates.*

4.4.3 State durations distributions

Another key aspect to an accurate representation of the observed activity patterns is the duration of the periods of engagement in the different activities considered. The modelling examples found in the literature that make use of the first-order Markov chain approach are mostly validated in terms of the state probability estimates. As in most of those examples, the goal is to develop a model for simulating activity patterns which can in turn be used to estimate residential electricity demand loads. The state probability profiles tell us how likely it is that a given individual would engage in any of the activity states considered. However, as we discussed in section 4.2, the timing and variability of the observed electricity demand patterns depends just as much on the relative frequency of the activities associated with the loads in question and the time these activities are performed for. Therefore, for the purpose of the activity pattern simulation, once a given person engages in a certain activity, being able to estimate the time said person spends performing such activity is just as relevant and desirable. This is particularly relevant in the case of activities associated with continuous use of electrical appliances such as TV watching or ICT use.

Figure 4.7 shows the distributions of the durations of the activity states corresponding to the eight activities considered by the model. The distributions observed in the data produced by the first-order and the variable memory length model are compared with the distribution observed in the empirical data.

The plots in Figure 4.7 would appear to indicate that the distributions obtained from the simulated data are well in agreement both with each other and with the empirical data. However, by comparing these distributions in terms of their RMS deviation (See Table 4.4), it is observed that in general the

distributions obtained from the activity data simulated by the variable memory length model are closer to the empirical distributions than those obtained from the first-order simulations. The relative percentage difference is 30% on average.

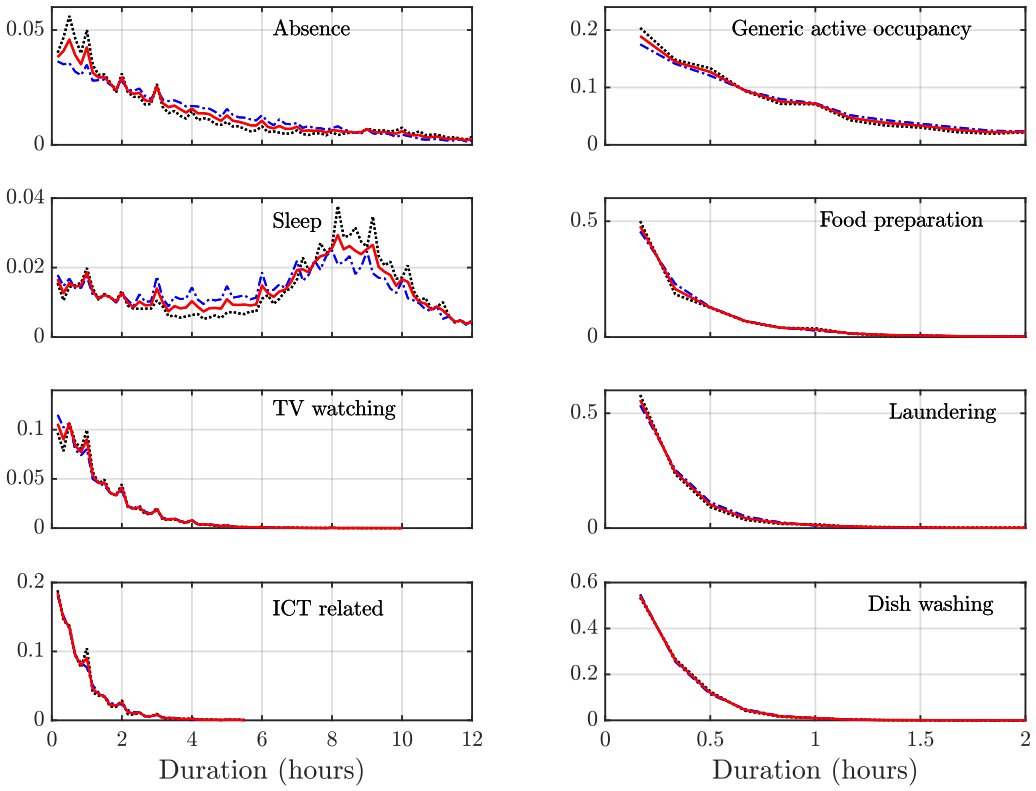


Figure 4.7: *Distribution of state durations. Dotted line - empirical distribution; Red line - Variable length model; Blue line - first-order model.*

Table 4.4: *Relative difference in RMS deviation of distributions of state durations by activity.*

Activity state	RMS deviation - 1^{st} order model	RMS deviation - Variable Memory Length model	Absolute difference	Relative percentage difference
Absence	0.0038	0.0029	0.0009	24%
Sleep	0.0042	0.0028	0.0015	34%
Generic active occupancy	0.0053	0.0038	0.0014	27%
Food preparation	0.0063	0.0042	0.0021	33%
Laundering	0.0052	0.0037	0.0016	29%
Dish washing	0.0147	0.0098	0.0048	33%
TV watching	0.0163	0.0117	0.0046	28%
ICT related	0.0069	0.0047	0.0022	32%

4.4.4 Transition Occurrence Frequency error profiles

The primary aim of the development of the activity modelling approach presented in this paper was to improve the simulation of daily activity scheduling. However, the lack of research on this particular area means that there are no standard metrics for quantifying this.

In order to assess the performance of a model in terms of the simulation of activity scheduling, we need to look closer at the behaviour in-between the states that make up the daily activity sequences. To this end, we developed a metric that allows us to estimate and compare the errors associated with the first-order model and the variable memory length model in terms of the occurrence frequency of pairwise state transitions.

Given set of (same-length) daily activity schedules, transition occurrence frequency matrices are extracted for each intra-day transition interval. We define transition occurrence frequency as the frequency with which the transition between two given activity states occurs. Thus, the entries of these matrices correspond to the observed frequencies of the pairwise activity state transitions at each transition interval.

The set of empirical activity sequences provides us with the reference values of the transition occurrence frequencies. The transition occurrence frequency matrices extracted from the simulated output of both first-order and variable memory length models are compared to those reference values.

Let us then define the Transition Occurrence Frequency (TOF) Error as the 2-norm of the matrix whose entries are the absolute differences of the transition

occurrence frequencies. If we denote this matrix as D^{abs}, we can express this mathematically as:

$$TOF_{err} = \left\| D^{abs} \right\|_2 , \tag{4.12}$$

The entries of the absolute difference matrix are given by $D^{abs}_{i,j} = \left| TOF^{sim}_{i,j} - TOF^{ref}_{i,j} \right|$, where TOF^{sim} and TOF^{ref} denote the simulated and empirical transition occurrence frequency matrices, respectively.

The 2-norm of a matrix is is also known as the spectral norm and is equivalent to the Euclidean norm of single vectors. For a given matrix A, this norm is defined as:

$$\|A\|_2 = \sqrt{\lambda_{max}\left(A^T A\right)}, \tag{4.13}$$

where λ_{max} is the maximum eigenvalue and A^T is the transpose of matrix A, respectively.

Both first-order and variable memory length models were, each, used to generate 50 sets of activity sequences with the same number of sequences as the empirical dataset. Average TOF errors were calculated for each of the simulated sets, as well as standard deviation ranges. The resulting average TOF error profiles can be observed in Figure 4.8.

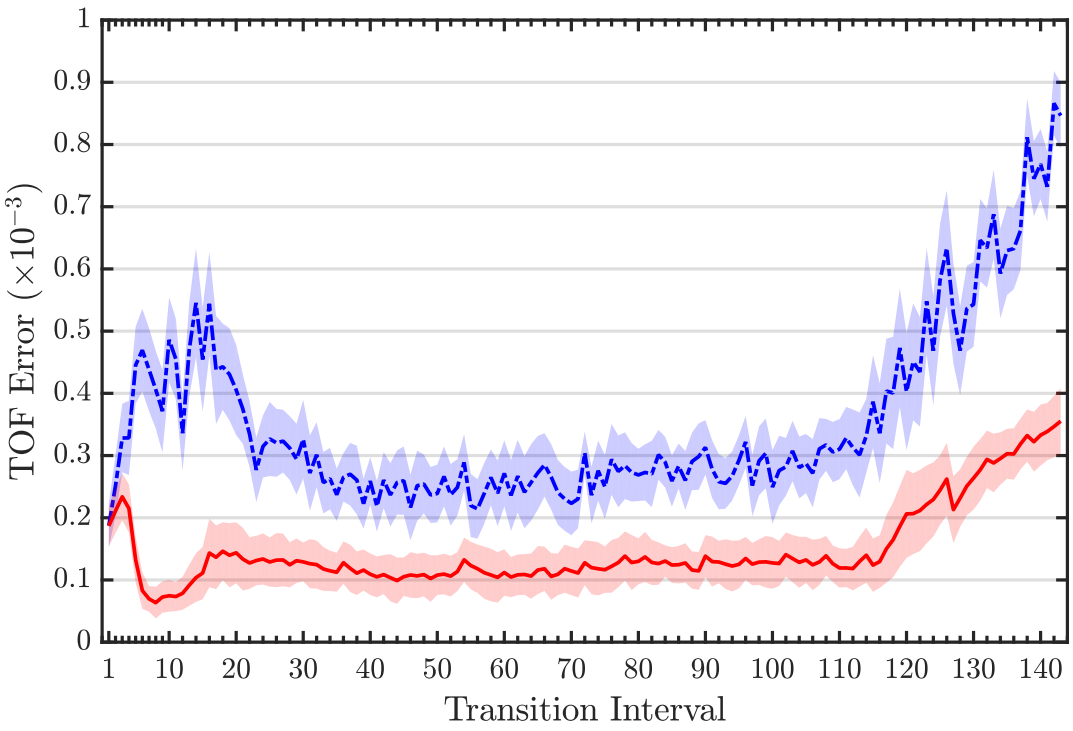


Figure 4.8: *Transition Occurrence Frequency (TOF) Error profiles: Blue line - first-order model error; Red line - Variable memory length error; the shaded regions along the profiles represent the standard deviation ranges.*

As we can observe from Figure 4.8, the TOF error of the variable memory length model is generally lower than that of the first-order Markov model. It is also observed that the TOF error associated with the first-order model would appear to have larger variability than the error associated with the variable memory length model. This might be attributed to the fact that the overall performance of the first-order model is not as stable, compared to the variable memory length model. Both of these results would appear to support the evidence for the improvement over the first-order Markov model provided by more standardised metrics such as the RMS error.

4.5 Discussion

In this paper we have presented a new approach to the stochastic simulation of residential activity patterns. There are a number of elements to the proposed approach which favour it over the current predominant approach based on first-order Markov chain models. In particular, this approach differs from the predominant one in that it focuses on exploiting characteristic behavioural patterns observed in the empirical activity data used to build the model. There are several reasons behind the interest in looking for an alternative approach to the simulation of residential activity patterns. Both the identified issues and the ways in which the proposed approach might help addressing them are outlined below.

- Compatibility with social practice theoretical frameworks

Human behaviour can be very diverse. However, research in the field of social practice theory shows that the restrictions imposed on people's behaviour by the social constructs of time result in a degree of synchronicity and the emergence of characteristic behavioural patterns.

The current predominant approach based on the Markov chain technique appears to perform reasonably well in terms of the representation of the overall activity profiles. Particularly when large simulated datasets are used to assess its performance. However, the inherent memorylessness of the stochastic processes represented by such models entails that any potential information on the presence of characteristic patterns and how they affect the overall evolution of the process is lost.

According to the branch of social practice theory that focuses on the study of daily schedules, the analysis of the temporalities and time dependences of the daily residential activity schedules requires that these are treated as a whole. However, the Markov chain technique by definition focuses exclusively on the transitions between any two adjacent states, paying no attention whatsoever to any potential structural dependencies present in the overall sequence of states.

The process for building the model based on the proposed approach focuses on the identification of characteristic patterns (relevant pasts) within the daily activity sequences observed in empirical datasets. However, in order to identify patterns that are meaningful relative to the entire *daily* empirical activity sequences, the analysis has to be carried out such that the daily activity sequences are processed as a unit. In that sense, the approach is consistent with the postulates of social practice theory outlined above.

- Improved understanding of the scheduling of activities throughout the day

Social practice theory acknowledges the existence of certain characteristic behavioural patterns that shape the overall daily activity sequences. However, research on this particular issue has thus far focused on the potential causes for this, and little attention has been paid to the analysis of empirical activity datasets with a view to identifying such characteristic patterns.

The data analysis carried out in order to identify the relevant pasts used to build the model is a good first step in learning more about the characteristic behavioural patterns observed in the residential sector.

The validation analysis of the model shows that the predictions for the transitions into subsequent activity states throughout the day are generally improved when the influence of these relevant pasts is taken into account. This is supported in particular by the results presented in Sections 4.4.2 and 4.4.4

It should also be noted that these results would appear to indicate that the approach to the analysis of empirical activity data presented in this paper has the potential to be used as a tool for a more in-depth analysis of such characteristic behaviours and their influence on the scheduling of activities throughout the day. An analysis of this sort would prove very informative and would broaden our understanding of the way daily activity schedules are shaped and maybe even provide further insight into the ways in which these can be effectively re-shaped. Such an in-depth analysis, however, falls beyond the scope of this paper.

- Improvement of the simulation of activity-related electricity demand

Electricity demand load modelling has found its most relevant application in the context of network design. Consequently, the requirements for the purpose of conventional network design procedures have been key drivers in the development of particular approaches to demand modelling. Conventional network design relies heavily on experience, and the current model for power systems ensures that the high variability of demand loads at the local level can be largely ignored. In this model, very large centralised generation units supply a very large number of users. Aggregating loads over such large numbers results in a demand profile where most of the variability that would be observed at the local level is smoothed out. Therefore, simplified representations of the

changes in demand throughout the day have been enough for the purposes of current network design and planning procedures, and so, many activity-based demand models have focused on reproducing the characteristics of these largely aggregated profiles.

However, in the context of the transition towards low-carbon energy systems, conventional network design procedures are increasingly becoming inadequate. In order to increase efficiency it is envisioned that generation units will be located closer to where power is needed, i.e. systems will be based on a more decentralised configuration where smaller generation units will be responsible for meeting the demands of smaller numbers of users. In addition, it is expected that generation technologies with highly variable outputs (e.g. solar and wind) will have an increasingly large share in the generation mix. In practice, this will pose serious new challenges to network balancing operations. Smaller numbers of users entail that the variability of the demand observed by the generation centres will be considerably larger. And the intermittent nature of renewables like solar and wind will make it much harder to manage this demand variability.

The way demand varies at the local level and how this might change in the future is less well understood. Therefore, it is necessary to develop approaches to demand modelling that help improve our understanding of this variability and help inform future network operation. In addition, in order to help alleviate the issues associated with the nature of renewables generation, adequate demand-side management strategies are being actively sought. However, our understanding of the way people go about their lives and how that reflects in terms of the observed demand patterns is still limited. Therefore, approaches to demand modelling that better capture the dynamics of the relationship between activity and demand patterns would prove really useful.

The current predominant approach to activity modelling for the purpose of demand modelling is based on the first-order Markov chain technique. The approach seems to perform well in terms of the representation of demand variability at the aggregate level. That is, for large numbers of users. However, as the results presented in Section 4.4.2 show, the quality of the estimates produced by this approach decreases as the number of simulated users does. This is also the case for the approach presented in this paper. However, the quality decline of the first-order estimates is faster relative to the quality of the estimates produced by the variable memory length model.

Attempts to improve the approach to activity modelling have been made through the development of semi-Markov models (Aerts et al. 2014, Wilke et al. 2013). The primary aim of these models is to improve the representation of the duration of the periods of engagement in the different activity states. In this approach, this is done by means of altering a conventional first-order Markov model by incorporating an extra step in-between the state transitions. In this intermediate step the duration of the subsequent states is altered rather artificially.

As the results presented in Section 4.4.3 show, the approach presented in this paper also provides a better representation of the duration of the periods of engagement in the different activity states. However, this is improved representation of the duration came about as a natural by-product of an improved simulation of the daily activity scheduling.

4.5.1 Potential for improvement

Given the size of currently available empirical activity datasets, it would appear that the improvement over the first-order model is modest, with an average of 28%. However, current and future activity data might offer the opportunity for a more thorough analysis of the characteristic activity patterns, and the construction of richer, more detailed models.

Time-use datasets are typically country-specific, and they are approximately the same size in terms of the number of independent activity sequences. However, the interest in time-use data has motivated the creation of standards that have allowed for the compilation of a homogenised time-use database which gathers time-use data from a number of countries.

Each individual dataset is meant to be representative of its country of origin. However, given that the temporal constraints a person is subject to might be the same (or very similar) even for people in different countries, it is reasonable to expect that activity sequences with similar characteristics will be found across many of these country-specific datasets. With the aid of data clustering techniques, this homogenised database offers the opportunity to create larger unified datasets which can then be studied using the data analysis approach proposed in this paper.

The source dataset is of great importance. However, as we have pointed out, the way the data is analysed is just as important. Therefore, further improvement might also come from the technical side. For instance, more detailed analysis of time-use data might help identify particular features of interest. This information

can then be used to tailor the algorithm implementation so that these specific features are highlighted during the simulation of activity sequences.

4.6 Conclusions

In this paper we presented a novel approach to the stochastic modelling of residential activity patterns based on processes with memory of variable length. The construction of the model is based on the analysis of time-use data with a view to identifying characteristic behavioural patterns present in empirical activity data. The primary aim of the proposed approach is to exploit said characteristic patterns for the simulation of residential activity sequences. The model performance is assessed based on the comparison of the statistical characteristics of the simulation outputs with those of the output of a typical first-order Markov chain model and the empirical activity data. The methods used for the analysis of the empirical activity data are sensitive to the size of the dataset. Therefore, the maximum performance improvement is bound by the restrictions posed by the size of the dataset. Given the amount of data available for the construction of the model, the improvement over the current predominant approach would appear to be modest. The results of the validation analysis show that the simulation output of the proposed model does represent an improvement over the output of the commonly used first-order Markov chain models. A comparison of the RMS deviations reveals that the average improvement across datasets with increasing numbers of simulated users is 28%. In addition to the use of standard error metrics, we also proposed a new metric to assess in more detail the quality of the simulated output in terms of the representation of the intra-day dynamics of the activity sequences. The comparison of the deviations

estimated by this new metric support the conclusions drawn by the use of the more standard metrics. Furthermore, the results of the analysis suggest that the proposed approach has the potential to be used as a new tool for a more detailed analysis of the structural complexity of empirical activity data. Such in-depth analysis would help broaden our understanding of the underlying behavioural patterns that shape the daily residential activity schedules. Therefore, an activity model based on this approach might prove really useful both in the context of time-use research and simulation of residential electricity demand loads.

5 | Simulating residential electricity demand: A modelling framework for activity-driven appliance usage of socio-demographically diverse households

Abstract

Several studies have shown that user behaviour is an important driver of demand for energy in the residential sector. As the transition towards low-carbon energy systems progresses, the importance of the role user behaviour plays will only increase. Active demand-side management may play an increasingly important role in future low-carbon systems. However, the extent of the effectiveness of these mitigation measures is heavily dependent on our understanding of user behaviour. To aid in this effort, a number of models have been developed with a view to further investigating the factors at play. However, there is still considerable room for improvement. In this paper we present a new approach to activity-driven residential electricity demand modelling that builds on the strengths of previously developed ones. The proposed modelling framework takes a hybrid approach that combines elements of stochastic modelling with agent-based modelling methods. The proposed improvements come at three different levels. Firstly, direct links between key socio-demographic factors and demand for electricity are embedded into the modelling framework so that the variability associated with them can be accounted for. Secondly, the simulation of user activity patterns makes use of a novel methodology that focuses on exploiting characteristic behavioural patterns. By taking these characteristic patterns into account it is achieved a certain level of self-adaptation in the evolution of the simulated daily activity sequences that results in a more realistic representation of empirical user behaviour. Thirdly, the approach to the conversion of user activity into power consumption is based on a detailed representation of the interaction between users and appliances that aims at capturing the particularities of the use of the different types of appliances considered throughout the day. The paper reports on the methodologies integrated within the proposed modelling framework and on the quality of the output of the simulations which is assessed in terms of factors relevant to informing the development of demand-side management strategies and network planning and operation.

5.1 Introduction

The residential sector is a major contributor to the final electricity consumption across countries. Among the 28 EU member states, this sector accounted on average for about 30% of the national total (European Environment Agency 2017). Understanding the way energy is consumed in this sector is key to the development of strategies seeking to solve efficiency and flexibility issues. In addition to total consumption estimates, load profiles also play an important role in the design, planning and control of power systems as they contain information regarding the variation of the level of demand throughout the day. This information is critical for determining system capacity requirements and developing adequate operation plans (Baetens et al. 2010, Piwko et al. 2005).

Conventional network design and planning have traditionally relied on estimates of the overall average daily load profile of a very large group of users. However, the plans for the transition towards a low-carbon economy would appear to indicate that power generation systems will evolve into systems based on a more decentralised configuration. As power systems undergo this transition, traditional approaches to planning and design are becoming obsolete; smaller generation units will supply smaller groups of users, which means that the variability observed at a more local level is becoming increasingly relevant. A better understanding of how the loads associated with the residential sector arise would allow for producing accurate estimates of the system loads at the local level, which will in turn allow for the development of more suitable strategies.

One of the main applications of residential electricity demand load modelling has been assisting in the design and planning of distribution networks. Therefore,

approaches to electricity demand modelling will also have to be improved so that they can continue to be effective tools for this purpose. There are several factors associated with the variability of demand for electricity across households, and these factors are related to each other through complex interdependencies. Moreover, the effects of these factors also vary with respect to the level of aggregation of individual loads. Therefore, particularly in the context of future network planning, it is desirable and indeed necessary to incorporate this complexity into the approaches to residential electricity demand modelling.

5.1.1 Drivers of residential electricity demand

A number of factors can be associated with the demand for electricity in the residential sector. However, research has found that the most notorious increases in demand for electricity are associated with increased use of energy consuming devices (Jones & Lomas 2016). In the UK, for instance, the proportion of electricity consumption associated with the use of appliances has more than doubled compared to the 1970 baseline (Jones & Lomas 2016). This large growth has been primarily attributed to increased ownership of certain types of appliances. However, it is also concluded that these developments are the result of wider societal changes such as increased living standards and life expectancy, lifestyle changes and increases in the number of smaller households (Boardman et al. 2005, Jones & Lomas 2016). The increases in demand for electricity associated with the use of appliances in the residential sector are widely expected to continue (Energy Saving Trust 2011, IEA 2009).

Socio-demographic factors such as household size and the age of the adult members of the household have been found to be good explanatory variables for

the levels of appliance usage (Huebner et al. 2016). Household size is probably the most commonly used socio-demographic variable. However, when this variable is taken into account, a common simplifying assumption about the relationship between household size and demand for electricity, or energy in general, is that the relationship is linear. That is, the larger the household the greater the associated demand. Nevertheless, it has been shown that this relationship is in fact more complex (Huebner et al. 2016, Ramírez-Mendiola et al. 2017). Regarding the direct effect of the age of the household members, the results have been rather ambiguous with some studies finding non-linear relationships and others reporting no effect (Huebner et al. 2016). Indirect effects, however, can be estimated through looking at the way age affects appliance ownership rates.

Appliance ownership was an explanatory variable included in several studies, with a general finding that owning more appliances or using them more intensively is associated with greater demand for electricity (Bedir et al. 2013, Wiesmann et al. 2011). In particular, the presence of certain appliances is associated with greater demands. These include entertainment devices, multiple cold appliances, multiple electric cooking appliances and clothes dryers (Huebner et al. 2016).

The effects of the drivers discussed above are those the modelling approach presented in this paper focuses on. Other studies based on different approaches have focused on different factors depending on their purpose. Comprehensive reviews on the divers of residential demand for electricity can be found elsewhere (Huebner et al. 2016, Yin et al. 2015).

5.1.2 Residential electricity demand modelling

It is widely acknowledged in the energy research literature that residential electricity demand modelling is a key tool for the analysis of power systems and the assessment of the potential of low-carbon strategies (Torriti 2017). The interest in these matters has driven the development of a number of residential electricity demand models over the last couple of decades. The approach to the simulation of residential electricity consumption patterns varies from one model to the next. However, the most successful models have one characteristic in common: they try to incorporate behavioural and social aspects.

Given the varying nature of people's behavioural patterns, these are bottom-up models which are commonly based on the stochastic simulation of household activity patterns. However, these models differ in terms of the level of detail in which household activity is modelled, the stochastic modelling approach used, and the conversion of the simulated activity patterns into power consumption patterns. One of the first examples of this kind of models was the approach developed by Capasso et al. (1994). The simulation of activity was carried out by a simple two-state model (available/unavailable) based on a Monte Carlo method. This was in turn used to produce 1 hour resolution daily load profiles. The approaches to residential demand modelling have evolved considerably since the approach developed by Capasso et al. was presented.

Over the last few years there has been an emerging approach dominated by the use of time-use data for the development of models for the simulation of household activity patterns. These models have been consistently based on the stochastic simulation method known as the Markov chain technique. One of the most representative examples of such models is the model developed by Richardson

et al. (2010). This model has found widespread applications in both academia and industry (Ballanti & Ochoa 2015, Ballanti et al. 2013, EA Technology 2012, Good et al. 2015, Navarro et al. 2013). The simulation of activity is carried out by a three-state Markov chain model that produces daily profiles of 'active occupancy'[1]. Power consumption is derived from these active occupancy profiles by the activation of certain power loads which are only observed during these periods of active occupancy. An attempt to include socio-demographic factors is made by taking the household size as one of the input parameters. However, this factor is only used by the occupancy sub-model and no attempt is made to link this factor with the production of the demand loads or overall energy requirements estimates. A number of other models which make use of this or similar approaches have been developed. A more comprehensive review of this kind of models can be found elsewhere (Torriti 2014).

5.1.3 Motivation

Based on the configuration of most current power generation systems, where generation is highly centralised, a large number of users must be considered at once for effective network design and planning. However, the plans for the transition towards a low-carbon economy would appear to indicate that power generation systems will evolve into a more decentralised configuration. In addition, it is also widely expected that renewables, most of which are highly variable power generation sources, will play a major role in the future generation mix.

[1]The term active occupancy refers to the periods where household members are present and awake, regardless of the specific activities they might be engaging in.

When large numbers of users are used to estimate the daily load profiles required for network analysis, this results in a smoothing effect. That is, the magnitude of the intra-day variability observed at the local level is lost as the aggregation level increases (Elombo et al. 2017). However, decentralised generation units would have to deal with the high levels of variability associated with smaller groups of users. In addition to the effects of changes to the configuration of the power system, the applications of demand side response will also add a dimension of complexity to network balancing operations.

Suitable balancing systems will need to be developed and network operators will have to adopt relevant strategies to accommodate for changes to the generation resources on the evolving networks (UK Power Networks 2014). To this end, detailed studies on system load profiles will need to be conducted to uncover the nature of the different system loads (Elombo et al. 2017).

Up until now, residential electricity demand modelling has focused on the ability to reproduce the smoothed out daily load profiles needed for conventional network planning. The approaches commonly used in widely implemented models have proven to be good enough for this purpose. However, partly because of this, the efforts to try and learn more about the intra-day variability observed at lower levels of aggregation and its potential implications in terms of performance requirements of local distribution networks have been limited. In addition to the focus on reproducing the overall patterns, other issues are likely to render some of the current modelling approaches obsolete in terms of future network planning.

As we discussed in section 5.1.1, a strong link exists between socio-demographic factors such as household size, household composition and age range of the adult members of the household, and the typical levels of

electricity consumption observed for a given household. Research has found that socio-demographic variables on their own account for about 21% of the variability observed (Huebner et al. 2016).

It is widely acknowledged in the literature that these links ought to be taken into account in order to achieve better representations of the differences observed across a given population. However, few attempts have been made to unravel the complex interdependencies of said links and incorporate them into the simulation of residential demand for electricity.

One of the most commonly used socio-demographic factors is the household size. Yet, when the households size has been considered in the modelling of residential electricity consumption, simplifying assumptions have been made regarding the relationship between this factor and overall annual electricity consumption estimates. These assumptions, be it implicitly or explicitly, have led to a representation of the associated variability that corresponds to a normal distribution. However, as it has been pointed out elsewhere (Ramírez-Mendiola et al. 2017) [Chapter 2], the relationship between the variability observed in total annual consumption estimates and the variation of driving factors such as the household size is much more complex than the normally distributed samples commonly assumed.

While household size has been found to be among the most important predictors, there are other equally important factors at play. For instance, it has been found that variations in appliance ownership and use account for approximately 34% of the variability in electricity consumption (Huebner et al. 2016). The ownership rates of different types of appliances are in turn dependent on factors such as household composition and age range of the adult members.

Therefore, it is essential that approaches to residential electricity demand modelling take these aspects into account.

Many of the currently used approaches have opted for the integration of behavioural factors through the simulation of households' activity. However, this activity modelling has largely focused on the simulation of active occupancy rather than engagement in activities themselves. In addition, the predominant approach to the simulation of households' activities is based on first-order Markov chain processes. This approach is good enough for providing the overall estimates resulting from aggregating and averaging data over a large number of users. However, by definition, it ignores any potential structural dependencies in the daily activity schedules, as the transitions into different activity states throughout the day are determined by the current state only.

Being able to represent the result of the interactions between users and appliances in terms of power consumption is also highly relevant. However, the approaches to the simulation of said interactions tend to oversimplify them. More specifically, it is commonly assumed that appliance power consumption only occurs during periods of active occupancy, and that the power consumption during the cycles of operation of said appliances is constant. Therefore, situations in which the appliance runs for longer than the time a given user is actively engaged in the associated activity and/or where the appliance presents variable power requirements throughout the cycle of operation are neglected.

The proposed modelling framework builds on the strengths of currently used approaches and attempts to address some of the issues discussed above by incorporating new information and making use of data which were not available when the existing models were developed. The proposed approach

differs from those currently used in that 1) it takes direct account of some of the key factors that characterise households in real life, 2) a more integral approach to the simulation of daily activity patterns is used; the approach looks at the activity profiles as evolving sequences of states in which some of the transition into subsequent states are affected by relevant portions of the past, and 3) the simulation of the interaction between users and appliances is based on monitored real-life events.

The paper reports on the different methodologies integrated within the proposed modelling framework and on the quality of the output of the simulations in terms of factors relevant to informing the development of demand-side management strategies and network planning and operation. The rest of the paper is structured as follows. In the next section the structure of the proposed modelling framework is detailed. This is followed by the validation analysis (Section 5.3), and a discussion of the advantages and shortcomings of the proposed approach (Section 5.4). Finally, we discuss the potential for improvement and present our concluding remarks.

5.2 Modelling Framework: Data and Methods

The proposed modelling framework was developed with a modular architecture in mind. Three are the modules responsible for the corresponding three main stages into which the overall modelling process is divided. The output of the first stage sets the basis for the simulation run corresponding to a particular household. Second and third stages are closely related and are iterated for the specified duration of the simulation period. Figure 5.1 offers a graphic representation of the framework's architecture.

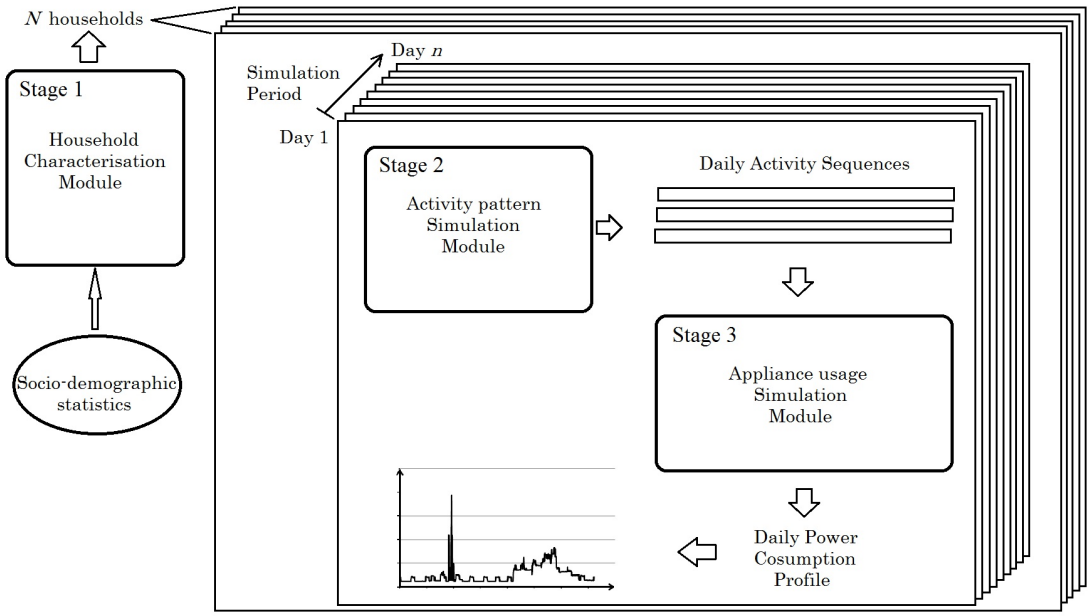


Figure 5.1: *Modelling framework architecture.*

In the following sections, each of the stages into which the modelling process is divided and the structure of the modules responsible for them will be detailed.

5.2.1 Stage 1: Household characterisation

The characterisation of the simulated households refers to the definition of the socio-demographic characteristics that distinguish a particular household, and how they are likely to be reflected in terms of the household's demand for electricity.

The purposes of the household characterisation module are four-fold:

- Providing an estimate for the average annual consumption
- Determining the size of the simulated households

- Determining the likely household composition type

- Determining the set of appliances likely available to the household

A more detailed account of the listed purposes is provided in the sections below.

- Providing an estimate for the average annual consumption

In terms of household electricity consumption, a common practice in the literature is to stratify households into groups corresponding to low, medium and high annual consumption ranges (Armstrong et al. 2009, Haines et al. 2010, Jones & Lomas 2016). However, the bounds of such ranges have been typically defined in rather arbitrary ways and tend to vary across studies, even when dealing with groups of households from the same country. As the analysis by Ramírez-Mendiola et al. (2017) [Chapter 2] confirms, there are indeed three clearly defined clusters in the distribution of households with respect to their total annual consumption. More importantly, this analysis provides us with a more rigorous quantification of what the low, medium and high annual consumption ranges are, and provides us with a statistical model for the observed distribution. The analysis was based on the set of households monitored as part of the UK Household Electricity Survey of 2010 (DECC & Intertek Testing and Certification Ltd. 2016). The sample of monitored households was selected so that its statistical characteristics matched those of the typical socio-economic mix (Zimmermann et al. 2012). The results of the analysis reveal that the distribution of households is well described by a tri-modal Gaussian mixture model (Ramírez-Mendiola et al. 2017) [Chapter 2]. Based on this model, it is possible to provide estimates for the average annual consumptions for a given set of households.

- Determining the size of the simulated households

The relationship between household size and electricity consumption has been consistently oversimplified in models that take this into account. In general, it is assumed that the relationship between these two factors is linear. That is, larger households are generally associated with greater demand for electricity (Huebner et al. 2016). Previous studies have already pointed out that this might not necessarily be the case (Wyatt 2013, Yohanis et al. 2008). However, this concern has not been addressed by the modelling approaches developed up until now.

Further analysis of the composition of each of the clusters identified by Ramírez-Mendiola et al. (2017) [Chapter 2] allowed for a more detailed breakdown of the groups of households in each cluster. In particular, it allowed for the extraction of the sub-distributions of households with respect to size based on the percent composition of each cluster. Using these distributions it is now possible to provide estimates in accordance with the non-linear relationship between household size and electricity consumption. Thus, it is possible to determine a likely household size which is in agreement with the previously determined annual electricity consumption estimate.

Figure 5.2 shows the distribution of households according to size and level of annual electricity consumption.

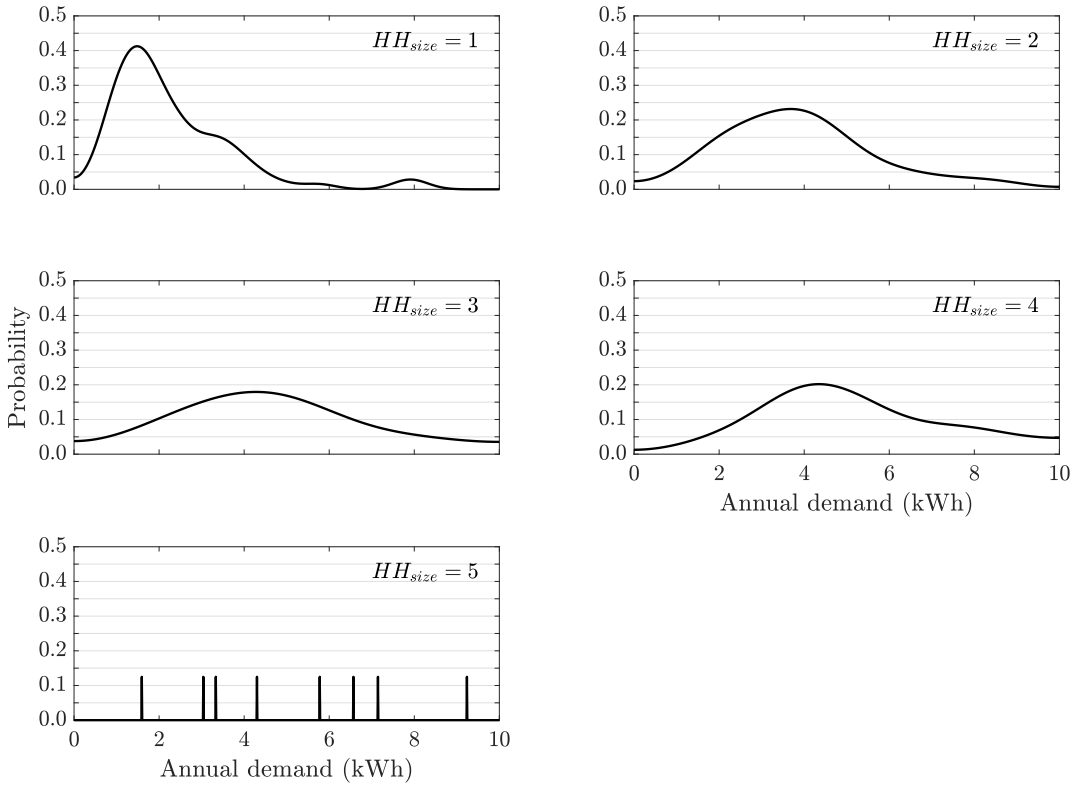


Figure 5.2: *Probability distributions of households with respect to total annual consumption. The households have been grouped by size to show the differences in variability within same-size groups.*

- Determining the likely household composition type

Another important socio-demographic factor in close relation with the household size is the household composition type. According to the terminology used in the 2011 UK census, a household is defined as one person living alone or a group of people (not necessarily related) living at the same address and sharing living spaces and common facilities [2].

There are three main household composition categories: one person

[2]According to the International Property Measurement Standards for residential buildings (International Property Measurement Standards 2016), common facilities are those parts of a building providing shared facilities that typically do not change over time, including for example, circulation areas, toilets and maintenance rooms.

households, one family households and other households. The 'one family households' category encompasses those households composed of a couple with or without children, or single parents with children. These sub-categories can be further broken down by making a distinction between dependent and non-dependent (adult) children. For the purposes of this sub-classification, any children aged 16 or over are deemed to be adult members of the household. The 'other households' category encompasses all other types of multi-person households including student households, multi-family households and households of one family and other unrelated adults.

One family households are the most common type, accounting for 62% of the total number of households, followed by one person households which account for 30% (ONS 2014). For the purposes of the characterisation stage of the proposed modelling framework, all members of a household falling into the 'other households' category are assumed to be adults. Even within the 'one family households' category, there are differences that need to be taken into account in order to describe the relationship between household size and household composition type. Namely, the number of children. The full breakdown of the percentage composition of households by type is shown below in Table 5.1.

In addition, the data from the 2011 census also provides information about the number of dependent children for those households with dependent children. The information concerning this can be found in Table DC1114EW of the census data (ONS 2011). A summary breakdown of the percentage composition of couple and lone parent families is shown in Table 5.2.

Table 5.1: *Breakdown of households by composition type.*

Composition type	Percentage share	Composition sub-type	Percentage share
One family	61.8	Couple with no children	32.8
		Couple with dependent children	35.9
		Couple with all children non-dependent	11.4
		Lone parent with dependent children	13.4
		Lone parent with all children non-dependent	6.5
One person	30.3		
Other	7.9		

Table 5.2: *Breakdown of family types by number of children.*

	Family type	
Number of children	Couple	Lone parent
One dependent child	41.2%	58.2%
Two dependent children	41.5%	28.3%
Three+ dependent children	17.3%	13.5%

- Determining the set of appliances likely available to the household

Before any simulation of energy consumption can start, it is necessary to determine the set of devices which are likely to be the origin of such demand loads. In order to do this, the information about the household obtained from the previous steps of the characterisation procedure (i.e. household size and composition type) is used to determine an adequate set of appliance ownership statistics. The importance of the use of ownership statistics has been highlighted by several studies, arriving at the general conclusion that owning more appliances, or more appliances of the same kind, is associated with greater demand for electricity (Huebner et al. 2016, Wiesmann et al. 2011, Yohanis et al. 2008).

Additional demographic studies are more concerned with aspects related to the ownership of the different types of household appliances. In the case of the UK, studies such as the General Lifestyle Survey (ONS 2013) and the Living Costs and Food Survey (ONS 2017), commissioned by the Office for National Statistics, and the Energy Follow-Up Survey (EFUS) (DECC 2013*a*), commissioned the Department of Energy and Climate Change, provide detailed information about the penetration rates of the different types of appliances commonly found in UK homes. In general, the ownership statistics of some of the most common appliances vary across household types. Moreover, ownership of multiple devices corresponding to the same type of appliance is also commonly observed. Taking this into account would allow for a better representation of the situation of real households.

In addition to the type of household, the age range of the adult members of the household also appears to play an important role. It is observed, for instance, that

appliances such as TVs are more common among households composed of adults over the retirement age (e.g. One-person households ownership rates: 98% for age 65+, 95% for age < 65). Conversely, devices such as computers are less common among households with adult members within this age range (e.g. One-person households ownership rates: 42% for age 65+, 78% for age < 65). As of 2014, 18% of the UK population was aged 65 years or over. According to projections from the Office for National Statistics, this proportion will reach 20% by 2024, and it is expected to keep increasing in the near future. Therefore, if demographic trends such as this one are to be taken into account for the simulation of energy consumption patterns, the differences in the ownership statistics corresponding to different sectors of the population will have to be integrated into the models.

Appliances are thus distributed among the simulated households so that the proportion of households that own the appliances matches the ownership statistics corresponding to the relevant household type. Based on this, a given household can be assigned up to 31 appliances. This includes cases of ownership of multiple devices falling within the same category. An example of a set of ownership statistics is shown in Table 5.3, and the percentage breakdown corresponding a case of multiple ownership of same-type appliances is shown in Table 5.4.

Table 5.3: *Set of appliance ownership statistics for one-person households, for people aged less than 65 years.*

Appliance	Cooker	Oven	Hob	Kettle	Microwave	Toaster	Dishwasher	Tumble dryer	Washing machine	TV set	Desktop PC	Laptop
Ownership rate	55%	66%	35%	97%	9%	99%	45%	49%	93%	95%	78%	79%

Table 5.4: *Percentage breakdown of households by number of TV sets owned.*

Number of devices	1	2	3	4
Proportion of households	38%	31%	20%	11%

**Based on the number of households that own at least one device.*

5.2.2 Stage 2: Simulation of household activity patterns

The simulation of household activity patterns is the core mechanism for the simulation of power consumption patterns. With the exception of base-load energy uses (e.g. cold appliances or internet routers), the activation of the installed appliances depends on the engagement of the members of the household in different activities associated with different sets of appliances.

A previously developed approach to the simulation of residential activity patterns (Ramírez-Mendiola et al. 2019) [Chapter 4] is used to produce daily activity profiles for the household. This approach makes use of a new method based on the use of stochastic processes with memory of variable length. The approach focuses on exploiting characteristic behavioural patterns observed in empirical time-use data[3] for the simulation of activity profiles. By taking these characteristic patterns into account a certain level of self-adaptation in the evolution of the simulated daily activity sequences is achieved. This results in a more realistic representation of empirical user behaviour (Ramírez-Mendiola et al. 2019) [Chapter 4]. The full details of the approach to the simulation of household activity schedules is set out in (Ramírez-Mendiola et al. 2019) [Chapter 4].

The activity model generates daily activity sequences composed of 144 activity states; one per every 10 min interval in a 24 hour period. Eight activity

[3]Time-use datasets stem from large-scale surveys that look specifically at how people spend their time, where do they spend it, and who they spend it with. In order to account for differences in time use for the different types of day, respondents are typically asked to complete time-diaries for one weekday and one weekend day, logging activities every 10 min. These datasets are the best currently available sources of empirical data on the activities carried out by people throughout the day. The latest major time-use dataset available for the UK became available in 2017. It comes from the national survey carried out between April 2014 and December 2015 (Gershuny, J. & Sullivan, O. 2017).

states are considered: absence, sleep, generic active occupancy, food preparation, dishwashing, laundering, TV watching and ICT related activities. With the exception of states corresponding to absence and sleep, the activity states are associated with a set of appliances that can be activated as a result of people's engagement in the relevant activity.

For the purposes of activity-based appliance usage modelling there are two key issues that need to be addressed: 1) the simultaneous use of appliances, which refers to the use of multiple devices (same or different type) at the same time of day, and 2) the shared use of appliances, which refers to the use of the same device by multiple members of the household at a given time of day. The following approach is taken to address these issues. Despite the potential for correlations between engagement in activities among household members, adult members are also bound to exhibit a certain degree of independence when it comes to the generation of demand loads. For instance, one person might be cooking while another one is doing laundry. In order to take this possibility into account, the activity models is used for generating individual activity patterns for each adult member of the household, for each day of the simulation period. In this way, we can explicitly identify the episodes of engagement in a certain activity that require attention concerning these issues. Particularly in the case of shared appliance use.

The empirical time-use dataset also contains information about the presence of other members of the household while engaging in a particular activity. We use this information to provide a way of determining whether the shared use of appliances is likely. For each 10 min period of the day we look at the episodes of activity engagement for which it was reported the co-presence of other household members. Based on this, an activity co-engagement likelihood factor is calculated

for each time interval, for each of the activities of interest. These factors are an estimate of the probability that the appliance(s) associated with the activity in question are being used by multiple members of the household. We thus define the daily activity co-engagement likelihood profiles as the series of co-engagement factors corresponding to a 24 hour period. When a potential appliance usage conflict arises during the simulations, the co-engagement factors are used to determine whether the conflicting users are likely to be sharing the appliance or not.

5.2.3 Stage 3: Simulation of daily power consumption

The purpose of the third module is the simulation of appliance-level power consumption. To this end, a distinction is made between activity-driven and non-activity-driven power usage. Non-activity-driven power usage refers to the power consumption associated with appliances such as cold appliances or other base-load energy uses (e.g. internet routers) which are not significantly affected by household activities. Activity-driven power usage refers to the power consumption that is a direct consequence of the interaction between households and the appliances available to them. To achieve a realistic representation of the power consumption associated with the different types of appliances considered, these are characterised by a set of factors extracted from appliance-level metered power consumption data. The appliance characterisation and handling of both activity-driven and non-activity-driven power usage is detailed below.

5.2.3.1 Characterisation of appliances

For the effective simulation of appliance usage three key elements need to be identified: 1) the likelihood of an appliance activation event; 2) the likely duration of the corresponding usage episode; and 3) the power consumption profile corresponding to the cycle of operation. The information that allowed for the characterisation of the different types of appliance considered in terms of these factors was sourced from the most comprehensive dataset available on UK appliance-level metered energy usage. Namely, the 2010 UK Household Electricity Survey (HES). HES was a large scale study during which the energy consumption of over 5000 individual appliances, distributed across 250 households was monitored in detail[4] (Zimmermann et al. 2012).

The data stemming from this study (DECC & Intertek Testing and Certification Ltd. 2016) is a collection of daily load profiles corresponding to the appliances monitored within the surveyed households. Therefore, it allows us to look at the daily usage patterns of the different types of appliances in detail. The surveyed sample is divided into two groups: a set of 25 households monitored for one year, with power consumption records at 10-min resolution, and a set of 225 households monitored for a month with records at 2-min resolution. For the purposes of appliances' characterisation we focused on those appliances with 2-min resolution records, as these provide a better picture of the variability of appliance power usage. A detailed account of the survey methodology and collected data can be found elsewhere (Zimmermann et al. 2012).

[4]The monitored households were chosen such that the statistical characteristics of the whole sample matched those of typical socio-economic mix (Zimmermann et al. 2012). Only owner-occupiers were recruited for the study. However, as the final report on the results of this study points out, the households were still fairly typical in terms of socio-demographic factors (Zimmermann et al. 2012).

In order to extract the elements need for the characterisation of the simulated appliances, the data was processed in the following manner. Appliances monitored within a given home are considered as a unit, with an associated stand-by consumption and a set of power consumption cycles corresponding to different usage episodes. For the appliances associated with non-activity-driven power usage each day of the monitoring period was considered a usage episode. All the monitored power consumption associated with each of the appliances of interest was grouped together. Based on each individual appliance sub-dataset, the value of the associated stand-by consumption is identified. This value is then used as a threshold for the identification of periods of active appliance operation. This criterion is applied to the extraction of both daily appliance switch-on probability profiles and appliances' power load traces corresponding to individual cycles of operation, making a distinction between the three types of day considered: Weekdays, Saturdays and Sundays. Based on the identified cycles of operation, along with the information about the start times, estimates of the joint probability distributions for the duration of a given cycle of operation started at a given time of day for the type of appliance in question were produced.

The sets of stand-by and cycles of operation power consumption associated with individual appliances are gathered into a pool of appliance profiles for the different types of appliances considered. In this context, the term appliance profile refers to the likely behaviour the appliance is expected to exhibit in terms of power consumption during stand-by and periods of activity. These sets of profiles are then mapped onto the sets of appliances assigned to the simulated households. To add some variability, the mapping of appliance profiles is done stochastically. Simulated appliances are thus characterised by the value of their power consumption during stand-by and a set of power consumption profiles

associated with different cycles of operation.

Figures 5.3 and 5.4 show, respectively, examples of the obtained probability distributions for switch-on events and duration of the cycles of operation, for the different types of day considered. Figure 5.5 shows examples of the variability observed in appliance power consumption traces corresponding to different cycles of operation.

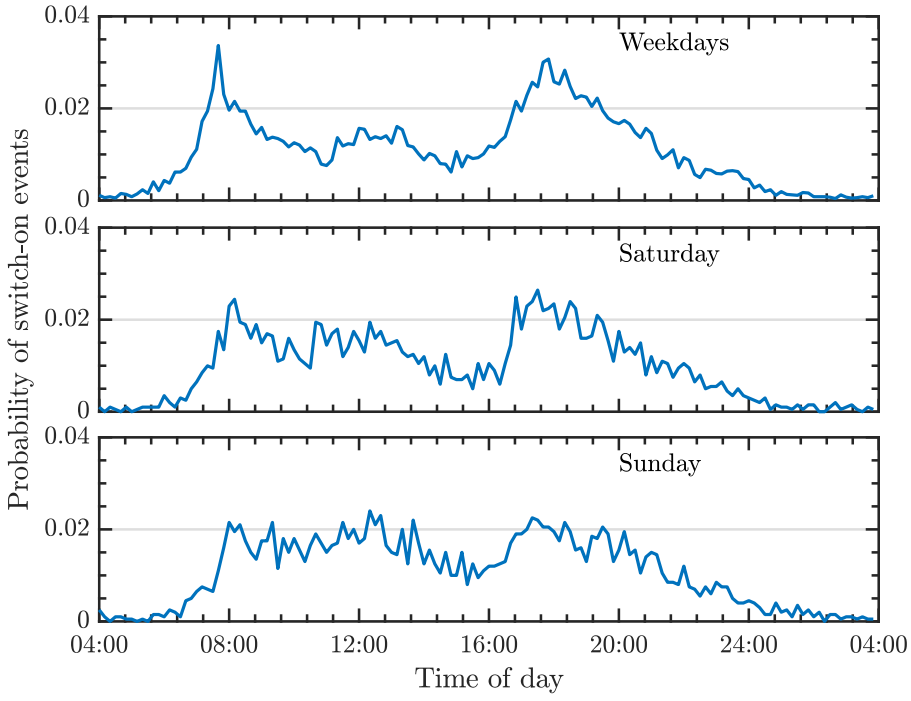


Figure 5.3: *Example of the daily switch-on probability profiles for the different types of day considered. Case: Microwave oven.*

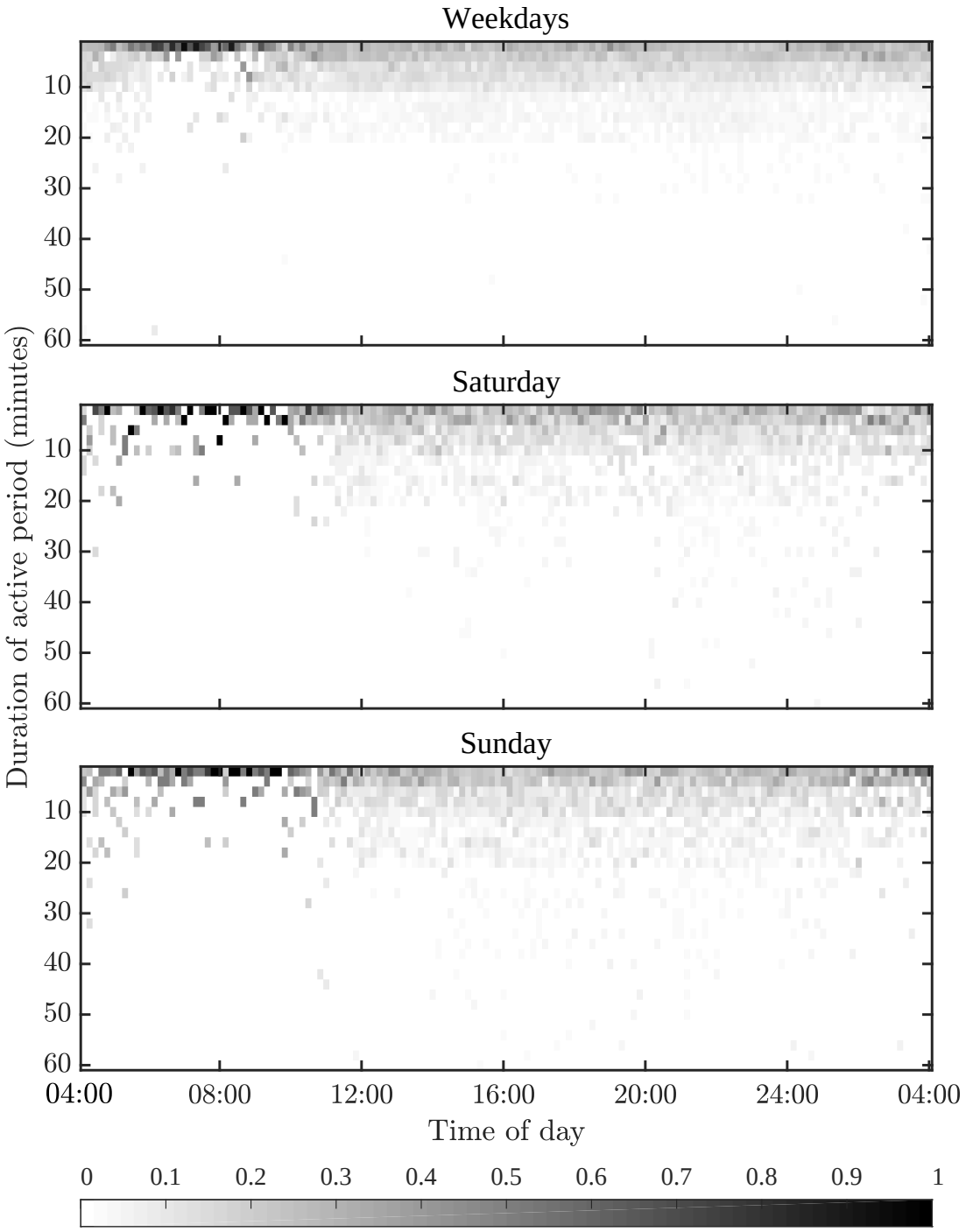


Figure 5.4: *Example of the joint probability distributions of duration of cycle of operation with respect to time of day for the different types of day considered. The shading intensity indicates the probability in accordance with the scale at the bottom of this figure. Case: Microwave oven.*

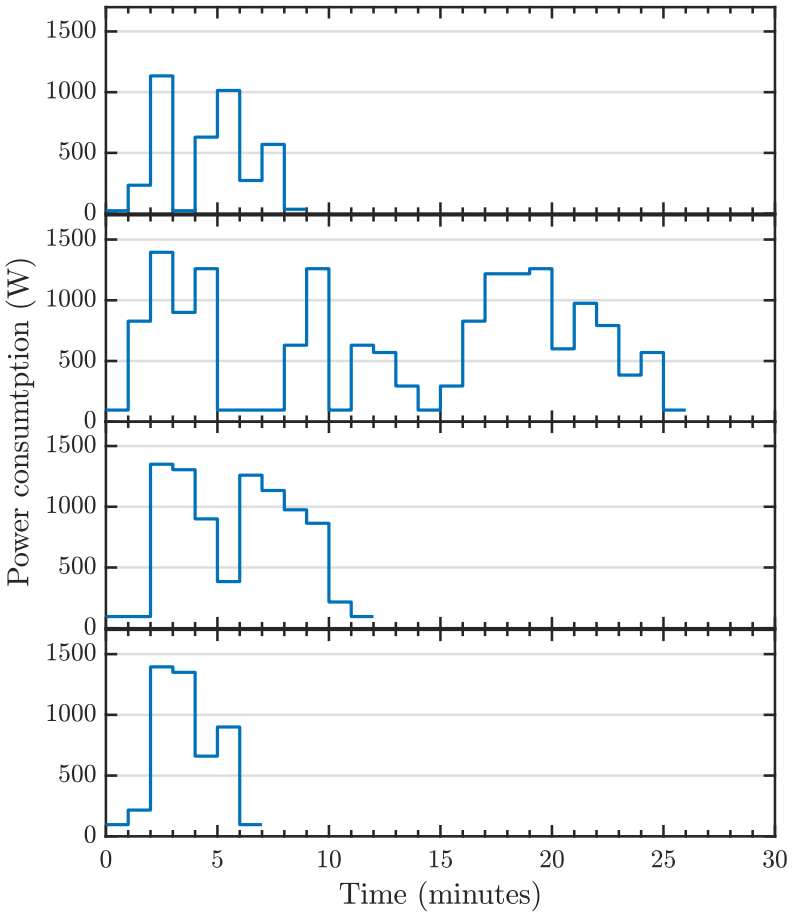


Figure 5.5: *Sample appliance power consumption traces for different durations. Case: Microwave oven.*

5.2.3.2 Simulating activity-driven power usage

The focus of the simulation of activity-driven power usage lies on providing an accurate representation of the interaction between the members of the households themselves and the interactions between household members and the appliances available to them. To this end, an agent-based interaction logic is defined to rule the interactions between these two types of agents: household members and appliances. Household members are characterised by their simulated activity schedules and the appliances are characterised by the consumption behaviour

profiles resulting from the appliance characterisation routine discussed in the previous section.

For the purpose of the simulations, the interaction between users and appliances determines whether there is an actual 'need' for the use of a given type of appliance associated with the activity the member of the household in question is engaged in; and if the 'need' exists, for what specific purpose the appliance is needed (i.e. programme cycle, duration of the usage episode, etc.). The interaction between household members determines whether there is co-engagement in a particular activity at a given point in time, and whether this involves appliance sharing or simultaneous use of different devices of the same kind.

The potential for user-appliance interactions is assessed at every time step in the daily activity sequences and it is determined by the factors associated with the type of activity and with the types of appliances associated with the activity in question. For each episode of engagement in a certain activity, the switch-on probability profiles of the associated appliances are used to determine whether these are activated. When an activation event does occur, the likely duration of the corresponding cycle of operation needs to be determined. This is done based on the type of appliance, the type of day and the time of day where the activation event occurs. The appropriate joint probability distribution of the durations of the cycles is sampled to determine the length of the period of active power consumption. Once the length of the cycle of operation is known, the set of power consumption profiles associated with the relevant appliance is sampled to select one that matches the corresponding duration of the cycle of operation.

Table 5.5 shows the association between the activities considered by the

activity pattern simulation module and the appliances considered for the simulation of activity-driven power usage.

As Table 5.5 shows, both food preparation and laundering activities are associated with multiple types of appliances. In the case of food preparation activity, the associated types of appliance are or can all be used to achieve the same goal; the typical home has a number of appliances that can all be used in the same cooking session, sometimes at once even. In the case of laundering activity, however, the multiple types associated are meant to achieve different purposes when used; when people engage in laundering activities, if they have access to a clothes dryer, they might need/decide to use it every so often.

Table 5.5: *Activities and associated appliance types.*

Activity	Associated appliances
Food preparation	Cooking appliances: cooker, oven, hob, microwave, toaster, kettle
Laundering	Laundering appliances: Washing machine, washer dryer, Tumble dryer
Dish washing	Dishwasher
TV watching	TV sets (1 - 4), set-top box(es), dvd(s),
ICT related	Desktop PC, Laptop(s)
Generic active occupancy	CD player, radio, kettle, iron, vacuum cleaner

The general approach to the identification of appliance switch-on events and the determination of the length of periods of operation has been detailed above. However, these special cases corresponding to activities associated with

multiple types of appliances need to be handled in a different manner. As the analysis by Ramírez-Mendiola et al. (2018) [Chapter 3] shows, the probability of using the different types of appliances with respect to the engagement in the associated activity varies throughout the day. In practice, this means that electricity consumption derived from the engagement in this kind of activities also exhibits some kind of variability associated with the use of the different types of appliances. In order to take this into account, special routines for the handling of cooking and laundering appliances were implemented. These routines incorporate the relationships observed between engagement in activity and the conditional probability of the activation of the different types of appliances associated.

- Handling of cooking appliances

Cooking appliances are sub-divided into minor and major cooking appliances. Major appliances are appliances such as cookers, electric hobs and ovens, and microwave ovens, which are typically central to cooking practices. Minor appliances are appliances such as toasters and kettles which can either be used independently or to occasionally assist in the main cooking procedures.

For the purpose of the simulations and for the sake of simplicity it is assumed that the use of major appliances for a given period of active cooking is mutually exclusive, thus avoiding the complications of determining the amount of overlap between cycles of operation. In order to be able to determine which one of the major appliances available is likely to be the main appliance in use, two considerations are made. Firstly, it is necessary to identify the combination of major appliances available to the household in question. Secondly, when

a member of the household engages in food preparation activities and it is determined that a cooking appliance needs to be activated, the conditional probability distribution for the activation of the particular combination of appliances present is used to determine the appliance likely needed. The analysis by Ramírez-Mendiola et al. (2018) [Chapter 3] was extended to extract the conditional probability distributions specific to the different combinations of major appliances observed.

The handling of minor appliances is more flexible.The activation of minor appliances might be required even when a major cooking appliance is already in use. For instance, if extra hot water is needed for the preparation of a meal for which the cooker is the main cooking appliance, if a kettle is available it is possible that this would be used to produce the hot water needed. These possibilities were taken into account for the development of the handling routine for cooking appliances, where the activation of the minor appliances is determined independently from major appliances.

- Handling of laundering appliances

Three possible arrangements of laundering appliances are considered for the simulations; the household owns either a washer dryer, a washing machine and a tumble dryer, or a washing machine only. The main distinction concerning the handling of laundering appliances relates to the presence of a tumble dryer.

When the only laundering appliance available is either a washing machine or a washer dryer, the activation of the appliance is determined based on the standard procedure detailed above. If both a washing machine and a tumble dryer are

present, every time a household member engages in laundering activities it has to be determined which one (if any) of the two appliances is likely to be required.

Similar to the case of cooking appliances, the likelihood of the activation of these two types of appliances varies with respect to the time of day. Therefore, a similar approach to the case of cooking appliances is taken. That is, conditional probability distributions for the activation of the two types of appliances were extracted based on the analysis presented in Ramírez-Mendiola et al. (2018) [Chapter 3]. These distributions are used to determine the type of appliance likely to be required depending on the conditions for the possible activation event, paying special attention to time of day. Once it has been determined which appliance is likely to be required, it is determined whether the appliance is actually required to perform the task at hand. In the case of the washing machine, this is determined based only on the corresponding switch-on probability profile. In the case of the tumble dryer, however, two other factors are taken into account. The activation of a tumble dryer is ruled out if 1) the washing machine is still in use by the time it is determined that it is likely that the tumble dryer might be activated, and/or 2) the washing machine has not been run before the time it was determined the tumble dryer might be required. The possibility of requiring the use of a tumble dryer without having previously used the washing machine exists. However, for the sake of simplicity, it is assumed that a tumble dryer run would only be (sensibly) required when a washing machine's load needs to be dried up.

- Conflicting use of appliances

During the simulation runs it is possible that more than one household member tries to engage in the same activity. Therefore, it needs to be determined whether the activity is likely to be carried out jointly or independently.

In order to determine whether shared use of appliances is likely or not, the activity co-engagement probability profiles described in Section 5.2.2 are used. The probability of activity co-engagement varies depending on the time of day and the activity in question. Therefore, every time a potential conflict in the use of the appliances available to the households arises, the relevant activity co-engagement factor is used to resolve the conflict. When shared use of appliances is ruled out, it is considered the possibility that the other user might decide to make use of another appliance associated with the activity in question. The probability of this happening is evaluated following the standard appliance handling procedure described above.

- Lighting

For the simulation of power consumption associated with lighting a previously developed approach was used (Richardson et al. 2009). The lighting model developed by Richardson et al. (2009) uses statistics from the lighting association to populate each home with different sets of lighting units which are meant to be representative of the typically observed residential lighting configurations. In addition, in order to account for the seasonality of lighting usage, the model incorporates a natural lighting factor. This factor depends on the outdoor irradiance and feeds directly into the likelihood of any individual lighting being turned on. For each month, an average daily irradiance profile is used to account for the variability of sunlight throughout the year. A calibration factor is also used to adjust the demand for energy derived from the use of artificial lighting so as to produce overall demand estimates in accordance with available statistics. During the simulation runs, the household activity profiles are fed into the lighting demand model in order to determine when it is likely that artificial

lighting would be needed.

5.2.3.3 Simulating non-activity-related power usage

Contrary to activity related energy uses, there are some other appliances/energy uses for which household activity has little or no effect over them. That is the cases of devices that fall into the category of base-load appliances. This includes cold appliances and other devices such as internet routers.

Despite of the little effect household activity may have over them, it is important to capture the variability of the power loads associated with them. To this end, power consumption profiles were also extracted from the appliance-level metered data. However, the appliance profiles constructed for this kind of devices are different from those of the activity-related energy uses. Instead of focusing on the periods of active power consumption, entire daily power consumption profiles were extracted. For each day of the simulation period, one of the available profiles is chosen to represent the power consumption associated with the appliance in question. This is then added to the power consumption profiles calculated for the activity-related energy uses.

5.3 Validation analysis

The aim of the validation analysis is to verify that the power consumption data generated by the model has similar characteristics to the metered power consumption data from the households considered in the validation dataset.

This validation dataset contains 10-min resolution power consumption profiles of 20 households that were monitored for a full year[5].

The most elementary components of the output of the simulations are the daily household-level power consumption profiles; a profile is produced for each day of the simulated period. For the purposes of the validation analysis the length of the simulation period was set to 1 year[6]. The individual household loads are then aggregated to look at the overall power demand for the corresponding group of households, and to produce estimates of overall energy requirements (e.g. total annual consumption). The model was used to generate power consumption data for several groups of 20 households with similar characteristics to the metered counterparts in the validation dataset.

The observed variability and the statistical characteristics of the power consumption data depend to a large extent on the level of aggregation. Therefore, it is important to look at these characteristics in detail, and assess the quality of the simulated output produced by the modelling framework in terms of the characteristics most relevant according to the level of aggregation. The validation analysis will thus take a sequential approach, analysing the output data from the lowest to the highest level of aggregation considered.

The similarity of metered and simulated data is evaluated making use of different metrics, which are relevant in terms of the statistical characteristics of power consumption data. Some of these metrics look specifically at characteristics which are required to perform an adequate analysis of the power flows in a local network, or for other standard network design and planning procedures.

[5]The households were monitored as part of the 2010 UK Household Electricity Survey (DECC & Intertek Testing and Certification Ltd. 2016).

[6]The length of the simulation period can be specified at the start of the simulation runs. However, if this value is not specified, the default length is 365 days.

Quick overviews of the qualities being assessed are provided in each subsection.

5.3.1 Simulated power consumption and average daily load profile

One of the key elements to an adequate simulation of the electricity consumption patterns is an accurate representation of the variability observed in the power consumption profile throughout the day. The observed variability changes with the level of aggregation. However, since households are considered as the basic unit for the purpose of the simulations, it is necessary to ensure that this variability is well reproduced at the household level.

As Figure 5.6 shows, the model produces very realistic daily power consumption profiles. The overall patterns are very similar, featuring recurring high spikes occurring predominantly around the morning and evening peak times, and lower demand variability for the periods in between. A two-person household was chosen as an example for the comparison between metered and simulated power consumption as this type of household is one of the most common. Power consumption profiles corresponding to a full week were extracted from the metered data of one of the two-person households available, and from the data generated for its simulated counterpart. As the sample metered and simulated households are not identical, there are some natural differences between the power consumption profiles, such as the exact times of peaks and their magnitude for a given day. The comparison shown in Figure 5.6 is merely illustrative.A more detailed comparison of the power consumption loads in terms of relevant statistical metrics is provided below in Section 5.3.2.

In addition to the need for an accurate representation of the intra-day variability at the household level, it is also important to ensure that the diversity of daily power consumption profiles is equally well represented. When the metered daily load profiles are averaged over a certain period and across multiple households, a profile with a characteristic shape is obtained. Therefore, it would be expected that a model that accurately represents this diversity would yield an average daily load profile with similar characteristics.

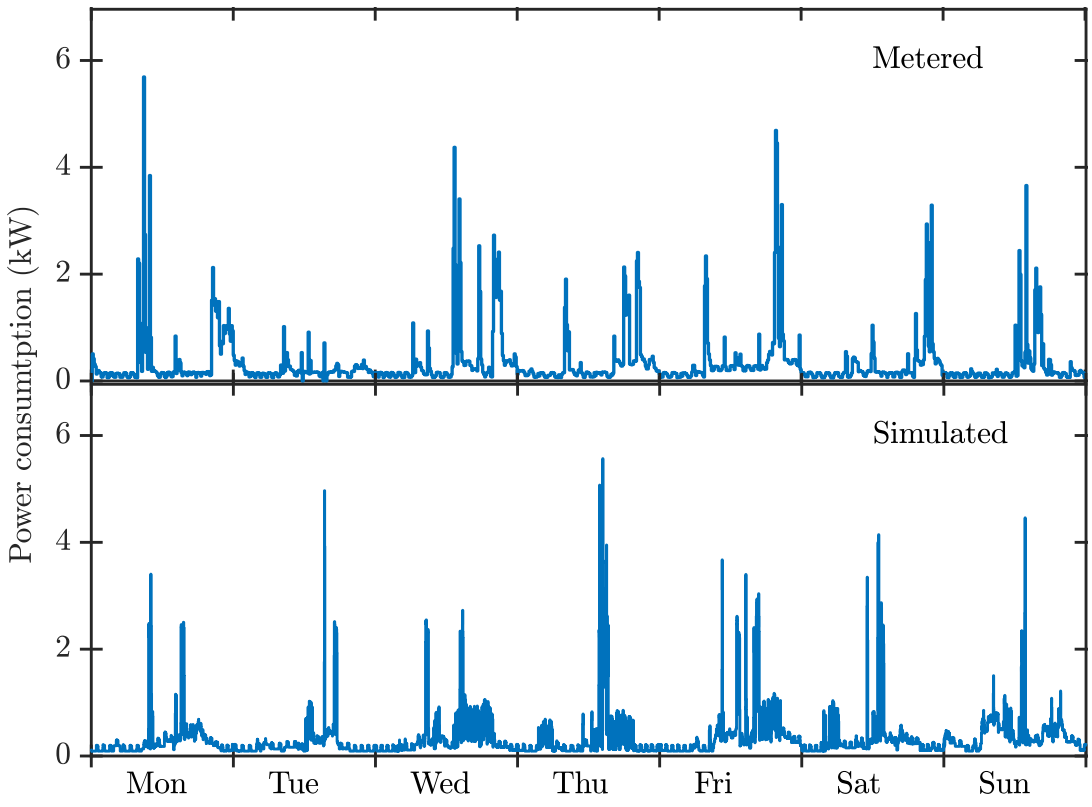


Figure 5.6: *Top: Sample weekly metered power consumption from a 2-person household. Bottom: Sample weekly simulated power consumption from the simulated counterpart.*

As the top plot in Figure 5.7 shows, the average daily load profile obtained from the simulated power consumption closely resembles the one obtained from the metered data. The RMS deviation between the metered and simulated average daily load profiles is of 4%. The continuous line is the smoothing spline fitted to the simulated average daily load profile and the shaded region corresponds to its confidence interval.

In order to further assess the quality of the simulated power consumption estimates, we also look at the level of variability throughout the day in terms of the standard deviation associated with the average daily load profile. The continuous lines in the bottom plot in Figure 5.7 correspond to the plots of the average standard deviation of the power consumption with respect to the time of day. The shaded regions around them correspond to the 1σ band. That is, the upper and lower bounds at one standard deviation as they vary throughout the day.

As we can observe, the daily variation of the standard deviation of both metered and simulated data follows roughly the same pattern, and there is substantial overlap between their corresponding shaded bands. It is also observed that the magnitude of the variability would appear to be proportional to the average power consumption for a given time of day, as the standard deviation profile follows a very similar pattern to the one observed in the average daily load profile.

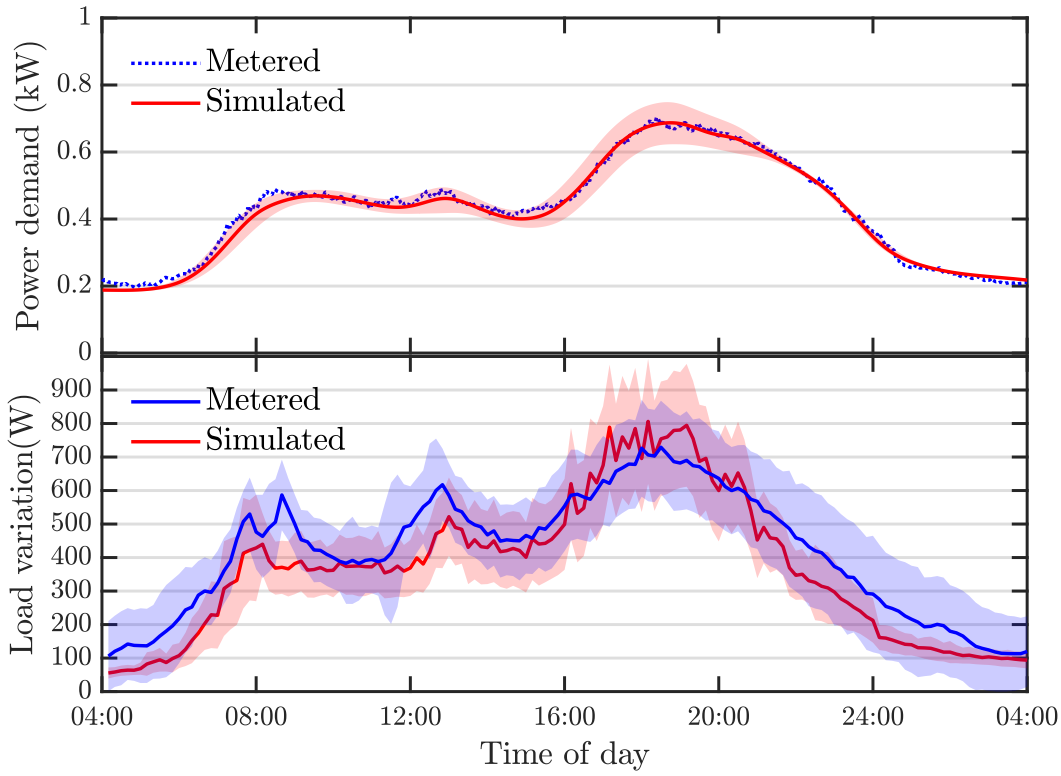


Figure 5.7: *Top: Average daily load profile for metered and simulated households. Bottom: Daily profile of load variability with respect to time of day.*

5.3.1.1 Comparison with previous approaches

As we discussed in Section 5.1.2, previous modelling approaches have tended to focus on the representation of the shape of the average daily load profile. Therefore, comparing the performance of different approaches regarding this would offer a means of assessing the potential improvements over the previously developed approach.

Among the examples of activity-driven electricity demand models for the UK residential sector, the model developed by Richardson et al. (2010) has found

widespread application. The modelling approach proposed by Richardson et al. (2010) is based on the simulation of periods of active home occupancy through first-order Markov chain. The daily active home occupancy profiles are then converted to power consumption profiles. We will refer to this model as the CREST model[7].

In order to compare to the metered power consumption profile and the one simulated using the approach proposed in this paper, the CREST model was used to simulate power consumption for a group of 20 households, for a simulation period corresponding to 1 year. Metered and simulated profiles are then compared using the time-series analysis metric known as Dynamic Time Warping (DTW) distance $[D_{DTW}]$[8].

As Figure 5.8 shows, the profile obtained from the data generated by the proposed model is closer to the 'optimal match' represented by the diagonal dotted line. In quantitative terms:

$$D_{DTW}(metered, simulated) = 22.76$$

$$D_{DTW}(metered, simulated_{CREST}) = 70.02$$

The relative percentage difference between these two distances is 67.5%.

[7]The model was named after the Centre for Renewable Energy Systems Technology (CREST), where it was developed.

[8]The Dynamic Time Warping distance is a time-series analysis metric commonly used to measure the similarity between two temporal sequences sampled at different speeds or resolutions.

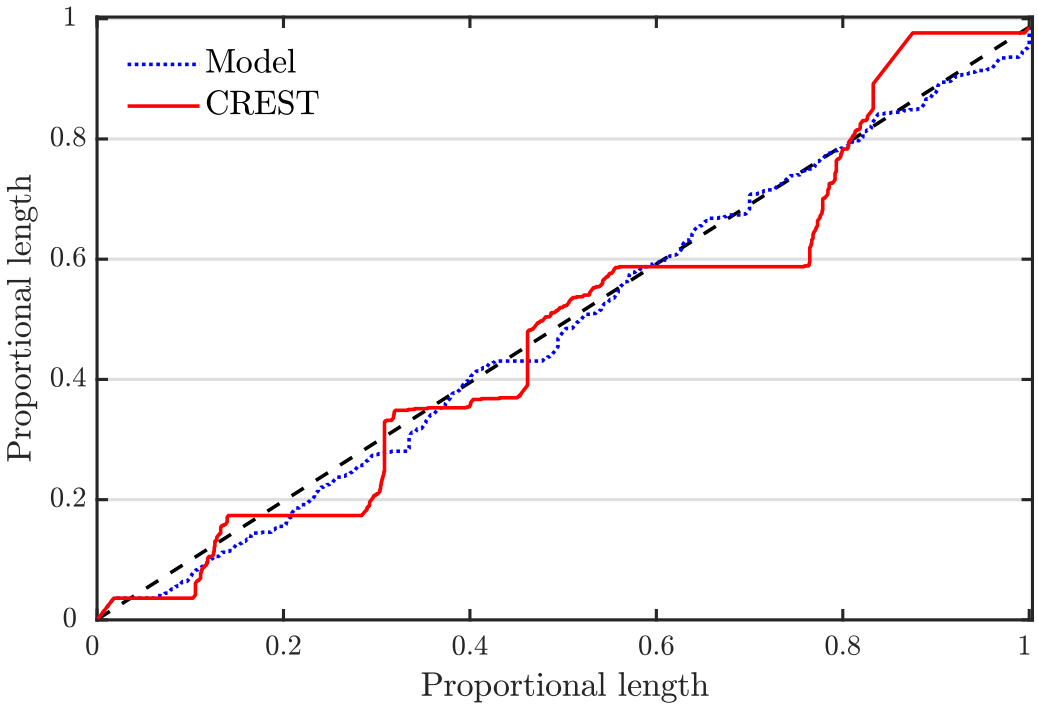


Figure 5.8: *Comparison of metered and simulated average daily load profiles. The diagonal dashed line indicates the optimal warping path. The continuous red line and dotted blue line associated with the simulated profiles correspond to their respective best warping path.*

5.3.2 Load curve metrics

In addition to the analysis of the variability of the power loads at the individual household level, it is also necessary to ensure that the effects of load aggregation observed in the metered data are well represented in the output of the simulations.

Standard statistical measures such as average values and standard deviation provide us with some information about the nature of the load variability. However, the characteristics of the observed load profiles can also be analysed in terms of more relevant metrics such as the Load Factor. For a given load profile, the load factor is defined as the ratio of maximum to mean power. Therefore, the value of this factor varies between 0 and 1.

Load aggregation tends to have a smoothing effect due to the random (non-)coincidence of the power loads associated with different households. A side effect of this overall smoothing is that the peak loads of the aggregate loads would appear less pronounced compared to the average load. In terms of the load factor, this would result in higher values. That is, the smoother the load profile, the higher the load factor. We can observe this effect in the top plot in figure 5.9.

To produce the curves observed in this sub-figure the following approach was taken. The daily load profiles of both metered and simulated households corresponding to the same day were increasingly aggregated and the corresponding load factor was calculated. Then the average over all the different combinations of days and number of households was calculated, along with the standard deviation of the observed values. The curves in the top plot in Figure 5.9 thus correspond to the variation of the (average) Load Factor with respect to the level of aggregation (number of aggregated households). The shaded bands around the curves correspond to the upper and lower bounds of the range of variability at one standard deviation. As we can observe, both curves follow the same trend and are very close to each other, which would suggest that the model produces a good representation of the change in variability as the number of simulated households increases.

Another effect observed when the loads associated with different households are aggregated is that the individual loads coincide more or less depending on, among other things, the activity patterns of the households. Therefore, a simulation of household power consumption based on the simulation of household activity patterns must be able to reproduce this effect.

The coincident power consumption behaviour of individual households can be quantified by another relevant load curve metric known as the Coincidence

Factor. The coincidence factor is defined as the ratio of the maximum value of the aggregate load to the sum of the maxima of the individual household loads. The coincidence factor also offers some information about the temporal nature of the power consumption; the magnitude of this factor gives an idea of how likely it is that the maximum power consumption of the individual households would be observed at the same time.

A closely related load curve metric is the Diversity Factor, which is the reciprocal of the Coincidence Factor. This factor, as the name would appear to suggest, quantifies the load diversity. The higher the value of the diversity factor, the more diverse the loads associated with individual households are in terms of the time where the maximum power consumption occurs.

In order to further assess the model's ability to represent the effects of coincident behaviour on the observed power consumption loads, we look at the coincidence factor of the aggregate metered and simulated loads at different times of day.

The bottom plot in Figure 5.9 shows the variation of the coincidence factor associated with metered and simulated power consumption with respect to the time of day. To produce the curves in this sub-figure, the individual household load profiles corresponding to a given day were aggregated. Then, the coincidence factors associated with hourly portions of the daily profiles were calculated, and averaged over the different daily aggregate loads. The curves in the bottom plot in Figure 5.9 correspond to the change in the (average) hourly coincidence factor with respect to the hour of day. The shaded bands are again the one-standard-deviation variability ranges.

With the exception of the first six hours of the day, the curves associated with

metered and simulated data would appear to remain within the same region. Moreover, the difference between the overall average coincidence factors is of less than 3%. This would suggest that the model produces a very realistic representation of the coincident behaviour of household power consumption loads, particularly during daytime.

Another feature of interest in terms of the representation of power consumption is the distribution of the demand from high to low levels observed over a given period. Figure 5.10 shows the hourly average load duration curves for metered and simulated households corresponding to the loads observed over the course of one year. The shaded bands around the load duration curves correspond to the range of variability across households at one standard deviation. It is observed that the overall distribution of the simulated loads closely resembles the distribution observed in the metered loads. This suggests that the representation of the spread of power loads between high and low levels produced by the model is realistic.

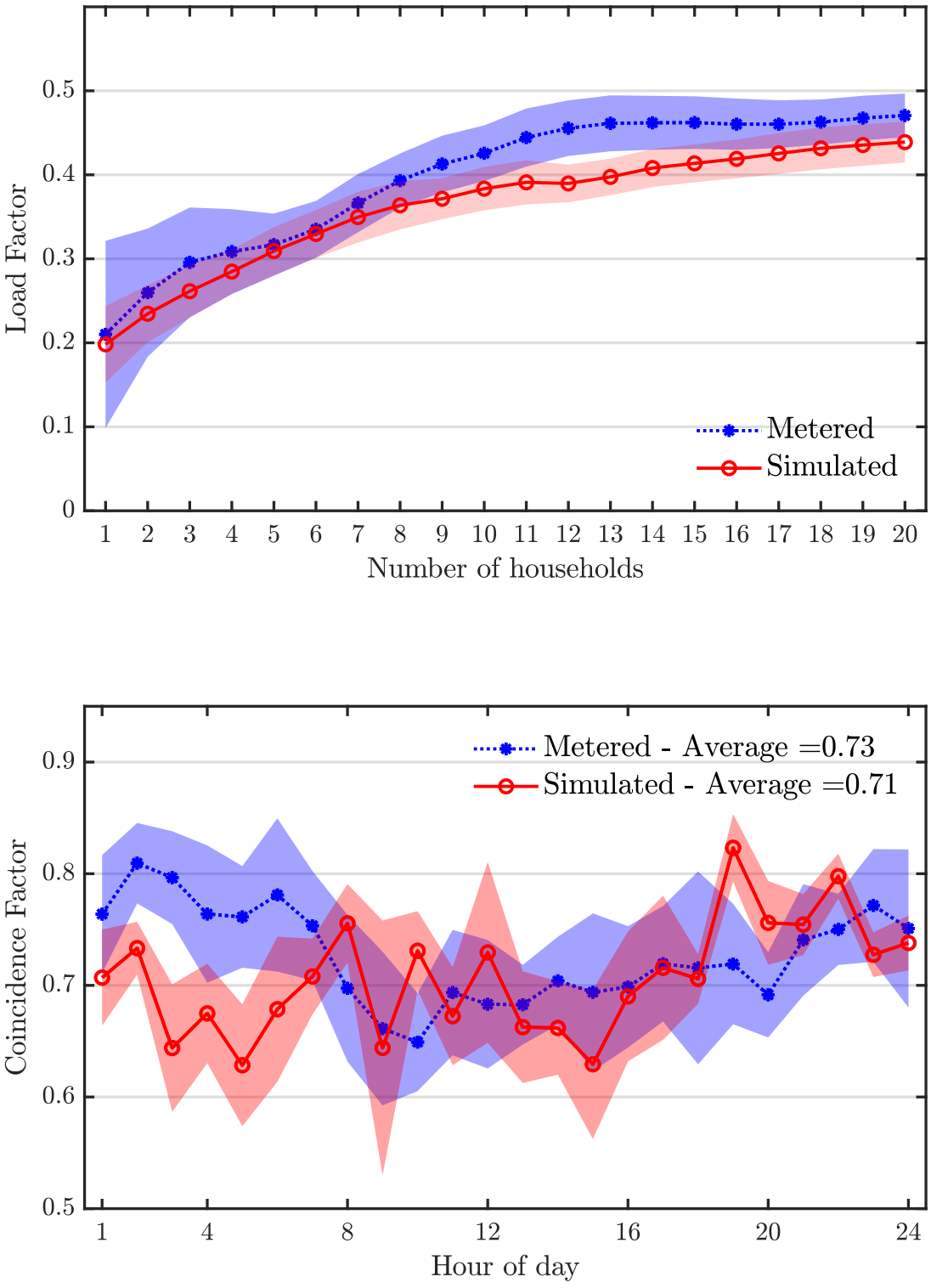


Figure 5.9: *Top: Change in Load Factor with respect to level of aggregation. Bottom: Hourly Coincidence factor profiles for metered and simulated loads.*

The variability observed across metered households, however, is generally higher than across the simulated counterparts. In addition, it is also observed that in the case of the simulated loads, the level of variability decreases as the level of demand does. This effect could potentially be attributed to the lack of diversity in the appliance profiles used for the purpose of the simulations compared to the diversity of the appliance stock found in real households. The demand loads at the lower end of the distribution are the demands associated with stand-by and other forms of inactive power consumption. As the range of variability of the metered load duration curve shows, this type of loads vary considerably more across real households.

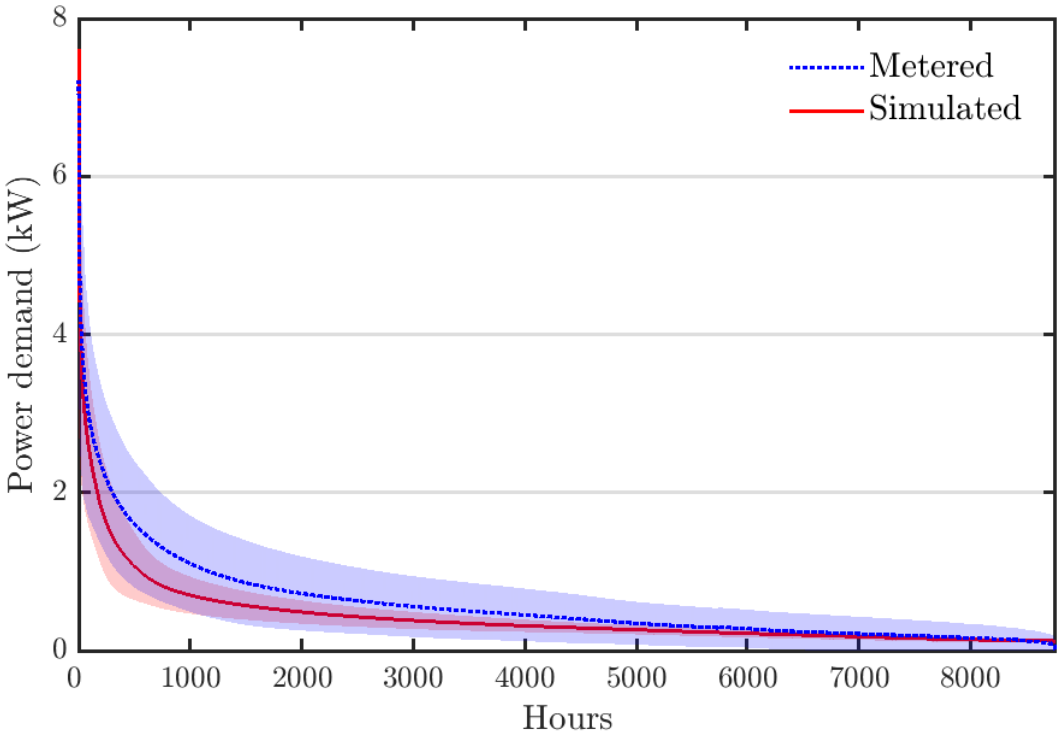


Figure 5.10: *Average load duration curves for a 1-year demand period.*

5.3.3 Monthly demand estimates

Another dimension of the variability observed in household's demand for electricity is the changes in their energy requirements with respect to the time of year. Figure 5.11 shows the variation in average household demand for electricity on a monthly basis. The shaded band around the curve for the simulated estimates corresponds to the upper and lower bounds of the range of variability at one standard deviation.

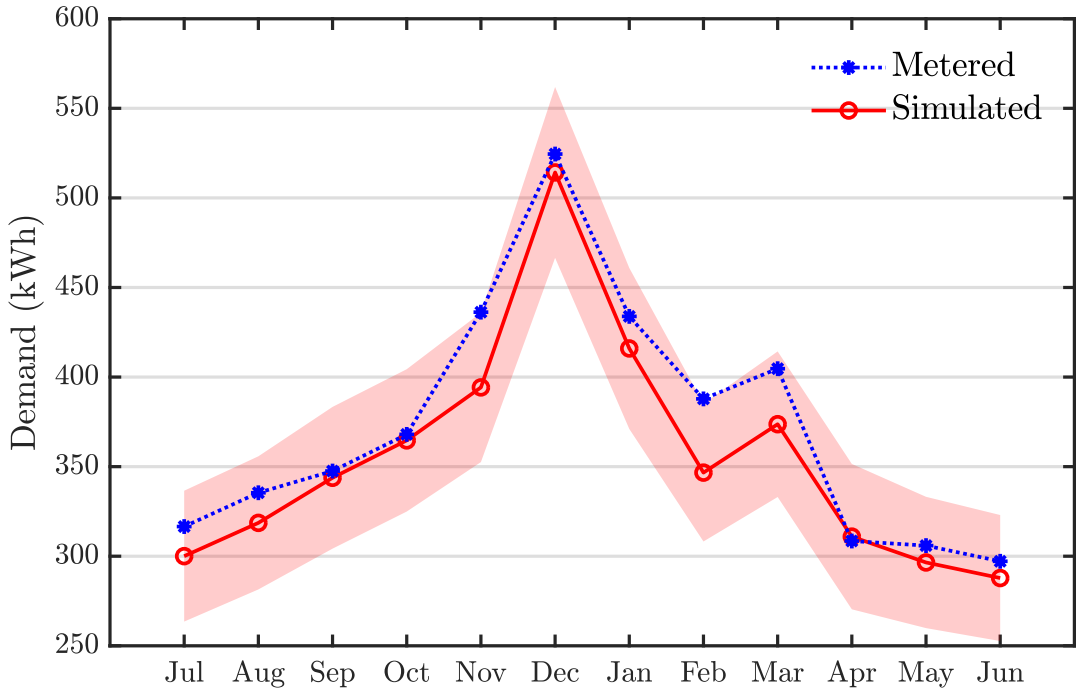


Figure 5.11: *Seasonal trend of the average total monthly demand for electricity.*

The seasonal trend observed in the power consumption simulated by the model follows the same pattern as the average consumption of the metered households. In general, it is also observed that the model slightly underestimates the monthly consumption compared to the metered average values. However,

the metered estimates are almost entirely within the standard deviation bands obtained for the simulated estimates. The largest difference between the metered and simulated average monthly demand estimates is just over 10%, and the correlation between metered and simulated curves is of 0.98, which suggests that the model's representation of the seasonal trend is sufficiently accurate.

5.3.4 Distribution of households with respect to total annual electricity consumption

The final level of aggregation we consider for this validation analysis is the total annual demand for electricity. Figure 5.12 compares the distribution of metered and simulated households with respect to the level of annual electricity consumption. The shaded regions indicate the 95% confidence interval for each of the empirical distributions. The vertical line indicates the average annual consumption.

As we can observe, the distribution of the simulated households is well within the confidence interval of the empirical distribution of metered households. However, in order to compare the two distributions in terms of their statistical characteristics, a Kolmogorov-Smirnoff test was performed. With a p-value of 0.41 at a 5% significance level, the test reveals that it is very likely that the set of simulated annual consumption estimates corresponds to a set of values drawn from a distribution with the same statistical characteristics as the distribution observed for the metered estimates. That is, that the two distributions in Figure 5.12 are virtually equivalent.

As it can be seen from the monthly consumption estimates in the previous sub-section, the model has a tendency to slightly underestimate total monthly

demand for electricity. Consequently, there is also a slight discrepancy between the average metered total annual consumption and the simulated estimate. However, the difference between these two is under 2 %. The small difference between the average total annual consumption estimates, along with the results of the statistical test used to compare the distributions of metered and simulated estimates, suggest that the model also produces a realistic representation of the spread of households across the different ranges of total annual consumption.

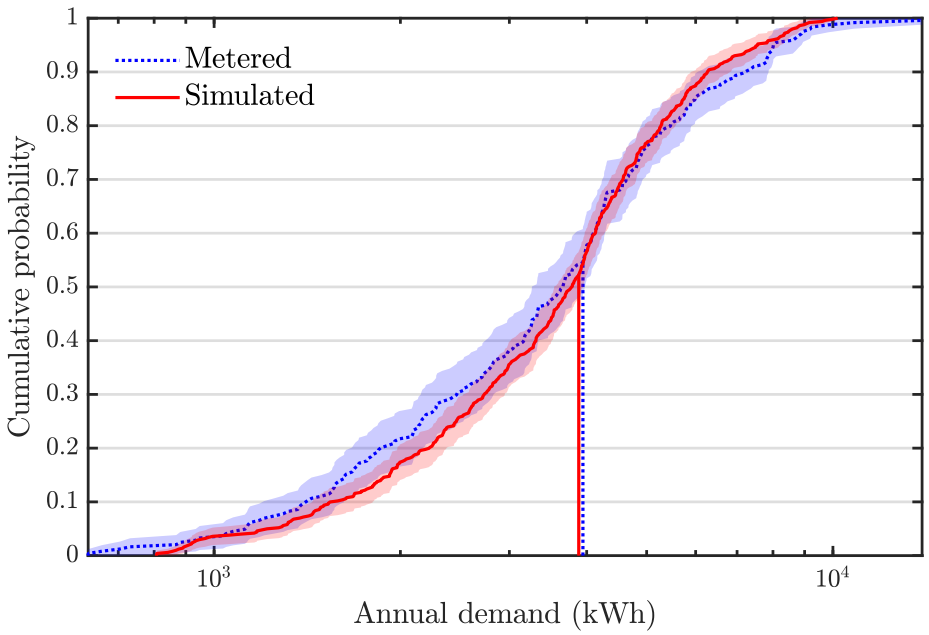


Figure 5.12: *Distribution of households with respect to total annual consumption.*

5.4 Discussion

We have presented a comprehensive approach to the simulation of household power consumption that aims at capturing the effects of the diversity of households on the variability observed in their demand for electricity. The modelling framework presented in this paper was developed with a modular structure in mind. The different modules work in combination to ensure that the variability of overall energy requirements and consumption behaviours observed in real households is well reproduced. The comparison of the simulated power consumption behaviour with the behaviour observed in a group of real households shows that the model is able to produce a realistic representation of the statistical characteristics of metered data. The simulated power consumption results from the simulation of the interaction between households and appliances. However, for the purpose of the validation of the model, the lowest level of aggregation considered corresponds to the daily load profile at the household level.

The demand for electricity is often metered and analysed on a half-hourly basis, as this level of resolution is generally good enough for the purposes of market trading or high-level network analysis. Therefore, many modelling efforts have focused on the reproduction of power consumption patterns at this level of resolution. Such coarse load profile estimates, however, are of little use for the purposes of analysing the operation of low-voltage networks in detail, or the potential of demand-side management measures. The modelling framework was configured to generate daily load profiles at a 1-min resolution. We consider that this resolution is a reasonable time-scale that allows for a detailed analysis of the fast intra-day variability of power consumption. Re-configuring the model

to produce higher-resolution load profiles, although possible, would require new datasets at higher-resolutions than the data used to build the model. However, the model can be readily re-configured to produce lower-resolution profiles (e.g. half-hourly) if desired.

5.4.1 Advantages of the proposed approach

5.4.1.1 Characterisation of simulated households

The household characterisation module of the proposed modelling framework ensures that the socio-demographic diversity of real households is adequately mapped onto the simulated households. Using current national socio-demographic statistics allows us to use the model to produce a snapshot of the current situation. This can be quite useful in itself. However, more often than not, we are also interested in learning more about how the situation might look like in the future. By establishing explicit links between detailed socio-demographic statistics such as the proportions of different types of households and differences in appliance ownership rates across different age ranges, the proposed modelling framework offers the opportunity to analyse how the situation might change based on observed and projected demographic trends. For instance, the increase in life expectancy is a major driving force of population growth and changes in its distribution (ONS 2012). In particular, in the UK the proportion of people aged 65 and over is expected to increase from about 18%[9] to about 20% by 2025 and about 23% by 2035 (ONS 2012). Another expected consequence of the rising population is that there will be a proportionally larger number of one person households (ONS 2012). Therefore,

[9]As of 2012, year of the last census.

the number of households will increase at a faster rate than the population. Both of these trends are likely to have an impact on the observed electricity demand patterns and overall energy requirements, and both factors would have to be taken into account in order to produce estimates for the changes in such patterns.

5.4.1.2 Approach to the simulation of household activity

The approach to the simulation of household activity patterns presented by Ramírez-Mendiola et al. (2019) [Chapter 4] has been shown to provide better results than the commonly used first-order Markov chain technique. The improved representation of the diversity of daily household activity schedules produced by this approach is reflected in the results of the validation of the modelling framework presented in this paper. As the results presented in Section 5.3 show, the model would appear to capture the diversity in daily power consumption patterns satisfactorily enough. Moreover, an approach to the simulation of household activity based on the representation of full daily schedules might also prove useful for other potential applications of the model. For example, being able to identify the loads associated with particular activities in detail might help in providing more realistic estimates of how much of the potential of demand response measures can actually be realised. This could be assessed, for instance, based on how strongly dependent are the activities associated with the power loads of interest on the time of day, and how sensitive these might be to changes in scheduling.

5.4.1.3 Approach to the simulation of appliance-level power consumption

The synthetic generation of the power consumption variations during appliances' periods of activity can become particularly complex, as even appliances of the same type can have considerably different power requirements during periods of active use. These requirements vary depending on, for instance, the chosen settings (e.g. washing machine programmes), and the technical specifications (i.e. nominal power ratings). The approach to the simulation of appliance power consumption used in the proposed modelling framework is based on sampling real appliance power consumption profiles. This approach is preferred over the synthetic generation of said profiles for the following reasons. Firstly, the way appliances are used, and particularly, the duration of the periods of active use, change with respect to the time of day. This issue is addressed by taking into account the distribution of the duration of the metered cycles of operation started at a given time of day to determine how long the appliance will be active for. Secondly, the time variability of the power consumption derived from actual user-appliance interactions depends to a large extent on what the appliance is being used for. This issue is addressed by taking actual power consumption profiles for the type of appliance in question, which are associated with the use of the appliance for different purposes (e.g. Figure 5.5). Taking these two factors into account allows for a more realistic representation of the use of different types of appliances for a variety of purposes, compared to the commonly used simplifying approach that ignores this diversity of uses.

The special handling routines for cooking and laundering appliances described in Section 5.2.3 were mainly developed in order to prevent unrealistic power

consumption peaks. In particular, these are meant to prevent situations in which, for instance, all cooking appliances are activated at once.While the probability of this happening in real life is greater than zero, it is in general an uncommon practice. In addition, making a distinction between the appliance likely being used for a particular purpose is also important in terms of the representation of the associated power consumption. The overall energy requirements and power consumption profiles associated with, for instance, a cooking session using the oven will differ from those of a session using a hob. Adding this dimension of variability also increases the model's ability to represent the diversity of power consumption behaviours across and within households.

5.4.2 Potential for improvement

As the results presented in Section 5.3 show, the model already produces a satisfactory representation of the aspects of interest. However, as with any other model, a number of potential extensions and refinements can be envisaged. For instance, given the relevant data, the model could be extended to include power loads associated with water and space heating, or charging of electric vehicles.

The bottom panel in Figure 5.7 shows that night time variance would appear to be under-represented by the model, compared to the variance of the metered power consumption which would appear to be constant throughout the day. This is most likely due to the fact that the idealised representations of the appliances used in the simulations exhibit a limited variability during periods of inactivity; in most cases, this is reduced to a constant stand-by consumption or nothing at all. However, while treating the metered data with the purpose of characterising these idealised appliances, the variability of the 'noise' filtered out was higher due to

several instances where the recorded meter readings exceeded the identified stand-by consumption threshold. Based on the information made available as part of the metered power consumption dataset, it was not possible to determine whether these variations were due to the appliances or the meters themselves.

In the context of the transition towards a more decentralised power system, it is becoming increasingly relevant to assess the potential effects of distributed generation resources on the overall load profiles. In the case of solar PV, for instance, it would be necessary to assess the potential contributions of the locally generated electricity to the changes in the demand profiles. This would require the integration of the demand modelling framework with a model that would produce estimates of the potential for solar PV generation such as the model presented in Bright et al. (2015). This model generates synthetic irradiance profiles based on weather observation data. The irradiance profiles generated by the model would appear to adequately capture the irradiance fluctuations and variability inherent in the available solar resource (Bright et al. 2015). Due to the present configuration of the modelling framework presented in this paper, an integration with such a model would also benefit the representation of the power consumption associated with artificial lighting, which is also based on the use of daily irradiance profiles to estimate how likely is artificial lighting needed.

Although all sorts of possible refinements and improvements could be suggested, one needs to bear in mind that these would almost inevitably increase the complexity of the modelling process. Therefore, the value of any potential extensions or refinements would have to be assessed in terms of their technical cost. In this context, the potential technical costs refer to the trade-off between the increases in complexity and the relative improvements in the output of the simulations. The higher the complexity of the modelling approach, the higher

the computational intensity of the simulation process. However, more powerful computational resources are increasingly available, so part of the technical costs of more complex simulations might be offset by the use of larger computational units.

5.5 Conclusions

A novel modelling framework for the simulation of residential demand for electricity has been presented.The modelling framework takes a hybrid approach that combines the strengths of stochastic modelling techniques and agent-based modelling methods to produce a realistic representation of household-appliances interaction and the resulting power consumption profiles. A detailed simulation of the interaction between users and appliances allows for the simulation of high-resolution power consumption profiles associated with different activity categories. The validation analysis shows that the model effectively captures the diversity observed in overall household energy requirements (e.g. total annual consumption) and daily power consumption behaviours. The simulated power consumption is analysed at different levels of aggregation, and the statistical characteristics of the simulated output compare well to the characteristics of the metered data. The RMS deviation between metered and simulated average daily load profiles is of 4%, and the difference between simulated and metered average total annual consumption is under 2%. The quality of the output of the simulations is further assessed in terms of metrics relevant to the design and planning of power systems. The positive results of the validation analysis suggest that the model could be effectively used for this purpose. In addition, the nature of approach to the simulation of household activity patterns would allow for the

use of the proposed modelling framework for a more detailed analysis of the potential of demand response measures, and to produce more realistic estimates of how much of this potential can actually be realised. The characterisation of residential energy consumption patterns could be further improved by doing a more geographically resolved analysis, as different types of households may exhibit substantially different patterns, and the composition of the residential sector in terms of the proportions of different types of households may vary significantly across different regions. Therefore, the use of national survey data imposes clear limitations to this kind of analysis. At the moment, the sample sizes of the surveys carried out thus far do not allow for a robust analysis based on regionally split sub-samples, and region-specific studies are rare. However, both the recruitment of larger samples and region-specific studies might become possible at some point in the near future as the use of more ICT-based surveying systems might allow for reaching considerably higher numbers of respondents without incurring in prohibitively higher costs and additional complications.

6 | Conclusions

6.1 Introduction

The need to decarbonise our energy systems has driven researchers and policy-makers to seek the most adequate alternatives to the centralised power generation paradigm. Over the last two decades it has become evident that a complete reworking of our power systems is necessary if we are to achieve the stringent decarbonisation goals that would allow us to meet the climate change mitigation targets, restricting global warming to 2°C or less (IPCC 2014). The integration of large shares of variable renewable power generation generally entails an overall reduction of the systems' flexibility (Bevrani & Raisch 2017, Ringkjøb et al. 2018, Tielens & Van Hertem 2016). Therefore, putting in place additional system balancing mechanisms is of the essence. One of the most fundamental changes to the current model entails enabling a more active participation of the demand-side in the overall system balancing (Altmann et al. 2010).

The potential for the implementation of demand-side mechanisms that would enhance overall system flexibility is increasingly recognised. Demand-side response is in fact often seen as a key component of future energy systems as it would allow for a more efficient use of power generation resources. However, in order to successfully implement such energy efficiency measures, it is necessary to develop a thorough understanding of the way the observed demand loads arise. This is especially true in the case of the residential sector, where the great diversity of end-users results in an equally diverse gamut of energy requirements.

The aim of this thesis is to investigate the relationship between key drivers of demand for energy, or electricity in particular, and observed demand patterns in the residential sector. With this aim in mind, a modelling framework was

developed that maps key socio-demographic factors and household activity patterns to residential energy demand loads.

The models developed in this thesis provide a new set of tools for appraising behaviour-oriented energy efficiency measures targeting the residential sector. This thesis thus makes a contribution towards our understanding of the potential of this type of demand-side management strategies. In particular, the approach to the analysis of residential activity and energy demand patterns offers new insights to inform the development of strategies tailored to specific groups of end-users.

The present chapter will review the thesis findings and offer conclusions. The knowledge gaps discussed in Chapter 1 will be revisited to clarify what has been learned about them. Finally, the chapter will be concluded with a summary of an outlook for future research.

6.2 Summary of findings

The investigation presented in this thesis has been arranged into two stages. The first stage was concerned with characterising residential electricity consumption. The findings of this analysis are presented in Papers 1 and 2 (Chapters 2 and 3). The second stage was concerned with using the insights gained from the analysis in the first stage for the design and development of a novel modelling framework for activity-driven residential electricity consumption. The further analysis carried out for this stage is presented in Papers 3 and 4 (Chapters 4 and 5). The findings of each paper are summarised and discussed below in relation to the knowledge gaps discussed in Chapter 1, and to the research questions introduced in Section 1.2.1.

6.2.1 Chapter 2: Characterising the diversity of energy requirements

6.2.1.1 Rationale and aim

A good understanding of the diversity of energy requirements across different segments of the consumer base is key to developing adequate energy systems planning strategies through better informed decision-making. The review of the literature showed that several studies acknowledge the importance of accounting for such diversity. However, the review also revealed that most current electricity demand models are blind to the extent of the variability of energy demands across households.

In the studies that have made attempts to take this variability into account, users have been loosely divided into groups referred to as low-, medium- and high-consumption. But in general, little effort has been put into studying the relationship between the socio-demographic characteristics of users and their corresponding energy consumption.

Identifying and (quantitatively) characterising the different segments/clusters of the consumer base is essential to the incorporation of this kind of information into the models assisting in analysing and planning energy systems. Therefore, it is necessary to develop a better understanding of these complex relationships.

The aim of this chapter is thus to provide an answer to the first research question that concerns this thesis: how does the diversity of socio-demographic characteristics map to the diversity of total energy demands? The analysis

presented in this chapter provides a characterisation of metered household electric energy demands and compares this to the representations of their diversity provided by current modelling approaches.

6.2.1.2 Findings and discussion

Chapter 2 has presented a comparative analysis of the statistical characteristics of metered and simulated electricity consumption datasets. The simulated data was generated using a widely implemented probabilistic modelling approach (see Section 2.3).

The findings of the analysis presented in Chapter 2 provide further evidence to support the consensus in the literature regarding the segmentation of the consumer base into low-, medium- and high-consumption groups. The analysis of the metered household annual consumptions revealed that these follow a complex distribution from which three clusters can be identified. And it was found that a trimodal Gaussian mixture model better captures the diversity of real life households and the complexity of their distribution with respect to total electric energy requirements.

Chapter 2 has also established that the underlying assumptions made in previously developed modelling approaches (e.g. Richardson et al. (2010)) result in much simpler distributions. When the distribution of the simulated households was compared to that of the metered ones, it was found that the misrepresentation of the observed variability can be as high as 40%. Moreover, statistical testing revealed that the probability of the distribution of simulated total energy requirements being consistent with the distribution of the population where the metered data stems from is of less than 5%.

The findings of the analysis presented in Chapter 2 have advanced the field of residential electricity demand modelling by providing a more rigorous characterisation of each of the identified segments in the consumer base. The statistical model presented in this chapter proves a useful tool for a consistent identification of the three highlighted segments. A quantification of the proportion of different types of users within each segment has also been provided.

In summary, the comparative analysis in Chapter 2 reveals that the simplifying assumptions commonly used in current modelling approaches result in a considerable loss of information in terms of the distribution of households across dimensions such as total energy demands and socio-demographic compositions. Thus, the need to incorporate the information about the diversity of energy demands into our residential electricity demand models is further emphasised.

6.2.2 Chapter 3: Characterising activity and appliance usage patterns

6.2.2.1 Rationale and aim

There are two key elements to the activity-driven modelling of energy demand loads: the simulation of the activity patterns driving the demand, and turning these patterns into the associated demand loads. The latter has been generally done by means of some kind of (fixed) activity-to-power conversion rule(s) that result in an almost direct conversion between the two.

The demand load estimates produced by models based on this approach had been sufficient to give an idea of the overall daily demand load profile. Recently,

however, the interest in analysing the intra-day variability of domestic electricity consumption in more depth has increased considerably. This calls for a more detailed study of the links between activity and appliance usage patterns.

The review of the literature shows that the need to better understand the timing of energy-related activities and the use of the associated appliances has been increasingly recognised (Flett & Kelly 2017, Torriti 2017). The review also revealed that advances in this line of research have been slow, primarily due to the lack of relevant data. The lack of studies looking at gathering the required data can in turn be attributed to the historically prohibitive costs of such an endeavour. However, recent developments and cost reductions in enabling technology have made it affordable to collect such data, and these newly available sources offer the opportunity to learn more about the subject.

The aim of this chapter is thus to provide an answer to the second research question that concerns this thesis: how do household activity patterns relate to domestic appliances usage patterns? The analysis presented in this chapter provides a characterisation of the usage patterns of domestic appliances commonly found in UK homes. Moreover, the analysis compares the identified appliance usage patterns to the patterns of engagement in the activities the appliances are associated with. This comparative analysis seeks to further unravel the relationship between household activity and electricity consumption patterns.

6.2.2.2 Findings and discussion

Chapter 3 has explored the links between residential activity patterns and the usage patterns of the appliances that can be reasonably associated with the

activities in question. The analysis presented in this chapter shows that the simplistic activity-to-power conversion rules on which the simulation of activity-driven residential electricity demand has been largely based mask the reality behind those links.

The case of activities such as laundering or food preparation, which are commonly associated with different types of appliances, is particularly complex. The analysis reveals that the usage patterns of the different types of appliances vary independently from one another with respect to the overall activity pattern. For instance, in the case of laundering activity, which can be associated with the use of washing machines and clothes dryers, the peak in activity engagement occurs in the morning. When both a washing machine and a clothes dryer are present in a given home, washing machines were found to be up to 80% more likely to be the appliance in use in the morning. Conversely, when present, clothes dryers are about 50% more likely to be the appliance in use in from the late afternoon onwards.

Dish washing was found to be a particularly good example of timing mismatches between appliance usage and reported activity engagement. The activity engagement profile follows a rather predictable pattern that features peaks after typical morning, midday and evening mealtimes. The dishwasher usage profile was found to show some resemblance to the activity profile, but the peaks occur approximately 30 to 40 minutes later.

Another key finding of the analysis presented in this chapter is the effect of 'background appliance usage'. The gap between the levels of activity engagement derived from reported time-use data and the levels of activity suggested by the active use of TV sets or ICT equipment varies with respect to the time of day.

During peak times, when the activities in question are more likely to be the primary activity, the gap is of approximately 25% only. During daytime, however, this gap can increase to up to 70%, which shows how much more likely is that the type of appliances associated with these activities might be left running in the background.

Chapter 3 has thus contributed to the field of electricity demand modelling by improving our understanding of the relationship between daily patterns of engagement in energy-related activities and the use of the appliances that give rise to the associated demand loads. The implications of the findings discussed above are in line with the conclusions drawn from previous studies on the importance of the timing of appliance use and energy-related activities (e.g. Flett & Kelly (2017), Torriti (2017)).

In summary, the findings presented in this chapter provide further evidence to support the need to improve the assumptions made in modelling approaches regarding user-appliance interactions and activity-to-power conversion rules. The analysis also provides a resolved temporal mapping between reported activity engagement profiles and the usage profiles of the appliance associated with the activity in question which further unravels the links between the two. This mapping thus may serve as the basis for the development of enhanced activity-driven electricity demand models.

6.2.3 Chapter 4: Improving the approach to residential activity modelling

6.2.3.1 Rationale and aim

Over the last decade, the importance of the integration of user behaviour into energy demand modelling has been increasingly recognised. The approaches to the simulation of residential activity patterns, however, have largely focused on the representation of the general features of the average activity patterns.

The techniques used thus far have proven effective for this purpose. Nevertheless, the need for a better understanding of the scheduling of daily activities calls for a more in-depth analysis of the user activity patterns and a more robust approach to the simulation of such patterns.

Modelling techniques such as the first-order Markov chain offer a simple mechanism for the representation of the features observed in aggregate activity patterns. However, the probabilistic properties that define this modelling technique mean that its adequacy for the study of the daily scheduling of activities is limited.

The features that characterise the overall activity patterns are an emergent property of the evolution of the daily individual activity sequences. A number of studies on energy demand through social science lenses have argued that social constructions of time impose certain restrictions on the scheduling of activities, and this results in the development of habitual or characteristic behavioural patterns that shape the way daily activity sequences develop in time (Shove &

Walker 2014, Walker 2014). Furthermore, it is argued that the timing of demand for energy is largely dependent on such characteristic behavioural patterns.

In spite of this, the integration of the insights from the study of energy demand from the social science perspective into the development of the activity modelling approaches that underpin energy demand models is yet to become a common practice. If models are to help us better understand the emergent properties of daily scheduling of activities and the evolution of daily activity sequences, we need to give more attention to the way empirical daily activity sequences develop in time.

The aim of this chapter is to explore the answers to the third research question that concerns this thesis: how can the influence of habitual behaviours on daily household activity scheduling be incorporated into household activity modelling? This chapter thus sets out develop a novel approach to the analysis of empirical activity data that focuses on the identification of the characteristic behavioural patterns with a view to developing new activity modelling approaches that take account of their influence.

6.2.3.2 Findings and discussion

Chapter 4 has presented the development of a new approach to activity modelling whose primary aim is to exploit characteristic behavioural patterns for the simulation of residential activity sequences. The use of said characteristic patterns to underpin the simulation of daily activity sequences offers the possibility of simulating household activity patterns by means of a self-adapting process that results in a more realistic representation of activity patterns, both at the individual and aggregate level.

The development of this model was supported by the development of an algorithm that allows for the identification of characteristic patterns as well as the points in the activity sequences at which the presence of such patterns might have significant influence on the evolution of the daily activity sequences.

The relevance of the identified patterns is assessed by means of a gain function. This function assigns a gain value to every identified pattern. Between 60 and 70% of the identified patterns were found to be on the high-influence region of the gain value scale.

Performance improvements over a conventional first-order Markov chain model were assessed based on two key factors: representation of the overall activity patterns at different levels of aggregation, and representation of the duration of the periods of activity. Overall activity patterns were calculated averaging over increasing numbers of individual activity sequences (from 10 to 10,000). The comparison of the RMS deviations with respect to the empirical patterns reveals that the average improvement across datasets with increasing numbers of simulated activity sequences is 28%. The comparison of the distribution of the durations of periods of activity for the different activities considered reveals that the deviation of the first-order Markov model is 30% higher on average.

In addition to the use of standard error metrics, we also proposed a new metric to assess in more detail the quality of the simulated output in terms of the representation of the intra-day dynamics of the activity sequences. The comparison of the deviations estimated by this new metric support the conclusions drawn based on the use of the more standard metrics.

Chapter 4 has thus contributed to the field of electricity demand modelling

by demonstrating that incorporating the insights from social science research into the development of modelling approaches does have the potential to improve our understanding of the relationship between activity and energy demand patterns. The identification of patterns of activity, and more importantly, those that might be relevant to daily activity scheduling can prove challenging. However, the analysis presented in Chapter 4 shows that said characteristic behavioural patterns are indeed present in the empirical activity data currently available. Furthermore, the results show that these can be effectively used to provide a more robust approach to activity modelling.

In summary, the results of the analysis suggest that the proposed approach does have the potential to be used as a new tool for a more detailed analysis of the structural complexity of empirical activity data. Such an in-depth analysis would help broaden our understanding of the underlying behavioural patterns that shape the daily residential activity schedules.

6.2.4 Chapter 5: Improving the approach to appliance usage modelling

6.2.4.1 Rationale and aim

Power systems are undergoing a radical transformation where distributed, intermittent power sources are playing an increasingly important role. In order to accommodate for this changes, additional ways of enhancing overall system flexibility are being sought. In addition to increased operating reserve, network reinforcements and storage solutions, the potential of increasing system flexibility through demand-side response mechanisms is being explored.

A detailed assessment of the potential of DSR mechanisms requires a thorough understanding of the demand loads. In the case of the residential sector, this calls for a detailed analysis of the relationship between residential activity and electricity consumption patterns.

Models can be useful in this regard. But if the goal is to produce an accurate representation of the demand loads arising from the residential sector, it is essential to capture the intricacies of the interaction between users and the domestic appliances commonly found in typical homes.

The review of the literature shows that there is a number of studies that look at the use of simulated activity patters to underpin the simulation of demand loads. However, the review also shows that in previously developed models the approach to the conversion of activity patterns into corresponding electricity demand loads has been largely ruled by simplifying assumptions that mask the complexity of user-appliance interactions. In some cases, for instance, it is assumed that power consumption is restricted to the periods of active occupancy only (e.g. Yao & Steemers (2005)). Some other examples try to take this approach one step further by considering special cases where particular appliances run during cycles of fixed duration, and therefore, might be active for longer periods than those of simulated active occupancy (e.g. Richardson et al. (2010), Widén & Wäckelgård (2010)). However, be it due to the lack of data or for the sake of simplicity, models based on this 'conventional' approach assume that appliances run for a fixed amount of time while consuming a constant amount of power.

Recent developments and cost reductions in enabling technologies have made it possible to collect new, highly-resolved residential electricity demand

data. These newly available sources offer the opportunity of both testing the assumptions made in previously developed approaches, and developing new enhanced approaches to the simulation residential electricity demand.

The aim of this chapter is thus to explore the answers to the fourth research question that concerns this thesis: how can the simulation of user-appliance interactions and activity-to-power conversion be improved to enhance load modelling performance? This chapter thus sets out develop a novel modelling framework that combines the approach to residential activity modelling developed in Chapter 4 with the insights from the analysis presented in Chapters 2 and 3 to account for the differences in the usage patterns of different appliances and the diversity of household energy demands.

6.2.4.2 Findings and discussion

Chapter 5 has presented a residential electricity demand modelling framework that combines the strengths of the stochastic approach to the simulation of residential activity patterns presented in Chapter 4 and a new approach to the simulation of the interactions between users and appliances commonly found in UK homes. This new approach draws on agent-based modelling methods, and was developed with a view to provide a more accurate mapping between user-appliance interactions and resulting electricity demand loads.

The different types of simulated appliance agents were characterised based on time-resolved metered power consumption data. This characterisation looked at identifying the different power consumption behaviours associated with different potential uses for each of the appliances considered. The temporalities of the

power consumption behaviours of the different appliances were characterised both in terms of the variation of power consumption levels during a given cycle of activity, and in terms of the occurrence of said cycles with respect to the time of day.

The quality of the simulated demand loads was assessed at different levels of aggregation, giving particular attention to the representation of the intra-day variability of power consumption. The analysis of the model output shows that the simulated demand load profiles closely resemble the profiles in the metered sample.

It was found that the modelling approach developed in Chapter 5 produced load profiles that closely resemble the profiles extracted from the metered dataset. The analysis shows that the RMS deviation between metered and simulated mean daily load profiles is 4%.

The assessment of the quality of the representation of the intra-day variability was based on the comparison of the evolution of the load coincidence factor throughout the day. The analysis of the hourly load coincidence profile reveals that the model adequately captures this variability. It was found that the difference between the metered and simulated overall average coincidence factors is of less than 3%.

In order to provide some perspective in terms of the improvements represented by the modelling approach proposed in Chapter 5, the output of a model based on the conventional approach was compared to both the metered load profiles and the output of the proposed model.

Making use of metrics commonly used in time-series analysis, it was found that there is a 67% relative percentage difference between the output of both

models. Thus, the analysis proves that the representation of residential power consumption behaviour is considerably improved by using the proposed modelling framework.

Chapter 5 has thus contributed to the field of electricity demand modelling by demonstrating that an approach that focuses on a more detailed representation of the interactions between residential users and electric appliances yield better results in terms of the representation of residential demand loads. As the literature points out, an approach that places the relationship between user activities and the use of appliances at the core of the modelling effort is key to an adequate assessment of the potential of DSR strategies targeting the residential sector, as these tend to be highly appliance specific.

The results of the analysis presented in Chapter 5 further highlight the benefits of such an approach. However, the development of this kind of approaches to residential electricity demand modelling has seen little progress. The lack of relevant data makes the development of this kind of approaches challenging. Nevertheless, the analysis presented in this thesis shows that the newly available sources of highly resolved data offer valuable opportunities to improve our residential electricity demand models.

6.3 Future outlook

Ten years after the European parliament agreed upon their most recent energy directive (European Parliament & Council of the European Union 2009), there is still much debate about what is the best way of ensuring a cleaner supply of energy while drawing on locally available clean energy resources. The overall

direction of the transition, however, is clear and it is increasingly evident that a complete reworking of our energy systems will be needed if we are to deliver on any acceptable target for global warming. Over the past years it has also been increasingly recognised that the role users will play in future low-carbon systems will be vital for achieving the stringent decarbonisation goals for our power systems. Models offer the opportunity to not only explore the current role of energy users but also to identify trends and the potential for interventions that would allow for a more efficient use of the available resources. However, if energy demand models are to inform the development of strategies that aim to influence the way energy is being used, the role of energy users must be at the core of the modelling efforts.

6.3.1 Contributions of this thesis

This thesis takes one step further the development of more holistic approaches to residential energy demand modelling that enable a more in-depth analysis of the influence of the social factors on the observed demands. A detailed characterisation of the overall energy demands and energy-related behaviours of UK residential users based on empirical activity and electricity consumption data results in a more realistic representation of electricity demand loads. The approach to the simulation of residential activity patterns and the interaction of residential users with the domestic appliances commonly found in UK homes incorporates some of the insights from the study of energy-related behaviours from the social science perspective. This seeks to offer a more comprehensive approach to the study of energy demand patterns that could potentially be used to inform the development of energy efficiency measures targeting the residential sector.

A comparison between empirical and simulated electricity consumption data generated using a model based on the current modelling paradigm reveals that conventional approaches fall short in terms of the representation of the diversity of energy demands. However, as this thesis shows, it is possible to make a better use of the available data to develop enhanced approaches and improve the representation of this diversity both at the intra-day and aggregate levels.

The findings of this thesis imply that a modelling approach that focuses on the role of users in producing the observed electricity demand loads is indeed more suitable for the analysis of the potential of residential DSM strategies. An approach such as the one developed in this thesis may prove particularly useful as the simulation of residential activity patterns is based on the identification of characteristic behavioural patterns that might require carefully designed interventions if they are to be changed for the better.

This thesis has also demonstrated that conventional approaches to residential electricity demand modelling fail to accurately represent the interaction between users and domestic appliances. Throughout this thesis, a distinction has been made between the variability of electricity demand associated with human activity *per se* and the variability associated with the power consumption behaviours of the appliances themselves (e.g. differences in power ratings, functions and programmes/cycles of operation). The analysis presented in this thesis has shown that an approach to activity-to-power conversion in which the outcome of user-appliance interactions is reduced to blocks of constant power consumption during fixed periods of time fails to represent the variability observed in demand load profiles. Thus, the need for improved approaches to the simulation of user-appliance interaction has been highlighted. Furthermore, this thesis suggests that more attention should be given to the representation of

the power consumption behaviours of the simulated energy-consuming devices considered in energy demand models.

The appliance characterisation methodology presented in this thesis (Chapter 5) is an attempt to capture the variety of energy consumption profiles associated with different uses of the same type of appliance. By taking into account the variety of functions the different domestic appliances are expected to fulfil, the simulation of user-appliance interactions results in demand profiles more akin to those observed in empirical data.

The relevance of such attention to detail in the development of this and future modelling approaches is highlighted by a key issue: the fact that introducing new technologies offering the same services, or the addition of completely new system components with largely unknown demand loads might lead to entirely different results in terms of the potential energy efficiency gains resulting from DSM interventions. Therefore, this thesis further emphasises that a meticulous, independent representation of the consumption behaviours of energy-consuming devices themselves is just as important as the representation of the user activity patterns that result in the observed demand load profiles. This is so that a clear distinction can be made between the two in terms of their potential for energy efficiency increases.

6.3.1.1 Relevant insights for Demand-Side Management

Demand-Side Response, often used interchangeably with demand-side flexibility, sits alongside remand reduction under the umbrella of Demand-Side Management practices, which may involve the use of technological, economic and policy

instruments to achieve their goals. The findings of this work and the tools developed based on them may prove useful in supporting the development of such instruments and the relevant approaches to manage residential electricity demand.

The development of DSR programmes requires the assistance of models based on a detailed understanding of the turndown potential of the different demand loads so that they can accurately forecast when load shifting may take place (Torriti 2017). While this might seem like a common practice in other sectors of the economy, the current participation of the residential sector in the demand response market is seriously underdeveloped, mainly due to the lack of information about residential load profiles and shiftable loads.

A recent study on the effectiveness of the models currently being used to inform UK policy-making found that they over-estimate the demand for electricity associated with the use of appliances by as much as one third, and with lighting by as much as one quarter (Ward & Darcy 2014). Therefore, estimates of household load which technically may be shiftable or reducible have, most likely, been exaggerated by these models. Conversely, the electricity demand associated with cooking is under-estimated by as much as one third (Ward & Darcy 2014). While loads associated with cooking are, presumably, rather inflexible, this still represents a problem in terms of the estimation of capacity requirements. The extent to which the penetration rates of residential DSR can be increased will depend on how accurate is the information about different demand loads; more specifically, their timings and the causes of the observed peaks.

Some practical examples of potentially direct applications of the findings and tools resulting from the research presented in this thesis include the

development of economic and technological DSM instruments such as dynamic pricing strategies and automated demand-side control systems. The development of any pricing system or control technology (e.g. appliance remote controllers) seeking to steer consumption patterns and shift demand loads in time will have to be designed based on the state of current consumption practices. Therefore, detailed characterisations of residential activity and appliance usage patterns and modelling tools that capture their diversity, such as those developed in this thesis, are an essential component. In addition to providing information on what people are likely to be doing on a given day, this can also indicate the times of day when they might actually be responsive to dynamic pricing systems, or when they might be affected by automated load shifting.

It is envisioned that automated load control systems will play a major role in managing the electricity consumption of different types of household appliances. These load controllers offer a number of opportunities, but they also present complex new challenges. The control systems are expected to be able to prioritise different loads depending on both system and users' requirements, which will require the development of suitable algorithms (Torriti 2017). And the development of these algorithms underpinning automated load control systems will need detailed information about the energy-consuming practices of different groups and types of users. Actively managing energy demand, however, effectively alters the diversity of loads on the system. And this can result in undesirable consequences such as "payback" or "rebound" effects, which increase or pose new challenges to the system. Models of residential electricity demand that produce a good representation of the diversity of end-users offer the opportunity to explore the effects of changes in such diversity. Therefore, they allow for a more adequate assessment of the effectiveness of interventions that might induce such changes, and assist in the identification of potential

shortcomings.

The modelling framework developed in this thesis might also be of use if the realm of policy-making. Analysing and understanding demand for electricity from the perspective of residential users is key to the development of policy means that can contribute to both increasing system efficiency and promoting sustainable everyday life (Ellegård & Palm 2011, Palm & Ellegård 2011). This understanding of how energy use contributes to current everyday life practices is a vital ingredient to connect household activity and energy policy targeting the residential sector.

There are two key barriers to the development of household-relevant policy which are closely intertwined. Firstly, capturing the uniqueness and complexity of the everyday life practices of different households, while at the same time demonstrating the existence of characteristic patterns and habitual behaviours common to many can prove really challenging (Palm & Ellegård 2011). But in doing so, the possibility of using such characteristic behavioural patterns as a criteria for clustering individuals or households opens up. This, in turn, could help researchers in identifying constraints affecting particular types of users and policy-makers in tailoring policies targeted to specific groups.

This relates to the second barrier, which is associated with the delivery of information and policy-related advice to consumers. There appears to be a disconnect between the language of policy-makers and the language of common households. It is often the case that when this kind of policy-related advice is provided to the households, they find it too general and, therefore, difficult to relate and respond to (Ellegård & Palm 2011). Thus, it is essential that policy-makers have access to the kind of tools that would allow them to relate the information they want to transmit to the households' activity patterns and

everyday life rhythms (Ellegård & Palm 2011). In this context, modelling tools that provide a good representation of the demand loads associated with the different activities households engage in may provide policy-makers with the necessary elements to better craft the required policy instruments.

6.3.1.2 Final overview of core chapters

While no part of the analysis presented in this thesis is superfluous, some of the contributions of this thesis may find more widespread uses and applications than others. Chapters 2 and 3 make use of fairly standard statistical and data analysis methods, and the findings reported in them are highly dependent on the data analysed. There is, of course, some scope for testing the generality of the findings reported in these chapters, but this is contingent on more the kind of data studied becoming available.

Chapters 4 and 5, on the other hand, detail the development of novel methodologies for the simulation of residential activity and electricity consumption patterns which may be readily applied to further study energy demand. The methods developed in Chapter 4 are particularly innovative, as they draw on advances in other disciplines, such as information theory and bioinformatics, which had not previously found a connection to energy research. The details about the development of the models described in this chapters have, or are soon to be, published. This means that they can now be replicated by other researchers interested in exploring their potential applications.

The nature of the data these models were based on entails a continued model development, if their validity over the years is to be ensured. However,

if the frequency of the studies where the data that underpins these models stems from remains the same, this might mean very long waiting times in between re-validations. With the advent of smart meters and other more sophisticated monitoring technologies, this might soon cease to be a problem. However, there remains a certain level of uncertainty about the extent to which these data might be available for research purposes.

6.3.2 Outlook for future research

To what extent can developments in electricity demand modelling enhance the informing process for the development of energy efficiency measures? This is a weighty question, a full discussion of which falls beyond the scope of this thesis. Models have proved very useful in giving an idea of the current state improving our understanding of the different factors at play. But as this thesis has shown, there is still room for improvement in the models' ability to help identify opportunities for targeted interventions and assess their potential. Based on the areas for improvement identified in this thesis, three issues are worthy of discussion.

The first issue relates to the overall approach to residential energy demand modelling. This thesis has shown that energy demand modelling approaches have focused on the representation of overall energy demand patterns at high levels of aggregation. Thus the potential to use the models based on conventional approaches to study the energy consumption associated with specific activities or end uses is limited. The comparison of the simulation output of a model based on the current paradigm with the approach developed in this thesis has proven that a focus on the representation of the interactions between users and appliances at the most basic level, rather than compromise the models' ability to represent

the overall demand load features, enhances it. The features that characterise aggregated demand load profiles are emergent properties of the daily evolution of people's engagement in a number of activities and their interactions with the electric appliances available to them. Therefore, if the goal is to reproduce the variability that results from the interactions occurring in 'real-life', then the simulation of residential demand for electricity should focus on capturing the particularities of such interactions.

The second issue relates to the data needed to underpin residential energy demand modelling. The mismatch between reported and 'metered' activity (e.g. see Figure 3.3, Chapter 3) as extracted from currently available data, makes it clear that better data sources are needed. Smart meters are expected to address this issue by providing a means of collecting high-resolution[1] energy demand data at the household level, and perhaps even at the appliance level through the use of data processing algorithms (e.g. NILM[2]). However, there are some concerns about the extent to which these data will actually be available for the purposes of energy demand modelling. Privacy concerns, for instance, are likely to be an issue. In the situation where the data does become widely available, it will be necessary to develop adequate data analysis algorithms to process the huge volumes of data that are expected to be collected by smart meters so that meaningful information can be extracted from it. The issue of finding better ways of coupling time-use and appliance usage remains, as this is vital to the development of improved approaches to residential energy demand modelling (Anderson 2016). Thus, it would also be important to encourage more resolved data collection on the activities of all household members to

[1]The technical specifications for the smart meters rolled out nationally establish that the metering systems should allow for the capture of active power consumption data at 5 second intervals (DECC 2013*b*).

[2]Non-intrusive appliance load monitoring.

better understand the range of behaviours within the same household. Further, material arrangements have a strong shaping influence on energy demand (Shove & Walker 2014). Studies looking at enhancing the base of empirical data would also need to make a clearer distinction between the material arrangements found in the surveyed homes so that models of energy demand can estimate more accurately the demand load implications of the engagement in the different activities considered.

The third concern is the applicability of residential energy demand models to informing the development of residential demand management strategies. As has been pointed out elsewhere, most studies looking at estimating the scale and potential of demand-side management measures have so far been largely speculative (Grünewald & Diakonova 2018). This, however, could be attributed in great measure to the lack of tools suitable for a more in-depth analysis. For instance, relatively recent estimates of the potential of residential demand-side response have been produced using modelling approaches developed nearly 60 years ago (Gyamfi & Krumdieck 2012). But even some contemporary developments in residential energy demand modelling are unlikely to engender substantial progress in the quantification of the potential for enhanced system flexibility derived from demand-response mechanisms. A more rigorous assessment of the potential of interventions looking at modifying or influencing energy-consuming behaviours requires a thorough understanding of the way household activities manifest themselves in terms of demand for energy. Energy demand models that focus on identifying critical energy consumption behaviours might provide better opportunities for estimating the potential energy efficiency gains that can be feasibly realised. Additionally, rather than focusing on identifying seemingly critical features of aggregated activity patterns, residential energy demand models could focus on the identification of such features within

particular groups or corresponding to particular types of users.

Perhaps the most useful role for researchers looking at energy use in the residential sector might be to improve our understanding of energy uses associated with the range of typical domestic activities and ways in which these can be influenced. A profound understanding of these factors is key to a precise identification of the opportunities for interventions what would allow for the full potential of the residential sector to contribute to overall power system flexibility to be realised. The lack of understanding of the relationship between household activity and appliance usage patterns has been demonstrated for the case of the UK, and a mapping between the two has been provided (Chapter 3). Similar investigation for other countries represents an important area for further research. The approaches developed in this thesis for the analysis and simulation of socio-demographic, activity and appliance usage data demonstrate the feasibility of developing increasingly better tools for the analysis of residential demand for energy. However, model development supported by larger, more comprehensive datasets would result in models with sounder foundations and the ability to provide new insights. Thus, data collection studies looking more closely at the relationship between activity and energy usage in combination with the development of enhanced modelling approaches would prove a very worthwhile investment to improve our understanding of the implications of the introduction of new technologies and the implementation of energy efficiency measures in the residential sector.

Appendices

A | Supplementary information – Paper 1 [Chapter 2]

A.1 Empirical probability density estimation

In Section 2.6, the Kernel Density Estimation (KDE) method was used to estimate the empirical probability density function of the distribution of households with respect to annual electricity consumption. This method is preferred over the alternative density estimation method based on the k-nearest neighbours (k-NN) algorithm for three reasons.

Firstly, as discussed in Section 2.5.1.2, studies about the convergence of different density estimation methods have shown that the rate of convergence of the Kernel Density Estimation to the true density is the highest among the different estimation methods.

Secondly, the primary applications of the k-NN algorithm are: 1) Statistical classification, which is a method used in supervised learning as an alternative to (automatic) clustering approaches; and 2) non-parametric regression analysis.

Thirdly, when the k-NN algorithm is used as a density estimation method, it yields the best results on empirical distributions made up of observations relatively far from each other. Both estimation methods take neighbour observations into account to determine the estimated distribution function values for new points. However, while the KDE method takes into account all those neighbour observations within a fixed radius, the method based on the k-NN algorithm adjusts the lenght of the radius for each observation so that only the k nearest neighbours are taken into account. Thus, the contributions of the nearest neighbours may be over-represented compared to the other observation points that are still reasonably nearby, resulting in a more fluctuating (less smooth) density estimate for regions with high concentration of observations. That is, the use of the k-NN method results in overfitted density estimates for sets of empirical observations concentrated in relatively close proximity to one another.

Figure A.1 illustratest this point.

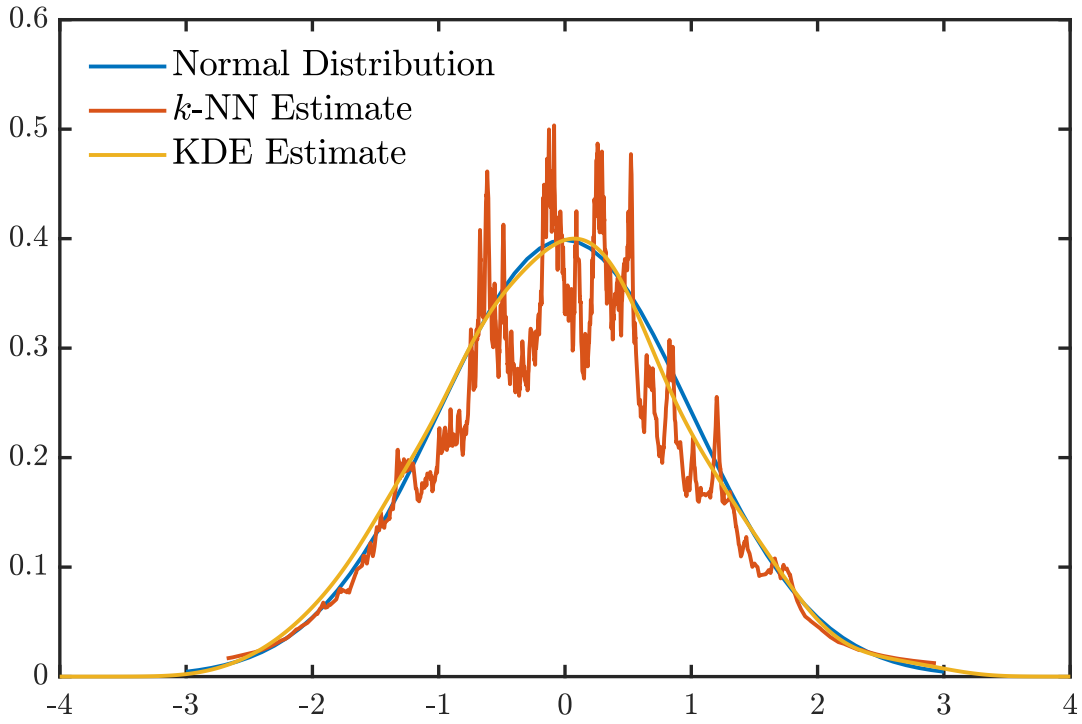


Figure A.1: *Comparison of Density Estimation methods. A sample of $N = 1000$ observations drawn from a Standard Normal Distribution is used to test the effectiveness of KDE and k-NN density estimation methods. In the KDE estimator, a Gaussian kernel is used with a bandwidth chosen to minimise the mean squared error. In the k-NN estimator, the conventional value $k = N^{1/2}$ is used.*

A.2 Trimodal Gaussian Mixture model

The statistical model presented in Section 2.6 provides a parametric estimate of the empirical probability density function associated with the distribution of households with respect to total annual consumption. In order to further assess the quality of the proposed Trimodal Gaussian Mixture model, alternative parametric estimates were tested.

The different types of models that would best fit the data were compared to

one another based on the Akaike Information Criterion (AIC). AIC is an estimator of the relative quality of statistical models for a given set of data. Thus, AIC provides an effective means for model selection; given a set of candidate models for the data, the preferred model is the one with the minimum AIC value.

AIC is given by the expression:

$$\mathrm{AIC} = 2k - 2\ln(\hat{\mathcal{L}})\,, \tag{A.1}$$

where k is the number of estimated parameters in the proposed model, and $\hat{\mathcal{L}}$ is the maximum value of the likelihood function for the model. Thus, AIC rewards the model's goodness-of-fit (as assessed by its likelihood function), but it also discourages overfitting by including a penalty term that is an increasing function of the number of estimated parameters. In most cases, increasing the number of parameters in the model improves the goodnes-of-fit. However, this also means that the chances of overfitting are higher. The penalty term takes this into account.

Table A.1 shows the alternative models fitted to the data, ranked by their AIC score. As the table shows, the "best" model corresponds to a Gamma distribution (parameters: $shape = 3.036$; $scale = 1.299 \times 10^3$). However, the proposed Gaussian Mixture model is the second best, and the difference between this model and the first, based on their AIC scores, is almost negilgible (percentage difference of 0.13%). This suggests that despite the difference in the number of parameters (9 for a Trimodal Gaussian Mixture vs 2 for a Gamma Distribution), the quality score of the proposed Gaussian Mixture model corresponds to that of a 'good fit' rather than an 'over-fit'.

It should be noted that the fact that this Trimodal Gaussian Mixture model offers a good quality estimate is entirely contingent on the data steming from the HES sample, on which the analysis presented in Chapter 2 was based. In order to test whether the trimodality observed in this data is indeed an inherent property of the distribution of households with respect to demand for electricity, a more comprehensive analysis supported by larger and even more representative data samples would have to be carried out. In this regard, smart meters offer an unprecedented opportunity as the would – at least in principle – provide us with detailed household-level electricity consumption data. Publicly available smart meter data could be used to test the generality of the results presented in Chapter 2. However, no such data was available at the time of the sudy presented in this thesis.

Table A.1: *Ranking of models based on their Akaike Information Criterion (AIC) score.*

Distribution Model	AIC score ($\times 10^3$)
Gamma	4.3950
Gaussian Mixture	4.4006
Log-normal	4.4040
Generalized Extreme Value	4.4062
Inverse Gaussian	4.4068
Log-logistic	4.4101
Weibull	4.4129
Rayleigh	4.4165
Nakagami	4.4167
Logistic	4.4422
Normal	4.4646
Exponential	4.5121
Extreme Value	4.6307

A.3 Limitations of Household Electricity Survey data

While the Household Electricity Survey is, to date, one of the best studies – perhaps even the most comprehensive one – on residential electricity demand in the UK, there are some obvious limitations to the data stemming from this study.

A study of this kind requires substantial amounts of monitoring equipment, as well as qualified personel to carry out the installation of the metering devices and posterior data collection. This translates into rather elevated costs for a study of this kind; and the longer the study runs for, the higher the associated costs. Therefore, budgetary limitations are reflected in terms of the extensiveness and representativeness of the study. The complications associated with gaining the potential participants' approval for taking part in the study for a given amount of time, and the upkeeping of the monitoring equipment throughout this period also impose some limitations.

The sampling methodology for the selection of the households that would take part in the study also imposed some key limitations. As discussed in Section 2.4, only owner-occupier households were monitored as part of this study. This includes people who own their home outright or who are buying it on a mortgage. This was mainly due to the complications associated with recruiting participants in different housing tenure categories (e.g. gaining permission from landlords). While owner-occupiers are the predominant housing tenure category in the UK ($\sim 65\%$) (English Housing Survey 2018), this is a clear limitation in terms of the portrait of national electricity consumption presented by the HES dataset.

The recruitment of participant households was outsourced to a private contractor, the market research company Ipsos MORI. Given the nature of the study, and the fact that there was no UK-based precedent to learn from, the research steering group decided that socio-demographic quotas would be set in order to guide the household sampling. However, to avoid making the recruitment impractical or unachievable, these quotas were set on life-stage of the heads of the household only. Life-stage groups were selected as follows:

- single pensioner household
- single non-pensioner household
- multiple pensioner household
- households with child(ren)
- multiple person households with no children living at home

Quotas were spread across the year to ensure the presence of the different groups throughout the year. That is, to prevent a subgroup from becoming the predominant one at a particular time of the year. However, given the size of the sample used for this study, this can only be controlled to a an extent.

A set of limitations more directly related to the budgetary limitations are concerned with the monitoring methodology. The total monitoring period only covers the extent of one year. However, only a few households were monitored for the entirety of this period. In order to reduce the costs associated with monitoring equipment, most households were monitored over the extent of one month, with different households being monitored at different points in the trial period so that a given set of monitoring devices could be used to monitor different households over the course of the trial. While in many cases there is considerable overlap

between the monitoring periods of the households that were monitored for a month only, this is another clear limitation as this restricts the effective sample size throughout the year of the study.

Despite the number of limitations this study might have been subject to, the results of the analysis of the data reveal that the sample would appear to be representative enough. The representativeness of the characteristics of the data from this study was assessed based on the following three aspects. The composition of the HES sample in terms of household sizes proportions are well in agreement with the proportions observed at the national level (ONS 2015). Moreover, the monthly consumption estimates derived from the HES data compare remarkably well with the estimates derived from the reported national residential consumption figures (BEIS 2019). That is, the seasonality effects are adequately represented[1]. Finally, and more importantly, as the results presented in Chapter 5 show, the representations of residentla electricity consumption based on elements extracted from this data compare well to the national load profile.

[1]This is discussed in more detail in Appendix D.

B | Supplementary information – Paper 2 [Chapter 3]

B.1 Misrepresentation of appliance usage

The activity engagement and appliance usage daily profiles presented in Chapter 3 are an attempt to provide a quantitative measure of the misrepresentation of the usage of appliances associated with the activities analysed in this chapter. However, it should be noted that this is but a step in the direction towards the characterisation of the use of appliances in relation to the engagement in activities.

Developing a robust statistical way of measuring the under- or over-representation of the use of appliances associated with particular activities would prove really useful. And it would have some weighty implications for the development of models and production of estimates of residential demand for electricity, as taking such measurements into account would be essential. This would, in turn, offer substantial benefits in terms of the possibilities for assessing the potential flexibility – or the lack of it – of the demand loads associated with residential activity. However, developing such a technique would require data

collection studies that would produce more adequate data for the purposes of such an endeavour.

In the following section, the limitations associated with the data collection methodologies of the studies discussed in Chapter 3 are further discussed.

B.2 Activity and appliance usage data collection

The kind of limitations discussed in the previous section are a direct result of the methodological limitations associated with the collection of the activity and appliance usage data used for the analysis presented in Chapter 3.

As discussed in Section 3.3, these data stem from two different studies: the Time-Use Survey and the Household Electricity Survey. And as the name of these studies would suggest, the focus of their efforts was different.

Time-use surveys are studies concerned with how people spend their time, and their primary aim is to quantify the total amount of time spent on certain activities. Their data collection method is based on the completion of daily time diaries that divide the day into intervals with fixed durations ranging from 10 to 30 minutes. This data collection method supposes two key issues for the purposes of the study of Chapter 3. Firstly, there is no direct indication of the usage of any particular energy-consuming device as a result of the engagement in the reported activities. More importantly, however, is the fact that time-use surveys carried out up to this point have relied entirely on the survey respondents' commitment

to fill out their daily activity diaries for the days they are taking part in the study. The latter point is a clear source of methodological bias, as it is very unlikely that every single survey respondent would report their daily activities with absolute fidelity.

Studies such as the Household Electricity Survey are rare, mainly due to the technical complications and high costs associated with conducting a study of such magnitude. The primary aim of this study was to produce a detailed characterisation of the residential electricity consumption patterns at the appliance level. And although it might be possible to infer the kind of activities the members of the household are engaging in based on the monitored appliance usage, there is no direct indication of actual engagement in any activities. The data collection method is based on electronic logging systems which can record the electricity consumption of the monitored appliances at the desired time resolution. These records, thus, picture the temporalities of appliance usage with relatively high fidelity.

The differences in the accuracy of the pictures presented by these two data collection methods result certain effects such as those discussed in Section 3.4.1, for instance, the differences in the occurrence times of the peaks in reported activity engagement and metered appliance usage.

Studies specifically aiming at monitoring simultaneously both engagement in activity and appliance usage may offer a more suitable data collection methodology. Recognising this, a few studies have attempted to capture such connections through targeted monitoring of certain activities and the appliances associated with them. An example of such studies is the one carried out by Durand-Daubin (2013), in which TV watching, computer usage and clothes

washing activities were monitored in detail along with the usage patterns of the relevant appliances in a sample of 60 households in Paris.

In this study, time-use and appliance-usage diaries were associated on an hour-by-hour basis in order to quantify the concurrence of activity engagement and appliance usage events. Effects such as the 'background appliance usage' reported in Section 3.4.1, (Figure 3.6) for TV watching and computer usage were also observed in this study. Durand-Daubin (2013) reports that the rate of concurrence of reported and actual appliance usage ranges from 70 to 80%. In addition, he also reports that the total durations reported are consistently lower than the durations inferred from the metered consumption.

B.3 Further limitations of the comparative analysis

Section 3.4.4 discusses the limitations of the analysis presented in Chapter 3 in terms of the generality of the results and their relevance to other countries. There are, however, a number of other limitations associated with the datasets used for the analysis themselves which are worth noting.

As noted in the previous section, the Time-use Survey and the Household Electricity Survey were two independent studies. That translates into differences in the sampling methodologies used to select the participant households, the time of the actual data collection, and the treatment of the data stemming from these studies.

While both studies were designed to provide a statistically representative picture of both time-use and residential electricity consumption in the UK, the nature of the data collection methods is fundamentally different; monitoring electricity consumption requires metering devices, whereas time-use surveys rely on self-reported engagement in activities. Since the collection of data for the Household Electricity Survey depends on the monitoring of individual appliances, the scope of the survey sample is limited by the availability of monitoring devices. In addition to these technical constraints, there are other practical constrains associated with the recruitment of participants. The organisations responsible for carrying out this study report that a decision was made to restrict the recruitment of participants to households that fell into the owner-occupier tenure category, so as to avoid the complications associated with requesting approval from landlords and letting agencies (Zimmermann et al. 2012). A more comprehensive discussion of the limitations of the Household Electricity Survey data is provided in Section A.3.

The biggest constraint of time-use studies is perhaps the willingness of potential participants to actively take part in the data collection. Contrary to the data collection of studies such as the household electricity survey, where the monitoring devices work in the background with virtually no interference with the everyday lives of households, the data collection of time-use surveys is based on time diary instruments that have to be filled out manually by the survey respondents. In addition to the methodological biases associated with this data collection method discussed in the previous section, this also poses restrictions in terms of the amount of days that survey respondents can be realistically expected to carry out their data collection duties. A common practice consists in asking survey respondents to record their activities during one work day and one weekend day, with the aim to capture the differences between the two types

of day. Studies that have attempted to collect data for longer periods of time exist, but in those cases, the constraints associated with participants' willingness to help are reflected in terms of smaller samples (e.g. (Durand-Daubin 2013)).

The differences in sample size between the Household Electricity Survey (HES) and the Time-Use Survey (TUS) would appear to indicate that the samples recruited for time-use studies are more representative because they are considerably larger ($\sim$ 4700 households for TUS vs 250 for HES). However, in terms of daily time diaries (TUS) and daily appliance usage monitoring records (HES), the two surveys compare really well ($\sim$ 16,000 time diaries (TUS) vs $\sim$ 15,000 household-day monitoring records (HES)).

Despite the fact that the primary aims of these two studies are not the same, the statistical treatment of the collected data shares some similarities. Both studies are ultimately concerned with estimating certain magnitudes; in the case of the Time-use Survey, this is the time spent on certain activities, in the case of the Household Electricity Survey, this is the amount of energy consumed by certain appliances. Moreover, both studies are concerned with the analysis of time series; while the resolution of these time series might not be the same, daily time diaries and appliance usage records allow for the analysis of the temporalities of everyday life in terms of their corresponding relevant magnitude.

A final issue worthy of discussion is the time gap that separates the data collection associated with these two studies. The underlying assumption for the analysis presented in Chapter 3 is that this time gap is small enough so as to preserve the overall temporal patterns observed. That is, were these two studies carried out simultaneously during the time of the former study (TUS) and then both repeated during the time of the latter study (HES), the differences observed

in the daily temporal patterns studied in Chapter 3 would not be statistically significant. Evidently, this assumption cannot be verified, as the hypothetical situation exposed here is unlikely to occur. Therefore, the generality of the conclusions drawn from the analysis presented in Chapter 3 is inherently limited. Opportunities to test this assumption might arise in the near future, where the complications and costs associated with these kind of studies are mitigated by technological developments. Until then, this will remain the subject of future work.

B.4 Individual appliance usage patterns and theories of social practice

The findings presented in Chapter 3 feed into a broader discussion on the dynamics of energy-related domestic social practices. Social practice theory places the ordering of people's activities at the centre of the study of social life in an attempt to offer a consistent ground to investigate the dynamic relations between people, time and energy demand. These intricate links suggest that understanding the timing of energy demand entails studying social practices in terms of their ordering and time dependence (Torriti 2017).

Understanding the timing of energy demand depends on understanding the ranges of practices and material arrangements and how they are temporally and spatially distributed. The term material arrangements refers to the physical objects and infrastructures which set the frameworks within which practices take place (Schatzki 2010, Torriti 2017). Such arrangements are only meaningful in

relation to the practices they enable, and in which they are involved (Shove & Walker 2014). Domestic appliances fall within this category.

The analysis presented in Chapter 3 focuses on a set of practices which can be unambiguously associated with the use of certain appliances. And this choice of focus poses some obvious limitations over the scope of this work.

However, in the broader context of the internal debates of theories of social practice, the limitations of this work are associated with the approach to the analysis taken. Social practice theory stipulates that practices are not performed in isolation; they tend to cluster together over time and space and have different levels of time dependence. By treating them – and the material entities associated with them (i.e. appliances) – in isolation, some of the dynamics of everyday life are inevitably lost.

This type of analysis has, nevertheless, helped in further untangling the links between the time dependence of practices and the time dependence of energy demand itself. In doing so, it has offered the opportunity for a more in-depth analysis of the heterogeneity of practices that give rise to peak demands.

C | Supplementary information – Paper 3 [Chapter 4]

C.1 Transition Occurrence Frequency Error profiles

The Transition Occurrence Frequency (TOF) Error metric was proposed as a metric to assess the performance of the proposed activity model specifically in terms of the simulation of activity scheduling, which requires looking more closely at the behaviour in-between the states that make up the daily activity sequences.

The error is calculated with respect to the set of empirical activity sequences which provide us with the reference values for the transition occurrence frequencies. For each transition interval the daily activity sequences are divided into, the transition occurrence frequency matrices extracted from the simulated output of both first-order and variable memory length models are compared to those reference values. We thus end up with TOF Error profiles such as those in Figure 4.8 in Chapter 4. To aid in the further discussion of these TOF Error profiles, this figure has been reproduced below in Figure C.1.

The transition intervals of the simulated daily activity sequences follow the same order as the intervals in the empirical activity sequences, meaning that the initial interval corresponds to the activity observed at 4:00 am on a given day.

As we can observe in Figure C.1, the TOF Error of the variable memory length model exhibits a counter-trend with respect to the TOF Error of the first-order model during the period immediately after 4:00 am. This initial counter-trend could be likely attributed to the improved ability of the variable memory length model to filter out non-significant variability in the transitions observed throughout the day. During these 'early hours', however, this would appear to be particularly notorious, which could be related to the inherent stability of the activity state predominantly observed during this period of the day: sleep. The variable memory length model is, by design, able to take the duration of the periods of sleep – or engagement in any other activity state – into account. In practice, this means that periods of relatively constant activity will be characterised by lower TOF Error, as the lower likelihood of observing transitions into different states results in less variability.

The variable memory length model's ability to filter out non-significant variability is further reflected by the fact that the variability of its TOF Error profile itself is generally much lower than the variability of the first-order model's TOF Error profile throughout the day.

Towards the end of the day (transition intervals 120+), both profiles would appear to follow a similar upward trend. In the case of the first-order model's TOF Error profile, this trend would appear to be exponential, and accompanied by an increase in the variability of the profile with respect to the variability observed during 'daytime'. In the case of the variable memory length model's profile,

however, this increase in the TOF Error would appear to be contained, following a seemingly asymptotic trend rather than an exponential one. Regardless of the characteristics of the trend, a considerable increase is observed, compared to any previously observed values. This may well be a data artifact resulting from cutting-off the activity diaries towards the end of the 24 hour periods the reported empirical activity sequences cover. However, in order to verify this, the models would have to be tested against sequences of activity that cover the overnight periods uninterrupted. Until such data becomes available, we might not be able to corroborate such possibilities.

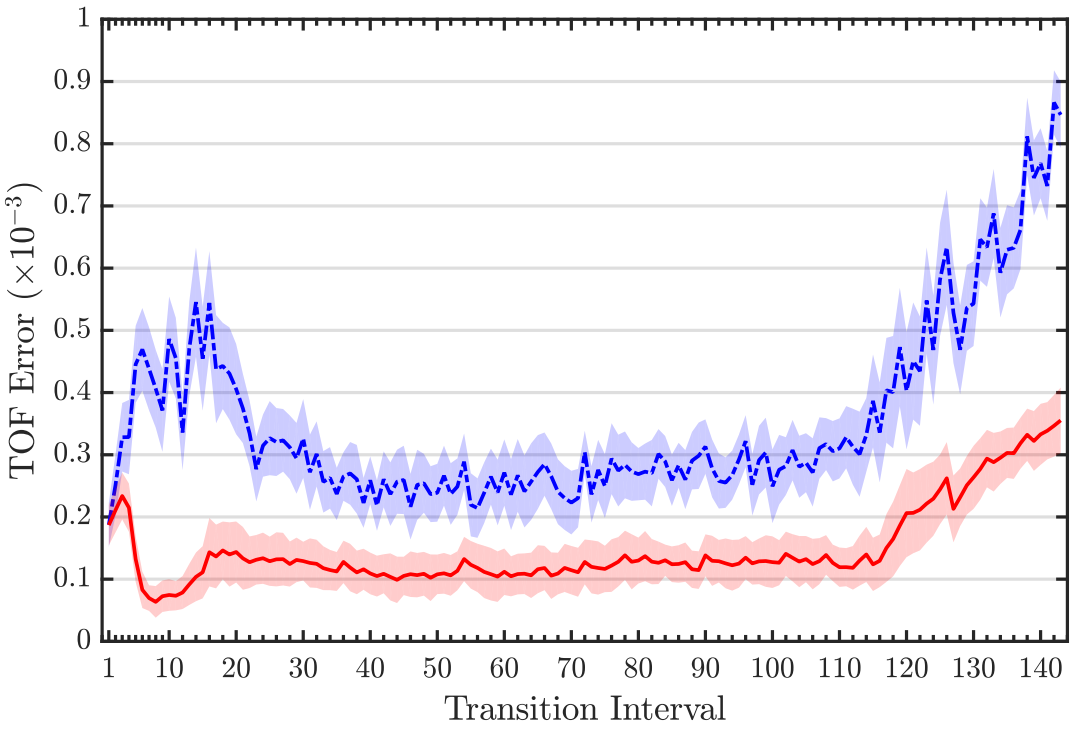


Figure C.1: *Transition Occurrence Frequency (TOF) Error profiles (Figure 4.8 in Chapter 4). Blue line - first-order model error; Red line - Variable memory length error; the shaded regions along the profiles represent the standard deviation ranges.*

C.2 Activity sequences and social science

The study of time use and the sequences of activities that characterise people's everyday lives has not been restricted to theories of social practice.

In the context of the study of residential activity and energy consumption patterns, the case of time geography is particularly noteworthy as it provides an ontological framework in which the abstract sequences that correspond to the evolution of daily human activity are projected on a space-time as a way of providing a visual language to describe the dynamics of the evolutive process.

Ellegård (1999) argues that taking a time-geographical approach to the study of people's everyday lives places the contexts in which such activities make up coherent routines and projects at the core of their study. She goes on further and proposes a method based on the time-geographical framework specifically designed for empirical studies of individuals' everyday life activities. Contrary to the social practice approach to the study of time use and energy consumption discussed in Chapter 4, the time-geographical approach focuses on individuals, and studies how people engage in certain activities in relation to the needs and wants of the individuals themselves.

Time geography establishes that, on a macro level, human societies are shaped by the continuous processes of individuals engaging in activities while living their everyday life (Ellegård 2018). Moreover, it states that engaging in activities results, be it directly or indirectly, on demand for certain resources; these could be natural or social. Time and spaces are also considered as a resource in the time-geographical framework, and it is assumed that time is the only resource

that is evenly distributed among all individuals, in the sense that every single person gets – at least in principle – 24 hours at their disposal at the start of every day, and all of this time is used through engaging in a series of activities (Ellegård 2018).

The time-geographical approach highlights the importance of the material resources and how the material restrictions people are subject to frame their everyday life (Ellegård & Palm 2011). Applying this combination of theoretical perspectives puts energy use into a material time-space framework which is regarded as embedded into all social processes (Ellegård & Palm 2011). In this sense, the study of time use and energy consumption through time-geographical methods is similar to the methods used in social practice theory, where the understanding of energy consumption is built on a detailed knowledge of the routines and methods established in a particular social context, and the communication between members of a particular household themselves and external entities.

The key difference, however, is that in theories of social practice the main constraints are the social rhythms and societal norms imposed on every member of a particular society rather than the net availability of time and other material resources. As discussed in Section 4.2.4, these type of constraints would appear to play a more important role in shaping the daily activity schedules of households. And perhaps more importantly, the quantification of the magnitude of the effects these constrains may have on the shaping of residential activity schedules may bring us one step closer to the quantification of the flexibility potential of the electricity demand loads associated with the residential sector.

D | Supplementary information – Paper 4 [Chapter 5]

D.1 Household characterisation variables

The development of models of household electricity demand has been primarily based on physical and economic variables such as the type of dwelling and the income bands of simulated households (Haines et al. 2010). However, previous studies that have focused on the analysis of the diversity of household energy demand have found that such variables are not strong determinant factors of household electricity use. For example, Firth et al. (2008) carried out a longitudinal study of 72 dwellings studied over a two year period. One of the key findings of this study is that the variation in the measured annual electricity consumption can be up to six-fold. And according to their results, even dwellings on the same site and with similar built form had notably different annual electricity consumptions. The authors go on to suggest that the number of occupants, number and type of appliances and occupancy patterns are in this case, more influential than the physical characteristics of the dwellings.

Recent studies have focused on analysing how much of the variability of residential energy consumption can be explained by different types of predictors such as physical variables, socio-demographics, and behaviours and attitudes towards energy (Huebner et al. 2015, 2016). Based on data collected from a representative sample of the English housing stock, Huebner et al. were able to contrast the explanatory power of such variables. The socio-demographic factors considered were: Household size and type (e.g. family with/without children), number of dependent children (if any), equivalised income range, age range of adult members of the household, dwelling tenure and length of residency, and employment status of the household. And this analysis revealed that socio-demographic variables on their own account for 21% to 24% of the variability observed in residential electricity demand.

A closely related factor is the presence of the different types of appliances. Both the appliance ownership rates and the appliance usage intensities vary across different socio-demographic groups. In practice, this means that there is considerable variability in terms of the total energy demands associated with households in each socio-demographic group. There are a few studies that have attempted to incorporate the effects of appliance ownership, and there are even fewer that look in detail at the usage of these appliances (Haines et al. 2010). But this is not surprising given the complexity of such an endeavour.

Table D.1 offers a brief review of the type of variables and methods used in 10 recent models of residential electricity demand.

Table D.1: *Review of predictor variables and modelling methods employed in 10 models recently developed.*

Model	Modelling method	Behavioural data	Predictor variables
Capasso et al. (1994)	Probabilistic: density functions	Yes (TUS) [Occupancy]	Socio-demographic: • Number and type of household members • Number of households in the area Technical: • Average appliance ownership rates • Average annual consumption of appliance types
Yao & Steemers (2005)	deterministic occupancy profiles	No	Technical: • Average appliance ownership rates • Average annual consumption of appliance types
Duffy et al. (2010)	first-order Markov chain	No	Technical: • Dwelling types (detached, semi-detached, terraced and apartment) • Load estimates
Richardson et al. (2010)	first-order Markov chain	Yes (TUS) [Occupancy]	Socio-demographic: • Household size Technical: • Average appliance ownership rates • Average annual consumption of appliance types
Widén & Wäckelgård (2010)	first-order Markov chain	Yes (TUS) [Activities]	Socio-demographic: • Household size Technical: • Dwelling type (detached house, apartment)
Wilke et al. (2013)	semi-Markov chain	Yes (TUS) [Activities]	Socio-demographic: • Sex • Tenure • Employment status
Santiago et al. (2014)	first-order Markov chain	Yes (TUS) [Occupancy]	Socio-demographic: • Household size
Fischer et al. (2015)	Probabilistic: density functions	Yes (TUS) [Activities]	Socio-demographic: • Household size and type • Age range • Employment status Technical: • Appliance power ratings
Shimoda et al. (2017)	Probabilistic: density functions	Yes (TUS) [Occupancy]	Socio-demographic: • Household size and type Technical: • Appliance power ratings

The household characterisation module presented in Chapter 5 is an attempt to map the inter-dependencies between the different socio-demographic variables themselves and their effects in terms of demand for electricity of particular households. The households that match a particular socio-demographic profile

are then equipped with a set of appliances that is consistent with the appropriate set of ownership statistics.

The socio-demographic statistics used to construct the module are based on data from a number of studies carried out by the UK Office for National Statistics. These include the reports on 'Household and household composition in England and Wales 2001-2011' (ONS 2014), the 2011 Census data tables (ONS 2011), and the online bulletin 'Overview of the UK population 2016' (ONS 2016).

The collection of ownership statistics for the different segments of the residential user base identified the following relevant sources:

- Energy follow-up survey 2011 (DECC 2013*a*),
- General Lifestyle Survey 2011 (ONS 2013),
- Living Costs and Food Survey 2016 (ONS 2017),
- Determinants of high electrical energy demand in UK homes: Appliance ownership and use (Jones & Lomas 2016),
- Understanding usage patterns of electric kettle and energy saving potential (Murray et al. 2016),
- How Trends in Appliances Affect Domestic CO2 Emissions: A Review of Home and Garden Appliances (Haines et al. 2010), with data from
- Measurement, modelling, mapping and management: a methodology for shrinking the urban carbon footprint (EPSRC 2010).

D.2 Monthly consumption estimates

The UK government's Energy Trends bulletin contains detailed information regarding the energy consumption of major economic sectors, including the

domestic (BEIS 2019). Among other things, the tables devoted to electricity consumption contain the monthly breakdowns for each sector extending from 1995 to date. Based on the figures reported in Table ET_5.5 of the most recent bulletin and the estimates for the number of households for the period between the years 2009 and 2012, it was possible to calculate the national average monthly consumption per household for every month between June 2009 and July 2012. Figure D.1 clearly shows that the maximum monthly consumption for the period June 2010 – July 2011 — the period during which the Household Electricity Survey took place — was observed in December, and this was considerably higher than the corresponding values for the same month in both the previous and subsequent years.

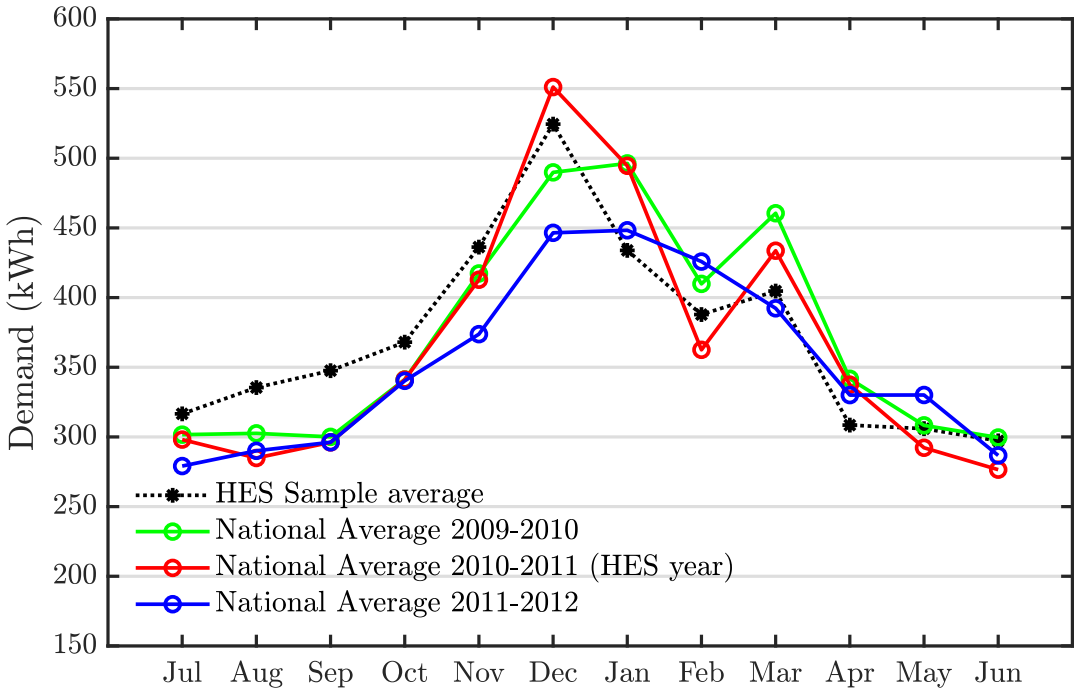


Figure D.1: *Comparison of the seasonal trend of the average total monthly demand for electricity between HES sample households and the national figures for the years 2009-2012.*

References

Aerts, D., Minnen, J., Glorieux, I., Wouters, I. & Descamps, F. (2014), 'A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison', *Building and Environment* **75**, 67–78.
URL: *https://doi.org/10.1016/j.buildenv.2014.01.021*

Aghaei, J. & Alizadeh, M.-I. (2013), 'Demand response in smart electricity grids equipped with renewable energy sources: A review', *Renewable and Sustainable Energy Reviews* **18**, 64–72.
URL: *https://doi.org/10.1016/j.rser.2012.09.019*

Alderson, H., Cranston, G. R. & Hammond, G. P. (2012), 'Carbon and environmental footprinting of low carbon UK electricity futures to 2050', *Energy* **48**(1), 96–107.
URL: *https://doi.org/10.1016/j.energy.2012.04.011*

Allen, S., Hammond, G. P., Harajli, H., Jones, C. I., Manus, M. C. & Winnett, A. B. (2008), 'Integrated appraisal of micro-generators: methods and applications', *Proceedings of Institution of Civil Engineers: Energy* **161**(2), 73–86.

Allen, S., Hammond, G. P. & McManus, M. (2008), 'Prospects for and barriers to domestic micro-generation: A United Kingdom perspective', *Applied Energy* **85**(6), 528–544.
URL: *https://doi.org/10.1016/j.apenergy.2007.09.006*

Altmann, M., Brenninkmeijer, A., Lanoix, J.-C., D., E., Crisan, A., Hugyecz, A., Koreneff, G. & Hanninen, S. (2010), Decentralised Energy Systems, Technical report, European Parliament's Committee on Industry, Research

and Energy (ITRE), Brussels, Belgium.
URL:
http://www.europarl.europa.eu/activities/committees/studies.do?language=EN

Andersen, P. D., Iversen, A., Madsen, H. & Rode, C. (2014), 'Dynamic modeling of presence of occupants using inhomogeneous Markov chains', *Energy and Buildings* **69**, 213–223.
URL: *https://doi.org/10.1016/j.enbuild.2013.10.001*

Anderson, B. (2016), 'Laundry, energy and time: Insights from 20 years of time-use diary data in the United Kingdom', *Energy Research & Social Science* **22**, 125–136.
URL: *https://doi.org/10.1016/j.erss.2016.09.004*

Anderson, B. & Torriti, J. (2018), 'Explaining shifts in UK electricity demand using time use data from 1974 to 2014', *Energy Policy* **123**, 544–557.
URL: *https://doi.org/10.1016/j.enpol.2018.09.025*

Anderson, D. F. (2007), 'A modified next reaction method for simulating chemical systems with time dependent propensities and delays.', *The Journal of chemical physics* **127**(21), 214107.
URL: *http://www.ncbi.nlm.nih.gov/pubmed/18067349*

Armstrong, M. M., Swinton, M. C., Ribberink, H., Beausoleil-Morrison, I. & Millette, J. (2009), 'Synthetically derived profiles for representing occupant-driven electric loads in Canadian housing', *Journal of Building Performance Simulation* **2**(1), 15–30.
URL: *https://doi.org/10.1080/19401490802706653*

Baetens, R., De Coninck, R., Helsen, L. & Saelens, D. (2010), 'The Impact of Load Profile on the Grid-Interaction of Building Integrated Photovoltaic (BIPV) Systems in Low-Energy Dwellings', *Journal of Green Building* **5**(4), 137–147.
URL: *http://doi.org/10.3992/jgb.5.4.137*

Bajada, J., Fox, M. & Long, D. (2013), Load modelling and simulation of household electricity consumption for the evaluation of demand-side management strategies, *in* 'IEEE PES ISGT Europe 2013', IEEE, pp. 1–5.
URL: *https://doi.org/10.1109/ISGTEurope.2013.6695250*

Ballanti, A. & Ochoa, L. F. (2015), Initial assessment of voltage-led demand response from UK residential loads, *in* '2015 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)', IEEE, pp. 1–5.
URL: *https://doi.org/10.1109/ISGT.2015.7131821*

Ballanti, A., Pilo, F., Navarro-Espinosa, A. & Ochoa, L. F. (2013), Assessing the benefits of PV var absorption on the hosting capacity of LV feeders, *in* 'IEEE PES ISGT Europe 2013', IEEE, pp. 1–5.
URL: *https://doi.org/10.1109/ISGTEurope.2013.6695423*

Barton, J., Huang, S., Infield, D., Leach, M., Ogunkunle, D., Torriti, J. & Thomson, M. (2013), 'The evolution of electricity demand and the role for demand side participation, in buildings and transport', *Energy Policy* **52**, 85–102.
URL: *https://doi.org/10.1016/j.enpol.2012.08.040*

Baruah, P., Eyre, N., Qadrdan, M., Chaudry, M., Blainey, S., Hall, J. W., Jenkins, N. & Tran, M. (2014), 'Energy system impacts from heat and transport electrification', *Proceedings of the ICE - Energy* **167**(3), 139–151.
URL:
http://www.icevirtuallibrary.com/content/article/10.1680/ener.14.00008

Bedir, M., Hasselaar, E. & Itard, L. (2013), 'Determinants of electricity consumption in Dutch dwellings', *Energy and Buildings* **58**, 194–207.
URL: *https://doi.org/10.1016/j.enbuild.2012.10.016*

BEIS (2016), Chapter 3: Domestic energy consumption, *in* 'Energy Consumption in the UK', Department of Business, Energy and Industrial Strategy, London, United Kingdom, chapter 3, p. 17.
URL:
https://www.gov.uk/government/statistics/energy-consumption-in-the-uk

BEIS (2018*a*), 2016 UK Greenhouse Gas Emissions, Final Figures, Technical report, UK National Statistics, London, United Kingdom.
URL: *https://www.gov.uk/government/statistics/final-uk-greenhouse-gas-emissions-national-statistics-1990-2016*

BEIS (2018*b*), Digest of United Kingdom Energy Statistics 2018, Technical report.

URL: *https://www.gov.uk/government/statistics/digest-of-uk-energy-statistics-dukes-2018-main-report*

BEIS (2019).

Bejerano, G. & Yona, G. (2001), 'Variations on probabilistic suffix trees: statistical modeling and prediction of protein families', *Bioinformatics* **17**(1), 23–43.
URL: *https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/17.1.23*

Bergman, N. & Eyre, N. (2011), 'What role for microgeneration in a shift to a low carbon domestic energy sector in the UK?', *Energy Efficiency* **4**(3), 335–353.
URL: *https://doi.org/10.1007/s12053-011-9107-9*

Bevrani, H. & Raisch, J. (2017), 'On Virtual inertia Application in Power Grid Frequency Control', *Energy Procedia* **141**, 681–688.
URL: *https://doi.org/10.1016/j.egypro.2017.11.093*

Boardman, B., Darby, S., Killip, G., Hinnells, M., Jardine, C. N., Palmer, J., Sinden, G., Lane, K., Layberry, R., Newborough, M., Peacock, A. D., Wright, A. & Natarajan, S. (2005), 40% House, Technical report, Environmental Change Institute, Oxford, United Kingdom.
URL:
http://www.eci.ox.ac.uk/research/energy/downloads/40house/40house.pdf

Bondu, A. & Dachraoui, A. (2015), Realistic and very fast simulation of individual electricity consumptions, *in* '2015 International Joint Conference on Neural Networks (IJCNN)', IEEE, pp. 1–8.
URL: *https://doi.org/10.1109/ijcnn.2015.7280339*

Boßmann, T. & Staffell, I. (2015), 'The shape of future electricity demand: Exploring load curves in 2050s Germany and Britain', *Energy* **90**, 1317–1333.
URL: *https://doi.org/10.1016/j.energy.2015.06.082*

Bradley, P., Leach, M. & Torriti, J. (2013), 'A review of the costs and benefits of demand response for electricity in the UK', *Energy Policy* **52**, 312–327.
URL: *https://doi.org/10.1016/j.enpol.2012.09.039*

Bright, J., Smith, C., Taylor, P. & Crook, R. (2015), 'Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data', *Solar Energy* **115**, 229–242.
URL: *https://doi.org/10.1016/j.solener.2015.02.032*

Brown, P. F., DeSouza, P. V., Mercer, R. L., Pietra, V. J. D. & Lai, J. C. (1992), 'Class-Based n-gram Models of Natural Language', *COMPUTATIONAL LINGUISTICS* **18**, 467—-479.
URL: *http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.9919*

Bühlmann, P. (2000), 'Model Selection for Variable Length Markov Chains and Tuning the Context Algorithm', *Annals of the Institute of Statistical Mathematics* **52**(2), 287–315.
URL: *http://doi.org/10.1023/A:1004165822461*

Bühlmann, P. & Wyner, A. J. (1999), 'Variable length Markov chains', *The Annals of Statistics* **27**(2), 480–513.
URL: *https://doi.org/10.1214/aos/1018031204*

Campillo, J., Wallin, F., Torstensson, D. & Vassileva, I. (2012), Energy demand model design for forecasting electricity consumption and simulating demand response scenarios in Sweden, *in* 'The Fourth International Conference on Applied Energy 2012', Suzhou, China.

Cao, S. & Sirén, K. (2016), 'The influence of simulation time-resolution on the matching between on-site micro-wind generation and building electric demand', *Journal of Building Performance Simulation* **9**(5), 449–468.
URL: *https://doi.org/10.1080/19401493.2015.1077270*

Capasso, A., Grattieri, W., Lamedica, R. & Prudenzi, A. (1994), 'A bottom-up approach to residential load modeling', *IEEE Transactions on Power Systems* **9**(2), 957–964.
URL:
http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=317650

CCC (2008), *Building a low-carbon economy – the UK's contribution to tackling climate change*, The Stationary Office, London, United Kingdom.
URL: *http://tinyurl.com/UKCCC2008*

Charniak & Eugene (2000), 'A maximum-entropy-inspired parser'.
URL: *https://dl.acm.org/citation.cfm?id=974323*

Chmutina, K. & Goodier, C. I. (2014), 'Alternative future energy pathways: Assessment of the potential of innovative decentralised energy systems in the UK', *Energy Policy* **66**, 62–72.
URL: *http://doi.org/10.1016/j.enpol.2013.10.080*

Coaffee, J. (2008), 'Risk, resilience, and environmentally sustainable cities', *Energy Policy* **36**(12), 4633–4638.
URL: *https://doi.org/10.1016/j.enpol.2008.09.048*

Darby, S. J. & McKenna, E. (2012), 'Social implications of residential demand response in cool temperate climates', *Energy Policy* **49**, 759–769.
URL: *https://doi.org/10.1016/j.enpol.2012.07.026*

DECC (2009), *The UK Low Carbon Transition Plan: National strategy for climate and energy*, number July 2009, DECC, London, United Kingdom.
URL: *http://tinyurl.com/UKDECC2009*

DECC (2011), The Carbon Plan: Delivering our low carbon future, Technical Report December, Department of Energy and Climate Change, London, United Kingdom.
URL: *http://www.decc.gov.uk/assets/decc/11/tackling-climate-change/carbon-plan/3702-the-carbon-plan-delivering-our-low-carbon-future.pdf*

DECC (2013*a*), Energy Follow-Up Survey (EFUS): 2011, Technical report, Department of Energy and Climate Change, London, United Kingdom.
URL:
https://www.gov.uk/government/statistics/energy-follow-up-survey-efus-2011

DECC (2013*b*), Smart metering equipment technical specifications: second version, Technical report, DECC, London, United Kingdom.
URL: *https://www.gov.uk/government/consultations/smart-metering-equipment-technical-specifications-second-version*

DECC (2014), Community Energy Strategy: Full Report, Technical report, London, United Kingdom.
URL:
https://www.gov.uk/government/publications/community-energy-strategy

DECC (2015), Chapter 3: Domestic energy consumption, *in* 'Energy consumption in the UK (2015)', Department of Energy and Climate Change, London, United Kingdom, chapter 3, p. 17.
URL:
https://www.gov.uk/government/statistics/energy-consumption-in-the-uk

DECC & Intertek Testing and Certification Ltd. (2016), *Household Electricity Survey, 2010-2011.* [data collection]. UK Data Service. SN: 7874.
URL: *https://doi.org/10.5255/UKDA-SN-7874-1*

Delzendeh, E., Wu, S., Lee, A. & Zhou, Y. (2017), 'The impact of occupants' behaviours on building energy analysis: A research review', *Renewable and Sustainable Energy Reviews* **80**(August), 1061–1071.
URL: *https://doi.org/10.1016/j.rser.2017.05.264*

Duffy, A., McLoughlin, F. & Conlon, M. (2010), The Generation of Domestic Electricity Load Profiles through Markov Chain Modelling, *in* '3rd International Scientific Conference on Energy and Climate Change', Athens, Greece, pp. 18 – 27.
URL: *http://arrow.dit.ie/dubencon2/9/*

Durand-Daubin, M. (2013), Household activities through various lenses: crossing surveys, diaries and electric consumption, *in* 'Behavior, Energy and Climate Change Conference (BECC) 2013', Berkeley, CA, US.
URL: *http://escholarship.org/uc/item/83z390h2#page-1*

EA Technology (2012), Assessing the impact of low carbon technologies on Great Britain's power distribution networks, Technical report, Ofgem.
URL: *https://www.ofgem.gov.uk/publications-and-updates/assessing-impact-low-carbon-technologies-great-britains-power-distribution-networks*

Ela, E., Kirby, B., Lannoye, E., Milligan, M., Flynn, D., Zavadil, B. & O'Malley, M. (2010), Evolution of Operating Reserve Determination in Wind Power Integration Studies, *in* '2010 IEEE Power & Energy Society General Meeting', National Renewable Energy Laboratory (U.S.), Minneapolis, Minnesota, USA.
URL: *https://digital.library.unt.edu/ark:/67531/metadc833051/*

Ellegård, K. (1999), 'A time-geographical approach to the study of everyday life of individuals – a challenge of complexity', *GeoJournal* **48**(3), 167–175.
URL: *https://doi.org/10.1023/A:1007071407502*

Ellegård, K. (2018), *Thinking Time Geography*, Routledge.
URL: *https://doi.org/10.4324/9780203701386*

Ellegård, K. & Palm, J. (2011), 'Visualizing energy consumption activities as a tool for making everyday life more sustainable', *Applied Energy* **88**(5), 1920–1926.
URL: *https://doi.org/10.1016/j.apenergy.2010.11.019*

Elombo, A. I., Morstyn, T., Apostolopoulou, D. & McCulloch, M. D. (2017), Residential load variability and diversity at different sampling time and aggregation scales, *in* '2017 IEEE AFRICON: Science, Technology and Innovation for Africa, AFRICON 2017', pp. 1331–1336.
URL: *https://doi.org/10.1109/AFRCON.2017.8095675*

Energy Saving Trust (2011), The elephant in the living room: How our appliances and gadgets are trampling the green dream, Technical report, Energy Saving Trust, London, United Kingdom.
URL:
https://tools.energysavingtrust.org.uk/Publications2/Corporate/Research-and-insights/The-elephant-in-the-living-room

Energy Saving Trust, DECC, Defra & Owen, P. (2012), Powering the nation: Household electricity-using habits revealed, Technical report, Energy Saving Trust, Department of Energy and Climate Change, Department for Environment Food and Rural Affairs, London, United Kingdom.
URL:
http://www.energysavingtrust.org.uk/Publications2/Corporate/Research-and-insights/Powering-the-nation-household-electricity-using-habits-revealed

Energy Saving Trust, Econnect & Element Energy (2005), Potential for Microgeneration Study and Analysis - Final Report, Technical Report November, Energy Saving Trust.
URL: *http://www.berr.gov.uk/files/file27558.pdf*

Engle, R. F., Mustafa, C. & Rice, J. (1992), 'Modelling Peak Electricity Demand', *Journal of Forecasting* **11**(3), 241–251.

English Housing Survey (2018), English Housing Survey 2017-18 - Headline Report, Technical report, Ministry of Housing, Communities & Local

Government.
URL: *https://www.gov.uk/government/statistics/english-housing-survey-2017-to-2018-headline-report*

EPSRC (2010), 'Measurement, modelling, mapping and management: a methodology for shrinking the urban carbon footprint'.
URL: *https://gtr.ukri.org/projects?ref=EP%2FF007604%2F1*

European Commision (2012), *Energy Roadmap 2050*.
URL: *https://doi.org/10.2833/10759*

European Environment Agency (2017), Final energy consumption by sector and fuel, Technical report, European Environment Agency, Copenhagen, Denmark.
URL: *https://www.eea.europa.eu/data-and-maps/indicators/final-energy-consumption-by-sector-9/assessment-1*

European Parliament (2009), 'Directive 2009/72/EC of the Council of 13 July 2009'.
URL: *https://eur-lex.europa.eu/eli/dir/2009/72/oj*

European Parliament & Council of the European Union (2009), 'Internal Market in Electricity Directive - Directive 2009/72/EC of the European Parliament and of the Council'.
URL: *https://eur-lex.europa.eu/eli/dir/2009/72/oj*

EUROSTAT (2009), *Harmonised European time use surveys - 2008 guidelines*, EUROSTAT, Luxembourg, Luxembourg.
URL:
http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-08-014-EN.pdf

Eyre, N., Darby, S. J., Grünewald, P., McKenna, E. & Ford, R. (2018), 'Reaching a 1.5°C target: Socio-technical challenges for a rapid transition to low-carbon electricity systems', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **376**(2119).
URL: *https://doi.org/10.1098/rsta.2016.0462*

Fan, S. & Hyndman, R. J. (2011), 'The price elasticity of electricity demand in South Australia', *Energy Policy* **39**(6), 3709–3719.
URL: *https://doi.org/10.1016/j.enpol.2011.03.080*

Filik, Ü. B., Gerek, Ö. N. & Kurban, M. (2011), 'A novel modeling approach for hourly forecasting of long-term electric energy demand', *Energy Conversion and Management* **52**(1), 199–211.
URL: *https://doi.org/10.1016/j.enconman.2010.06.059*

Finn, P. & Fitzpatrick, C. (2014), 'Demand side management of industrial electricity consumption: Promoting the use of renewable energy through real-time pricing', *Applied Energy* **113**, 11–21.
URL: *https://doi.org/10.1016/j.apenergy.2013.07.003*

Firth, S., Lomas, K., Wright, A. & Wall, R. (2008), 'Identifying trends in the use of domestic appliances from household electricity consumption measurements', *Energy and Buildings* **40**(5), 926–936.
URL: *http://www.sciencedirect.com/science/article/pii/S0378778807002022*

Fischer, D., Härtl, A. & Wille-Haussmann, B. (2015), 'Model for electric load profiles with high time resolution for German households', *Energy and Buildings* **92**, 170–179.
URL: *https://doi.org/10.1016/j.enbuild.2015.01.058*

Flett, G. & Kelly, N. (2015), Household-differentiated demand modelling for communities, *in* '14th International Conference of the International Building Performance Simulation Association', BS2015, Hyderabad, India, pp. 1821 – 1828.
URL: *http://tinyurl.com/Flett-Kelly2015*

Flett, G. & Kelly, N. (2017), 'A disaggregated, probabilistic, high resolution method for assessment of domestic occupancy and electrical demand', *Energy and Buildings* **140**, 171–187.
URL: *http://doi.org/10.1016/j.enbuild.2017.01.069*

Fouquet, R. (2011), 'Divergences in Long-Run Trends in the Prices of Energy and Energy Services', *Review of Environmental Economics and Policy* **5**(2), 196–218.
URL: *https://doi.org/10.1093/reep/rer008*

Fujimoto, T., Yamaguchi, Y. & Shimoda, Y. (2017), 'Energy management for voltage control in a net-zero energy house community considering appliance operation constraints and variety of households', *Energy and Buildings*

147, 188–199.
URL: *https://doi.org/10.1016/j.enbuild.2017.05.009*

Gates, J. E. (1942), 'Forecasting the Demand for Electric Energy', *The Journal of Land and Public Utility Economics* **18**(1), 77–81.
URL: *http://www.jstor.org/stable/3158780*

Gershuny, J. & Sullivan, O. (2017), *United Kingdom Time Use Survey, 2014-2015* [data collection]. Centre for Time Use Research, University of Oxford. UK Data Service. SN: 8128,.
URL: *https://doi.org/10.5255/UKDA-SN-8128-1*

Good, N., Zhang, L., Navarro-Espinosa, A. & Mancarella, P. (2015), 'High resolution modelling of multi-energy domestic demand profiles', *Applied Energy* **137**, 193–210.
URL: *https://doi.org/10.1016/j.apenergy.2014.10.028*

Gottwalt, S., Ketter, W., Block, C., Collins, J. & Weinhardt, C. (2011), 'Demand side management—A simulation of household behavior under variable prices', *Energy Policy* **39**(12), 8163–8174.
URL: *https://doi.org/10.1016/j.enpol.2011.10.016*

Grahn, P., Munkhammar, J., Widén, J., Alvehag, K. & Soder, L. (2013), 'PHEV Home-Charging Model Based on Residential Activity Patterns', *IEEE Transactions on Power Systems* **28**(3), 2507–2515.
URL: *https://doi.org/10.1109/TPWRS.2012.2230193*

Grandjean, A., Adnot, J. & Binet, G. (2012), 'A review and an analysis of the residential electric load curve models', *Renewable and Sustainable Energy Reviews* **16**(9), 6539–6565.
URL: *https://doi.org/10.1016/j.rser.2012.08.013*

Grünewald, P. & Diakonova, M. (2018), 'Flexibility, dynamism and diversity in energy supply and demand: A critical review', *Energy Research & Social Science* **38**, 58–66.
URL: *https://doi.org/10.1016/j.erss.2018.01.014*

Grünewald, P. & Torriti, J. (2013), 'Demand response from the non-domestic sector: Early UK experiences and future opportunities', *Energy Policy*

61, 423–429.
URL: *https://doi.org/10.1016/j.enpol.2013.06.051*

Gyamfi, S. & Krumdieck, S. (2012), 'Scenario analysis of residential demand response at network peak periods', *Electric Power Systems Research* **93**, 32–38.
URL: *https://doi.org/10.1016/j.epsr.2012.07.004*

Haines, V., Lomas, K., Murray, T., Ian, R., Tracy, N. & Monica, G. (2010), How Trends in Appliances Affect Domestic CO2 Emissions : A Review of Home and Garden Appliances, Technical report, Department of Energy and Climate Change.
URL: *https://tinyurl.com/Haines2010*

Haldi, F. & Robinson, D. (2011), 'The impact of occupants' behaviour on building energy demand', *Journal of Building Performance Simulation* **4**(4), 323–338.
URL: *https://doi.org/10.1080/19401493.2011.558213*

Hammond, G. P., Howard, H. R. & Jones, C. I. (2013), 'The energy and environmental implications of UK more electric transition pathways: A whole systems perspective', *Energy Policy* **52**, 103–116.
URL: *http://doi.org/10.1016/j.enpol.2012.08.071*

Hammond, G. P. & Pearson, P. J. G. (2013), 'Challenges of the transition to a low carbon, more electric future: From here to 2050', *Energy Policy* **52**, 1–9.
URL: *http://doi.org/10.1016/j.enpol.2012.10.052*

Haney, A. B., Jamasb, T., Platchkov, L. M. & Pollit, M. G. (2011), Demand-side management strategies and the residential sector: lessons from the international experience, *in* 'The Future of Electricity Demand: Customers, Citizens and Loads', Cambridge University Press, chapter 14.

Hargreaves, T., Nye, M. & Burgess, J. (2010), 'Making energy visible: A qualitative field study of how householders interact with feedback from smart energy monitors', *Energy Policy* **38**(10), 6111–6119.
URL: *https://doi.org/10.1016/j.enpol.2010.05.068*

Harris, A., Rogers, M. M., Miller, C. J., McElmurry, S. P. & Wang, C. (2015), 'Residential emissions reductions through variable timing of electricity consumption', *Applied Energy* **158**, 484–489.
URL: *https://doi.org/10.1016/j.apenergy.2015.08.042*

Hastie, T., Tibshirani, R. & Friedman, J. (2009), *The Elements of Statistical Learning*, Vol. 27 of *Springer Series in Statistics*, Springer New York, New York, NY.
URL: *https://doi.org/10.1007/978-0-387-84858-7%0A*

Hoes, P., Hensen, J., Loomans, M., de Vries, B. & Bourgeois, D. (2009), 'User behavior in whole building simulation', *Energy and Buildings* **41**(3), 295–302.
URL: *https://doi.org/10.1016/j.enbuild.2008.09.008*

Howard, B. & Modi, V. (2017), 'Examination of the optimal operation of building scale combined heat and power systems under disparate climate and GHG emissions rates', *Applied Energy* **185**, 280–293.
URL: *https://doi.org/10.1016/j.apenergy.2016.09.108*

Howard, B., Saba, A., Gerrard, M. & Modi, V. (2014), 'Combined heat and power's potential to meet New York City's sustainability goals', *Energy Policy* **65**, 444–454.
URL: *https://doi.org/10.1016/j.enpol.2013.10.033*

Huebner, G. M., Hamilton, I., Chalabi, Z., Shipworth, D. & Oreszczyn, T. (2015), 'Explaining domestic energy consumption – The comparative contribution of building factors, socio-demographics, behaviours and attitudes', *Applied Energy* **159**, 589–600.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261915010983*

Huebner, G. M., Shipworth, D., Hamilton, I., Chalabi, Z. & Oreszczyn, T. (2016), 'Understanding electricity consumption: A comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes', *Applied Energy* **177**, 692–702.
URL: *https://doi.org/10.1016/j.apenergy.2016.04.075*

Hughes, T. P. (1983), *Networks of power: electrification in Western society, 1880-1930*, Johns Hopkins University Press, Baltimore, USA.
URL: *https://hdl.handle.net/2027/fulcrum.w6634365g*

IEA (2009), Gadgets and Gigawatts: Policies for Energy Efficient Electronics, Technical report, International Energy Agency.
URL: *https://www.iea.org/publications/freepublications/publication/gadgets-and-gigawatts-policies-for-energy-efficient-electronics.html*

IEA (2011), *Harnessing Variable Renewables*.
URL: *https://webstore.iea.org/harnessing-variable-renewables*

International Property Measurement Standards (2016), International Property Measurement Standards (IPMS): Residential Buildings, Technical report, International Property Measurement Standards, London, United Kingdom.
URL: *https://ipmsc.org/standards/residential/*

IPCC (2007), *Climate Change 2007: The Physical Science Basis*, Cambridge University Press, Cambridge, United Kingdom.
URL: *http://tinyurl.com/IPCCreport2007*

IPCC (2014), *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.
URL: *http://doi.org/10.1017/CBO9781107415416*

IPCC (2018), Global Warming of 1.5 °C - an IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate chan, Technical report, Intergovernmental Panel on Climate Change.
URL: *http://www.ipcc.ch/report/sr15/*

Ipsos-RSL & ONS (2003), United Kingdom *Time Use Survey, 2000* [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor], SN: 4504,.
URL: *https://doi.org/10.5255/UKDA-SN-4504-1*

Jones, R. V., Fuertes, A. & Lomas, K. J. (2015), 'The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic buildings', *Renewable and Sustainable Energy Reviews* **43**, 901–917.
URL: *https://doi.org/10.1016/j.rser.2014.11.084*

Jones, R. V. & Lomas, K. J. (2016), 'Determinants of high electrical energy demand in UK homes: Appliance ownership and use', *Energy and Buildings* **117**, 71–82.
URL: *https://doi.org/10.1016/j.enbuild.2016.02.020*

Kannan, R. & Strachan, N. (2009), 'Modelling the UK residential energy sector under long-term decarbonisation scenarios: Comparison between energy systems and sectoral modelling approaches', *Applied Energy* **86**(4), 416–428.
URL: *https://doi.org/10.1016/j.apenergy.2008.08.005*

Kelly, N., Samuel, A. & Hand, J. (2015), 'Testing integrated electric vehicle charging and domestic heating strategies for future UK housing', *Energy and Buildings* **105**, 377–392.
URL: *https://doi.org/10.1016/j.enbuild.2015.07.044*

Lampoudi, S., Gillespie, D. T. & Petzold, L. R. (2009), 'The multinomial simulation algorithm for discrete stochastic simulation of reaction-diffusion systems.', *The Journal of chemical physics* **130**(9), 94104.
URL:
http://scitation.aip.org/content/aip/journal/jcp/130/9/10.1063/1.3074302

Lee, Y. S. & Malkawi, A. M. (2014), 'Simulating multiple occupant behaviors in buildings: An agent-based modeling approach', *Energy and Buildings* **69**, 407–416.
URL: *https://doi.org/10.1016/j.enbuild.2013.11.020*

Leonardi, F. G. (2006), 'A generalization of the PST algorithm: modeling the sparse nature of protein sequences', *Bioinformatics* **22**(11), 1302–1307.
URL: *https://doi.org/10.1093/bioinformatics/btl088*

López-Rodríguez, M., Santiago, I., Trillo-Montero, D., Torriti, J. & Moreno-Munoz, A. (2013), 'Analysis and modeling of active occupancy of the residential sector in Spain: An indicator of residential electricity consumption', *Energy Policy* **62**, 742–751.
URL: *https://doi.org/10.1016/j.enpol.2013.07.095*

MacKay, D. J. C. (2003), *Information theory, inference, and learning algorithms*, Cambridge University Press.
URL: *https://tinyurl.com/MacKay2003*

Matsumoto, S. (2016), 'How do household characteristics affect appliance usage? Application of conditional demand analysis to Japanese household data', *Energy Policy* **94**, 214–223.
URL: *https://doi.org/10.1016/j.enpol.2016.03.048*

McKenna, E., Higginson, S., Grünewald, P. & Darby, S. J. (2017), 'Simulating residential demand response: Improving socio-technical assumptions in activity-based models of energy demand', *Energy Efficiency* pp. 1–15.
URL: *https://doi.org/10.1007/s12053-017-9525-4*

McKenna, E., Krawczynski, M. & Thomson, M. (2015), 'Four-state domestic building occupancy model for energy demand simulations', *Energy and Buildings* **96**, 30–39.
URL: *https://doi.org/10.1016/j.enbuild.2015.03.013*

McKenna, E. & Thomson, M. (2016), 'High-resolution stochastic integrated thermal–electrical domestic demand model', *Applied Energy* **165**, 445–461.
URL: *https://doi.org/10.1016/j.apenergy.2015.12.089*

Meyabadi, A. F. & Deihimi, M. H. (2017), 'A review of demand-side management: Reconsidering theoretical framework', *Renewable and Sustainable Energy Reviews* **80**(January 2016), 367–379.
URL: *https://doi.org/10.1016/j.rser.2017.05.207*

Miele, V., Bourguignon, P.-Y., Robelin, D., Nuel, G. & Richard, H. (2005), 'seq++: analyzing biological sequences with a range of Markov-related models', *Bioinformatics* **21**(11), 2783–2784.
URL: *https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bti389*

Moghaddam, M. P., Abdollahi, A. & Rashidinejad, M. (2011), 'Flexible demand response programs modeling in competitive electricity markets', *Applied Energy* **88**(9), 3257–3269.
URL: *https://doi.org/10.1016/j.apenergy.2011.02.039*

Moura, P. S. & de Almeida, A. T. (2010), 'The role of demand-side management in the grid integration of wind power', *Applied Energy* **87**(8), 2581–2588.
URL: *https://doi.org/10.1016/j.apenergy.2010.03.019*

Murray, D., Liao, J., Stankovic, L. & Stankovic, V. (2016), 'Understanding usage patterns of electric kettle and energy saving potential', *Applied Energy* **171**, 231–242.
URL: *http://www.sciencedirect.com/science/article/pii/S0306261916303579*

Navarro, A., Ochoa, L. F., Mancarella, P. & Randles, D. (2013), Impacts of photovoltaics on low voltage networks: a case study for the north west of England, *in* '22nd International Conference and Exhibition on Electricity Distribution (CIRED 2013)', Institution of Engineering and Technology, pp. 1419–1419.
URL: *https://doi.org/10.1049/cp.2013.1229*

O'Brien, W., Gunay, H. B., Tahmasebi, F. & Mahdavi, A. (2017), 'A preliminary study of representing the inter-occupant diversity in occupant modelling', *Journal of Building Performance Simulation* **10**(5-6), 509–526.
URL: *https://doi.org/10.1080/19401493.2016.1261943*

Ofgem (2013), Creating the right environment for demand-side response: next steps, Technical report, Ofgem.
URL: *https://www.ofgem.gov.uk/publications-and-updates/creating-right-environment-demand-side-response-next-steps*

Ofgem (2015), 'Typical Domestic Consumption Values for gas and electricity 2015'.
URL: *https://www.ofgem.gov.uk/gas/retail-market/monitoring-data-and-statistics/typical-domestic-consumption-values*

ONS (2003*a*), The United Kingdom 2000 Time Use Survey: Technical Report, Technical report, Office for National Statistics, London, United Kingdom.
URL: *http://www.ons.gov.uk/ons/rel/lifestyles/time-use/2000-edition/time-use-survey–technical-report.pdf*

ONS (2003*b*), The United Kingdom 2000 Time Use Survey: Technical Report, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://discover.ukdataservice.ac.uk/series/?sn=2000054*

ONS (2011), '2011 Census - Nomis - Official Labour Market Statistics [website]'.
URL: *https://www.nomisweb.co.uk/census/2011*

ONS (2012), Social Trends, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://data.gov.uk/dataset/social_trends*

ONS (2013), General Lifestyle Survey 2011, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://tinyurl.com/UKONS2013*

ONS (2014), Households and Household Composition in England and Wales: 2001-2011, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://tinyurl.com/UKONS2014*

ONS (2015), Families and Households: 2015, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://tinyurl.com/UKONS2015*

ONS (2016), Overview of the UK population, Technical report, UK Office for National Statistics.
URL: *https://tinyurl.com/UKpopulationOverview*

ONS (2017), Living costs and food survey 2016, Technical report, Office for National Statistics, London, United Kingdom.
URL: *https://tinyurl.com/UKONS2017*

Paatero, J. V. & Lund, P. D. (2006), 'A model for generating household electricity load profiles', *International Journal of Energy Research* **30**(5), 273–290.
URL: *https://doi.org/10.1002/er.1136*

Page, J., Robinson, D., Morel, N. & Scartezzini, J.-L. (2008), 'A generalised stochastic model for the simulation of occupant presence', *Energy and Buildings* **40**(2), 83–98.
URL: *https://doi.org/10.1016/j.enbuild.2007.01.018*

Palm, J. & Ellegård, K. (2011), 'Visualizing energy consumption activities as a tool for developing effective policy', *International Journal of Consumer Studies* **35**(2), 171–179.
URL: *https://doi.org/10.1111/j.1470-6431.2010.00974.x*

Palmer, J., Terry, N., Kane, T., Firth, S., Hughes, M., Pope, P., Young, J., Knight, D. & Godoy-Shimizu, D. (2013), Electrical appliances at home: tuning in to energy savings, Technical report, Cambridge Architectural Research — Element Energy — Loughborough University, Cambridge, United Kingdom.
URL:
https://www.gov.uk/government/publications/household-electricity-survey–2

Palmintier, B. & Webster, M. (2011), Impact of unit commitment constraints on generation expansion planning with renewables, *in* '2011 IEEE Power and Energy Society General Meeting', IEEE, pp. 1–7.
URL: *http://doi.org/10.1109/PES.2011.6038963*

Parliament of the United Kingdom (2004), 'Energy Act 2004'.
URL: *https://www.legislation.gov.uk/ukpga/2004/20/contents*

Parliament of the United Kingdom (2008), Climate Change Act 2008, Technical report, London, United Kingdom.
URL: *http://www.legislation.gov.uk/ukpga/2008/27/contents*

Peacock, A. D. & Owens, E. H. (2014), 'Assessing the potential of residential demand response systems to assist in the integration of local renewable energy generation', *Energy Efficiency* **7**(3), 547–558.
URL: *http://doi.org/10.1007/s12053-013-9236-4*

Pearce, R. (2019), 'Interactive Map of UK power sources - Carbon Brief'.
URL: *https://carbonbrief.github.io/uk-power-sources/#all*

Pfeiffer, A., Millar, R., Hepburn, C. & Beinhocker, E. (2016), 'The '2°C capital stock' for electricity generation: Committed cumulative carbon emissions from the electricity generation sector and the transition to a green economy', *Applied Energy* **179**(2016), 1395–1408.
URL: *https://doi.org/10.1016/j.apenergy.2016.02.093*

Pfenninger, S. (2017), 'Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability', *Applied Energy* **197**, 1–13.
URL: *https://doi.org/10.1016/j.apenergy.2017.03.051*

Pfenninger, S., Hawkes, A. & Keirstead, J. (2014), 'Energy systems modeling for twenty-first century energy challenges', *Renewable and Sustainable Energy Reviews* **33**, 74–86.
URL: *https://doi.org/10.1016/j.rser.2014.02.003*

Pfenninger, S., Hirth, L., Schlecht, I., Schmid, E., Wiese, F., Brown, T., Davis, C., Gidden, M., Heinrichs, H., Heuberger, C., Hilpert, S., Krien, U., Matke, C., Nebel, A., Morrison, R., Müller, B., Pleßmann, G., Reeg, M., Richstein, J. C., Shivakumar, A., Staffell, I., Tröndle, T. & Wingenbach, C. (2018), 'Opening the black box of energy modelling: Strategies and lessons learned', *Energy Strategy Reviews* **19**, 63–71.
URL: *https://doi.org/10.1016/j.esr.2017.12.002*

Piwko, R., Osborn, D., Gramlich, R., Jordan, G., Hawkins, D. & Porter, K. (2005), 'Wind Energy Delivery Issues - Transmission planning and competitive electricity market operation', *IEEE Power and Energy Magazine* **3**(6), 47–56.
URL: *https://doi.org/10.1109/MPAE.2005.1524620*

Poncelet, K., Delarue, E., Six, D., Duerinck, J. & D'haeseleer, W. (2016), 'Impact of the level of temporal and operational detail in energy-system planning models', *Applied Energy* **162**, 631–643.
URL: *https://doi.org/10.1016/j.apenergy.2015.10.100*

Powells, G., Bulkeley, H., Bell, S. & Judson, E. (2014), 'Peak electricity demand and the flexibility of everyday life', *Geoforum* **55**, 43–52.
URL: *https://doi.org/10.1016/j.geoforum.2014.04.014*

Ramírez-Mendiola, J. L. (2018), Stochastic Chain with Memory of Variable Length Activity Model [Access Permission upon request to the corresponding author].
URL: *https://github.com/JoseLuisRaMen/scvml-activity-model*

Ramírez-Mendiola, J. L., Grünewald, P. & Eyre, N. (2017), 'The diversity of residential electricity demand – A comparative analysis of metered and simulated data', *Energy and Buildings* **151**, 121–131.
URL: *https://doi.org/10.1016/j.enbuild.2017.06.006*

Ramírez-Mendiola, J. L., Grünewald, P. & Eyre, N. (2018), 'Linking intra-day variations in residential electricity demand loads to consumers' activities: What's missing?', *Energy and Buildings* **161**, 63–71.
URL: *https://doi.org/10.1016/j.enbuild.2017.12.012*

Ramírez-Mendiola, J. L., Grünewald, P. & Eyre, N. (2019), 'Residential activity pattern modelling through stochastic chains of variable memory length', *Applied Energy* **237**, 417–430.
URL: *https://doi.org/10.1016/j.apenergy.2019.01.019*

Reiss, P. C. & White, M. W. (2005), 'Household Electricity Demand, Revisited', *Review of Economic Studies* **72**(3), 853–883.
URL: *https://doi.org/10.1111/0034-6527.00354*

Rezaee Jordehi, A. (2016), 'Allocation of distributed generation units in electric power systems: A review', *Renewable and Sustainable Energy Reviews* **56**, 893–905.
URL: *https://doi.org/10.1016/j.rser.2015.11.086*

Richardson, I., Thomson, M. & Infield, D. (2008), 'A high-resolution domestic building occupancy model for energy demand simulations', *Energy and Buildings* **40**(8), 1560–1566.
URL: *https://doi.org/10.1016/j.enbuild.2008.02.006*

Richardson, I., Thomson, M., Infield, D. & Clifford, C. (2010), 'Domestic electricity use: A high-resolution energy demand model', *Energy and Buildings* **42**(10), 1878–1887.
URL: *https://doi.org/10.1016/j.enbuild.2010.05.023*

Richardson, I., Thomson, M., Infield, D. & Delahunty, A. (2009), 'Domestic lighting: A high-resolution energy demand model', *Energy and Buildings* **41**(7), 781–789.
URL: *https://doi.org/10.1016/j.enbuild.2009.02.010*

Ringkjøb, H.-K., Haugan, P. M. & Solbrekke, I. M. (2018), 'A review of modelling tools for energy and electricity systems with large shares of variable renewables', *Renewable and Sustainable Energy Reviews* **96**, 440–459.
URL: *https://doi.org/10.1016/j.rser.2018.08.002*

Rissanen, J. (1983), 'A Universal Data Compression System', *IEEE Transactions on Information Theory* **29**(5), 656–664.
URL: *http://doi.org/10.1109/TIT.1983.1056741*

Robinson, D., Wilke, U. & Haldi, F. (2011), Multi agent simulation of occupants' presence and behaviour, *in* 'Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association', pp. 2110–2117.
URL: *http://www.scopus.com/inward/record.url?eid=2-s2.0-84870228555&partnerID=tZOtx3y1*

Ron, D., Singer, Y. & Tishby, N. (1997), 'The power of amnesia: Learning probabilistic automata with variable memory length', *Machine Learning* **25**(2-3), 117–149.
URL: *http://link.springer.com/10.1007/BF00114008*

Santiago, I., Lopez-Rodriguez, M., Trillo-Montero, D., Torriti, J. & Moreno-Munoz, A. (2014), 'Activities related with electricity consumption in the Spanish residential sector: Variations between days of the week, Autonomous Communities and size of towns', *Energy and Buildings* **79**, 84–97.
URL: *https://doi.org/10.1016/j.enbuild.2014.04.055*

Schatzki, T. (2010), 'Materiality and Social Life', *Nature and Culture* **5**(2), 123–149.
URL: *https://doi.org/10.3167/nc.2010.050202*

Sekar, A., Williams, E. & Chen, R. (2016), 'Heterogeneity in time and energy use of watching television', *Energy Policy* **93**, 50–58.
URL: *https://doi.org/10.1016/j.enpol.2016.02.035*

Sense (2019).

Shimoda, Y., Taniguchi-Matsuoka, A., Inoue, T., Otsuki, M. & Yamaguchi, Y. (2017), 'Residential energy end-use model as evaluation tool for residential micro-generation', *Applied Thermal Engineering* **114**, 1433–1442.
URL: *https://doi.org/10.1016/j.applthermaleng.2016.09.066*

Shove, E. (2004), 'Efficiency and Consumption: Technology and Practice', *Energy & Environment* **15**(6), 1053–1065.
URL: *http://doi.org/10.1260/0958305043026555*

Shove, E., Pantzar, M. & Watson, M. (2012), *The dynamics of social practice: Everyday life and how it changes*, SAGE, London, United Kingdom.
URL: *https://doi.org/10.4135/9781446250655*

Shove, E. & Walker, G. (2014), 'What Is Energy For? Social Practice and Energy Demand', *Theory, Culture & Society* **31**(5), 41–58.
URL: *https://doi.org/10.1177/0263276414536746*

Solomon, B. D. & Krishna, K. (2011), 'The coming sustainable energy transition: History, strategies, and outlook', *Energy Policy* **39**(11), 7422–7431.
URL: *https://doi.org/10.1016/j.enpol.2011.09.009*

Stankovic, L., Stankovic, V., Liao, J. & Wilson, C. (2016), 'Measuring the energy intensity of domestic activities from smart meter data', *Applied Energy* **183**, 1565–1580.
URL: *https://doi.org/10.1016/j.apenergy.2016.09.087*

Strbac, G., Aunedi, M., Pudjianto, D., Djapic, P., Gammons, S. & Druce, R. (2012), Understanding the balancing challenge, Technical report, NERA Economic Consulting, London, United Kingdom.
URL: *http://www.nera.com/publications/archive/2012/understanding-the-balancing-challenge.html*

Strbac, G., Konstantelos, I., Aunedi, M., Pollitt, M. & Green, R. (2016), Delivering future-proof energy infrastructure, Technical report, National Infrastructure Commission, London, United Kingdom.
URL: *https://www.nic.org.uk/publications/delivering-future-proof-energy-infrastructure-goran-strbac-et-al/*

Strengers, Y. (2011), 'Beyond demand management: co-managing energy and water practices with Australian households', *Policy Studies* **32**(1), 35–58.
URL: *https://doi.org/10.1080/01442872.2010.526413*

Strengers, Y. (2012), 'Peak electricity demand and social practice theories: Reframing the role of change agents in the energy sector', *Energy Policy*

44, 226–234.
URL: *https://doi.org/10.1016/j.enpol.2012.01.046*

Swan, L. G. & Ugursal, V. I. (2009), 'Modeling of end-use energy consumption in the residential sector: A review of modeling techniques', *Renewable and Sustainable Energy Reviews* **13**(8), 1819–1835.
URL: *https://doi.org/10.1016/j.rser.2008.09.033*

Tielens, P. & Van Hertem, D. (2016), 'The relevance of inertia in power systems', *Renewable and Sustainable Energy Reviews* **55**, 999–1009.
URL: *https://doi.org/10.1016/j.rser.2015.11.016*

Torriti, J. (2014), 'A review of time use models of residential electricity demand', *Renewable and Sustainable Energy Reviews* **37**, 265–272.
URL: *https://doi.org/10.1016/j.rser.2014.05.034*

Torriti, J. (2015), *Peak Energy Demand and Demand Side Response*, Routledge.
URL: *https://doi.org/10.4324/9781315781099*

Torriti, J. (2017), 'Understanding the timing of energy demand through time use data: Time of the day dependence of social practices', *Energy Research & Social Science* **25**, 37–47.
URL: *https://doi.org/10.1016/j.erss.2016.12.004*

Torriti, J. & Grünewald, P. (2014), 'Demand Side Response: Patterns in Europe and Future Policy Perspectives under Capacity Mechanisms', *Economics of Energy & Environmental Policy* **3**, 69–88.
URL: *https://www.jstor.org/stable/26189265*

Torriti, J., Hanna, R., Anderson, B., Yeboah, G. & Druckman, A. (2015), 'Peak residential electricity demand and social practices: Deriving flexibility and greenhouse gas intensities from time use and locational data', *Indoor and Built Environment* **24**(7), 891–912.
URL: *https://doi.org/10.1177/1420326X15600776*

Torriti, J., Hassan, M. G. & Leach, M. (2010), 'Demand response experience in Europe: Policies, programmes and implementation', *Energy* **35**(4), 1575–1583.
URL: *https://doi.org/10.1016/j.energy.2009.05.021*

Torriti, J. & Santiago, I. (2016), 'Simultaneous activities in the household and residential electricity demand in Spain', *Time & Society* .
URL: *https://doi.org/10.1177/0961463X16656867*

UK National Infrastructure Commission (2016), Smart Power, Technical report, UK National Infrastructure Commission, London, United Kingdom.
URL: *https://www.nic.org.uk/publications/smart-power/*

UK Power Networks (2014), Development of new network design and operation practices, Technical report.
URL:
https://innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-2-projects/Low-Carbon-London-(LCL)/

UKERC Energy Data Centre, Electricity Association & Elexon Ltd. (2013), Electricity user load profiles by profile class, Technical report, Electricity Association, London, United Kingdom.
URL: *http://data.ukedc.rl.ac.uk/browse/edc/Electricity/LoadProfile/doc*

Unruh, G. C. (2000), 'Understanding carbon lock-in', *Energy Policy* **28**(12), 817–830.
URL: *https://doi.org/10.1016/S0301-4215(00)00070-7*

Vassileva, I., Wallin, F. & Dahlquist, E. (2012), 'Analytical comparison between electricity consumption and behavioral characteristics of Swedish households in rented apartments', *Applied Energy* **90**(1), 182–188.
URL: *http://doi.org/10.1016/j.apenergy.2011.05.031*

Verbong, G. P., Beemsterboer, S. & Sengers, F. (2013), 'Smart grids or smart users? Involving users in developing a low carbon electricity economy', *Energy Policy* **52**, 117–125.
URL: *https://doi.org/10.1016/j.enpol.2012.05.003*

Vogler-Finck, P. J. & Früh, W.-G. (2015), 'Evolution of primary frequency control requirements in Great Britain with increasing wind generation', *International Journal of Electrical Power & Energy Systems* **73**, 377–388.
URL: *https://doi.org/10.1016/j.ijepes.2015.04.012*

Wahba, G. (1975), 'Optimal Convergence Properties of Variable Knot, Kernel, and Orthogonal Series Methods for Density Estimation', *The Annals of Statistics* **3**(1), 15–29.
URL: *https://doi.org/10.1214/aos/1176342997*

Walker, G. (2014), 'The dynamics of energy demand: Change, rhythm and synchronicity', *Energy Research & Social Science* **1**, 49–55.
URL: *https://doi.org/10.1016/j.erss.2014.03.012*

Walker, G., Hunter, S., Devine-Wright, P., Evans, B. & Fay, H. (2007), 'Harnessing community energies: Explaining and evaluating community-based localism in renewable energy policy in the UK', *Global Environmental Politics* **7**(2), 64–82.
URL: *https://doi.org/10.1162/glep.2007.7.2.64*

Ward, J. & Darcy, S. (2014), The household electricity demand-side and participation in the GB electricity markets, Technical report, Sustainability First, London, United Kingdom.
URL: *http://www.sustainabilityfirst.org.uk/gbelec.html*

Weinberger, M., Rissanen, J. & Feder, M. (1995), 'A universal finite memory source', *IEEE Transactions on Information Theory* **41**(3), 643–652.
URL: *http://doi.org/10.1109/18.382011*

Welsch, M., Deane, P., Howells, M., Ó Gallachóir, B., Rogan, F., Bazilian, M. & Rogner, H.-H. (2014), 'Incorporating flexibility requirements into long-term energy system models – A case study on high levels of renewable electricity penetration in Ireland', *Applied Energy* **135**, 600–615.
URL: *https://doi.org/10.1016/j.apenergy.2014.08.072*

Welsch, M., Howells, M., Hesamzadeh, M. R., Ó Gallachóir, B., Deane, P., Strachan, N., Bazilian, M., Kammen, D. M., Jones, L., Strbac, G. & Rogner, H. (2015), 'Supporting security and adequacy in future energy systems: The need to enhance long-term energy system models to better treat issues related to variability', *International Journal of Energy Research* **39**(3), 377–396.
URL: *https://doi.org/10.1002/er.3250*

Widén, J., Lundh, M., Vassileva, I., Dahlquist, E., Ellegård, K. & Wäckelgård, E. (2009), 'Constructing load profiles for household electricity and hot water

from time-use data—Modelling approach and validation', *Energy and Buildings* **41**(7), 753–768.
URL: *https://doi.org/10.1016/j.enbuild.2009.02.013*

Widén, J., Molin, A. & Ellegård, K. (2012), 'Models of domestic occupancy, activities and energy use based on time-use data: deterministic and stochastic approaches with application to various building-related simulations', *Journal of Building Performance Simulation* **5**(1), 27–44.
URL: *https://doi.org/10.1080/19401493.2010.532569*

Widén, J. & Wäckelgård, E. (2010), 'A high-resolution stochastic model of domestic activity patterns and electricity demand', *Applied Energy* **87**(6), 1880–1892.
URL: *https://doi.org/10.1016/j.apenergy.2009.11.006*

Wiesmann, D., Lima Azevedo, I., Ferrão, P. & Fernández, J. E. (2011), 'Residential electricity consumption in Portugal: Findings from top-down and bottom-up models', *Energy Policy* **39**(5), 2772–2779.
URL: *https://doi.org/10.1016/j.enpol.2011.02.047*

Wilke, U., Haldi, F., Scartezzini, J.-L. & Robinson, D. (2013), 'A bottom-up stochastic model to predict building occupants' time-dependent activities', *Building and Environment* **60**, 254–264.
URL: *http://doi.org/10.1016/j.buildenv.2012.10.021*

Willems, F., Shtarkov, Y. & Tjalkens, T. (1996), 'Context weighting for general finite-context sources', *IEEE Transactions on Information Theory* **42**(5), 1514–1520.
URL: *http://doi.org/10.1109/18.532891*

Wyatt, P. (2013), 'A dwelling-level investigation into the physical and socio-economic drivers of domestic energy consumption in England', *Energy Policy* **60**, 540–549.
URL: *https://doi.org/10.1016/j.enpol.2013.05.037*

Yamaguchi, Y. & Shimoda, Y. (2017), 'A stochastic model to predict occupants' activities at home for community-/urban-scale energy demand modelling', *Journal of Building Performance Simulation* **10**(5-6), 565–581.
URL: *https://doi.org/10.1080/19401493.2017.1336255*

Yamaguchi, Y., Yilmaz, S., Prakash, N., Firth, S. K. & Shimoda, Y. (2018), 'A cross analysis of existing methods for modelling household appliance use', *Journal of Building Performance Simulation* .
URL: *https://doi.org/10.1080/19401493.2018.1497087*

Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F. & Mahdavi, A. (2015), 'Occupant behavior modeling for building performance simulation: Current state and future challenges', *Energy and Buildings* **107**, 264–278.
URL: *https://doi.org/10.1016/j.enbuild.2015.08.032*

Yao, R. & Steemers, K. (2005), 'A method of formulating energy load profile for domestic buildings in the UK', *Energy and Buildings* **37**(6), 663–671.
URL: *https://doi.org/10.1016/j.enbuild.2004.09.007*

Yin, Y., Mizokami, S. & Aikawa, K. (2015), 'Compact development and energy consumption: Scenario analysis of urban structures based on behavior simulation', *Applied Energy* **159**, 449–457.
URL: *https://doi.org/10.1016/j.apenergy.2015.09.005*

Yohanis, Y. G., Mondol, J. D., Wright, A. & Norton, B. (2008), 'Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use', *Energy and Buildings* **40**(6), 1053–1059.
URL: *https://doi.org/10.1016/j.enbuild.2007.09.001*

Zheng, M., Meinrenken, C. J. & Lackner, K. S. (2014), 'Agent-based model for electricity consumption and storage to evaluate economic viability of tariff arbitrage for residential sector demand response', *Applied Energy* **126**, 297–306.
URL: *https://doi.org/10.1016/j.apenergy.2014.04.022*

Zimmermann, J.-P., Evans, M., Griggs, J., King, N., Harding, L., Roberts, P. & Evans, C. (2012), Household Electricity Survey: A study of domestic electrical produc usage - Final report, Technical report, Intertek, London, United Kingdom.
URL: *https://tinyurl.com/Zimmermann2012*